Journal of Health Economics 104 (2025) 103050

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/jhe

Impact of primary care market mergers on quality: Evidence from the English NHS[☆]

Yuan Lyu [a,b,c,*], Zhaocheng Zhang [d]

[a] School of Applied Economics, Renmin University of China, Beijing, China
[b] Cambridge Judge Business School, University of Cambridge, Cambridge, CB2 1AG, UK
[c] Institute for Global Health and Development, Peking University, Beijing, China
[d] Faculty of Economics, University of Cambridge, Cambridge, CB3 9DD, UK

ARTICLE INFO

Keywords:
Primary care
Mergers
Quality
Staggered roll-out design

ABSTRACT

The primary care market has experienced a growing trend of provider consolidation through mergers and acquisitions, yet the implications of this concentration remain unclear. 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. Examining all provider mergers from 2014 to 2018, we find that mergers improve certain aspects of clinical quality management, but they do not translate into broader population-level clinical quality gains, and patient satisfaction declines significantly. Importantly, the effects vary by merger motivation and the size of the merging parties, rather than their geographic proximity. Survival-driven mergers help sustain care quality and patient access, whereas efficiency-driven mergers lead to greater quality deterioration. Mergers between larger practices also lead to more negative outcomes than those involving smaller practices. In contrast, we find no significant difference between within-market and cross-market mergers. An exploration of the mechanism reveals that changes in market concentration do not explain the observed quality outcomes. Instead, shifts in workforce composition, driven by the underlying merger motivations, play a key role.

1. Introduction

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). Importantly, there has been a persistent trend of provider concentration through mergers and acquisitions 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] Yet, from a patient perspective, whether this trend offers advantages remains largely unexplored, mainly due to a lack of data. We address this gap by providing, to our knowledge, the first empirical

☆ We thank the editor and two anonymous referees for providing valuable comments on the paper. We are grateful for helpful suggestions from Christos Genakos, Paul Kattuman, Stefan Scholtes, Nicolas de Roos, Rita Santos, Rhys Williams, Wenchuan Zhao, Eleonora Brandimarti, and seminar participants at the RES Annual Conference, EEA-ESEM Conference, LBS TADC Conference, University of York, Cambridge Judge Business School, and University of Cambridge Applied Micro Workshop. All errors are our own.

* Corresponding author at: School of Applied Economics, Renmin University of China, Beijing, China.
E-mail address: y.lyu@ruc.edu.cn (Y. Lyu).

[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).

https://doi.org/10.1016/j.jhealeco.2025.103050
Received 30 July 2024; Received in revised form 9 April 2025; Accepted 31 July 2025
Available online 27 August 2025
0167-6296/

evidence on the effects of provider mergers on quality in the primary care market. Understanding the merger impact in this market is particularly important since much of the concentration in the primary care market remains unnoticed by regulators (Gravelle et al., 2019).

We study how mergers affect quality in the English primary care market. To do so, we assemble the first comprehensive dataset documenting the universe of English practice mergers between 2014 and 2018. We focus on the quality impact as primary care is provided free at the point of use in England, making quality a significant aspect. The English primary care market provides rich data on various quality measures, allowing us to offer new insights into the multiple dimensions of quality affected by mergers.

In theory, the effect of general practice mergers on quality is ambiguous. On one hand, mergers could enhance quality by yielding economies of scale and scope. Merged practices could realize economies of scale by employing both physician and non-physician staff to implement quality improvement processes and by using information technology to support these initiatives (Mehrotra et al., 2006). Also, merged practices can leverage physicians' specialized skills to achieve economies of scope and improve clinical quality (Casalino, 2006). These factors suggest potential quality improvements. However, mergers might also reduce the incentive to provide high-quality care by increasing market concentration and undermining competition. Therefore, the direction and magnitude of the merger effect on quality remain empirical questions.

To answer this, we use a Difference-in-Differences (DiD) strategy. Mergers occur in different years for different merged practices. We address this staggered roll-out design by using a *stacked DiD regression* approach following Deshpande and Li (2019) and Cengiz et al. (2019). To mitigate endogeneity concerns, we include a comprehensive set of time-varying practice- and local-level covariates in the regression. Practice and time fixed effects are also included to account for unobserved time-invariant differences across general practices as well as common time shocks. The control group is selected using the propensity score matching (PSM) method from the pool of never-merged practices. We show that merged practices exhibit similar trends to our chosen control group, supporting the parallel trends assumption. For robustness, we also use not-yet-merged practices as an alternative comparison group, as the timing of mergers is potentially random. We conduct numerous robustness checks to ensure the effectiveness of our estimates.

We find that mergers improve certain aspects of clinical quality management but do not translate into broader population-level clinical quality gains. In fact, we observe a decline in clinical quality assessed across various chronic illnesses over the long run. At the same time, patient satisfaction declines dramatically. Overall patient satisfaction decreases by approximately three percentage points, which equates to about a 4% decline for the average practice. Furthermore, we find suggestive evidence of potential financial gains from the mergers, implying that mergers may lead to financial benefits but at the expense of quality.

However, not all mergers have the same impact. First, merger motivation matters. Survival-driven mergers, which help struggling practices remain open, tend to preserve access to care and care quality, whereas efficiency-driven mergers often result in greater quality decline. Second, mergers between larger practices tend to produce more negative outcomes than those involving smaller practices. Third, we find no significant difference between within-market mergers, where practices in the same geographical market merge, and cross-market mergers, where practices across different markets merge. An exploration of the mechanism shows that changes in market concentration do not drive the quality outcomes after mergers. Instead, differences in merger motivation play a crucial role. Survival-driven mergers help maintain care by preventing closures, while efficiency-driven mergers, particularly those involving already overburdened practices, increase doctor workload without adequate staffing adjustments, leading to a sharper decline in quality.

This paper contributes to several areas of literature. First, it relates to literature evaluating the effects of mergers and acquisitions on non-price outcomes, both in health care and in broader contexts. Existing empirical industrial organization 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 modeling approaches to simulate the effect of hypothetical mergers and quantify the welfare effects of mergers.[2]

As health care markets have become increasingly concentrated through mergers and acquisitions (Gaynor et al., 2015), a growing body of research has examined the impact of mergers in the healthcare sector. However, most empirical studies focus on price effects,[3] while research on the quality impact remains limited.[4] The small number of studies that do examine quality largely focus on hospitals and specialized care providers, with findings that are mixed and context-dependent. For instance, Ho and Hamilton (2000) and Capps (2005) find no significant effect of hospital mergers on most quality indicators in the U.S. Gaynor et al. (2012) report only limited evidence of quality improvements following hospital mergers in the English NHS. Romano and Balan (2011) analyze a consummated hospital merger in the Chicago suburbs and find mixed effects, with some quality indicators improve, while others worsen or remain unchanged. Beyond hospitals, studies of specialist care mergers also yield mixed results. For instance, in the U.S. dialysis market, Eliason et al. (2020) find that acquisitions of independent facilities by large chains lead to worse patient outcomes, while Cutler et al. (2017) find no significant impact of concentration due to mergers on quality. Despite the growing trend of mergers in the primary care market, no previous studies, to our knowledge, have examined their impact, mainly due to the challenge of identifying merger events in this market. This paper fills this gap by compiling a comprehensive dataset of merger

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

3 For example, studies of U.S. hospital mergers find evidence of price increases for insurers after mergers (e.g., Dafny, 2009; Gowrisankaran et al., 2015; Dafny et al., 2019).

4 See Gaynor et al. (2015) for a comprehensive literature review on market concentration and its effects in the healthcare market.

events in the English primary care market and presenting the first empirical evidence on the impact of provider mergers in this context. We take a retrospective approach and evaluate the outcomes of actual mergers using a reduced-form method.

