THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Essays on Learning and Information-processing in Financial Markets

SIMONA RISTESKA

A thesis submitted to the Department of Finance of the London School of Economics and Political Science for the degree of *Doctor of Philosophy*

July 2021

Declaration of Authorship

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

Statement of conjoint work
I confirm that Chapter 2 is jointly co-authored with Francesco Nicolai. I contributed 50% of the work for Chapter 2.
I confirm that Chapter 3 is jointly co-authored with Francesco Nicolai and Marco Pelosi. I contributed 33% of the work for Chapter 3.

I declare that my thesis consists of 45,857 words.

Acknowledgements

I am deeply grateful to my advisors Christian Julliard and Michela Verardo for their encouragement and inspiration throughout this endeavour. The numerous occasions on which I have reached out to them for advice have greatly helped me carry this dissertation through and grow as a researcher. I would also like to extend my sincere gratitude to Daniel Ferreira and Amil Dasgupta for their invaluable mentorship and support.

I thank Ashwini Agrawal, Ulf Axelson, Cynthia Balloch, Mike Burkart, Thummim Cho, Vicente Cuñat, Juanita Gonzalez-Uribe, Dirk Jenter, Ralph Koijen, Paula Lopes-Cocco, Dong Lou, Igor Makarov, Ian Martin, Martin Oehmke, Daniel Paravisini, Cameron Peng, Christopher Polk, Walker Ray, Andrea Tamoni, Huan Tang, Dimitri Vayanos, Jean-Pierre Zigrand, and many others for their insightful comments.

Over the past few years at LSE I have certainly met some great friends who inspired me and made my PhD journey an experience to cherish. I am grateful to Lukas Kremens, Marco Pelosi, Karamfil Todorov, Andreea Englezu, Zhongchen Hu, Bruce Iwadate, Yue Yuan, Agnese Carella, Fabrizio Core, Alberto Pellicioli, Lorenzo Bretscher, Juan Chen, Kornelia Fabisik, Jesús Gorrin, James Guo, Brandon Han, Guido Maia, Olga Obizhaeva, Dimitris Papadimitriou, Bernardo Ricca, Gosia Ryduchowska, Petar Sabtchevsky, Amirabas Salarkia, Una Savic, Ran Shi, Arthur Taburet, Su Wang, Xiang Yin, for their insights and friendly chats.

I would like to thank the London School of Economics, and especially the Department of Finance, for the studentship that allowed me to pursue my PhD degree.

Most importantly, I express my biggest thanks to my family for their belief in me. I would have not been able to turn my ambitions into reality without the motivation and help of my grandparents, my parents Goran and Radica, and the unwavering support of my sister Tamara. Thank you for all the sacrifices you have made for me.

Last but not least, I thank my partner in life and crime Francesco for his tremendous understanding, patience and unconditional love. It's a joy to explore, debate and learn economics with you.

Abstract

This dissertation consists of three chapters studying how economic agents learn, form their beliefs and make economically relevant decisions. The main theme of the thesis is to infer beliefs from observable actions and test whether agents process information in a way that is consistent with various theoretical models.

In the first chapter I use price data from the real-estate market to infer agent beliefs that are consistent with their pricing behaviour. In particular, I study the way agents make inference from the observable actions of others. In the housing market, where the use of comparables for pricing is most common, I show that the inability to fully grasp the structure of information flows leads agents to overweight stale news due to their repeated use by intermediate agents. The findings are inconsistent with a fully Bayesian model and might instead be reconciled with a model of naïve learning.

The second chapter is conjoint work with Francesco Nicolai. We use data on mutual fund portfolio holdings to extract fund managers' stock return expectations in a fairly general model of portfolio formation. We employ panel regressions to partial out the effect of time-varying stock and manager characteristics and show that subjective expected returns are significantly affected by personal experience. In particular, we provide evidence that professional managers are more strongly influenced by recent returns and those experienced at the early stages of their holding period.

The third chapter, co-authored with Francesco Nicolai and Marco Pelosi, provides evidence of the disparity in the incidence of property taxes levied at different points in time. We show that housing demand is significantly less elastic to taxes deferred to the future relative to taxes levied at the moment of purchase, even after accounting for liquidity constraints. We attribute these findings to lack of salience, implying that the burden of deferred taxes will be borne in the future when they are levied. We develop a model to show that the trade-off between lack of salience and liquidity constraints gives rise to an optimal tax mix.

