DOI: 10.1111/1540-6229.70023

WILEY

ORIGINAL ARTICLE

Online real estate agencies and their impact on the housing market

Cigdem Gedikli[1] | **Robert Hill**[2] | **Oleksandr Talavera**[3] | **Okan Yilmaz**[1]

[1]School of Management, Swansea University, Swansea, UK

[2]Department of Economics, University of Graz, Graz, Austria

[3]Department of Economics, University of Birmingham, Birmingham, UK

Correspondence
Robert Hill, Department of Economics, University of Graz, Graz 8010, Austria.
Email: robert.hill@uni-graz.at

Abstract

Online platforms have transformed many markets, as evidenced by the rise of firms such as Amazon, Uber, and Airbnb. However, the recent emergence of online real estate agencies has not yet received much attention. We investigate the impact of online agencies on the housing market. Our dataset consists of 1,274,792 properties in England and Wales, for which we have matched Zoopla listings with actual transactions from the Land Registry. Using an instrumental variable approach, we find that time on market (TOM) is shorter by about 80 days and the sale-list price ratio is smaller by about 2.4% for properties listed with online agencies. These findings, combined with an average fee of less than one-third of that charged by traditional agencies, explain why online agencies have rapidly gained market share. Their share has risen particularly for properties in the mid-price range and in regions with younger demographics. Moreover, we find that the rise of online agencies has caused traditional agencies to change their behavior—TOM and the sale-list price ratio are lower for traditional agencies in regions with a higher share of online agencies.

KEYWORDS
diffusion of new technologies, digital disruption, online platforms, real estate market, sale-list price ratio, time on market

1 | INTRODUCTION

Online platforms are transforming markets around the world. Amazon's impact on the retail market and Uber's on the ride-hailing market are particularly striking examples in this regard.[1] By contrast, the impact of online platforms on the real estate market has been more nuanced. In one sense, online platforms have already completely transformed the real estate market. For example, Zillow and Redfin are widely used in the United States (Fairweather et al., 2024; Gindelsky et al., 2019), whereas Zoopla is widely used in the United Kingdom (B. France, 2021). Similarly, the Airbnb platform has established itself as the world leader for short-term rental listings (Oskam & Boswijk, 2016). Mortgages are also being increasingly processed online (Fuster et al., 2019). Indeed, the term PropTech has been coined to refer to the disruptive impact of technology on the real estate market (Anderson et al., 2023; Braesemann & Baum, 2020). In recent years, the market share of online real estate agencies has started rising rapidly (see Table 2). This shift toward online agencies has so far been largely ignored by researchers. This is the first paper to consider how the emergence of online agencies is transforming the housing market.

The distinction between online and offline in the real estate market is perhaps more blurred than in most other markets. Almost all real estate agencies have an online presence and, as has already been noted, most for-sale properties are listed on online websites, such as Zoopla in the United Kingdom. However, there are two key differences between online and offline real estate agencies: First, although traditional agencies have local physical (bricks and mortar) offices, online agencies operate completely online without any local branches.[2] Accordingly, the communication between online agencies and sellers is done mostly online, and the seller typically does not benefit from dealing with an agent with connections in the local community. Second, online agencies adopt a different business model in terms of the services they provide and how they price them. Online agencies in the real-estate market generally charge a fixed fee (irrespective of whether the property actually sells or not), whereas traditional agencies charge a percentage fee on the actual transaction price.[3] Additionally, online agencies offer more flexible service packages. In basic service packages, sellers can take photos of the property themselves or organize and host their viewings. As a result, online agencies cost sellers less. During the period that our sample covers, the standard package fee of online agencies, which is paid upfront, was around £1000 regardless of the property's value. The average traditional agency fee is 1.42% of the sale value (ranging from 0.9% to 3.6%), including value added tax. Accordingly, with the average fee rate on a £231,000 property (the average paid price in our sample), the traditional agency fee is £3280, which is more than three times higher than the standard package online agency fee.

[1] As of June 2023, Amazon had a market cap of $1.274 trillion, making it the fifth largest firm in the world (by market cap). Uber has a market cap of $80 billion (https://tinyurl.com/y9a6789d).

[2] Even an online agent will still have a bricks-and-mortar head office. For example, during our sample period, Purplebricks emerged as the biggest online agent in the United Kingdom with about a 77% share of the online market (and with its headquarters in Birmingham) (Table 1). Purplebricks was recently bought by Strike (Hughes, 2023).

[3] See for details https://tinyurl.com/5b6a6wu5, accessed July 19, 2025.

We focus here on the United Kingdom as it is a market leader in online real estate agencies. The share of total transactions managed by online agencies in the United Kingdom has increased from 0.15% in 2012 to 26.8% in 2018 (Table 2). Given such a rapid expansion, the United Kingdom is an ideal laboratory for exploring the disruptive effect of online agencies. We measure the impact of online agencies using a dataset consisting of 1,274,214 properties in the United Kingdom for which we have Zoopla listings matched with actual transactions from the Land Registry.

Controlling for property characteristics using regression and matching approaches, we find that the average time on market (TOM) is between 23 and 25 days shorter and the sale-list price ratio is about 0.47% lower for online agencies. However, these methods may not properly control for omitted variables. To address this issue, we consider a two-stage least squares instrumental variable (IV) approach. With this approach, we find that the impact of using an online agent is much larger. The average TOM is shorter for online agencies by about 80 days and the sale-list price ratio is 2.4% smaller.

To understand this finding, we can distinguish between two types of omitted variables that may be affecting the standard ordinary least squares (OLS) results. The first type is omitted variables relating to the characteristics of sellers (such as age, education, income, and wealth) and not observed in the Zoopla dataset. The second type is omitted variables relating to the property itself (such as the quality of the structure and finishes, general ambiance and noise). The role played by these omitted variables in determining our results is discussed in Section 5.5.

The combination of lower fees (on average less than a third), shorter TOM and a lower sale-list price ratio is attractive for sellers, perhaps explaining why the market share of online agencies has risen rapidly. A faster selling process is also attractive for buyers.

Our article is related to Sing and Zou (2024) who compare real estate agencies in Singapore with online listings with those that do not post their listings online. They find that agencies with online listings sell at a higher price and with a shorter time on the market. Hence online listing platforms improve search efficiency as compared to traditional advertising channels. Our setting is different in that all the agencies (in the United Kingdom) in our analysis have online listings. Our focus is on comparing traditional agencies (with bricks-and-mortar offices) and online agencies that do everything online. We likewise find that online agencies have a shorter TOM but do not find a significant price effect. Overall, our results are broadly consistent with Sing and Zou (2024) in the sense that moving more services online increases search efficiency.

Our findings provide new perspectives on a number of existing research areas. First, we contribute to the literature on the determinants of TOM and the sale-list price ratio. Generally, a longer TOM is explained in terms of the characteristics of properties, the relative bargaining power of buyers and sellers and the state of the economy. Properties with more atypical characteristics take longer to sell, other things being equal (Han & Strange, 2016; Haurin et al., 2010; Kolbe et al., 2021). In a market downturn, both TOM and the sale-list price ratio rise as the relative bargaining power of sellers falls (Genesove & Mayer, 2001). Other factors that can affect bargaining power include wealth, gender, and age (Harding et al., 2003; Hayunga & Pace, 2019) and whether the buyers/sellers are local or foreign (Cvijanović & Spaenjers, 2021). Our results here show that the type of agent also affects both TOM and the sale-list price ratio. Both are lower for properties sold through online agencies.

Second, our results relate to the literature exploring the disruptive impact of new technologies on incumbent firms and stakeholders. For example, Barron et al. (2021) find that in the United States, on average, a 1% increase in Airbnb listings increases long-term rents by 0.018% and house prices by 0.026%. Chen et al. (2022) observe similar results. Zervas et al. (2017) find that Airbnb has reduced hotel revenue in Austin, Texas, by between 8% and 10%. Similarly, various authors have shown how Uber has reduced fares and improved the quality of service through greater monitoring of drivers and customer ratings (Angristet al., 2021; Berger et al., 2018; Cramer & Krueger, 2016; Hall, 2018; Liu et al., 2021). Moreover, Berger et al. (2018) find that the income of taxi drivers fell by about 10% after Uber's entry into US cities. Similarly, we find that both TOM and the sale-list price ratio of traditional agencies are lower in regions with a greater penetration of online agencies. More specifically, our regression analysis indicates that a 1 percentage point increase in online agency market share is associated with 0.9 days shorter TOM and a 0.02% decrease in the sale-list price ratio for traditional agencies. In other words, increased competition from online agencies is forcing traditional agencies to adapt and speed up the selling process. Traditional agencies have also responded by lowering their fees (Da Silva, 2023).

Third, our article contributes to the literature on the diffusion of new technologies. A substantial part of this literature focuses on how new technologies diffuse through the labor market, with low-skilled workers typically faring worse. For example, Autor et al. (2003) focus on the impact of computers, Akerman et al. (2015) on broadband internet, Agrawal et al. (2019) on machine learning prediction methods, whereas Bloom et al. (2021) consider the impact of new technologies in general. Bloom et al. (2021) find that new technologies tend to emerge first in highly concentrated locations and then only gradually diffuse spatially. We likewise find significant differences in the rate of penetration of online agencies. The penetration is highest in regions of the United Kingdom with lower average age and education (Figure 1). Looking specifically at the relationship between price and online penetration, we find that online penetration is highest in the middle part of the market. This is because properties in the middle of the distribution are easier to sell and hence more suitable for listing with an online agent.

Fourth, our findings on the emergence of online real estate agencies relate to the broader trend toward reduced demand for commercial real estate office space. Our results show that the cost savings reaped by real estate agencies moving online are not offset by lower quality of service (as measured by TOM and the sale-list price ratio). This may explain why the number of properties being sold by traditional agencies is declining (Table 2). To the extent this finding translates to other businesses, this suggests that the problem of excess supply in the office market is only going to get worse in the future. Van Nieuwerburgh (2022) and Gupta et al. (2022) note how this trend risks causing an urban doom loop as the falling value of commercial real estate reduces tax revenue for cities, causing them to cut services or raise taxes, potentially causing a further fall in property values as residents and firms move out of the city. Furthermore, given the close links between the commercial real estate market and the banking sector, declines in the value of commercial real estate could damage bank balance sheets and financial stability.

The remainder of this article is structured as follows. Section 2 provides an overview of online agencies in the UK real estate sector, whereas Section 3 discusses the dataset. Section 4 explains the regression, matching, and IV approaches we use to estimate the impact of online agencies on TOM and the sale-list price ratio. Section 5 presents the results, and Section 6 concludes.

2 | ONLINE AGENCIES IN THE UK REAL ESTATE SECTOR

The advent of online agencies in the United Kingdom traces its origins back to the launch of HouseNetwork in 2003. Out of them, Purplebricks was the largest online seller, with 76.9% of listings in 2014–2018 (Table 1). In this sense, Purplebricks exerts a similar dominance over the online UK real estate market as Amazon and Uber do over their respective markets. Over the years, there has been a remarkable evolution in the real estate landscape, as depicted in Table 2 and Figure A1.

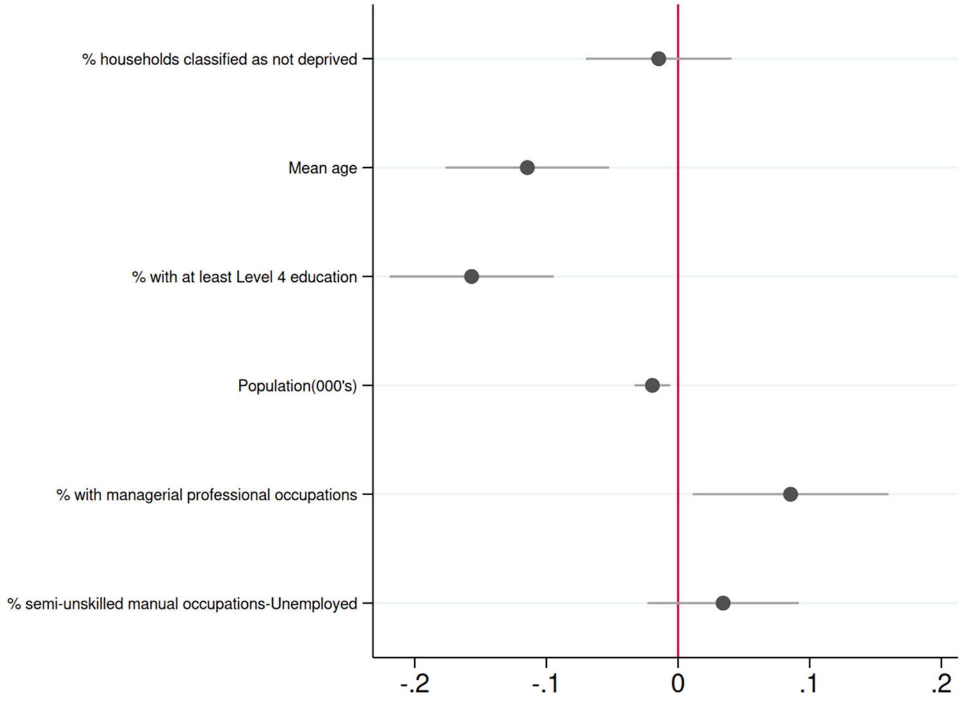

FIGURE 1 Online agency market share regression on local demographic characteristics (at outcode level). This figure shows the regression coefficients from the linear regression of the share of online real estate agencies in each outcode (first three digits of the postcode; 2269 in total) against various outcode-level demographic characteristics. These include the average age of residents, the percentage of residents with at least Level 4 education (equivalent to university education), the percentage of households not classified as deprived (i.e., not deprived in any dimensions of education, employment, health, or housing), total population, the share of persons in managerial/professional occupations, and the share of persons in semi/unskilled manual occupations. These data are obtained from the UK 2011 Census. All outcode (also known as postcode district) demographic characteristics are obtained from the Office for National Statistics via https://www.nomisweb.co.uk/query/select/getdatasetbygeog.asp.