Furthermore, this paper adds to the literature examining the impact of market competition on quality in the primary care market. Several empirical studies have explored whether increasing competition improves quality, but findings remain mixed.[5] Some studies suggest that general practitioners (GPs) facing competitive pressure become more responsive to patient preferences. For instance, assuming that patients always prefer a sick note irrespective of their illness severity, Brekke et al. (2019) find that GPs issue more sick notes under greater competition. Similarly, Schaumans (2015) shows that Belgian GPs prescribe more medication when facing more competition. However, these outcomes do not directly reflect clinical quality (Gravelle et al., 2019). Other studies examine direct quality indicators but reach different conclusions. For example, Dietrichson et al. (2020) study both clinical quality and patient satisfaction measures and find that competition has no significant impact on quality in the Swedish primary care market. In England, Pike (2010) reports a positive correlation between competition and quality measured by both avoidable hospitalization rates and patient satisfaction scores. However, their study is based on a cross-sectional design and thus raises endogeneity concerns. A more rigorous study by Gravelle et al. (2019) address this by using a stronger econometric specification and eight years of panel data from over 8000 English general practices. They measure competition as the number of rival GPs within a small radius and find that increased local competition is associated with higher patient satisfaction and, to a lesser extent, improvements in clinical quality. While promoting competition has historically been viewed as a tool to improve quality, this has occurred alongside a broader trend of provider consolidation through mergers in the general practice market (Siciliani et al., 2017; Gravelle et al., 2019). Given this ongoing shift, further research is needed to explicitly assess the impact of consolidation on quality.[6] Our paper contributes by directly analyzing the effects of mergers on quality in the general practice market. By providing empirical evidence on primary care mergers, we inform ongoing policy discussions on whether the long-term trend toward consolidation in the primary care market should be encouraged or regulated.

While no empirical studies directly examine the impact of mergers on quality in the primary care market, one closely related study is Gravelle et al. (2022).[7] Using data on all English general practices from 2005 to 2016, they study the relationship between the size of general practice and a broad set of quality indicators. Their findings are mixed. While larger practices show no significant association with some clinical quality measures, they are associated with lower patient satisfaction. Based on these findings, they conclude that simply encouraging practices to form larger groups may not improve quality outcomes. While their study provides valuable insights, it does not directly assess the impact of mergers. Although one might infer that if mergers increase practice size, they could lead to similar outcomes, this interpretation does not capture the full range of potential effects associated with mergers. Our paper differs in several key respects. First, we study actual merger events by analyzing the universe of general practice mergers in England from 2014 to 2018. Unlike their findings, we find that mergers improve certain aspects of clinical quality management, rather than unchanged as their study shows. Second, we document significant heterogeneity in merger impacts depending on the type of merger, a dimension that cannot be inferred from the analysis of practice size alone. Third, while Gravelle et al. (2022) do not investigate why practice size might influence quality, we go one step further by examining the channels driving merger effects. We find that workforce composition changes, driven by different merger motivations, play a key role in explaining post-merger quality changes. By explicitly discussing the mechanisms, we hope to offer a deeper understanding of the merger impact in the primary care market.

The rest of this paper proceeds as follows. Section 2 provides the institutional background. Section 3 describes the data. Section 4 outlines the research design. Section 5 presents and discusses the results. Section 6 tests the sensitivity of the findings. Section 7 concludes.

2. Institutional background

2.1. Primary care in England

In England, primary care is provided by general practices. All residents in England are entitled to freely choose and register with one general practice. For most people, this registered practice is the first and most frequently used point of contact for physical or mental health concerns. Once registered, patients are assigned a general practitioner (GP) who becomes their primary GP. Primary care services are free at the point of use, making quality an essential factor.

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 rather than with individual GP partners. The practice receives most of its income from the NHS through these contracts.[8] There are four primary sources of

[5] Most research on healthcare market competition focuses on hospitals, while the general practice market is much less researched. See Gaynor and Town (2011) for a review.

[6] While market competition and mergers are related, they are not identical mechanisms. Competition is one possible channel through which mergers may affect quality, but not all mergers significantly change competitive dynamics. Some small mergers between nearby practices may have minimal impact on local market competition but still influence quality through other channels. Therefore, it is important to assess merger impacts directly rather than inferring their effects from competition studies alone.

[7] Another related paper is Kelly and Stoye (2014), who find that larger practices are associated with better quality of care in the English general practice market. However, their fixed-effects model controls only at the Primary Care Trust (PCT) level, making it difficult to make a causal claim.

[8] Practices may also generate supplementary revenue by providing certain services, such as issuing private prescriptions or medical certificates.

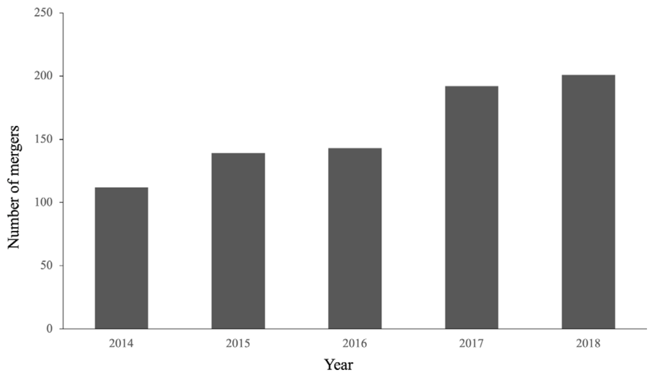

Fig. 1. Number of GP practice mergers in England.
Notes: This figure displays the annual count of GP practice mergers in England. Each bar represents the number of merger events occurring in a given year. A merger event refers to the amalgamation of an acquiring practice and a target practice. The data used in this figure is collected and calculated by the author.

income payments: a global sum payment, which is an annual amount calculated for each practice based mainly on the number of registered patients; quality incentive rewards from the Quality and Outcomes Framework (QOF) (detailed further in Section 3); payment schemes for providing a range of enhanced services, such as vaccination programs; and payments for specific purposes, such as dispensing services for authorized practices.[9] General practices are reimbursed for their premises costs but must fund all other expenses, such as hiring practice nurses and clerical staff, from their revenue (Santos et al., 2017).

2.2. Why GP practices merge?

There are three main reasons behind practice mergers.[10] 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. We 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.[11] Failing practices with poor quality ratings may be motivated to merge to improve their ratings and ensure their survival. We try to differentiate between these motivations and explore the heterogeneous effects in Section 5.3.3.

3. Data

3.1. GP practice merger data

To our knowledge, no existing dataset documents all instances of general practice mergers in England. Therefore, we manually assemble this information. We collect data on English practice mergers annually from 2014 to 2018. We choose these five years because 2014 is the earliest year for which pre-merger outcomes are available, and we want to focus on the period without disruptions of the pandemic. We compile a list of potential mergers and manually verify each one. Whenever possible, we identify the motivation behind each merger. Detailed procedures are explained in Appendix A.

We define a merger event as the amalgamation of two practices: the target practice and the acquirer practice. The acquirer practice serves as the main site following the merger, while the target practice becomes a branch site. If an acquirer takes over multiple targets in a year, each pairwise combination is treated as a separate event.[12] In total, we document 787 practice mergers during the study period. Fig. 1 plots the number of mergers by year. There is a noticeable increasing trend in the frequency of mergers, indicating that practice mergers have become more prevalent over time.

[9] Practices authorized to provide dispensing services receive payments to cover the costs of drugs and appliances, as well as a dispensing fee per item dispensed.
[10] This follows discussions with insider experts.
[11] For more information, see the guidelines published by CQC.
[12] We observe only 89 cases (less than 15% of the total sample) where multiple surgeries merge into one in the same year. We perform a robustness check excluding these cases, and our findings remain robust. Results are available upon request.

3.2. Outcome variables

Quality is multi-dimensional. To capture its various aspects, we examine both objective clinical quality measures and subjective patient experience measures. Below, we describe our outcome variables and provide a summary in Table C1.