Contents

List of Tables

List of Figures

1. Learning from Past Prices: Evidence from the UK Housing Market

SIMONA RISTESKA[1]

The main objective of this paper is to investigate how economic agents make inference from the observable actions of others. In the face of uncertainty and imperfect knowledge about the state of the world, social learning allows individuals to learn private information that is embedded in other people's actions. However, in settings where individuals have limited knowledge about the structure of information flows, this may lead to inferential mistakes whenever agents fail to account for the repetitive use of stale information. A particular form of social learning commonly used in financial markets is referred to as pricing by comparables. When determining the value of an asset, individuals frequently draw on past observations of similar transactions to guide their decision-making. This valuation method is employed in a variety of settings, from firm valuation in corporate finance to the pricing of illiquid assets such as corporate bonds and loans and, perhaps most notably, in the housing market when assessing the value of commercial and residential properties.

In this paper, I show that sellers in the housing market overweight old information when setting prices. In particular, I use the UK market for residential housing as a laboratory to investigate how agents make inference when exposed to the release of new information. The institutional setting in the UK serves as an ideally suited natural experiment in these regards: beginning in March 2012, the UK Land Registry has been regularly publishing, on the twentieth working day of each month, data on all the transactions of residential properties that have taken place in the previous month. The regular release of price data provides me with a shock to the information set of prospective sellers around the latest publication date. Combined with rich data on listings from a large property website in the UK and data on house characteristics, this allows me to analyse the causal effect of recent transactions on property

[1]I benefited from helpful comments from Ulf Axelson, Cynthia Balloch, Daniel Ferreira, Dirk Jenter, Christian Julliard, Lukas Kremens, Dong Lou, Francesco Nicolai, Daniel Paravisini, Marco Pelosi, Cameron Peng, Walker Ray, Huan Tang, Michela Verardo, and the seminar participants at LSE, Copenhagen Business School, Adolfo Ibanez Business School, the New Economic School Moscow, BI Norwegian Business School, ESSEC Business School, Durham Business School, Nova Business School, Amsterdam Business School, Warwick Business School and WHU – Otto Beisheim School of Management. The paper contains HM Land Registry data © Crown copyright and database right 2019. The data is licensed under the Open Government Licence v3.0.1. I thank the University of Glasgow - Urban Big Data Centre for providing Zoopla property data. Zoopla Limited, © 2019. Zoopla Limited. Economic and Social Research Council. Zoopla Property Data, 2019 [data collection]. University of Glasgow - Urban Big Data Centre. Any errors or omissions are the responsibility of the author.

listings. I first supply empirical evidence that prospective sellers use data on past transactions to inform their decisions. Transactions from the previous month have a significantly larger effect on listings posted in the period after these have been made publicly available. I, therefore, confirm the well-known fact that pricing by comparables is widely used by sellers, as one should expect given the fact that this is an approach openly recommended by real-estate agents, property websites and other housing market professionals. Any given comparable transaction has about 0.45% incremental effect on listings that observe it relative to those that do not, even though the latter are closer in time. I then proceed to show that this is a lower bound on the true effect by conducting a difference-in-differences analysis where I benchmark the incremental effect described above to data before the first publication date in March 2012: the results indicate that the actual response is almost twice as large than the previously estimated one.

I then proceed to the main results of the paper: by looking at how the influence of any given transaction evolves with its repeated use, I demonstrate that sellers in the housing market fail to recognise potential duplication of information and are, therefore, prone to overweight stale news at the expense of more recent information. Specifically, I find that the effect of recently published transaction prices increases monotonically with the number of redundant channels of influence. The ability to observe the date and price of new property listings (hereafter also referred to as quote) allows for a detailed analysis of the way information flows from past prices to subsequent listings, potentially also via intermediate comparable listings. The results show that the incremental effect that recent transactions have on future quotes can be more than 3% when the number of intermediate comparables grows beyond three relative to the case where no such redundant channels are present. This incremental effect is added to the baseline influence that recent prices have on future listings of about 82-84% implying that housing market fundamentals are quite persistent and, consequently, even small pricing mistakes can have significant long-run effects. The above findings cannot be squared with Bayesian inference. In particular, a Bayesian agent would take into account the fact that recent comparables have been influenced by earlier ones and should, therefore, adjust the relative weights placed on observables accordingly in order to avoid double-counting stale information. This implies that the effect of a given comparable cannot increase as the number of interim listings grows. I, therefore, reject the null hypothesis that agents in the housing market behave in a Bayesian fashion.