Starting at about 1% in 2014, the number of properties being advertised by online agencies soared to nearly 30% by the close of 2018. However, it is important to note that online agencies are not evenly spread throughout the country. Figure A2 offers insights into this geographical distribution, revealing that within England and Wales, the highest concentration of online agencies can be found in the West Midlands and Yorkshire regions, whereas the lowest prevalence is observed in Cornwall, Norfolk, and the North East. This non-uniform dispersion underscores the dynamic nature of the online real estate market in the United Kingdom.

Although traditional agencies charge a percentage fee (typically about 2%), there are several different approaches when it comes to how online estate agencies charge their fees in the real estate market. Purplebricks, for instance, charges a fixed fee of £999 upfront, with additional charges for extras. Yopa follows a similar upfront fee structure, with clients having the option of paying £999 initially, plus extras, or a higher fee of £2000 at the time of sale. Housesimple, on the other hand, offers a no-upfront-fee model, but charges £700 for extras if requested. Express Estate Agency employs the traditional method of charging a percentage commission at the time of sale. Doorsteps presents a flexible pricing range, asking for £200–£1000 upfront.

TABLE 1 Online agencies by number of properties (2014–2018).

Agent name	Freq.	Percent	Cum.
Purplebricks	57,550	76.9	76.9
Emoov national	6692	8.94	85.84
Housesimple	5465	7.3	93.14
Yopa	2633	3.52	96.66
Springbok properties, nationwide	1665	2.22	98.89
Doorsteps.co.uk, national	653	0.87	99.76
Emoov	180	0.24	100
Total	74,838	100	

TABLE 2 Traditional versus online agency market shares by number of properties.

	2012	2013	2014	2015	2016	2017	2018	Total
Traditional	287,380	332,877	403,902	300,775	246,013	201,770	48,353	1,821,070
%	99.85	99.7	99.25	97.22	93.16	88.56	73.25	96.05
Online	440	987	3045	8592	18,052	26,056	17,656	74,828
%	0.15	0.3	0.75	2.78	6.84	11.44	26.75	3.95
Total	287,820	333,864	406,947	309,367	264,065	227,826	66,009	1,895,898

Meanwhile, Springbok has a higher price point, with fees set at £3500 to be paid at the time of sale. Emoov adopts a relatively lower upfront fee of £895. These diverse fee structures cater to various preferences and financial situations, allowing homeowners to choose an option that best suits their needs when selling their property. It should be noted that at the same time competition from Purplebricks has caused traditional agencies to lower their fees (Da Silva, 2023).

3 | DATA

In our analysis, we use Zoopla property data (ZPD) as our main data source.[4] Zoopla is the second most popular online property portal in the United Kingdom, with more than 8 million sale/rental property listings. In this article, we focus on residential properties that were listed for sale between 2014 and 2018, as online agencies gained considerable market share after 2013. ZPD provides a wide range of characteristics for each property, such as the list price (the price listed in the advertisement), number of bedrooms, bathrooms, and property type (e.g., terraced, detached, and flat). The locations of the property in terms of address (first line), postcode, and town (1120 unique values) are also provided, along with written descriptions by the seller or the agent of the property.

The essential element for our study is that ZPD provide information on the name of the real estate agency that markets the listing. As Zoopla does not allow private sellers to list their properties on the website, each property on Zoopla is marketed by an agency. Using this information, we identified the listings that are advertised by online agencies. Purplebricks was the market leader among online agencies during our sample period. In our classification, we also included listings that are marketed by the five major competitors of Purplebricks, which employ similar business models.[5]

[4] The data are provided by the Urban Big Data Centre (UDBC) licensed by Zoopla, https://tinyurl.com/efw2p2oy.

[5] These are Emoov, Housesimple, Yopa, Springbok, Doorsteps.

ZPD contain information on a rich set of characteristics for each property. However, the date on which the property is sold and the transaction prices are not provided. To obtain this information, we match the ZPD with the UK Land Registry's Price Paid Data (PPD).[6] PPD provide information on all property sales in England and Wales, including the paid price and the date of sale. Both PPD and ZPD contain detailed information on the location of the property.

Using the postcode (unique to each street) and property number (recorded as Primary Addressable Object Name in PPD and extracted from the first line of address in ZPD), we matched listings in ZPD with records in PPD. As a property can be listed and sold more than once during the period that we are interested in, we matched each listing on ZPD with the sale record in the PPD that is closest to the date the listing was published on the Zoopla website.[7] Using this approach, we were able to match around 80% of the listings on ZPD with a sale record from PPD. Although some of the unmatched listings can be due to missing or inaccurate address information, the majority of unmatched listings likely represent properties that were withdrawn from the market by the owners.

During the data cleaning, if there were multiple listings for the same property within the same month, we kept the first entry and deleted the remaining entries, as these are likely to be duplicates created by the agencies. Then, if there were listings for the same property with different entry dates on Zoopla having the same sale record (in terms of date of sale), we deleted the earlier listing, as this typically refers to a property that was withdrawn and then relisted. To remove outliers, we dropped listings in the top and bottom 1% of the price distribution. Finally, we dropped listings with more than five bedrooms and more than three bathrooms (each corresponds to less than 1% of listings with respect to the relevant characteristic). As a result, our final sample included 1,274,792 listings (of which 73,401 are marketed by online agencies) from England and Wales for the period 2014–2018.

Tables 3 and 4 present the descriptive statistics of our sample. Table 3 shows the average values of property characteristics across the whole dataset, whereas Table 4 provides separate average values for traditional and online agencies (both raw and adjusted for location and time of listing).[8] On average, TOM (time between the date of the first listing on Zoopla and the legal completion date recorded in PPD) is around 8 months (Table 3), and it is 29 days shorter for properties marketed by online agencies (Table 4). Both list and paid prices are higher for properties marketed by online agencies (Table 4). Table 3 also shows that around 66% of the listings were sold at a price lower than the original list price, whereas around 18% were sold at a higher price. A paid price less than the list price is likely to be a result of price adjustments for properties that did not receive the original list price. A raised price most likely stems from competing offers from potential buyers. Looking at the whole sample, we observe that, on average, the sale price is 2% lower than the list price. This reduction is smaller for properties marketed by online agencies (1.7%).

It can be seen in Table 4 that properties listed with online agencies tend to have slightly more bedrooms (2.9 vs. 2.8) and bathrooms (1.0 vs. 0.6) than properties listed by traditional agencies. The final column of Table 4 provides a preliminary adjustment for selection by computing differences in the value of characteristics and outcomes across traditional and online agencies separately for each town in a given year and then averaging these differences across the whole dataset. The

[6] PPD is publicly available.

[7] ZPD data include the ID number for each property and listing. There can be several listings for each property if the property is marketed more than once on the website. Hence, in this article, we use the term "listing" to refer to each entry on the Zoopla website.

[8] The presentation format in Table 4 is similar to that in Levitt and Syverson (2008) and Ash and Ben-Shahar (2024).

TABLE 3 Descriptive statistics.

	Mean	sd	p25	p50	p75
Sale-list price ratio	2.027	6.893	0.000	2.000	4.444
Tom	238.228	281.233	103.000	149.000	241.000
Online	0.058	0.233	0.000	0.000	0.000
Ln(list price)	12.211	0.567	11.813	12.196	12.608
Ln(paid price)	12.188	0.576	11.791	12.170	12.578
Discounted	0.662	0.473	0.000	1.000	1.000
Raised	0.181	0.385	0.000	0.000	0.000
Exact	0.157	0.364	0.000	0.000	0.000
Dop	0.000	0.263	−0.157	−0.005	0.148
Number of bedrooms	2.786	0.877	2.000	3.000	3.000
Number of bathrooms	0.643	0.733	0.000	1.000	1.000
Furbished (other)	0.028	0.165	0.000	0.000	0.000
New windows	0.001	0.033	0.000	0.000	0.000
New bathroom	0.004	0.066	0.000	0.000	0.000
New kitchen	0.008	0.089	0.000	0.000	0.000
New roof	0.001	0.037	0.000	0.000	0.000
Garden	0.690	0.462	0.000	1.000	1.000
Garage	0.302	0.459	0.000	0.000	1.000
Fireplace	0.123	0.329	0.000	0.000	0.000
Conservatory	0.076	0.265	0.000	0.000	0.000
Hard-flooring	0.247	0.431	0.000	0.000	0.000
Detached	0.203	0.403	0.000	0.000	0.000
Flat	0.137	0.344	0.000	0.000	0.000
Semi-detached	0.308	0.462	0.000	0.000	1.000
Terraced	0.348	0.476	0.000	0.000	1.000
Other	0.004	0.060	0.000	0.000	0.000
Observations	1,274,792				

Note: *The sale-list price ratio* is calculated by subtracting the paid price from the list price and dividing the difference by the list price (multiplied by 100). TOM refers to the number of days between the date the property is listed on Zoopla and the date the sale is recorded in the Land Registry. Online is a dummy variable that equals 1 if the property is advertised by an online agency, and 0 otherwise. Ln(list price) is the log of price asked on Zoopla. Ln(paid price) is the log of the paid price recorded in the Land Registry. Discounted price is a dummy variable that takes the value of 1 if the paid price is lower than the list price, and 0 otherwise. Raised price is a dummy variable that takes the value of 1 if the paid price is higher than the list price, and 0 otherwise. Exact price is a dummy variable that takes the value of 1 if the paid price is equal to the list price and, 0 otherwise. Dop represents the degree-of-overpricing which is the residuals of the Ln(list price) regression on all of the explanatory variables as in Table 5. Garden, Garage, Conservatory, and Fireplace are dummy variables representing the presence of these characteristics in the property (equal to 1 if these characteristics exist and 0 otherwise). Hard-flooring is a dummy variable that takes the value 1 for the presence of wooden/laminate flooring in the property, and 0 otherwise. *New windows, New bathroom, New kitchen*, and *New roof* are dummy variables equal to 1 indicating whether the properties undergo these types of refurbishments, and 0 otherwise. Refurbished(other) is a dummy capturing any other refurbishments.

results in the final column are generally reasonably similar to those in the third column, which does not attempt to control for selection. Adjusting for location and time of listing reduces the difference for some variables of interest (e.g., TOM falls from 29 days shorter for online agencies to 21 days shorter) while increasing the difference for other variables (e.g., the sale-list price ratio

TABLE 4 Sample characteristics for properties marketed by traditional and online.