3.2.1. Clinical quality measures

We measure clinical quality using the official QOF data.[13] The QOF is a voluntary scheme introduced in 2004 that financially incentivizes practices to meet quality targets for their registered patients. Initially, the QOF included four domains: clinical, organizational, patient experience, and additional services,[14] where for each domain, there are indicators to assess the performance of general practices. 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.[15] 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. Although participation in the QOF is voluntary, participation rates exceed 95%.

QOF targets high-priority disease areas where primary care plays a principal role and there is evidence that improved primary care can yield significant health benefits.[16] We use QOF data to measure the clinical quality of each practice. Following Gravelle et al. (2019) and Gravelle et al. (2022), we 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. Only about two-thirds of the points are related to clinical quality indicators for specific health conditions, while the rest are for general tasks like record-keeping (Gravelle et al., 2019). Additionally, there might be gaming of exception reporting to earn points. To illustrate, achievement 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 can exclude certain patients from individual clinical indicators (exception reporting) for reasons such as clinical inappropriateness or patient dissent. To address this, we construct a second clinical measure, *ClinPA*, using QOF data, focusing only on clinical indicators and practice population achievement, measured as $100 \times A/T$. We calculate *ClinPA* as the weighted average of these performance measures over consistent clinical indicators between 2013 and 2019. The weights are the maximum points available for the indicators. There are 20 consistent clinical indicators over these six years. Some indicators are excluded when calculating *ClinPA* because the QOF underwent a major change in 2019, with several indicators redesigned.[17] Therefore, we 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*.

3.2.2. Patient experience measures

We 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).[18]

We focus on 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). Additionally, the survey includes detailed questions about different aspects of patient experiences, such as the likelihood of seeing their main GPs, appointment booking experiences, waiting times, and satisfaction with practice opening hours. These aspects allow us to explore which areas of patient experience are affected (refer to Table C1 for detailed variable definitions).

3.2.3. Financial performance measures

While our focus is on the quality effects, we also present suggestive evidence on how mergers affect the financial performance of the practice. Ideally, we would assess financial performance using profit data, but such information is confidential to the practice and unavailable. We thus use the amount of NHS payments to each practice as a proxy. These payments represent the bulk of the practice's income and thus provide information on revenue.

Using revenue data obtained from NHS Digital, we construct two measures: *RevPerPatient*, which denotes the revenue per patient in logarithm form, and *RevPerGP*, which represents revenue per full-time equivalent (FTE) of GPs in logarithm form.

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

[14] From 2014 onwards, the QOF has been revised to three domains: clinical, public health, and public health for additional services. While the domains have changed, most indicators remain, with some reorganized.

[15] For example, a blood pressure control indicator for patients with hypertension sets a threshold of a last-time reading of less than 150/90 mmHg for at least 45% of patients. Performance is assessed by the percentage of patients meeting this target.

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

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

[18] During our sample years, the survey was administered twice a year for 2014–2016, with results published in January and July, and once a year for 2017–2019, with results published in July. For our main analysis, we follow Gravelle et al. (2022) and use the survey results published in July for 2014–2016.

3.3. Additional practice-level and local-level data

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

For local area characteristics, we trace each practice's location to the Lower Layer Super Output Area (LSOA) level and assign the corresponding LSOA-level characteristics to it. We use the Index of Multiple Deprivation (IMD) to account for the socioeconomic status of each area. The IMD data is publicly available from the Office for National Statistics. A higher IMD rank implies a less deprived status. We also account for the rural or urban classification of each LSOA using data from the Office for National Statistics. Table C1 shows the full list of practice and local characteristics used in our analysis.

3.4. Summary statistics

Our final dataset consists of all general practices in England from 2013 to 2019. This includes 8039 practices, with 52,164 practice-year observations.[19] Table 1 presents the mean values of the variables used in our analysis, for the full sample (Column (1)) and separately for merged and never-merged practices (Columns (2) and (3)). For merged practices, we consider observations from the year before the merger for both the acquirer and target practices. Column (4) provides the t-statistic comparing the difference in means between the two subsamples in Columns (2) and (3). We find that, prior to merging, 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 administrative staff. Additionally, mergers are more likely to occur in more deprived areas. These observations motivate us to use PSM to select a comparable group of practices for comparison with the merged practices. We explain our methodology next.

4. Empirical methodology

We study the effect of general practice mergers by comparing practices that experience mergers with those that do not. As there is only one merged entity after the merger, to ensure consistency, we construct pseudo-merged values for the pre-merger period by combining data from both merging parties, weighted by the number of patients from each practice.[20]

4.1. Stacked difference-in-difference regression method

We have the issue of variation in treatment timing as mergers occur in different years. Recent econometrics literature shows that the standard two-way fixed effects (TWFE) DiD estimators fail to produce valid estimations in such settings (see, e.g., Callaway and Sant'Anna 2021, Sun and Abraham 2021, Goodman-Bacon 2021, Borusyak et al. 2021).[21] To address this issue, we adopt a stacked DiD regression approach following Cengiz et al. (2019) and Deshpande and Li (2019) and use the method proposed by Callaway and Sant'Anna (2021) as a robustness check.

Specifically, we proceed as follows. First, we categorize all merged units into five cohorts based on the year of the merger, from 2014 to 2018. For each cohort, we then use PSM to construct a comparison group from the set of never-merged practices (details in Section 4.2). This results in five separate datasets/groups, each containing a treatment group (merged entities formed in a specific year) and a corresponding comparison group (never-merged practices selected through matching). Finally, we stack these five datasets and perform a standard TWFE DiD regression on the stacked dataset, with the dataset-specific unit and time fixed effects.[22]

The main regression equation is as follows:

$$y_{igt} = \gamma_{ig} + \gamma_{tg} + \beta(Treat_{ig} \times Post_{gt}) + X_{igt}\delta + \varepsilon_{igt} \tag{1}$$

where y_{igt} denotes the outcome for merged unit i from group g in year t. $Treat_{ig}$ is an indicator equal to 1 for merged units of group g. $Post_{gt}$ is an indicator equal to 1 for post-merger years, specified separately for each of the five datasets. γ_{ig} represent group-specific unit fixed effects and γ_{tg} are group-specific year fixed effects. X_{igt} are a set of time-varying practice-related covariates that we control

[19] Note that online GP practices are excluded from our sample. Also, our data is not a balanced panel. Results using only balanced panel remain robust and are available upon request.

[20] Gaynor et al. (2012) also construct pseudo-merged values prior to the merger to study hospital mergers in England. Unlike their method, which constructs pseudo-merged values only for outcome variables and performs PSM separately for each merging party, we construct pseudo-merged values for both outcome and control variables before the merger and perform PSM for the single merged entity. Our results remain robust when conducting separate matching for the target and acquirer practices. Results are available upon request.

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

[22] This approach can be applied using either a static or a dynamic specification. The estimates of dynamic treatment effects in Section 5.2 follow this approach.

Table 1
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, we 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.

for.[23] The year of merger is dropped altogether for treated practices to prevent measurement errors. Standard errors are clustered at the group-specific unit level. β is the key coefficient of interest that captures the impact of mergers.

To examine the persistence of merger effects, we also estimate the dynamic specification of the stacked DiD regression:

$$y_{igt} = \gamma_{ig} + \gamma_{tg} + \sum_{\tau=-4}^{5,\tau\neq 0} \delta_\tau (Treat_{ig} \times D_{ig}^{\tau}) + X_{igt}\alpha + \varepsilon_{igt} \tag{2}$$

where D_{ig}^{τ} are indicators equal to 1 for practice i of group g that are τ years after or before the merger year. $\tau = -5$ is left out as the reference year.[24] As in the main specification, the year of merger is dropped for treated practices to prevent measurement errors.[25] δ_τ captures the difference in outcomes between merged practices and never-merged controls, for τ years after (or before) the merger versus five years before the merger. These estimates shed light on how merger effects evolve over time.