The above results can be reconciled with a different learning model where agents fail to recognise potential duplication of information in prior observables, practice known in the theoretical literature on social learning as naïve herding or persuasion bias (Eyster and Rabin, 2010; DeMarzo et al., 2003). These papers show that, in order to make correct inference as implied by Bayesian updating, one needs to engage in a very complex process of discerning all the channels through which a given signal might have already exercised an indirect influence on their actions or else they would be inclined to overweight that piece of information. Specifically, the agent needs to be able to disentangle all determinants of a given observation, namely: (a) the private signal of that individual; (b) the part that is influenced by the observability of prior actions and; (c) the public information about fundamentals observed by everyone. This is not an easy task even in a setting with fully rational agents and common knowledge of the structure of the information network, nevertheless introducing uncertainty about information flows significantly exacerbates the problem. Agents who instead, due to bounded rationality, attempt to make approximate inference from past actions, by assuming that these are driven purely by distinct signals, are subject to naïve learning and risk placing too much weight on stale information.

Since sellers in the housing market have difficulty recognising potential duplication of information, it would be interesting to study whether this leads to inferior market outcomes. In particular, one could analyse if part of the mistakes made by sellers are corrected upon matching with buyers. In the final set of tests, I provide some suggestive evidence that this is indeed the case by showing that sellers who eventually sell their properties at the largest percentage difference to listed price are those that have been most highly influenced by past prices.

The empirical results provide solid evidence that learning and pricing behaviour in the real-estate market cannot be reconciled with Bayesian inference. They are, however, unable to demonstrate what the economic impact of such behaviour is on housing market dynamics in the long run. For this reason, I finally develop a simple model of learning to simulate the response of naïvely formed prices to various shocks and benchmark this to the rational case. The results indicate that in a world with naïve agents, prices are much more sensitive to noisy signals about demand as they overreact to this information for a long time. The deviation from fundamental values can be very large at 35% of the shock at a twenty-year horizon. On the other hand, naïve prices exhibit underreaction to true changes in the value of the underlying state due to the fact that real shocks get suppressed by stale news. These results are of particular importance once we consider that the decision to purchase a (new) home is typically one of the biggest financial decisions households need to make and, therefore, pricing mistakes can have large effects on their welfare.

In this paper, I provide empirical evidence regarding learning mistakes of sellers in the housing market. I, therefore, contribute to the literature on the behaviour of real-estate market participants and the way it affects pricing dynamics (Merlo and Ortalo-Magné, 2004; Brunnermeier and Julliard, 2008; Piazzesi and Schneider, 2009; Head et al., 2014; Ngai and Tenreyro, 2014; Merlo et al., 2015; Anenberg, 2016; Burnside et al., 2016; Davis and Quintin, 2017; Glaeser and Nathanson, 2017; Guren, 2018; Andersen et al., 2019; Giacoletti and Parsons, 2019; Bracke and Tenreyro, 2020). In particular, I expand on the results of Glaeser and Nathanson (2017) who calibrate a model where house market participants extrapolate from past prices by failing to adjust for the fact that past actions reflect beliefs about future demand. More broadly, this paper relates to the work of Murfin and Pratt (2019) who show that lenders in the market for corporate loans similarly overweight old information by treating past transactions as independent signals.

The paper proceeds as follows: in Section 1.1 I provide a theoretical foundation of naïve learning and outline the natural experiment that guides the empirical analysis; Section 1.2 provides a survey of the existing literature on naïve learning and housing market dynamics; Section 1.3 describes the data and shows summary statistics; Section 1.4 presents the results of the empirical analysis; Section 1.5 develops a model in order to convey the economic magnitude of the long-run effects arising from pricing mistakes, and; Section 1.6 concludes.

1.1 Theoretical Motivation and Methodology

When there is uncertainty about the state of the world and the amount of knowledge that other actors possess, agents are naturally inclined to use observable actions and outcomes as a way to make better informed decisions. One of the most obvious examples of this is the widespread use of comparables for pricing financial assets. Under this approach agents looking to determine the value of a given asset make use of available data on prices and transactions of similar securities[2]. When agents have less than full knowledge regarding the path through which information propagates, they are likely to incur in mistakes if they apply the comparables approach blindly. In particular, if agents do not account for common drivers among the set of observed comparables and, instead, treat these observations as independent from each other, they might overweight some signals at the expense of others. This practice is known in the literature on social learning as naïve herding. The theoretical literature on this topic, pioneered by DeMarzo et al. (2003) and Eyster and Rabin (2010), shows that agents who fail to account for common signals embedded in past actions are likely to make suboptimal choices and even herd on the wrong decision in the long run with positive probability. Even more surprisingly, Eyster and Rabin (2014) show that agents are required to anti-imitate, i.e., apply negative weight on the observable actions of some agents, in order to perform correct inference. My goal is to provide empirical evidence of the way that economic agents learn from past actions and examine whether there is any indication of naïve herding. I use the housing market as the setting for my analysis as this is one of the areas where the use of comparables