	Full sample Traditional agent Mean (st. dev.)	Full sample Online agent Mean (st. dev.)	Full sample Mean of traditional Minus online agency	Within town and year Mean of traditional Minus online agency
Sale-list price ratio	2.044	1.744	0.301***	0.749***
	(6.983)	(5.195)	0.026	0.067
Tom	239.915	210.622	29.293***	20.999***
	(284.771)	(213.328)	1.069	2.848
Ln(list price)	12.209	12.243	−0.034***	−0.086***
	(0.569)	(0.536)	0.002	0.009
Ln(paid price)	12.186	12.224	−0.038***	−0.095***
	(0.578)	(0.539)	0.002	0.009
Discounted	0.664	0.627	0.037***	0.037***
	(0.472)	(0.484)	0.002	0.005
Raised	0.180	0.189	−0.009***	−0.003
	(0.384)	(0.391)	0.001	0.003
Exact	0.156	0.184	−0.029***	−0.034***
	(0.362)	(0.388)	0.001	0.003
Dop	−0.000	0.000	−0.000	−0.017***
	(0.265)	(0.242)	0.001	0.002
Number of bedrooms	2.780	2.884	−0.104***	−0.161***
	(0.879)	(0.835)	0.003	0.008
Number of bathrooms	0.620	1.017	−0.396***	−0.141***
	(0.733)	(0.620)	0.003	0.010
Furbished (other)	0.027	0.041	−0.014***	−0.018***
	(0.162)	(0.199)	0.001	0.002
New windows	0.001	0.002	−0.001***	−0.001**
	(0.032)	(0.047)	0.000	0.000

(Continues)

TABLE 4 (Continued)

	Full sample Traditional agent Mean (st. dev.)	Full sample Online agent Mean (st. dev.)	Full sample Mean of traditional Minus online agency	Within town and year Mean of traditional Minus online agency
New bathroom	0.004	0.011	−0.007***	−0.006***
	(0.063)	(0.102)	0.000	0.001
New kitchen	0.007	0.016	−0.009***	−0.010***
	(0.086)	(0.127)	0.000	0.001
New roof	0.001	0.002	−0.001***	−0.001**
	(0.036)	(0.050)	0.000	0.000
Garden	0.688	0.734	−0.046***	−0.012*
	(0.463)	(0.442)	0.002	0.005
Garage	0.301	0.322	−0.021***	−0.022***
	(0.459)	(0.467)	0.002	0.005
Fireplace	0.122	0.148	−0.027***	−0.029***
	(0.327)	(0.355)	0.001	0.004
Conservatory	0.075	0.098	−0.023***	−0.033***
	(0.263)	(0.297)	0.001	0.003

(Continues)

TABLE 4 (Continued)

	Full sample Traditional agent Mean (st. dev.)	Full sample Online agent Mean (st. dev.)	Full sample Mean of traditional Minus online agency	Within town and year Mean of traditional Minus online agency
Hard-flooring	0.243	0.305	−0.061***	−0.055***
	(0.429)	(0.460)	0.002	0.005
Detached	0.204	0.197	0.007***	−0.009*
	(0.403)	(0.397)	0.002	0.004
Flat	0.140	0.096	0.044***	0.035***
	(0.347)	(0.295)	0.001	0.003
Semi-detached	0.306	0.342	−0.036***	−0.023***
	(0.461)	(0.474)	0.002	0.004
Terraced	0.347	0.361	−0.014***	−0.006
	(0.476)	(0.480)	0.002	0.005
Other	0.004	0.004	−0.000	0.003***
	(0.059)	(0.063)	0.000	0.001
Observations	1,201,381	73,411		

Note: Column 1 (2) shows the mean values (and standard deviations in parenthesis) of the variables for the sample of properties marketed by traditional agencies (online agencies). Column 2 shows the difference between mean values (mean of traditional agency minus online agency) of the variables between traditional and online agencies. The last column also shows the same difference within the same year and town. The *sale-list price ratio* is calculated by subtracting the paid price from the list price and dividing the difference by the list price (multiplied by 100). TOM refers to the number of days between the date the property is listed on Zoopla and the date the sale is recorded in the Land Registry. Online is a dummy variable that equals to 1 if the property is advertised by an online agency, and 0 otherwise. Ln(list price) is the log of price asked on Zoopla. Ln(paid price) is the log of the paid price recorded in the Land Registry. *Discounted price* is a dummy variable that takes the value of 1 if the paid price is lower than the list price and 0 otherwise. *Raised price* is a dummy variable that takes the value of 1 if the paid price is higher than the list price, and 0 otherwise. *Exact price* is a dummy variable that takes the value of 1 if the paid price is equal to the list price and, 0 otherwise. *Garden, Garage, Conservatory*, and *Fireplace* are dummy variables representing the presence of these characteristics in the property (equal to 1 if these characteristics exist and 0 otherwise). *Hard-flooring* is a dummy variable that takes the value 1 for the presence of wooden/laminate flooring in the property and 0 otherwise. *New windows, New bathroom, New kitchen*, and *New roof* are dummy variables equal to 1 indicating whether the properties undergo these types of refurbishments and 0 otherwise. Refurbished(other) is a dummy capturing any other refurbishments.

*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

rises from 3.0% smaller for online agencies to 7.5% smaller). In later sections, we explore the role of selection in more detail.

4 | EMPIRICAL STRATEGY

4.1 | An OLS-based approach

4.1.1 | Time on market

TOM could serve as an indicator of the performance of agencies and their marketing quality. Among other factors that affect TOM, a better advertised and accurately valued property is likely to sell faster. To investigate how TOM differs for properties marketed by online and traditional agencies, we estimate OLS regression as follows:[9]

$$\mathrm{TOM} = \rho + \lambda \mathrm{Online}_i + \mathrm{X}_i \nu + R_i r + T_i \tau + Z_i \psi + e_i \cdot \tag{1}$$

In Equation (1), the dependent variable is the TOM for property i, measured by the number of days between the date on which the listing was advertised on Zoopla and the official date of the completion of the sale. Online, our key variable of interest is a dummy variable that takes the value of 1 if the property is marketed by an online agency and the value of 0 otherwise. X_i is a vector of property characteristics such as the number of bedrooms, bathrooms, property type (detached, semi-detached, terraced, flat, and others—see Table 4 for further details), and an indicator dummy variable showing if the property is refurbished (this information is extracted from the textual description provided in the advertisement). Following Rutherford, Springer and Yavas (2005), we include a variable R_i in the time-on-market regression model, which measures the degree of overpricing. R_i is calculated as the residual of a hedonic regression estimated with asking price as the dependent variable. A positive residual suggests that the asking price was too high (although it could also be that this property performs well on the omitted variables) typically implying a longer TOM. T_i is a vector of time dummies for each year-quarter (e.g., 2016Q1, 2016Q2) in which the property was listed. Z_i is a vector of dummy variables representing the location of the properties at the post-town level (there are 1120 post-towns in our sample). The key parameter of interest is λ, which captures the average TOM difference between properties listed by online agencies and traditional agencies after controlling for the various property characteristics.

4.1.2 | Sale-list price ratio

As a further indicator of agency performance, we compare the ratio of the list price on Zoopla to the eventual paid price for online and traditional agencies. As seen in Table 3, only 16% of the properties are sold at their original list price in our sample. When there are competing offers from buyers, the price can rise above the initial list price and the seller can benefit from the positive net gain over the list price. However, a lack of offers and longer TOM may lead to discounts and hence a sale price that is further below the list price. In both cases, the marketing quality and strategies

[9] In Section 5.2, we also undertake a survival analysis of time on market to take account of properties that list but never sell.

offered by agencies could play an important role in attracting potential buyers. In order to explore how online agencies perform against traditional agencies in terms of the sale-list price ratio, we estimate the following OLS regression model:

$$\Delta P_i = \varphi + \theta \text{Online}_i + X_i \eta + T_i \tau + Z_i \psi + \xi_i \tag{2}$$

In Equation (2), the dependent variable ΔP_i, is the sale-list price ratio for the listing i. This is defined as the list price minus the paid price, divided by the list price (all multiplied by 100).

$$\Delta P_i = 100 \left(\frac{P_{list} - P_{paid}}{P_{list}} \right)$$

The vectors X_i, T_i, and Z_i in Equation (2) are the same as in Equation (1). In addition to our estimate of the model in Equation (2) for the full sample, we perform separate price change estimations for two subsamples: (i) a sample of properties that were sold at a price lower than the list price (i.e., discounted price sample), (ii) a sample of properties that were sold at a price equal to or higher than the list price (i.e., exact or raised price sample). This subsample analysis enables a comparison to be made between the performance of online and traditional agencies in the positive and negative domains of price change.

4.1.3 | Sale price

Next we compare the actual sale price for online and traditional agencies. The explanatory variables here are exactly the same as in Equation (2):

$$P_i = \varphi + \theta \text{Online}_i + X_i \eta + T_i \tau + Z_i \psi + \xi_i$$

4.1.4 | The effect of online agent market share on traditional agency performance

As mentioned in Section 1, there is well-established empirical evidence on how online platforms (such as Uber and Airbnb) affect their traditional counterparts. To investigate whether the presence of online agencies has an impact on the performance of traditional agencies, we estimate the association between the local market share of online agencies and the TOM and the sale-list price ratio for properties marketed by traditional agencies. We, therefore, restrict our sample to properties, i, marketed by traditional agencies and estimate the following models:

$$\text{TOM}_i = \rho + \lambda \textit{Onlineshare}_i + X_i \upsilon + R_i r + T_i \tau + Z_i \psi + e_{2i} \tag{3}$$

$$\Delta P_i = \varphi + \theta \text{OnlineShare}_i + X_i \eta + T_i \tau + Z_i \psi + \xi_i \tag{4}$$

The key variable of interest in Equations (3) and (4), OnlineShare, represents the percentage of properties marketed by online agencies in the same quarter and post-town as the listing i. The vectors X_i, T_i, and Z_i in Equations (3) and (4) contain the same set of variables as in Equations (1) and (2). Accordingly, the coefficients λ and θ indicate the change in TOM and the sale-list price ratio for listing i with respect to a 1 percentage point change in the market share of online agencies within the local market.

In addition to a listing analysis at the level of individual properties, we also explore the impact of online agency penetration at the level of individual agencies. For each traditional agency, we calculate average TOM, the sale-list price ratio, and the number of listings each month. Then, we re-estimate Equations (3) and (4) at the level of individual agencies, excluding the individual property characteristics vector X_i. We also estimate the regression model in Equation (5), which has number of listings per traditional agency as the dependent variable.

$$\text{Listings}_j = \phi + \mu \text{OnlineShare}_j + T_j \tau + Z_j \psi + \varepsilon_j \tag{5}$$

Finally, we again re-estimate Equations (3) and (4), this time focusing on the average TOM and sale-list price ratio of traditional agencies at the post-town level (and again excluding the individual property characteristics vector X_i). In other words, in this case, i indexes the post-towns. Likewise, we estimate the regression model in Equation (6), which has the number of traditional agencies in each post-town as the dependent variable.[10]

$$\text{TradAgents}_k = \pi + \gamma \text{OnlineShare}_k + T_i \tau + Z_k \psi + \varepsilon_k \tag{6}$$

4.2 | Coarsened exact matching

The share of online agencies in the housing market increased rapidly over time and across different parts of England and Wales (see Figure A2). However, due to the variations in market shares across years and locations, a non-random selection problem may occur if sellers cannot access online agencies (mainly due to not being aware of them). Additionally, if online agencies are more popular among the owners of certain types of properties (such as flats vs. houses), the distribution of property characteristics may not be random between online and traditional agencies. If this is the case, estimated results from linear regression models may be unreliable. Therefore, in additional models, we estimate the impact of online agencies on TOM and price change, drawing on a matched sample.

To achieve a more similar distribution of property characteristics between the treatment group (properties marketed by online agencies) and the control group (properties marketed by traditional agencies), we use coarsened exact matching (CEM). As with other matching methods, CEM is a data pre-processing technique. CEM trims the data to ensure that the remaining data are better balanced in terms of covariates between the treatment and control groups (Blackwell et al., 2010). CEM is a relatively new method and is shown to be more efficient than traditional matching methods, such as propensity score matching, in providing lower levels of imbalance, model dependence, and bias (Blackwell et al., 2010; Iacus et al., 2012). Moreover, as King and Nielson (2019) note, CEM is computationally fast which makes it a practical alternative given our large sample size.

While matching the listings marketed by the online and traditional agencies, we use the same set of covariates in Equations (1) and (2). Specifically, CEM initially coarsens the covariates that we use in the models by recoding and generating a set of strata that contain the same coarsened values of the covariates. Then, it assigns the generated strata to the original data. Finally, it prunes

[10] The elements j in Equation (5) denote traditional agencies, whereas the elements k in (6) denote post-towns.

the data by only retaining the observations whose stratum contains at least one treatment and one control unit. After the procedure is done, these strata are used to calculate the average treatment effect on the treated.

4.3 | Two-stage least squares instrumental variable (IV) approach

Although measuring the treatment effect on a matched sample may provide more reliable estimates than a standard regression approach, it does not resolve the problem of omitted variables. Hence, we also employ an IV approach. Drawing on insights from studies that utilize spatial instruments to measure learning and spillover effects, we employ local indicators of online agencies' market shares as instruments for choosing online agencies over traditional ones.[11]

As outlined previously, online real estate agencies entered the UK property market around 2012 and have gained popularity over the years. Unlike their traditional counterparts, online agencies do not have physical offices spread across neighborhoods, which limits their visibility. Nonetheless, as their market share has increased, more online agency signs have begun appearing on the streets. The growing number of properties marketed by online agencies and their increased advertisements in local areas are likely to heighten sellers' awareness of these agencies and, hence, increase their local popularity and preference over traditional agencies. Based on this potential link, we exploit the market share of online agencies in local markets as an instrument for sellers' choices between traditional and online agencies.