To make a causal claim, the main identification challenge is that practice mergers may not occur randomly. We use a rich set of covariates in the regression to control for time-varying market-level and practice-level characteristics. Practice and time fixed effects adjust for all unobserved permanent differences across practices as well as common time shocks. Our selection of a comparison group using PSM ensures that the control group closely resembles the treated units. We show later that pre-merger estimates from the event

[23] At the practice level, we control for the competitive environment, the total number of registered patients, a set of prevalence rates, the workforce composition (including FTE of GPs, nurses, and admins), and dispensing status. At the local level, we incorporate the local IMD rank and a dummy variable indicating whether the area is urban or rural.

[24] We choose $\tau = -5$ as the reference year to allow for possible anticipation effects. The main findings remain consistent when $\tau = -1$ or $\tau = -2$ is used as the reference year.

[25] In Fig. D1, we plot estimates from an alternative specification where we do not exclude $\tau = 0$. The main findings remain robust.

study specification Eq. (2) are insignificant, suggesting the parallel trends assumption holds. We also perform a robustness check using not-yet-merged practices as an alternative control group, as the timing of mergers is potentially random.

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. We 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 estimated by logit or probit regression on matching variables that determine the treatment assignment.

We use practice-level and local-level characteristics as matching variables. At the practice level, we consider the prevalence of nine disease conditions and the number of registered patients to account for the demand for each practice. We also include the FTE of practice staff to capture the workforce of each practice. The dispensing status of each practice is also included. At the local level, we use the IMD rank and an urban indicator as proxies for the socioeconomic status of the local market. We also include the number of competing GP surgeries within a 2 km radius to account for the competitive environment.

As shown in Table C2, the pool of merged practices changes over the years, and variables affecting the likelihood of mergers differ across years. Therefore, we perform matching separately for each cohort of treated practices. We match with replacement, allowing never-merged practices to be resampled every year. We estimate the propensity score using the pre-treatment values of the matching covariates, and we select the three closest matches as the comparison group for each treated practice.[26,27]

A suitable matching procedure should balance the distribution of the matching variables between the treatment and control groups. We assess the matching quality by evaluating the standardized bias between the matched and unmatched samples and performing a two-sample t-test. Standardized bias is a commonly used indicator proposed by Rosenbaum and Rubin (1985).[28] A standardized bias greater than 20 is typically considered large (see, for instance, Gaynor et al., 2012). Moreover, we perform a two-sample t-test on the sample means between the merged and matched never-merged practices. If matching is suitable, there should be no significant differences between the covariate means of the two groups.

Results of the balance tests are presented in Table C3. For each covariate, we perform the tests twice: once using the raw, unmatched sample, and once with the matched sample. Without matching, significant differences exist between the treated and unmatched controls in variables measuring patient numbers, staff FTEs, practice dispensing status, and local market conditions. These differences are statistically significant at the 5% level. However, after matching, the differences become insignificant, suggesting that our matching procedure effectively identifies a comparable control group. We show later that treated practices evolve similarly to our selected control practices, supporting the parallel trends assumption.

5. Results

5.1. Effect of GP practice mergers

Table 2 presents the estimates from Eq. (1).[29,30] Column (1) shows that merged practices experience an improvement in QOF performance post-merger, with QOF points increase by around 1 percentage point.[31] However, the positive impact in QOF performance diminishes when assessing the practice population achievement in clinical quality indicators, as shown in Columns (2) and (3). These results suggest that while mergers may enhance overall GP practice management in achieving QOF targets, they do not lead to meaningful improvements in clinical quality for the patient population. Column (4) shows that patient satisfaction decreases by around 3 percentage points post-merger.[32] This decline equates to about a 4% decrease on an average satisfaction

[26] In the robustness checks (see Section 6), we vary the number of matches, selecting one, five, and seven closest matches as the control group, and our findings remain robust.

[27] Contamination of spillover effects may occur if matched practices are in the same local market as the treated practice. However, we find that only two matches are in the same market, so this concern is minor. Nevertheless, we perform a robustness check using matches only from outside-market practices, and our main findings hold. Results are available upon request.

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

[29] Table C4 shows baseline results from Eq. (1) without additional time-varying covariates, and the results are robust.

[30] The dependent variables in Columns (1) to (5) in Table 2 are fractional in nature, meaning they take values strictly between 0 and 1 (or, when expressed as percentages, between 0 and 100). This raises a potential issue for the standard linear regression model as described in Eq. (1), as they assume an unbounded dependent variable. More importantly, when a large share of observations clusters near the boundaries (e.g., close to one, as seen for *qofOutcome* in Table 1), linear models may not fully capture the underlying distribution of the data. To address this concern, we follow Papke and Wooldridge (2008) and apply the fractional response regression to Eq. (1). The model details and estimation results are discussed in Appendix B. Comparing the results in Table B1 with those in Table 2, we find that our linear specification in Eq. (1) remains suitable. It produces consistent signs and significance levels for the treatment effects, with magnitudes closely matching those from the fractional response model. Given this robustness, we proceed with the linear specification in Eq. (1) throughout the paper.

[31] This increase in QOF performance may be due to GP practices selecting and treating healthier patients after the merger. To investigate this possibility, we use the prevalence rates of long-term illnesses as the dependent variables in the regression. Results in Table C5 indicate that the patient pool, in terms of long-term illness prevalence, remains stable after the merger. This implies that the positive impact on QOF performance does not result from merged practices selecting healthier patients.

[32] For the main analysis, we use survey data from the year the survey was taken. However, acknowledging that merged practices may need time to adapt, we also lag the survey outcomes by one year and re-estimate the regressions. The results remain robust and are available upon request.

Table 2
Effect of GP practice mergers.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	qofOutcome	ClinPA	ClinPA2	OverallSat	Recommend	RevPerPatient	RevPerGP
Treat × Post	0.865***	0.054	−0.240	−2.725***	−2.730***	0.026	0.215***
	(0.310)	(0.161)	(0.288)	(0.414)	(0.631)	(0.016)	(0.028)
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	15,536	18,144	15,536	18,144	18,144

Notes: Coefficients on the additional time-varying covariates are not shown to save space. These covariates include the following: at the practice level, the number of competitors within a 2 km 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 effects. The year of the merger is excluded for merged practices. The sample size is calculated as: 6 years (2013–2019, excluding the merger year) × (1 treated practice + 3 control practices) × 787 mergers, resulting in approximately 18,888 observations. However, the actual number is lower due to two reasons: (1) not all practices have a full set of observations for every year in the sample, and (2) matching with replacement allows the same control practice to be used for multiple treated practices. 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. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

rate of around 85%. Similarly, Column (5) shows 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 for an average practice. These results indicate that merged practices perform worse from a patient perspective. To explore possible reasons for patients' disappointment, we analyze additional survey questions that provide more detailed insights into patient experiences. Results in Table C6 show that patients are less likely to see their main GPs, are less satisfied with practice opening hours, report longer waiting times and have worse experience accessing care.

Our estimates on patient satisfaction are broadly consistent with Gravelle et al. (2022), who estimate that if two average-sized practices merged, patient recommendation rates would decline by 2.89%. However, their calculation assumes mergers between similarly sized practices, whereas in reality, mergers vary in scale, leading to heterogeneous effects as we demonstrate in Section 5.3. In terms of clinical quality indicators, our findings diverge somewhat. Gravelle et al. (2022) find that larger practices do not significantly improve QOF performance and are associated with worse clinical quality at the practice population level, while we find that mergers lead to an increase in QOF performance but no significant change in clinical quality at the population level. These differences suggest that mergers may affect quality through mechanisms not captured by practice size alone. We will investigate the channels in Section 5.4.

Columns (6) and (7) in Table 2 present the financial outcomes. Column (6) shows that revenue per patient remains unchanged after the merger. This may not be surprising, as the payment structure-the global sum payment that consists of a fixed per capita payment to practice-remains unchanged. However, Column (7) demonstrates a significant increase of around 24% in revenue per FTE GP after the merger, which translates to an additional yearly revenue of around £75,108 for the average practice.[33] As general practices are typically small businesses owned by GP partnerships, 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 implies potential financial gains from the merger.[34] In Table C7, we provide suggestive evidence that the increase in revenue per GP could be due to merged practices stretching resources and allocating a greater patient load to doctors.

5.2. Dynamic effects

We plot the estimated coefficients from the dynamic specification Eq. (2) in Fig. 2.[35] The pre-merger estimates support the parallel trends assumption, as the confidence intervals include zero for the pre-merger years. Fig. 2(a) shows no significant change in QOF achievement, and Fig. 2(b) shows some evidence of a decline in population performance in clinical indicators in the long run. Fig. 2(c) and Fig. 2(d) illustrate a substantial and long-lasting decrease in patient satisfaction following mergers. Fig. 2(e) reveals no significant change in revenue per patient. However, there is an immediate increase in revenue per FTE GP after the merger as shown in Fig. 2(f), and this positive effect persists over time.

One limitation of the event study approach for exploring the long-run effect is the reduced sample size. For example, only mergers occurring in 2014 are used when studying the effect five years after the merger. Therefore, we perform an additional analysis where we generate two dummy variables: one for one to two years post-merger and another for three to five years post-merger. These are

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

[34] It is possible that GP partners hire salaried GPs who focus only on medical roles. Unfortunately, due to a lack of data, we are unable to differentiate between revenue per FTE GP partners and revenue per salaried GPs. Regardless, we believe this issue does not strongly affect our interpretation, as salaried GPs usually constitute a smaller portion of the FTE GP workforce. For instance, a recent report by GPOnline at https://www.gponline.com/general-practice-lost-one-30-partners-last-year-official-data-show/article/1724274 highlights that GP partners still make up 61% of the FTE GP workforce despite a drop in this number in 2021.

[35] The estimated coefficients are presented in Table C8.

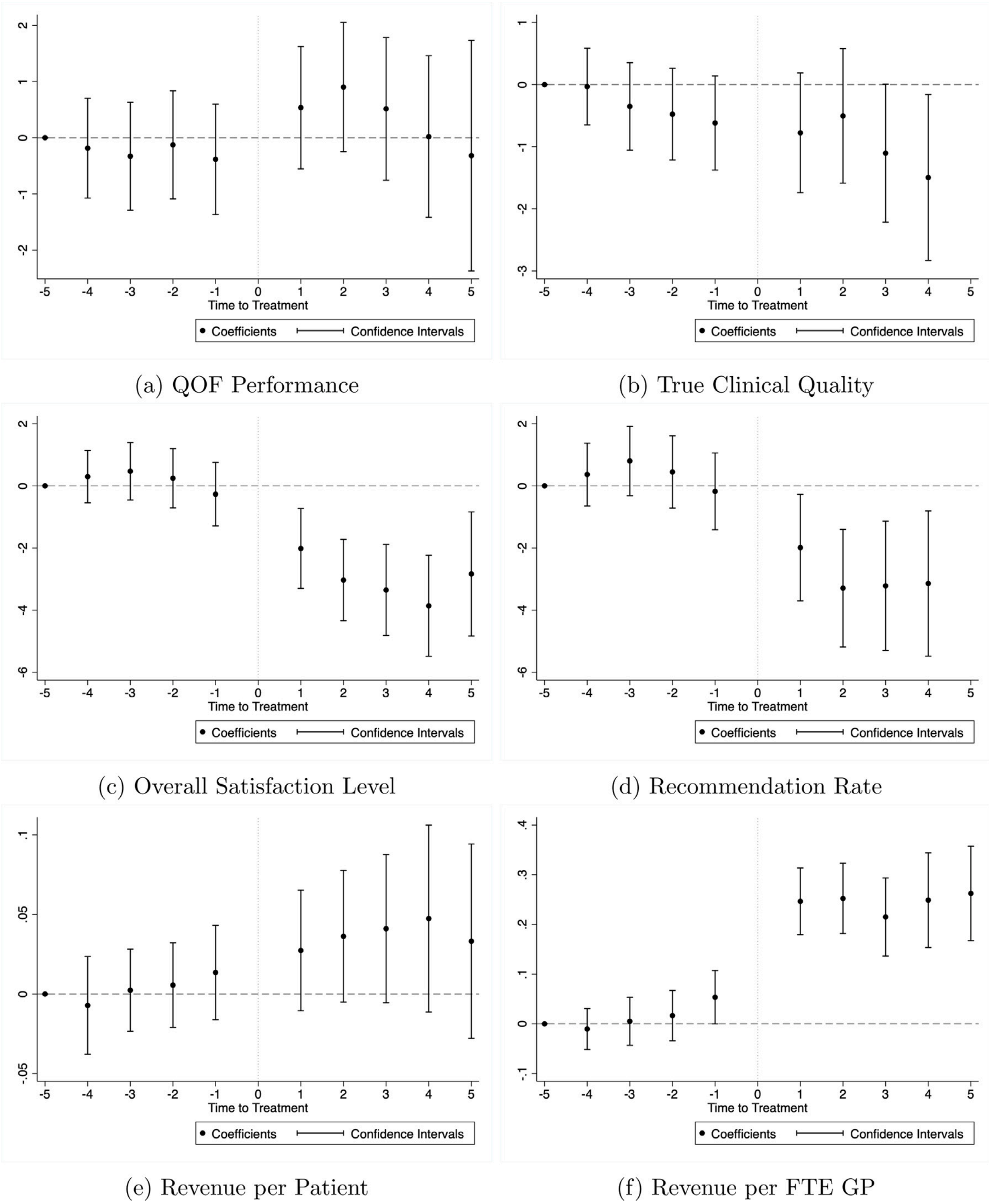

Fig. 2. Estimation of dynamic treatment effect.

Notes: This figure plots the estimated coefficients with 95% confidence intervals from Eq. (2). From Fig. 2(a) to Fig. 2(f), the dependent variables are qofOutcome, ClinPA2, OverallSat, Recommend, RevPerPatient, and RevPerGP, respectively. The reference period is represented by -5. The year of merger (i.e. time period 0) is dropped for treated practices to prevent measurement errors.

Table 3
Heterogeneous 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-small 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-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-large 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-small Merge = Small-large Merge	0.717	0.037	0.150	0.347	0.430	0.881	0.300
Small-small Merge = Large-large Merge	0.955	0.319	0.009	0.915	0.984	0.039	0.010

Notes: We compare three types of mergers: between small practices, between large practices, and between small and large practices. A practice is considered small if its pre-merger 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, which are the same as those used in the main analysis, 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. All specifications include cohort-specific practice and year fixed effects. The year of the merger is excluded for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

interacted with the treatment status to assess short- and long-term effects. Results in Table C9 confirm enduring financial gains from mergers and a long-term decline in quality.

5.3. Heterogeneity in the impact of GP practice mergers

We explore heterogeneous effects along three dimensions: the pre-merger size of merging parties, whether the two merging parties are within the same geographical market, and the stated merger motivations.

5.3.1. Heterogeneity of merger effects based on the size of merging parties

We categorize merging parties as small or large based on patient numbers one year before the merger compared to the mean size across all practices in England. Mergers are then classified into three types: small-small, large-large, and small-large. We generate dummy variables for each type and re-estimate Eq. (1), substituting $Treat_{ig} \times Post_{gt}$ with three interaction variables, each interacting one of the three dummies with $Post_{gt}$. Results are shown in Table 3. Column (2) shows financial gains across all merger types. Column (3) indicates improvements in clinical quality for small-small mergers, with a significant 1.2 percentage point increase in QOF performance. This estimate is statistically similar to those of small-large mergers but not to those of large-large mergers. Patient satisfaction declines across all merger types, with the largest drop observed in large-large mergers (Columns (6) and (7)). A robustness check using the 25th and 75th percentiles to define practice size confirms these findings (see Table C10).