Two possible instruments based on the online agency market share in the local markets are considered. The first instrument defines the local market by using the first part of the postcodes (referred to as the *outcode*) and the first digit of the second part (referred to as the *incode*), and it captures the share of listings marketed by online agencies in this narrowly defined local market during the previous quarter before the property enters the market.[12, 13] This narrow-scale local market share is likely to capture awareness of and preference for online agencies through word-of-mouth communication among neighbors, as it only covers one's own and neighboring streets. This instrument satisfies the relevance condition (as can be seen from the p-values on the KP LM statistics in Table 9). However, as argued by Betz et al. (2018), spatial instruments may suffer from simultaneity in the first-stage and in satisfying the exclusion restriction (Manski, 1993; Moffitt, 2001). More specifically, the community's preference for online agencies and the sellers' own

[11] The so-called leave-one-out or spatial instruments for each unit (i) are calculated by summing or averaging the treatment variable for a set of other units ($-i$), such as peers defined by regions, neighbors, friends, or colleagues (Sundquist, 2021). Spatial instruments have been used in various contexts, including measuring the impact of democracy on growth, estimating female labor supply, environmental outcomes, examining the effects of norms and religiosity on outcomes such as mental health, and in political science studies (see, e.g., Acemoglu et al., 2008; Pfeiffer & Lin, 2012; Robalino & Pfaff, 2012; Betz et al., 2018; Cavapozzi et al., 2021; Fruehwirth et al., 2019; Nicoletti et al., 2018).

[12] The current quarter is not used because the decision to choose online agencies is likely made before the property is advertised on the market.

[13] Postcodes in the United Kingdom consist of a combination of five to seven letters and numbers, defining four different levels of geographic units. Each postcode includes two parts: the outcode and the incode. The outcode contains the area and the district, whereas the incode includes a number followed by two letters. The number identifies the sector within the postal district (https://tinyurl.com/4s87zm9h). In our narrow-scale local market, we use the outcode and the first digit (the number) of the incode. For example, the outcode B11 covers Sparkhill, Sparkbrook, Tyseley, and Greet in Birmingham, whereas the more specific B11 2 district covers the areas of Sparkhill and Tyseley. It is the latter, the narrower local market, which we use in creating our first instrument.

choice of online agencies may arise from similar unobserved characteristics, or because they live in similar environments leading to similar choices. Although we include detailed locational controls in our models, there may remain some shared unobserved characteristics of the community and the individual sellers that are correlated with the decision to opt for online agencies. To acknowledge this possibility, our second instrument uses the market proportion of online agencies in the same outcode (e.g., B11 in Birmingham) during the previous quarter, while excluding the listings in the narrow-scale local market (e.g., B11 2 in Birmingham). Although this instrument still measures overall awareness, as it is possible to see the advertisements in the wider local area, for example, while shopping or commuting to work, it reduces the chances of unobserved correlated effects.

The first and second stage regression equations of the IV estimation for Tom are as follows:[14]

$$\text{Online}_i = \alpha_1 + \beta_1 \text{Postshare}_i + X_i \upsilon + T_i \tau + Z_i \psi + e_{1i} \tag{7}$$

$$\text{TOM}_i = \alpha_2 + \beta_2 \widehat{\text{Online}}_i + X_i \upsilon + R_i r + T_i \tau + Z_i \psi + e_{2i} \tag{8}$$

Online_i is the endogenous dummy variable which takes the value 1 if a property is listed with an online agent. PostShare is the instrument. The predicted value of online from Equation (7) is used as an explanatory variable in Equation (8). For the case of the sale-list price ratio, the dependent variable TOM_i is replaced by ΔP_i in Equation (8).

5 | RESULTS

5.1 | Time on market (OLS regression)

Table 5 provides the estimation results for TOM based on the regression equation in Equation (1). The results indicate that properties marketed by online agencies are associated with a shorter TOM. This finding remains when we control for location, time of entry, and property characteristics. Properties with online agencies are sold 24.7 days earlier than those marketed by traditional agencies (Column 3, Table 5). Additionally, the results show that TOM is longer for larger properties, whereas detached houses and semis sell more quickly than flats and terraces.

5.2 | Survival analysis

It is possible that a potential selection bias arises from the exclusion of listings that do not result in a sale, as these properties—likely withdrawn from the market before selling—may systematically differ from those that do sell. To address this concern, we estimated a survival analysis in which we include both matched (sold) and unmatched (withdrawn) listings.

For the survival analysis, we constructed a measure of marketing duration called time on website (TOW), defined as the number of days between a property's initial appearance on the Zoopla platform and the date it is withdrawn. This measure focuses on all listings, regardless of whether they have a corresponding transaction in the Land Registry. Using TOW also avoids potential distortion in time-on-market calculations that could arise if we relied solely on official sale dates, which are often recorded well after a sale agreement is reached.

[14] Equation (7) is estimated using the linear probability model.

TABLE 5 Time on market (TOM) (in days) regression.

	(1) Tom	(2) Tom	(3) Tom
Online	−23.896***	−26.829***	−24.062***
	(0.890)	(0.891)	(0.890)
Number of bedrooms		6.930***	17.183***
		(0.310)	(0.362)
Number of bathrooms		11.856***	9.958***
		(0.391)	(0.394)
Degree of overpricing			57.440***
			(0.999)
Furbished (other)			0.153
			(1.516)
New windows			−16.685*
			(6.709)
New bathroom			−4.872
			(3.527)
New kitchen			−7.547**
			(2.691)
New roof			−4.428
			(6.332)
Garden			−10.673***
			(0.599)
Garage			−12.586***
			(0.583)
Fireplace			−3.075***
			(0.773)
Conservatory			−4.946***
			(0.927)
Hard-flooring			6.813***
			(0.607)
Flat			41.550***
			(1.169)
Semi-detached			−26.826***
			(0.747)
Terraced			−17.705***
			(0.798)
Other			176.139***
			(6.640)
Location FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	1,274,792	1,274,792	1,274,792
Adjusted R^2	0.020	0.022	0.033

(Continues)

TABLE 5 (Continued)

Note: The dependent variable TOM refers to the number of days between the date the property is listed on Zoopla and the date the sale is recorded in the Land Registry. See Table 3 for definitions of explanatory variables. Robust standard errors in parentheses. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

We estimated a Weibull survival model where the event of interest is the successful sale of a property. Listings matched to Land Registry sales are coded as 1 (event observed), whereas unmatched listings are treated as right-censored (event not observed), reflecting properties that exited the market without a recorded sale. We repeated our estimation using two samples: the full sample (including properties likely withdrawn from the market before the sale) and the sample of matched properties only.

Table 6 presents the results from the survival analysis. Column 1 reports estimates for the full sample (including both matched and withdrawn properties), whereas Column 2 focuses on the matched properties only. The results are qualitatively similar across both samples, indicating that properties marketed by online agencies exhibit higher hazard rates, other things equal—that is, they have higher (72%) conditional sale (hazard) rates. This higher hazard rate implies that properties listed with online agents on average have a shorter TOW. It is also worth noting that this difference is even slightly bigger for the sample including withdrawn properties than the sample that excludes them.

5.3 | Sale-list price ratio (OLS regression)

The results for the sale-list price ratio from the regression equation in Column 2 are presented in Table 7. The first column in Table 7 shows that, after controlling for location and the time listings enter the market, the sale-list price ratio is around 0.45% lower for properties marketed by online agencies. This indicates that properties listed with online agencies experience a smaller reduction from the original list price. The difference remains when additional controls, including property characteristics, are added to the model (see Columns 2 and 3, Table 7). However, as argued earlier, the price change can be positive or negative depending on whether or not there are competing offers. The final column in Table 7 displays the results for the absolute value of the price change (i.e., we utilize the absolute value of the price change for each listing), which can be interpreted as the absolute deviation from the original list price, irrespective of the direction. The findings indicate that listings marketed by online agencies exhibit a smaller absolute deviation.

Table 8 presents the results separately for the discounted price sample (properties whose paid price is lower than the list price) and the raised price sample (properties whose paid price is greater than or equal to the list price). The results show that, amongst the discounted price sample, the reduction from the list price is, on average, 0.9% smaller for properties marketed by online agencies. On the contrary, amongst the raised price sample, the increase from the list price is 1.2% greater for properties marketed by online agencies. Overall, on average, being marketed by an online agency is associated with a slightly lower list price and slightly higher (significant at 1% level) paid price (see Columns 3 and 4 in Table 8).

The sales price results are shown in Table 9. While the online coefficient is negative and significant when the full set of controls is included, it is very small (−0.004). Hence, we conclude that the choice of agent-type does not have an economically meaningful impact on the sale price.

TABLE 6 Survival analysis (controlling for post town and quarter of entry).

	(1) Sample of all listings Hazard ratio	(2) Sample of listings that has a sale record Hazard ratio
Online	1.724***	1.673***
	(0.009)	(0.009)
Number of bedrooms	1.019***	0.987***
	(0.001)	(0.001)
Number of bathrooms	0.893***	0.917***
	(0.001)	(0.001)
Furbished (other)	1.057***	1.034***
	(0.006)	(0.005)
New windows	1.062*	1.030
	(0.028)	(0.026)
New bathroom	1.058***	1.011
	(0.015)	(0.014)
New kitchen	1.075***	1.025*
	(0.011)	(0.010)
New roof	1.070**	1.040
	(0.025)	(0.024)
Garden	1.157***	1.040***
	(0.002)	(0.002)
Garage	1.050***	1.018***
	(0.002)	(0.002)
Fireplace	1.061***	1.028***
	(0.003)	(0.003)
Conservatory	1.063***	1.041***
	(0.004)	(0.003)
Hard-flooring	1.005*	1.007***
	(0.002)	(0.002)
ln_*p*	1.672***	1.698***
	(0.001)	(0.001)
Observations	1,869,249	1,390,926
Adjusted R^2		

Note: This table shows the result of the survival analysis where the event is the sale of the property and time to event is time on website, which measures the number of days between when a listing first appears on the Zoopla website and when it is withdrawn. The first column shows the results for all listings on Zoopla regardless they have a sale record or not. The second column shows the results of the Zoopla listings that have a sale record in Land registry. Online is a dummy variable that equals to 1 if the property is advertised by an online agency, and 0 otherwise. *Garden, Garage, Conservatory*, and *Fireplace* are dummy variables representing the presence of these characteristics in the property (equal to 1 if these characteristics exist and 0 otherwise). *Hard-flooring* is a dummy variable that takes the value 1 for the presence of wooden/laminate flooring in the property and 0 otherwise. *New windows, New bathroom, New kitchen*, and *New roof* are dummy variables equal to 1 indicating whether the properties undergo these types of refurbishments and 0 otherwise. Refurbished(other) is a dummy capturing any other refurbishments.

***$p < 0.01$. **$p < 0.05$. *$p < 0.1$.

TABLE 7 Sale-list price regression.

	(1) Sale-list price ratio	(2) Sale-list price ratio	(3) Sale-list price ratio	(4) Sale-list price ratio: absolute value
Online	−0.465***	−0.498***	−0.470***	−1.041***
	(0.022)	(0.022)	(0.022)	(0.019)
Number of bedrooms		−0.015*	0.074***	−0.218***
		(0.008)	(0.009)	(0.008)
Number of bathrooms		0.179***	0.149***	−0.139***
		(0.008)	(0.008)	(0.007)
Furbished (other)			−0.144***	−0.381***
			(0.033)	(0.027)
New windows			−0.142	−0.548***
			(0.142)	(0.120)
New bathroom			0.039	−0.346***
			(0.075)	(0.062)
New kitchen			−0.114	−0.274***
			(0.061)	(0.051)
New roof			0.124	−0.238*
			(0.141)	(0.117)
Garden			0.023	−0.390***
			(0.015)	(0.013)
Garage			−0.015	−0.338***
			(0.013)	(0.011)
Fireplace			0.165***	0.155***
			(0.017)	(0.014)
Conservatory			−0.177***	−0.289***
			(0.018)	(0.015)
Hard-flooring			−0.071***	−0.414***
			(0.014)	(0.011)
Flat			0.404***	0.755***
			(0.030)	(0.025)
Semi-detached			−0.482***	−0.192***
			(0.016)	(0.013)
Terraced			−0.182***	0.222***
			(0.018)	(0.015)
Other			0.218	5.473***
			(0.258)	(0.213)
Location FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	1,274,792	1,274,792	1,274,792	1,274,792
Adjusted R^2	0.041	0.041	0.043	0.053

(Continues)

TABLE 7 (Continued)

Note: The dependent variable, *sale-list price ratio*, is calculated by subtracting the paid price from the list price and dividing the difference by the list price (multiplied by 100). The dependent variable in the last column is the absolute value of the sale-list price ratio. See Table 3 for definitions of explanatory variables. Robust standard errors in parentheses.
$^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

5.4 | Coarsened exact matching (CEM)

Table 10 presents the TOM, sale-list price ratio and sale price treatment effects based on the sample matched by the CEM method. These results are broadly consistent with the OLS regression results we present in Tables 5, 7, and 9. The TOM for properties listed with online agencies is now 22.7 days shorter, instead of 24.7 days shorter as with linear regression in Table 5. The sale-list price ratio is now 0.63% less for online agencies as compared with 0.45% less, according to the OLS results in Table 7. The impact of agent type on sale price is now zero.