5.3.2. Heterogeneity of merger effects: Within-market vs. Cross-market mergers

In the English primary care market, mergers can occur within the same geographical market (within-market mergers) or across different markets (cross-market mergers). Results in Table 4 indicate no significant difference in quality outcomes or financial performance between these two types. This suggests that market concentration may not be the main factor affecting quality, as we would expect worse outcomes for within-market mergers if it were. We formally test the market concentration mechanism in Section 5.4.1.

5.3.3. Heterogeneity of merger effects based on different motivations to merge

We explore the effects of mergers based on their claimed motivations: achieving efficiency, survival, or neutral (no clear motivation). We identify 453 efficiency-driven mergers, 143 survival-driven mergers, and 191 neutral mergers. We analyze whether the effects differ among these types.[36] Results in Table 5 show consistent financial gains across all motivations (Column (2)). However, the impact on quality varies. Efficiency-driven mergers have the most detrimental effect on quality, showing no improvement in QOF performance (Column (3)) and a significant deterioration in population-level clinical quality (Column (5)), along with the largest drop in patient satisfaction (Column (6)). In contrast, survival-driven mergers improve clinical quality and experience a much smaller decline in patient satisfaction. This distinction between survival-driven and efficiency-driven mergers offers key insights into the mechanisms at play, which we explore in detail in Section 5.4.2.

[36] Specifically, we combine the owner-retirement and failing types of mergers as they both involve practices facing difficulties. This category differs from the efficiency-driven mergers, where the merger occurs between normally functioning practices without significant operational difficulties. Our main findings hold if we examine mergers caused by owner retirement and failure separately (Table C11).

 Journal of Health Economics 104 (2025) 103050

Table 4
Heterogeneous 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: We compare within-market (mergers within the same geographical market) and cross-market mergers (mergers across different markets). The omitted group consists of never-merged practices selected by PSM. Coefficients on the additional time-varying covariates, which are the same as those used in the main analysis, 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. All specifications include cohort-specific practice and year fixed effects. The year of the merger is excluded 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 5
Heterogeneous effects: Different claimed merger motivations.

Merger motivations	(1) RevPerPatient	(2) RevPerGP	(3) qofOutcome	(4) ClinPA	(5) ClinPA2	(6) OverallSat	(7) Recommend
Survival-driven Mergers	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)
Efficiency-driven Mergers	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)							
Survival-driven = Efficiency-driven	0.368	0.454	0.006	0.018	0.062	0.002	0.120

Notes: Mergers motivated by failure and owner retirement are grouped as "survival-driven mergers". Mergers aimed at achieving efficiency are labeled as "efficiency-driven 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, which are the same as those used in the main analysis, 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. All specifications include cohort-specific practice and year fixed effects. The year of the merger is excluded for merged practices. Standard errors clustered at the cohort-practice level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

5.4. Discussions

5.4.1. Test of 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. However, mergers may decrease the incentive to maintain high quality in the absence of competitive pressures.

To investigate this, we study the subsample of within-market mergers, comparing merged entities 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. If quality declines only in already concentrated markets, it suggests that changes in market concentration might be one potential mechanism. We define a relevant geographic market as a 2 km radius around a practice and classify markets as highly competitive if the number of competitors exceeds the national average.[37,38] Results in Table 6 show no difference in quality impact based on pre-merger market concentration, suggesting that market concentration changes may not be the driving force behind the quality decline. We reinforce this finding by performing several robustness checks. Results in Table C12 define market competition levels based on the 25th and 75th percentiles of the distribution. Table C13 measures

[37] The treated unit is defined as the combined merged entity. As the acquirer practice serves as the main site after the merger, we focus on the acquirer's market to define the market competitive level for the merged entity.

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

Table 6
Test the 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	9994	11,685	9994
Test equality of coefficients (*p*-value)					
Low comp = High comp	0.206	0.141	0.270	0.536	0.187

Notes: Focusing on the subsample of within-market mergers, we 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 2 km radius around each practice. The omitted group consists of never-merged practices selected by PSM. Coefficients on the additional time-varying covariates, which are the same as those used in the main specification, 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. 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.

market concentration changes using the Herfindahl–Hirschman Index (HHI) and categorizes mergers into those with no concerns, unlikely to be challenged, and potentially raising significant concerns.[39] Our results hold.

This finding suggests that competition in the English general practice market may not function as strongly as economic theory would predict. Gravelle et al. (2019) show that greater competition is associated with higher patient satisfaction and better clinical quality in the English general practice market, but they note that the magnitude of these effects is small. In reality, patients are required to register with a GP practice, which then becomes their primary point of contact for healthcare. This can create significant switching costs for patients dissatisfied with the quality of care. Official statistics show that only around 8% of patients switch practices each year, with over 90% of these moves driven by patient relocation (Empel et al., 2023).[40] This suggests that patient responses to quality are limited, which weakens competitive pressures on GP practices. If practices were not strongly competing on quality to begin with, then changes in market concentration following a merger would not necessarily translate into predictable changes in quality.

5.4.2. Alternative explanations

As changes in market concentration are not the primary mechanism behind quality changes, we explore alternative explanations in this section. Our main analysis finds that mergers lead to improvements in QOF performance, no changes in clinical quality indicators, but a significant decline in patient experience. Importantly, these effects differ by merger type. The heterogeneous analysis sheds light on the underlying mechanisms at play, particularly when comparing survival-driven and efficiency-driven mergers.[41]

We begin by exploring the potential reasons behind changes in QOF performance and *ClinPA* measures. QOF performance reflects overall GP practice management in maintaining clinical quality, while *ClinPA* is derived from a subset of QOF indicators that specifically assess clinical quality outcomes at the population level. Heterogeneous analysis in Table 5 shows that survival-driven mergers improve QOF performance and maintain population clinical quality, whereas efficiency-driven mergers show no improvement in QOF and signs of deterioration in clinical quality. To investigate why, we first assess whether changes in patient mix explain these differences. We re-run the heterogeneous analysis using long-term illness prevalence rates as the outcome variable. Results in Table C14 show that regardless of merger types, the patient pool, in terms of long-term illness prevalence, remains stable after the merger,[42] ruling out this explanation.

Instead, we find that staff workload adjustments are the more likely mechanism affecting QOF performance. Efficiency-driven mergers change staff composition in ways that negatively affect clinical quality. To study this, we focus on the subsample of

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

[40] While some studies suggest that patients prefer high quality practices in England, this effect appears limited. For instance, Santos et al. (2017) find that patients are more likely to choose practices with higher QOF scores, but their study is based on a regional sample of around 1000 practices in the East Midlands. Moreover, distance remains the first-order factor in patient choice, with 40% of patients registered at the nearest practice. Similarly, Empel et al. (2023) and Nagraj et al. (2013) find that better quality is associated with lower switching rates, but this effect is observed only within the small subset of patients who switch practices without relocating. Given that such patients represent a minor fraction of the population, their findings do not generalize to the broader level, especially as most patients do not switch GP practices at all.

[41] Our discussion focuses on heterogeneity in stated merger motivations. Differences in merger size are likely to follow a similar intuition, as survival-driven mergers typically involve smaller practices, while efficiency-driven mergers are more common among larger ones. Additionally, the finding of no significant differences between within-market and cross-market mergers reinforces our conclusion that changes in market concentration are not the primary driver of quality outcomes.