5.5 | Instrumental variables (IV) results

IV estimates provide statistically significant coefficients for both market share instruments (Table 11). As the market share of online agencies increases in the local market, both on a narrow scale and a wider scale, the probability of a seller selecting these agencies also rises.[15] The second-stage results are consistent with our main findings, although the coefficients are larger. Compared to OLS (CEM) estimation results, TOM difference increases from 25 days (23 days) to 77 days when only 4-digits of postcode online share is used as instrument and to 82 days when 4-digits of postcode area online share and 3-digits of postcode area online share used simultaneously. Similarly, the sale-list price ratio is now 2.3% smaller for online agencies as compared with 0.45% (0.63%) smaller for OLS (CEM).

To further test the validity of our instrument—specifically the exclusion restriction—we follow the procedure suggested by Angrist et al. (2010), which provides a practical method for evaluating the plausibility of the exclusion restriction. According to this approach, if the exclusion restriction holds, then in subsamples where the first stage is statistically insignificant (i.e., the instrument does not significantly predict the treatment), the reduced form should also be insignificant. We apply this method using postcode-area-level subsamples, defined by the first letters of postcodes. Our results are presented in Table 12. In 13 out of 105 local areas, the first-stage regression p-value for the instrument—the share of listings marketed by online agencies—was greater than 0.05, indicating statistical insignificance. For these areas, we estimated reduced-form regressions in which the outcome variables—TOM and the sale-to-list price ratio—were regressed on the instrument, controlling for the same covariates as in our main specification. As shown in Table 12, for the sale-to-list price ratio, the coefficient on the instrument was insignificant in 9 (11) out of 13 local areas at the 5% (1%) significance level. For TOM, the instrument was statistically significant in only one area at the 5% level. These findings support the validity of the instrument in satisfying the exclusion restriction.

IV estimation often generates larger coefficients than OLS (see, e.g., Card, 2001). In our case, the larger impact of online agencies in the two-stage IV results can be attributed to a combination of seller and property omitted variables in the OLS regressions in Equations (1) and (2).

[15] See Table A1 for the first stage regression results.

TABLE 8 Sale-list price for subsamples and price regressions.

	(1) Discounted price sample Sale-list price ratio	(2) Exact or raised price sample Sale-list price ratio	(3) Full sample Ln(list price)	(4) Full sample Ln(sale price)
Online	−0.867***	1.142***	−0.010***	−0.004***
	(0.020)	(0.038)	(0.001)	(0.001)
Number of bedrooms	−0.157***	0.375***	0.199***	0.199***
	(0.007)	(0.018)	(0.000)	(0.000)
Number of bathrooms	−0.084***	0.295***	0.034***	0.033***
	(0.007)	(0.016)	(0.000)	(0.000)
Furbished (other)	−0.353***	0.412***	0.073***	0.075***
	(0.028)	(0.057)	(0.002)	(0.002)
New windows	−0.459***	0.628**	0.039***	0.041***
	(0.134)	(0.225)	(0.007)	(0.007)
New bathroom	−0.230***	0.567***	0.008*	0.008*
	(0.069)	(0.119)	(0.003)	(0.003)
New kitchen	−0.293***	0.220*	0.001	0.003
	(0.050)	(0.112)	(0.003)	(0.003)
New roof	−0.073	0.582*	0.021***	0.020**
	(0.126)	(0.230)	(0.006)	(0.006)
Garden	−0.366***	0.469***	0.022***	0.022***
	(0.012)	(0.028)	(0.001)	(0.001)
Garage	−0.318***	0.412***	0.070***	0.070***
	(0.010)	(0.024)	(0.001)	(0.001)
Fireplace	0.182***	−0.060	0.052***	0.050***
	(0.014)	(0.033)	(0.001)	(0.001)
Conservatory	−0.262***	0.297***	0.016***	0.018***
	(0.015)	(0.034)	(0.001)	(0.001)
Hard-flooring	−0.397***	0.483***	0.016***	0.017***
	(0.011)	(0.026)	(0.001)	(0.001)
Flat	0.906***	−0.520***	−0.490***	−0.495***
	(0.024)	(0.056)	(0.001)	(0.001)
Semi-detached	−0.222***	−0.072*	−0.270***	−0.265***
	(0.012)	(0.032)	(0.001)	(0.001)
Terraced	0.251***	−0.316***	−0.411***	−0.409***
	(0.014)	(0.035)	(0.001)	(0.001)
Other	4.838***	−6.862***	−0.302***	−0.325***
	(0.229)	(0.432)	(0.007)	(0.008)
Location FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	843,865	430,927	1,274,792	1,274,792

(Continues)

TABLE 8 (Continued)

	(1) Discounted price sample	(2) Exact or raised price sample	(3) Full sample	(4) Full sample
	Sale-list price ratio	**Sale-list price ratio**	**Ln(list price)**	**Ln(sale price)**
Adjusted R^2	0.095	0.029	0.784	0.784

Note: The dependent variables in Columns 1 and 2 are the *sale-list price ratios* as in Table 7. The dependent variable in Column 3, Ln(list price), is the log of the asking price on Zoopla. The dependent variable in Column 4, Ln(sale price), is the log of the paid price recorded in Land Registry. The results in the first (second) column are estimated using the subsample of properties whose paid price is lower (higher or equal to) than the asking price. See Table 3 for definitions of explanatory variables. Robust standard errors in parentheses.

$^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

From Figure 1, it can be seen that richer sellers are relatively more likely to select traditional agencies. Harding et al. (2003) and Steegmans and Hassink (2017) find that richer sellers face higher opportunity costs of their time. This implies that they will invest less time and effort in bargaining—leading to a shorter TOM.[16] As the income/wealth of sellers is an omitted variable, it follows that OLS will underestimate how much online agencies reduce TOM.

Our results in Figure 1 also show that younger sellers are relatively more likely to select online agencies. However, there is no clear evidence that younger sellers tend to have a different TOM or sale-list price ratio than other sellers. For example, Hayunga and Pace (2019) find that a first-time seller dummy variable is not significant for explaining TOM. Hence, the omission of the age of the seller in Equation (1), other things equal, should not cause systematic bias in the OLS results.

Omitted variables relating to the property itself (such as the quality of the building and finishes, ambiance and noise levels) may also affect the OLS results. Automated valuation models (AVMs) tend to overestimate the market price of properties that perform worse on the omitted variables. Given that online agencies are more likely to use AVMs to value properties rather than rely on a human appraiser, this may tilt sellers of properties that perform poorly on the omitted variables toward online agencies if they recommend a higher list price. These properties will tend to have a longer TOM and sale-list price ratio (as a result of the list price being set too high by the AVM). This perspective is consistent with the claim of Taylor (1999) and Chen and Rutherford (2013) that a longer TOM signals a lower quality property. Again, the implication is that OLS will underestimate how much online agencies reduce TOM and the sale-list price ratio.

Haurin et al. (2010) have argued that atypical properties will tend to have a longer TOM. Indeed, TOM is a measure of liquidity. To put it another way, more standard properties are more liquid and hence should have a shorter TOM. However, the standardness or atypicality of properties can be mostly detected from the observed variables in the OLS model. Hence this is not an omitted variables problem and should not cause bias in the OLS results.

In conclusion, our IV results confirm our earlier findings based on OLS and matching approaches that properties marketed by online agencies are sold quicker and with a smaller deviation from the original list price. However, once we correct for omitted variables, the impact of

[16] It is not clear what investing less time in bargaining means for the sale-list price ratio, as it could be either that sellers set a lower list price or that they accept a higher sale-list price ratio. These authors do not clarify which scenario is more applicable.

TABLE 9 Hedonic sale price regression (controlling for quarter and post town).

	(1) Ln(paid price)	(2) Ln(paid price)	(3) Ln(paid price)
Online	0.038***	−0.001	−0.004***
	(0.002)	(0.001)	(0.001)
Number of bedrooms		0.297***	0.199***
		(0.000)	(0.000)
Number of bathrooms		0.047***	0.033***
		(0.000)	(0.000)
Furbished (other)			0.075***
			(0.002)
New windows			0.041***
			(0.007)
New bathroom			0.008*
			(0.003)
New kitchen			0.003
			(0.003)
New roof			0.020**
			(0.006)
Garden			0.022***
			(0.001)
Garage			0.070***
			(0.001)
Fireplace			0.050***
			(0.001)
Conservatory			0.018***
			(0.001)
Hard-flooring			0.017***
			(0.001)
Flat			−0.495***
			(0.001)
Semi-detached			−0.265***
			(0.001)
Terraced			−0.409***
			(0.001)
Other			−0.325***
			(0.008)
Observations	1,274,792	1,274,792	1,274,792
Adjusted R^2	0.503	0.712	0.784

Note: The dependent variable in column Ln(sale price) is the log of the paid price recorded in Land Registry. Online is a dummy variable that equals to 1 if the property is advertised by an online agency, and 0 otherwise. *Garden, Garage, Conservatory*, and *Fireplace* are dummy variables representing the presence of these characteristics in the property (equal to 1 if these characteristics exist, and 0 otherwise). *Hard-flooring* is a dummy variable that takes the value 1 for the presence of wooden/laminate flooring in the property, and 0 otherwise. *New windows, New bathroom, New kitchen*, and *New roof* are dummy variables equal to 1, indicating whether the properties undergo these types of refurbishments, and 0 otherwise. Refurbished(other) is a dummy capturing any other refurbishments.

$^{***}p < 0.01$. $^{**}p < 0.05$. $^{*}p < 0.1$.

TABLE 10 Average treatment effects (ATT) based on coarsened exact matching (CEM).

	(1) Tom	(2) Sale-list price ratio	(3) Sale price
Online	−22.746***	−0.630***	−0.000
	(0.924)	(0.026)	(0.002)
Matched obs.	960,160	960,160	960,160
Unmatched obs.	314,632	314,632	314,632

Note: The same set of covariates in Table 6 is used in the matching process. Standard errors in parentheses ***$p < 0.01$.

TABLE 11 IV results.

	(1) Sale-list price ratio	(2) Sale-list price ratio	(3) Tom	(4) Tom
Online	−2.501***	−2.367***	−77.049***	−82.237***
	(0.241)	(0.203)	(8.443)	(6.907)
Number of bedrooms	0.044***	0.043***	16.069***	15.987***
	(0.010)	(0.010)	(0.374)	(0.378)
Number of bathrooms	0.200***	0.195***	9.935***	9.883***
	(0.010)	(0.009)	(0.428)	(0.421)
Furbished	−0.396***	−0.395***	−13.988***	−14.094***
	(0.039)	(0.039)	(1.312)	(1.324)
Flat	0.241***	0.238***	43.672***	42.955***
	(0.030)	(0.030)	(1.151)	(1.162)
Semi-detached	−0.505***	−0.513***	−25.559***	−25.918***
	(0.016)	(0.017)	(0.766)	(0.773)
Terraced	−0.227***	−0.238***	−14.693***	−15.183***
	(0.018)	(0.018)	(0.783)	(0.790)
Other	0.275	0.281	166.920***	164.758***
	(0.260)	(0.264)	(6.477)	(6.539)
Location FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Obs.	1,167,393	1,143,019	1,167,393	1,143,019
F-stat	5427.53	3851.11	5427.53	3851.11
Hansen J stat *p*-value		0.04		0.92
KP LM test statistic *p*-value	0.00	0.00	0.00	0.00

Note: The dependent variable in Columns 1 and 2, sale-list price ratio, is calculated by subtracting the paid price from the list price and dividing the difference by the list price (multiplied by 100). The dependent variable in Columns 3 and 4 is time on market. Columns 1 and 3 show the results estimated using the narrow-scale local market share (4-digits of postcode online share) as the sole instrument, whereas Columns 2 and 4 show results obtained by using both 4- and 3-digit postcode area market shares as instruments simultaneously. Refurbished is a dummy variable that takes the value of 1 if the description of the listing on Zoopla includes the term "refurbished," and 0 otherwise. Robust standard errors in parentheses.
***$p < 0.01$. **$p < 0.05$. *$p < 0.1$.

TABLE 12 First stage and reduced form regression *p*-values for the instrument.

Region	(1) First stage *p*-value for IV	(2) Reduced form_*p*-value for IV (Sale-list price ratio)	(3) Reduced form_*p*-value for IV (Tom)
Cambridge	0.238	0.045	0.838
Darlington	0.094	0.006	0.668
Exeter	0.071	0.004	0.049
Ilford	0.499	0.070	0.285
Lancaster	0.865	0.001	0.290
Llandrindod Wells	0.432	0.489	0.630
North West London	0.899	0.071	0.741
Oxford	0.092	0.273	0.946
Slough	0.709	0.759	0.819
Galashiels	0.462	0.157	0.171
Telford	0.080	0.159	0.421
Southall (West London)	0.202	0.083	0.541
Watford	0.113	0.616	0.778

Note: This table shows *p*-values from first-stage (Column 1) and reduced-form regressions for 13 postcode-area-level subsamples where the instrument is insignificant at the 5% level in the first stage. Reduced-form regressions show the effect of the instrument on sale-to-list price ratio (Column 2) and time on market (TOM) (Column 3), controlling for covariates.
Abbreviation: IV, instrumental variable.

using an online agent is substantially higher than indicated by the OLS and matching approaches. It should also be noted that OLS and IV address slightly different questions (Imbens & Angrist, 1994). IV measures the local average treatment effect. In our setting, this focuses on sellers whose decision whether to list with an online agent is sensitive to the share of sellers in the neighborhood that are using online agencies. By contrast, OLS focuses on the average treatment effect for all sellers.

5.6 | Share of online agencies and traditional agency performance

Table 13 presents the results from estimating Equations (3) and (4). These equations measure the relationship between the share of online agencies in the local markets and the performance of traditional agencies in terms of TOM and the sale-list price ratio. A 1 percentage point increase in the market share of online agencies leads to a 0.9 days shorter TOM (Column 1) and 0.02% reduction in the sale-list price ratio (Column 2) for properties listed by traditional agencies. In addition, results in Columns 3 and 4 show that, for properties whose paid price is lower (higher) than the list price, the magnitude of discount (increase) from the list price decreases (increases) with the market share of online agencies in the local area.

Table 14 displays the results at the agency and post-town levels obtained from estimating Equations (3)–(6). Although the coefficient magnitudes are smaller at the agency level, the signs of the coefficients align with the findings at the listing level. Specifically, a 1 percentage point increase in the market share of online agencies is linked to a 0.7-day decrease in the average TOM (Column 1),

TABLE 13 Online agency market share effect regression (at listing level).

	(1) Tom	(2) Sale-list price ratio (full sample)	(3) Sale-list price ratio (discounted price sample)	(4) Sale-list price ratio (exact or raised price sample)
Online agency market share	−0.862***	−0.024***	−0.012***	0.015***
	(0.060)	(0.002)	(0.001)	(0.004)
Number of bedrooms	17.471***	0.070***	−0.163***	0.383***
	(0.376)	(0.010)	(0.007)	(0.019)
Number of bathrooms	9.330***	0.153***	−0.091***	0.314***
	(0.406)	(0.009)	(0.007)	(0.017)
Degree of overpricing	58.466***			
	(1.034)			
Furbished (other)	−0.337	−0.149***	−0.379***	0.452***
	(1.612)	(0.035)	(0.030)	(0.062)
New windows	−17.405*	−0.153	−0.495***	0.587*
	(7.068)	(0.154)	(0.146)	(0.243)
New bathroom	−6.620	0.055	−0.257***	0.671***
	(3.891)	(0.082)	(0.075)	(0.133)
New kitchen	−9.610***	−0.094	−0.323***	0.325**
	(2.904)	(0.066)	(0.054)	(0.123)
New roof	−8.098	0.150	−0.058	0.646**
	(6.770)	(0.152)	(0.137)	(0.251)

(Continues)

TABLE 13 (Continued)

	(1) Tom	(2) Sale-list price ratio (full sample)	(3) Sale-list price ratio (discounted price sample)	(4) Sale-list price ratio (exact or raised price sample)
Garden	−10.549***	0.005	−0.373***	0.458***
	(0.623)	(0.016)	(0.013)	(0.030)
Garage	−12.638***	−0.019	−0.319***	0.416***
	(0.610)	(0.013)	(0.011)	(0.026)
Fireplace	−3.050***	0.162***	0.186***	−0.083*
	(0.816)	(0.018)	(0.015)	(0.035)
Conservatory	−5.160***	−0.182***	−0.282***	0.329***
	(0.978)	(0.019)	(0.015)	(0.036)
Hard-flooring	7.385***	−0.084***	−0.412***	0.494***
	(0.640)	(0.014)	(0.012)	(0.027)
Flat	42.683***	0.390***	0.901***	−0.530***
	(1.208)	(0.031)	(0.025)	(0.059)
Semi-detached	−26.650***	−0.490***	−0.232***	−0.066*
	(0.778)	(0.017)	(0.013)	(0.033)
Terraced	−17.188***	−0.184***	0.251***	−0.317***
	(0.830)	(0.019)	(0.014)	(0.037)
Other	185.361***	0.189	5.010***	−7.225***
	(6.995)	(0.273)	(0.238)	(0.459)
Location FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	1,201,381	1,201,381	797,861	403,520
Adjusted R^2	0.033	0.044	0.098	0.028

Note: All four models are estimated using the sample of properties advertised by traditional agencies. Online agency market share refers to the percentage of properties marketed by online agencies in the same quarter and county. The sale-list price ratio is calculated by subtracting the paid price from the list price and dividing the difference by the list price (multiplied by 100). The results in the third (fourth) column are estimated using the subsample of properties whose paid price is lower (higher or equal to) than their asking price. Robust standard errors in parentheses.
$^{***}p < 0.01$. $^{**}p < 0.05$. $^{*}p < 0.1$.

TABLE 14 Online agency market share effect regression (at agency and post-town levels).

	(1) Mean tom	(2) Mean sale-list price ratio	(3) Number of listings	(4) Number of traditional agencies
Online agency market share	−0.690***	−0.013***	−0.043***	−0.028***
	(0.094)	(0.003)	(0.002)	(0.005)
Location FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	217,318	217,318	217,318	19,056
Adjusted R^2	0.032	0.048	0.072	0.907

Note: The sampling unit in Columns 1–3 is each traditional agency in our data. The dependent variables in Columns 1–3 are the average sale-list price ratio, average tom, and the number of listings for traditional agencies in each quarter over the span of our data at the post-town level. The sampling unit in Column 4 is each post-town in our sample, and the dependent variable is the number traditional agencies operating in each quarter over the span of our data at post-town level. Robust standard errors in parentheses.

$^{***}p < 0.01$. $^{**}p < 0.05$. $^{*}p < 0.1$.

a 0.013 percentage point smaller reduction in the sale-list price ratio (Column 2), and 0.04 fewer listings (Column 3) for traditional agencies. Lastly, a 1 percentage point increase in the market share of online agencies corresponds to 0.03 fewer agencies in local markets.

5.7 | How the share of online agencies depends on local demographics

The relationship between the market share of online agencies and local demographic characteristics is explored in Figure 1 in the context of a regression model. We find that online market share is decreasing in average age, level of education, and increasing in the percentage of managerial professional occupations in an area.[17] The age effect is probably attributable to young people being more willing to try new online technologies. The other results show that sellers in richer regions are more likely to want a traditional agent that can devote more time and individual expertise to selling their property.

5.8 | How the share of online agencies varies with price

Figure 2 shows how the market share of online agencies varies across the cheaper, middle and more expensive parts of the market. For each year in our sample, we observe an inverted-U shaped curve, implying that online agencies achieve the highest market share in the middle part of the market. This finding can be attributed to a combination of two factors. First, there is the pricing structure. Online agencies charge a fixed fee, whereas traditional agencies charge a percentage fee. For properties at the cheap end of the market, therefore, the percentage fee may be more

[17] The result for managerial professional experience seems somewhat at odds with the other results. It should be noted, though, that this finding is controlling for all other variables. If we just compare how the percentage of managerial professional occupations correlates with the online market share without any controls, we find that the correlation is negative.

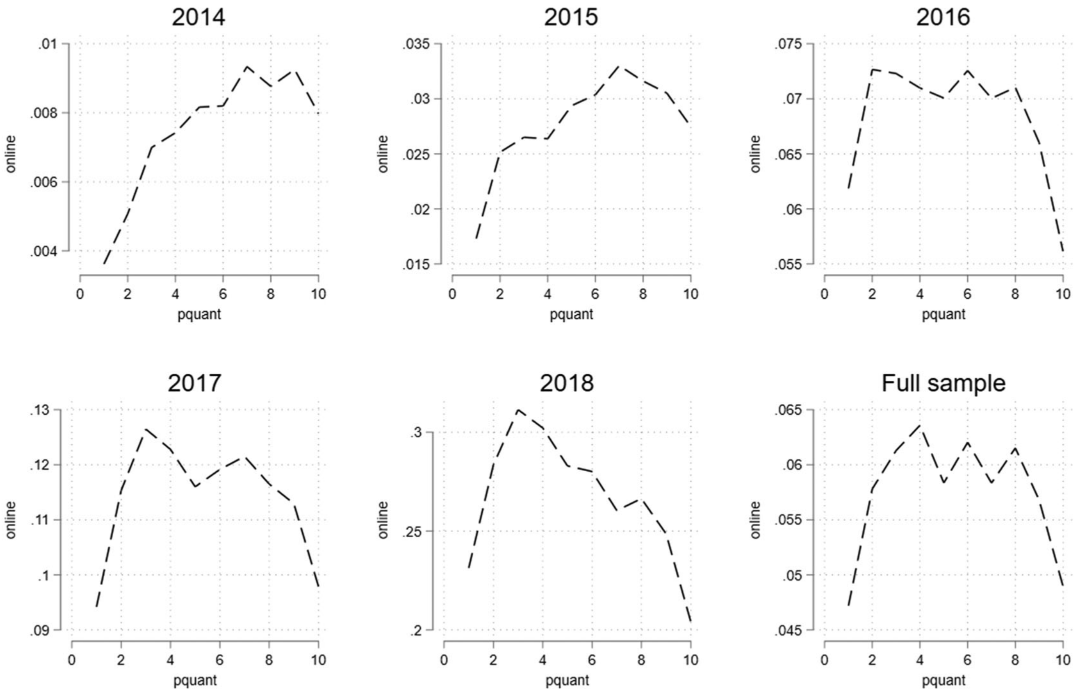

FIGURE 2 Market share of online agencies across price quintiles and years. This figure illustrates the share of listings marketed by the online agencies at each quintiles of the price distribution for each year and the full sample period.

attractive, causing sellers in this part of the market to prefer traditional agencies. For more expensive properties, a second factor comes into play. Such properties are more unique, and, as noted above, sellers may want an agent who can devote more time to marketing. Hence, sellers at the high end of the market are also more likely to select a traditional agent.

We formalize these ideas in a simple model in Appendix B that generates an inverted U-shaped curve for the relationship between price and online agency market share similar to that in Figure 2.

5.9 | Online agent performance at different quintiles of price, market size, and competition pressure

We further investigate the performance of online agencies at different price quintiles, market size and competition distribution in Table 15 by including interaction terms in Equations (1) and (2). Price quintile dummies are created for each year separately, considering the changes in house prices in this period. Market size is captured as the number of listings in the same outcode (first three digits of the postcode), and quintiles are created separately for each year. Finally, competition is defined by the count of traditional real estate agencies that have posted at least one listing within the same outcode.

In Panel A in Table 15, the reduction in TOM of online agencies is smaller for higher quintiles of the price distribution (first column), larger markets (second column) and in markets with more competition (third column).

TABLE 15 Interaction effects.

	Price quintiles	Market size	Competition
Panel A: Quintile dummies and online dummy interactions marginal effects on time on market			
Online × 1st quintile	−32.792***	−29.367***	−35.666***
	(1.976)	(1.959)	(1.924)
Online × 2nd quintile	−33.350***	−25.955***	−24.294***
	(1.592)	(1.889)	(2.008)
Online × 3rd quintile	−23.008***	−25.914***	−24.630***
	(1.750)	(1.836)	(1.836)
Online × 4th quintile	−16.727***	−20.442***	−20.634***
	(1.839)	(1.874)	(1.790)
Online × 5th quintile	−7.690***	−18.849***	−14.491***
	(2.235)	(1.768)	(1.836)
Observations	1,274,792		
Panel B: Quintile dummies and online dummy interactions marginal effects on the sale list price ratio			
Online × 1st quintile	−1.468***	−0.652***	−0.765***
	(0.069)	(0.047)	(0.048)
Online × 2nd quintile	−0.760***	−0.607***	−0.642***
	(0.044)	(0.046)	(0.047)
Online × 3rd quintile	−0.291***	−0.452***	−0.397***
	(0.040)	(0.046)	(0.045)
Online × 4th quintile	0.071*	−0.329***	−0.319***
	(0.036)	(0.047)	(0.044)
Online × 5th quintile	0.205***	−0.297***	−0.226***
	(0.035)	(0.046)	(0.048)
Observations	1,274,792		

Note: Price quintile dummies are created for each year separately. Market size is the number of listings in the same outcode (the first three digits of the postcode), and quintiles are created for each year separately. Competition is defined as the number of real estate agencies which posted at least one listing in the same outcode and quintiles that are created for each year separately. All estimations include the same set of controls as in Table 4. Robust standard errors are in parenthesis.
***$p < 0.01$. **$p < 0.05$. *$p < 0.1$.