[42] While some illness prevalence rates slightly change following mergers, the overall patient mix remains largely stable across merger types.

efficiency-driven mergers, and compare practices with different GP workload. We classify mergers as *LowBurden* if the patients-per-FTE-GP ratio of the merging parties falls within the 25th percentile of the national average one year prior to the merger, and *HighBurden* if within the 75th percentile. Table C15 shows that the increase in GP workload is less detrimental for efficiency-driven mergers with a modest GP workload to begin with compared to those that were already overburdened. Importantly, *LowBurden* mergers offset the increased GP workload by hiring more nurses (Column (2)). Nurses play a crucial role in supporting QOF performance, particularly in managing chronic conditions and preventive care (Griffiths et al., 2010; Khan and Peckham, 2024). Many QOF indicators, such as those related to regular blood pressure monitoring and record keeping of patients with chronic conditions, can be well handled by nurses. By taking on these responsibilities, nurses alleviate pressure on GPs while ensuring achieving QOF targets. As a result, QOF performance improves (Column (3)), and population achievement in clinical quality indicators are maintained (Column (4)). On the other hand, for those efficiency-driven mergers with already high GP work burden, the further increase of GP workload is much more detrimental (Column (1)). Without a corresponding increase in nursing staff (Column (2)), this heightened strain leads to a decline in both QOF performance and population clinical quality (Columns (3) and (4)).[43] By contrast, survival-driven practices tend to merge with better-functioning practices, benefiting from learning and achieving economies of scale (Mehrotra et al., 2006). This allows them to improve chronic disease management, leading to improved QOF performance.

In terms of overall patient experience, Table 5 shows that while patient satisfaction declines for both merger types, the decline is significantly worse for efficiency-driven mergers. We argue that survival-driven mergers perform better in patient experience because they prevent practice closures. If patients prioritize access to care and do not frequently switch practices, maintaining a local practice through a merger can be beneficial. To test this, we focus on survival-driven mergers and analyze detailed patient experience measures related to access to care. Results in Table C16 confirm that after the merger, patient satisfaction with access to care (Column (1)), waiting times (Column (2)), and practice opening hours (Column (3)) are maintained. Although continuity of care slightly declines (Column (4)), this is likely due to the increased GP workload (Column (5)). Efficiency-driven mergers, however, are not aimed at maintaining access to care but rather at improving operational efficiency. As a result, they often stretch resources further by increasing GP workload, leading to a larger decline in patient experience. To explore this, we again examine efficiency-driven mergers by GP workload. Table C17 shows that efficiency-driven mergers with already overburdened practices experience significantly worse patient experience outcomes. In contrast, those with *LowBurden* practices experience a much smaller decline. In particular, patient satisfaction with continuity of care and access to care (Columns (2) and (3)) are maintained. These findings suggest that increased GP workload is a key driver of deteriorating patient experience. Survival-driven mergers may mitigate the negative effects by preserving access to care, whereas efficiency-driven mergers, particularly those with already overburdened practices, further exacerbate resource constraints, leading to a more severe decline in patient experience.

6. Robustness checks

6.1. Robustness to potential policy confounder

In July 2019, the NHS introduced Primary Care Networks (PCNs) in England to promote collaboration among general practices (Gravelle et al., 2022, Morciano et al., 2020, and Checkland et al., 2023). PCNs bring together groups of GP practices alongside other healthcare providers, including community health services, pharmacies, hospitals, and voluntary organizations, to deliver more integrated care to local populations.[44] PCNs are typically organized geographically, with each network usually serving a population of 30,000 to 50,000 patients.[45] They were introduced as an optional add-on to the national GP contract, providing additional funding for services delivered through these networks. While GP practices were not mandated to join, those that did not would lose out on this extra funding. As a result, nearly all GP practices joined a PCN when the policy came into effect (Morciano et al., 2020 and Checkland et al., 2023).[46]

While the establishment of PCNs represents an important policy change that may have influenced GP practices and the service they provide, it is unlikely to be relevant to our study for several reasons. First, PCNs were officially established in July 2019, while our study relies on yearly data spanning from 2013 to 2019, with 2019 as the final year. Given that PCNs were only introduced midway through 2019, their impact is unlikely to be substantial within our study window. Second, nearly all GP practices in England joined a PCN when the policy was introduced. This means that almost all practices in our sample were subject to the same policy environment, and any potential effect of PCNs should be accounted for by the year fixed effects in our regression. By cross-referencing the national PCN lists published by NHS England Digital,[47] we identified only 30 GP practices (representing 207 out of 18,144 total practice-year observations) that did not join a PCN at the time of its rollout. Given the small number of non-participating practices, we believe our results should not be affected by the forming of PCNs. To further ensure that our findings are not confounded by the PCN policy, we conduct two robustness checks. First, we re-estimate the main regression excluding the 30 GP practices that did not join a PCN (Table C18). Second, we re-estimate the main regression while excluding data from 2019 (Table C19). Results remain robust in both cases.

43 The estimated coefficient for *qofOutcome* is negative, though not statistically significant.

44 For further details on PCNs, see https://www.england.nhs.uk/long-read/primary-care-networks-pcns/.

45 NHS England designed PCNs at this scale to balance local population needs with the efficiency gains of collaborative working, though exceptions may be granted with commissioner approval.

46 For further discussion on the characteristics and impact of PCNs, see Morciano et al. (2020) and Checkland et al. (2023).

47 The national list of PCNs and their GP practice members was published by NHS England Digital in November 2019 and updated monthly thereafter. This data is publicly available at https://digital.nhs.uk/services/organization-data-service/data-search-and-export/csv-downloads/gp-and-gp-practice-related-data.

6.2. Robustness to model specification

First, Table C20 uses 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. Second, we select alternative control groups. We re-construct the control group by (i) selecting the closest 1, 5, 7 never-merged practices by PSM (Table C21) (ii) using all never-merged practices (Table C22), and (iii) using future mergers occurring at least two years later as the control group (Table C23). The rationale behind using not-yet mergers as the control group is that the timing of mergers is potentially random (as shown in Table C24), which helps ensure a plausible causal estimate. Finally, there might be potential reverse causality for practices where poor quality is the primary motivation to merge. Therefore, Table C25 excludes mergers involving practices rated as inadequate by CQC. Our main findings remain robust.

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.

We find that on average, merged practices experience an increase in revenue by approximately £75,108. However, this financial improvement is not matched by maintaining the same level of patient outcomes. Although mergers improve certain aspects of clinical quality management, they do not translate into significant population-level clinical quality gains. In fact, we find evidence of a decline in clinical quality in the long run, along with a significant drop in patient satisfaction. Future research is needed to examine the overall welfare implications of the mergers as increasing revenue through mergers may be a potentially desirable outcome to address the doctor shortage issue in England.

Importantly, we find that the effect of mergers differs by merger motivations. Survival-driven mergers appear to preserve clinical quality and access to care, while efficiency-driven mergers, especially those involving already overburdened practices, exacerbate resource constraints and lead to worse quality outcomes. Further mechanism analyses indicate that changes in market concentration are not the primary driver behind quality outcomes. Instead, shifts in workforce composition, driven by the underlying merger motivations, play a key role.

These findings have important policy implications, particularly since much of the concentration in physician markets remains unnoticed by regulators (Gravelle et al., 2019). Our findings highlight the need for a more nuanced policy approach to general practice mergers in England. Supporting struggling practices through mergers can be a viable policy tool, whereas efficiency-driven mergers may require safeguards to prevent potential quality declines.

Although this paper focuses on the English general practice market, the findings have broader implications for other fixed-price healthcare systems experiencing similar consolidation trends. Our findings suggest that bigger is not always better. Mergers may involve complex restructuring processes that require careful management to avoid unintended consequences. In particular, consolidation efforts should be accompanied by necessary resources to support appropriate staffing levels for the delivery of high-quality care.

CRediT authorship contribution statement

Yuan Lyu: Writing – review & editing, Investigation, Writing – original draft, Validation, Methodology, Resources, Formal analysis, Supervision, Data curation, Project administration, Conceptualization. **Zhaocheng Zhang:** Validation, Investigation, Methodology, Data curation.

Appendix. Supplementary materials

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jhealeco.2025.103050.

References

American Medical Association, 2022. Policy Research Perspectives. American Medical Association.
Baker, A.C., Larcker, D.F., Wang, C.C., 2022. How much should we trust staggered difference-in-differences estimates? J. Financ. Econ. 144 (2), 370–395.
Beckert, W., Christensen, M., Collyer, K., 2012. Choice of NHS-funded hospital services in England. Econ. J. 122 (560), 400–417.
Berry, S.T., Waldfogel, J., 2001. Do mergers increase product variety? Evidence from radio broadcasting. Q. J. Econ. 116 (3), 1009–1025.
Borusyak, K., Jaravel, X., Spiess, J., 2021. Revisiting event study designs: Robust and efficient estimation. arXiv preprint arXiv:2108.12419.
Brekke, K.R., Holmås, T.H., Monstad, K., Straume, O.R., 2019. Competition and physician behaviour: Does the competitive environment affect the propensity to issue sickness certificates? J. Heal. Econ. 66, 117–135.
Caliendo, M., Kopeinig, S., 2008. Some practical guidance for the implementation of propensity score matching. J. Econ. Surv. 22 (1), 31–72.
Callaway, B., Sant'Anna, P.H., 2021. Difference-in-differences with multiple time periods. J. Econometrics 225 (2), 200–230.
Capps, C.S., 2005. The Quality Effects of Hospital Mergers. (6), US Department of Justice, Antitrust Division.
Casalino, L.P., 2006. Which type of medical group provides higher-quality care? Ann. Intern. Med. 145 (11), 860–861.
Cengiz, D., Dube, A., Lindner, A., Zipperer, B., 2019. The effect of minimum wages on low-wage jobs. Q. J. Econ. 134 (3), 1405–1454.
Checkland, K., Bramwell, D., Warwick-Giles, L., Bailey, S., Hammond, J., 2023. Primary care networks as a means of supporting primary care: findings from qualitative case study-based evaluation in the english NHS. Br. Med. J. Open 13 (11), e075111.

Cutler, D., Dafny, L., Ody, C., 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., 2009. Estimation and identification of merger effects: An application to hospital mergers. J. Law Econ. 52 (3), 523–550.

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

Deshpande, M., Li, Y., 2019. Who is screened out? Application costs and the targeting of disability programs. Am. Econ. J.: Econ. Policy 11 (4), 213–248.

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

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

Empel, G., Gravelle, H., Santos, R., 2023. Does quality affect choice of family physician? Evidence from patients changing general practice without changing address. Econ. Model. 126, 106395.

Fan, Y., 2013. Ownership consolidation and product characteristics: A study of the US daily newspaper market. Am. Econ. Rev. 103 (5), 1598–1628.

Fan, Y., Yang, C., 2022. Estimating Discrete Games with Many Firms and Many Decisions: an Application to Merger and Product Variety. Technical Report, National Bureau of Economic Research.

Fulton, B.D., 2017. Health care market concentration trends in the United States: evidence and policy responses. Health Aff. 36 (9), 1530–1538.

Gaynor, M., 2006. What do we know about competition and quality in health care markets? NBER Working Paper No. 12301.

Gaynor, M., Ho, K., Town, R.J., 2015. The industrial organization of health-care markets. J. Econ. Lit. 53 (2), 235–284.

Gaynor, M., Laudicella, M., Propper, C., 2012. Can governments do it better? Merger mania and hospital outcomes in the english NHS. J. Heal. Econ. 31 (3), 528–543.

Gaynor, M., Town, R.J., 2011. Competition in health care markets. Handb. Heal. Econ. 2, 499–637.

George, L.M., 2002. Ownership concentration and product variety in daily newspaper markets. Commun. Policy Inf. Technol.: Promises, Probl. Prospect. 235–253.

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

Gowrisankaran, G., Nevo, A., Town, R., 2015. Mergers when prices are negotiated: Evidence from the hospital industry. Am. Econ. Rev. 105 (1), 172–203.

Gravelle, H., Liu, D., Propper, C., Santos, R., 2019. Spatial competition and quality: Evidence from the english family doctor market. J. Heal. Econ. 68, 102249.

Gravelle, H., Liu, D., Santos, R., 2022. How do clinical quality and patient satisfaction vary with provider size in primary care? Evidence from english general practice panel data. Soc. Sci. Med. 301, 114936.

Griffiths, P., Murrells, T., Maben, J., Jones, S., Ashworth, M., 2010. Nurse staffing and quality of care in UK general practice: cross-sectional study using routinely collected data. Br. J. Gen. Pract. 60 (570), e36–e48.

Ho, V., Hamilton, B.H., 2000. Hospital mergers and acquisitions: does market consolidation harm patients? J. Heal. Econ. 19 (5), 767–791.

Jeziorski, P., 2014. Effects of mergers in two-sided markets: The US radio industry. Am. Econ. J.: Microeconomics 6 (4), 35–73.

Kelly, E., Stoye, G., 2014. Does GP Practice Size Matter? GP Practice Size and the Quality of Primary Care. R101, IFS Report.

Khan, N., Peckham, S., 2024. Advanced nurse practitioner (ANPs) experiences of the quality and outcomes framework (QOF) scheme: a UK case study. Br. Med. J. Open 14 (11), e087492.

McCauley, L., Phillips, R.L.J., Meisnere, M., Robinson, S.K., 2021. Implementing High-Quality Primary Care: Rebuilding the Foundation of Health Care. National Academies Press (US).

Mehrotra, A., Epstein, A.M., Rosenthal, M.B., 2006. Do integrated medical groups provide higher-quality medical care than individual practice associations? Ann. Intern. Med. 145 (11), 826–833.

Morciano, M., Checkland, K., Hammond, J., Lau, Y.-S., Sutton, M., 2020. Variability in size and characteristics of primary care networks in England: observational study. Br. J. Gen. Pract. 70 (701), e899–e905.

Nagraj, S., Abel, G., Paddison, C., Payne, R., Elliott, M., Campbell, J., Roland, M., 2013. Changing practice as a quality indicator for primary care: analysis of data on voluntary disenrollment from the english GP patient survey. BMC Fam. Pract. 14, 1–8.

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

Papke, L.E., Wooldridge, J.M., 2008. Panel data methods for fractional response variables with an application to test pass rates. J. Econometrics 145 (1), 121–133.

Pike, C., 2010. An empirical analysis of the effects of GP competition. 1, (2), Cooperation and Competition Panel-Working Paper Series.

Roland, M., Guthrie, B., 2016. Quality and outcomes framework: what have we learnt? Br. Med. J. 354.

Roland, M., Olesen, F., 2016. Can pay for performance improve the quality of primary care? Br. Med. J. 354.

Romano, P.S., Balan, D.J., 2011. A retrospective analysis of the clinical quality effects of the acquisition of highland park hospital by Evanston Northwestern healthcare. Int. J. Econ. Bus. 18 (1), 45–64.

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

Santos, R., Gravelle, H., Propper, C., 2017. Does quality affect patients' choice of doctor? Evidence from England. Econ. J. 127 (600), 445–494.

Schaumans, C., 2015. Prescribing behavior of general practitioners: competition matters. Heal. Policy 119 (4), 456–463.

Siciliani, L., Chalkley, M., Gravelle, H., 2017. Policies towards hospital and GP competition in five European countries. Heal. Policy 121 (2), 103–110.

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

Sweeting, A., 2010. The effects of mergers on product positioning: evidence from the music radio industry. Rand J. Econ. 41 (2), 372–397.