The results in Panel B of Table 15 are similar for the sale-list price ratio. Online agencies even perform slightly worse than traditional agencies in the top two quintiles of the price distribution. In summary, the reduction in both TOM and the sale-list price ratio of online agencies is more pronounced for lower priced properties, smaller markets, and in the presence of weaker competition.

5.10 | Discussion

Our results show that properties listed with online agencies have shorter times on market and lower sale-list price ratios. Additionally, we distinguish between the properties whose paid prices are lower than their list prices (discounted price sample) and those whose paid prices are equal to or higher than the list prices (exact or raised price sample). Results from this subsample analysis show that the magnitude of the deviation from the list price is smaller for the properties marketed by online agencies in the discounted price sample, whereas the increase from the list price is larger amongst properties that are sold at prices higher than the list price. In this sense, sellers working with online agencies not only sell their properties quicker but also receive a closer price to the initial valuation of their property (higher prices if competing offers push the price upward).

We also find that the performance of online agencies is likely to have a spillover effect on their traditional counterparts, as our results indicate that the performance of traditional agencies, both in the form of TOM and price changes, improves when the market share of online agencies in the local market increases. As can be seen from Table 13, a 1 percentage point increase in online agency market share is associated with 0.9 days shorter TOM (Column 1) and a 0.02% decrease in the sale-list price ratio (Column 2). In addition, there is growing evidence that the emergence of online agencies has caused traditional agencies to substantially reduce their fees (Da Silva, 2023).

The difference in performance between online and traditional real estate agencies likely stems from their business models. The fee received by online agencies does not depend on the sale price, whereas traditional agencies receive a percentage of the sale price. This incentivizes traditional agencies to push for a higher list price and to accept a longer TOM than their online counterparts.[18]

Another important difference in business models is that most online agencies leave viewings to the seller while traditional agencies arrange a viewing by communicating between sellers and potential buyers. The simpler logistics of arranging viewings directly rather than through an agent may also reduce TOM.

More generally, the differences in incentives faced by traditional and online agencies, resulting from the way fees are charged, could also affect the effort exerted by agencies. Sellers need to trade-off the fees charged by agencies against the effort exerted. In Appendix B, we sketch a simple model in which for the middle part of the market, the lower fees of online agencies outweigh the loss from lower effort. However, beyond a certain point, the extra effort exerted by traditional agencies more than compensates for the extra fee they charge. This leads to the inverted U-shaped graphs for online market share as a function of price in Figure 2.

6 | CONCLUSION

In this article, we consider the impact of online real estate agencies on the housing market. Given the substantial market share achieved by online platforms such as Uber and Airbnb, there has been considerable interest in studying the performance of online firms and their effect on their traditional counterparts. However, the presence of online real estate agencies has not been explored in this literature.

In order to analyze the performance of online agencies, we focus on two measures, namely, TOM and the sale-list price ratio (the percentage point difference between the list price and the paid price). TOM can serve as an indicator of liquidity in the real estate market. For sellers, TOM shows how fast they can convert their real estate into cash, so it plays an important role in the decision to sell a property. The sale-list price ratio measures the deviation from the expected gain from the sale of a property. Naturally, sellers would like to minimize the difference between their list price and the paid price (except when competing offers push the paid price higher than the list price). Our results show that online agencies outperform traditional agencies in both of these indicators. Properties marketed by online agencies are sold quicker and reductions (rises)

[18] Denton (2015) provides some anecdotal evidence for this hypothesis.

in price are smaller (larger) for them. Apart from advantages in TOM and the sale-list price ratio, online agencies typically also charge lower fees.

Our article contributes to the growing literature on online platforms by providing evidence from the real estate market. The results of our work have important implications, especially for sellers and real estate agencies. For sellers, our results show that online agencies are not only cost-effective but also perform better in selling properties quicker and setting a more accurate list price. For traditional agencies, our results highlight the importance of accurate valuation and providing more flexible service packages, which can help them compete with online agencies, which have an obvious cost advantage due to operating exclusively online.

The penetration of online agencies varies across regions and in different parts of the market. Furthermore, an increased presence of online agencies in a region causes traditional agencies to raise their performance in terms of lower TOM, lower sale-list price ratios and lower fees. As in other disrupted markets, incumbents have been forced to adapt to survive. Even so, the number and share of transactions handled by traditional agencies is in decline (Table 2). Real estate agencies moving online, just like office workers increasingly working from home, is part of the broader trend toward declining demand for commercial real estate office space (Gupta et al., 2022; Van Nieuwerburgh, 2022).

ACKNOWLEDGMENTS

Open Access funding provided by Universitat Graz/KEMÖ.

REFERENCES

Acemoglu, D., Johnson, S., Robinson, J., & Yared, P. (2008). Income and democracy. *American Economic Review*, *98*(3), 808–842.

Agrawal, A., Gans, J., & Goldfarb, A. (2019). *The economics of artificial intelligence: An agenda*. University of Chicago Press.

Akerman, A., Gaarder, I., & Mogstad, M. (2015). The skill complementarity of broadband internet. *Quarterly Journal of Economics*, *130*, 1781–1824.

Anderson, J. T., Fuerst, F., Peiser, R. B., & Seiler, M. J. (2023). iBuyer's use of proptech to make large-scale cash offers. *Journal of Real Estate Research*, *46*(1), 114–135.

Angrist, J. D., Caldwell, S., & Hall, J. V. (2021). Uber versus taxi: A driver's eye view. *American Economic Journal: Applied Economics*, *13*(3), 272–308.

Angrist, J. D., Lavy, V., & Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, *28*, 773–824.

Ash, D., & Ben-Shahar, D. (2024). Zero-price effect and consumer welfare: Evidence from online classified real estate service. *Real Estate Economics*, *52*(5), 1165–1196.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, *118*, 1279–1334.

Barron, K., Kung, E., & Proserpio, D. (2021). The effect of home-sharing on house prices and rents: Evidence from Airbnb. *Marketing Science*, *40*(1), 2347.

Berger, T., Chen, C., & Frey, C. B. (2018). Drivers of disruption? Estimating the Uber effect. *European Economic Review*, *110*, 197–210.

Betz, T., Cook, S. J., & Hollenbach, F. M. (2018). On the use and abuse of spatial instruments. *Political Analysis*, *26*(4), 474–479.

Blackwell, M., Iacus, S., King, G., & Porro, G. (2010). CEM: Coarsened exact matching in stata. *Stata Journal*, *9*(4), 524–546.

Bloom, N., Hassan, T. A., Kalyani, J., Lerner, A., & Tahoun, A. (2021). *The diffusion of disruptive technologies* (Working Paper 28999). NBER.

Braesemann, F., & Baum, A. (2020). *PropTech: Turning real estate into a data-driven market? Oxford future of real estate initiative* (Working Paper, Version 1.3). Oxford Institute of Real Estate Initiative.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, *69*(5), 1127–1152.

Cavapozzi, D., Francesconi, M., & Nicoletti, C. (2021). The impact of gender role norms on mothers' labor supply. *Journal of Economic Behavior and Organization*, *186*, 113–134.

Chen, J., & Rutherford, R. C. (2013). Quality and time-on-the-market in residential markets. *Journal of Real Estate Finance and Economics*, *44*, 414–428.

Chen, W., Wei, Z., & Xie, K. (2022). The battle for homes: How does home sharing disrupt local residential markets? *Management Science*, *68*(12), 8589–8612.

Cramer, J., & Krueger, A. B. (2016). Disruptive change in the taxi business: The case of Uber. *American Economic Review*, *106*(5), 177–182.

Cvijanović, D., & Spaenjers, C. (2021). 'We'll always have Paris': Out-of-country buyers in the housing market. *Management Science*, *67*(7), 4120–4138.

Da Silva, M. (2023). *Purplebricks' 'failed' model blamed for lower estate agency standards*. Property Industry Eye. https://propertyindustryeye.com/purplebricks-failed-model-to-blame-for-lowering-estate-agency-standards-says-spicerhaart-boss/

Denton, J. (2015, April 9). Estate agents intentionally overvaluing homes as sellers ask for around £74k more than property is worth. *This is Money*. https://tinyurl.com/2kub297e

Fairweather, D., Kahn, M. E., Metcalfe, R. D., & Sandoval-Olascoaga, S. (2024). *Expecting climate change: A nationwide field experiment in the housing market* (Working Paper 33119). NBER.

France, B. (2021). May–June 5. Covid-19 and house prices in the UK: What can be learned from web-scraping data? Bulletin de al Banque de France, 235.

Fruehwirth, J. C., Iyer, S., & Zhang, A. (2019). Religion and depression in adolescence. *Journal of Political Economy*, *127*(3), 1178–1209.

Fuster, A., Plosser, M., Schnabl, P., & Vickery, J. (2019). The role of technology in mortgage lending. *Review of Financial Studies*, *32*(5), 1854–1899.

Genesove, D., & Mayer, C. (2001). Loss aversion and seller behavior: Evidence from the housing market. *Quarterly Journal of Economics*, *116*(4), 1233–1260.

Gindelsky, M., Moulton, J. G., & Wentland, S. A. (2019). Valuing housing services in the era of big data: A user cost approach leveraging Zillow Microdata. In K. G. Abraham, R. S. Jarmin, B. Moyer, & M. D. Shapiro (Eds.), *Big data for twenty-first-century economic statistics* (pp. 339–370). NBER, University of Chicago Press.

Gupta, A., Mittal, V., & Van Nieuwerburgh, S. (2022). *Work from home and the office real estate apocalypse* (Working Paper 30526). NBER.

Hall, J. D. (2018). Is Uber a substitute or complement for public transit? *Journal of Urban Economics*, *108*, 36–50.

Han, L., & Strange, W. C. (2016). What is the role of the asking price for a house? *Journal of Urban Economics*, *93*, 115–130.

Harding, J. P., Rosenthal, S. S., & Sirmans, C. F. (2003). Estimating bargaining power in the market for existing homes. *Review of Economics and Statistics*, *85*(1), 178–188.

Haurin, D. R., Haurin, J. L., Nadauld, T., & Sanders, A. (2010). List prices, sale prices and marketing time: An application to U.S. housing markets. *Real Estate Economics*, *4*(38), 659–685.

Hayunga, D. K., & Pace, P. K. (2019). The impact of TOM on prices in the US housing market. *Journal of Real Estate Finance and Economics*, *58*, 335–365.

Hughes, C. (2023, June 2). Purplebricks' £1 sale to rival strike backed by shareholders. *The Guardian*.

Iacus, S., King, G., & Porro, G. (2012). Causal inference without balance checking: coarsened exact matching. *Political Analysis*, *20*(1), 1–24.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467–475.

King, G., & Nielsen, R. (2019). Why propensity scores should not be used for matching. *Political Analysis*, *27*(4), 435–454.

Kolbe, J., Schulz, R., Wersing, M., & Werwatz, A. (2021). Real estate listings and their usefulness for hedonic regressions. *Empirical Economics*, *61*, 3239–3269.

Levitt, S. D., & Syverson, C. (2008). Market distortions when agents are better informed: The value of information in real estate transactions. *Review of Economics and Statistics*, *90*(4), 599–611.

Liu, M., Brynjolfsson, E., & Dowlatabadi, J. (2021). Do digital platforms reduce moral hazard? The case of Uber and Taxis. *Management Science*, *67*(8), 4665–4685.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, *60*(3), 531–542.

Moffitt, R. A. (2001). Policy interventions, low-level equilibria, and social interactions. In S. N. Durlauf, & H. P. Young (Eds.), *Social Dynamics* (pp. 45–82). MIT Press.

Nicoletti, C., Salvanes, K. G., & Tominey, E. (2018). The family peer effect on mothers' labor supply. *American Economic Journal: Applied Economics*, *10*(3), 206–234.

Oskam, J., & Boswijk, A. (2016). Airbnb: The future of networked hospitality businesses. *Journal of Tourism Futures*, *2*(1), 22–42.

Pfeiffer, L., & Lin, C.-Y. C. (2012). Groundwater pumping and spatial externalities in agriculture. *Journal of Environmental Economics and Management*, *64*(1), 16–30.

Pope, A. (2017). *GB postcode area, sector, district*. University of Edinburgh. https://doi.org/10.7488/ds/1947<./bib>

Robalino, J. A., & Pfaff, A. (2012). Contagious development: Neighbor interactions in deforestation. *Journal of Development Economics*, *97*(2), 427–436.

Rutherford, R. C., Springer, T. M., & Yavas, A. (2005). Conflicts between principals and agents: evidence from residential brokerage. *Journal of Financial Economics*, *76*(3), 627–665.

Sing, T. F., & Zou, Y. (2024). Improving search efficiency in housing markets using online listings. *Real Estate Economics*, *52*(4), 1020–1044.

Steegmans, J., & Hassink, W. (2017). Financial position and house price determination: An empirical study of income and wealth effects. *Journal of Housing Economics*, *36*, 8–24.

Sundquist, J. (2021). *How to use spatial instruments*. (Working Paper), Department of Political Science, Yale University. https://jamessundquist.github.io/website/docs/How_to_Use_Spatial_Instruments.pdf

Taylor, C. R. (1999). Time-on-the-market as a sign of quality. *Review of Economics Studies*, *66*, 555–578.

Van Nieuwerburgh, S. (2022). *The remote work revolution: Impact on real estate values and the urban environment* (Working Paper No. w30662). NBER.

Zervas, G., Proserpio, D., & Byers, J. W. (2017). The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. *Journal of Marketing Research*, *54*(5), 687–705.

How to cite this article: Gedikli, C., Hill, R., Talavera, O., & Yilmaz, O. (2025). Online real estate agencies and their impact on the housing market. *Real Estate Economics*, 1–40. https://doi.org/10.1111/1540-6229.70023

APPENDIX A
ADDITIONAL FIGURES AND TABLES

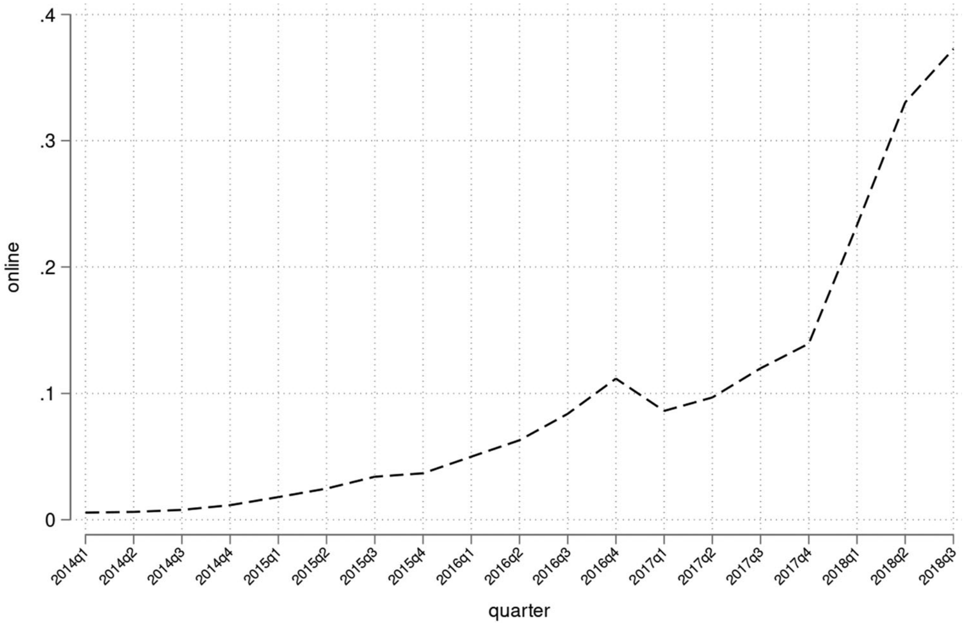

FIGURE A1 Market share of online agencies over time. This figure illustrates the share of listings marketed by the online agencies in each quarter between 2014 (when online agencies commenced their nationwide service in the United Kingdom) and 2018.

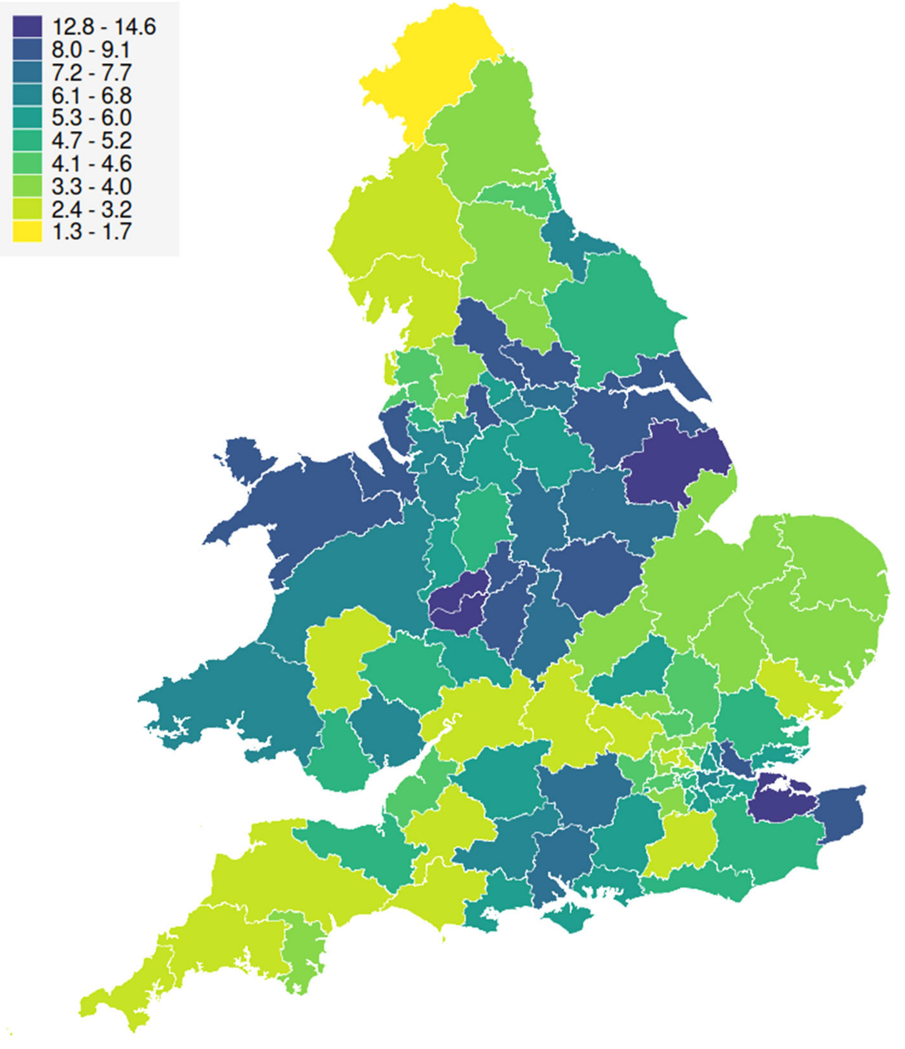

FIGURE A2 Market share of online agencies across England and Wales. This figure shows the share of listings marketed by the online agencies across England and Wales for the sample period from 2014 to 2018. The Map is created by Stata "spmap" command and by matching ZPD with shapefiles provided by Pope (2017) at postal area level (first two digits of the postcode). *K*-mean clusters are used for the cut-off points.

TABLE A1 2SLS first stage regression results.

	(2) Online	(6) Online
4-Digits of postcode online share	0.006***	0.004***
	(0.000)	(0.000)
3-Digits of postcode online share		0.005***
		(0.000)
Number of bedrooms	0.003***	0.003***
	(0.000)	(0.000)
Number of bathrooms	0.020***	0.020***
	(0.000)	(0.000)
Furbished	0.012***	0.012***
	(0.001)	(0.001)
Flat	−0.003***	−0.004***
	(0.001)	(0.001)
Semi-detached	0.009***	0.009***
	(0.001)	(0.001)
Terraced	0.009***	0.008***
	(0.001)	(0.001)
Other	−0.024***	−0.023***
	(0.004)	(0.004)
Location FE	Yes	Yes
Time FE	Yes	Yes
Obs.	1,167,393	1,143,019

Note: Columns 1 and 2 exhibit the results of the first stage regressions of the estimates presented in Table 9 where dependent variable online dummy takes the value of 1 if a property is marketed by online agencies and 0 otherwise. *4-Digits of postcode online share* is the share of listings marketed by online agencies in the area corresponding to the first part of the postcodes and the first digit of the second part of the postcode. *3-digits of postcode online share* is the share of listings marketed by online agencies in the area corresponding to the first part of the postcodes, excluding the listings in the 4-digits of postcode area. Both instruments are calculated using existing listings from the previous quarter. Refurbished is a dummy variable that takes the value of 1 if the description of the listing on Zoopla includes the term refurbished and 0 otherwise. Robust standard errors in parentheses.

*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

APPENDIX B
MODELLING THE SELLER'S DECISION OF WHERE TO LIST THEIR PROPERTY

We assume sellers choose to list their properties with either traditional or online agencies and that a seller's utility in each case is as follows:

$$U\,(online) = q_O[P_n\ -\ k\ -\ b\ TOM_O\ P_n]\ -\ (1\ -\ q_O)\,(k + Z)\,; \tag{B1}$$

$$U\,(traditional) = q_T[(1\ -\ t)\,(1 + E_n)\,P_n\ -\ b\ TOM_T\ P_n]\ -\ q_T Z. \tag{B2}$$

where the terms in the utility functions are defined as follows: q_O is the probability that a property listed with an online agent sells; q_T is the probability that a property listed with a traditional agent sells; P_n is the price at which property n would sell if listed with an online agent (assuming it actually sells); k is the fixed fee charged by an online agent irrespective of whether the property actually sells or not; TOM_T is the expected time on market for a property sold by a traditional agent; TOM_O is the expected time on market for a property sold by an online agent; b measures the implicit cost to the seller of 1 more day of time on market as a share of the (online) property price; t is the fraction of the paid price that a traditional agent charges when the property sells (e.g., $t = 0.01$ would mean that the agent charges a fee equal to 1% of the paid price). If there is no sale, the traditional agent does not charge a fee. Z is the loss of utility incurred by the owner when the property fails to sell. It is assumed Z is independent of whether the property was listed with an online or traditional agent. E_n measures the return in terms of selling price from the effort a traditional agent puts into trying to sell the property. That is, $P_s = (1 + E_n)P_n$, where P_s is the selling price achieved by a traditional seller, whereas P_n is the selling price of an online agent.

The real estate agent is assumed to be motivated to exert more effort by the prospect of a higher commission as P_n rises. We assume that E_n increases linearly with P_n as specified in the following equation:

$$E_n = -0.1 + P_n/2000. \tag{B3}$$

It follows that E_n is negative for properties with a price below the average level of 200 in Table 3 (i.e., 20,000 lb). In other words, our assumption is that for the average property (where $E_n = 0$) it makes no difference to the sale price whether it is sold through a traditional or online agent. Properties that are better than average sell at a higher price through a traditional agent, whereas properties that are below average sell at a higher price through an online agent.

Now, let ΔU denote the difference in utility of a seller from using an online agent instead of a traditional agent.

$$\Delta U = U\,(online)\ -\ U\,(traditional) \tag{B4}$$

Substituting Equations (B1)–(B3) into (B4), we obtain that ΔU is a quadratic in P_n. Under the assumption that the probability of a sale is independent of the choice of agent (i.e., $q_O = q_T = q$), this function takes the following form:

$$\Delta U = \alpha\ {P_n}^2 + \beta\ P_n + \gamma, \tag{B5}$$

where $\alpha = -q\,(1 - t)/2000 <,\ \beta = q\,[1 - 0.9\,(1 - t) + b\,(TOM_T - TOM_O)] > 0, \gamma = -k < 0.$

$\beta > 0$ follows from the fact that $\mathrm{TOM_T} > \mathrm{TOM_O}$ (in Table 9).

The function ΔU in Equation (B5) has an interior maximum and two positive solutions for P_n where $\Delta U = 0$. By implication, the prediction from the model is that owners' of properties at the low end and high end of the market will prefer traditional agencies, whereas those in the middle will prefer online agencies. Intuitively, at the low end of the market, owners prefer traditional agencies because the gain from a lower fee outweighs the lower effort exerted by the agent. At the high end of the market, the gain from the effort exerted by traditional agencies outweighs the cost of the higher fee. In the mid range, the trade-off between fee and effort favors online agencies.

From our data we can provide indicative values for the parameters of the model so as to provide estimates of the two thresholds at which $\Delta U = 0$. For example, we know that $k = 1$ (in thousands of pounds) and $t = 0.0142$. From Table 9 we obtain that $\mathrm{TOM_T} - \mathrm{TOM_O} = 80$. In addition, we assume that $q_\mathrm{O} = q_\mathrm{T} = 0.75$ (i.e., a seller listing a property stands a 75% chance of actually selling it). We set $b = 0.0004$. This implies that 80 days less on market is valued at £6400 for a £200,000 property.

Substituting these values into Equation (B5), using the quadratic formula we find that online agencies are preferred by owners selling properties in the price range (£9500, £284,200). Even though all properties would be above the lower bound of £9500, this example still provides an indication of why we observe an inverted U-curve in Figure 2 (i.e., online agencies are most preferred in the middle part of the price distribution).