[2] Consider a simple asset with a periodic cash flow C, growth rate g and discount rate r. Its price P is then determined by the standard Gordon growth formula: $P = \frac{C}{r-g}$. Re-arranging, we obtain the price-to-cash flow ratio: $\frac{P}{C} = \frac{1}{r-g}$. This formula implies that assets with the same discount rate and growth rate (or difference thereof) should have the same value multiple. The approach of pricing by comparables thus relies on the availability of assets with similar risk and growth characteristics to the asset in question.

for pricing is most common. Moreover, the market for residential properties is, unlike most other financial markets, largely populated by households who might be less sophisticated compared to major actors in other security markets. As a result, the challenge of extracting the correct signals and avoiding any learning mistakes could be more difficult to overcome in the housing market where agents are present only temporarily and with possibly limited time and information resources.

To motivate the empirical analysis of this paper, I provide a simple stylised model that illustrates the key features of naïve learning and contrast it with Bayesian updating. Consider an environment where agents learn about the state of demand denoted by D. Prospective sellers looking to determine the listing price for their property receive a noisy signal s_n with a normally distributed error, $\varepsilon_n \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$ identically and independently distributed across agents and time:

$$s_n = D + \varepsilon_n \tag{1.1}$$

Agents act sequentially and every period each seller observes the entire history of prices: $I_n = \{p_{n-1}, p_{n-2}, ..., p_0\}$. For simplicity, suppose that sellers have a diffuse prior and set prices equal to their best estimate of the state of demand: $p_n(D) = E[D|s_n, I_n]$. The first agent $n = 0$ receives a signal s_0 and, not observing any prior actions, sets the price equal to the signal:

$$p_0 = E[D|s_0, I_0] = s_0 \tag{1.2}$$

Agent 1 receives a signal and observes the action of agent 0. Given the equal precisions, he assigns the same weight to both signals:

$$p_1 = E[D|s_1, I_1] = \left(1 - \frac{\sigma^{-2}}{(\sigma^{-2} + \sigma^{-2})}\right) \times \underbrace{E[D|I_1]}_{=p_0} + \frac{\sigma^{-2}}{(\sigma^{-2} + \sigma^{-2})} \times s_1 = \frac{1}{2} \times (s_0 + s_1) \tag{1.3}$$

The difference between Bayesian and naïve updating arises with the arrival of the third agent. A Bayesian approach would require the agent to calculate the posterior belief as the average of his prior and the new signal weighted by the signal precision, where the prior is equal to the expectation of demand given the history of observed actions:

$$p_2 = E[D|s_2, I_2] = \left(1 - \frac{\sigma^{-2}}{(\sigma^{-2} + \sigma^{-2} + \sigma^{-2})}\right) \times \underbrace{E[D|I_2]}_{=p_1} + \frac{\sigma^{-2}}{(\sigma^{-2} + \sigma^{-2} + \sigma^{-2})} \times s_2 \tag{1.4}$$

Note that agent 2's prior belief is equal to agent 1's posterior, i.e., under common knowledge of rationality and the informational structure, agent 2 simply sets his prior equal to agent 1's best estimate. Plugging in the expression for the posterior of agent 1 into equation (1.4), we obtain the following expression:

$$p_2 = \frac{2}{3} \times p_1 + \frac{1}{3} \times s_2 = \frac{1}{3} \times (s_0 + s_1 + s_2) \tag{1.5}$$

The key take-away from the last equation is that under Bayesian updating, agent 3 assigns appropriate weights to all previous signals in proportion to their respective precisions. More generally, the posterior of agent n is equal to the precision-weighted average of all $n+1$ available signals, i.e.:

$$p_n = \frac{1}{n+1} \times (s_0 + s_1 + ... + s_n) \tag{1.6}$$

Crucially, note that a simpler way to achieve this is by using the previous agent's posterior belief and disregarding all prior actions, i.e., the posterior belief of agent $n-1$ is a sufficient statistic for all previously observed information[3]:

$$p_n = \frac{n}{n+1} \times p_{n-1} + \frac{1}{n+1} \times s_n \tag{1.7}$$

The approach above, however, implies knowledge of the full history of actions and, most importantly, the way they have influenced each other. This might not be feasible in many real-world scenarios and might therefore lead to suboptimal decision-making in environments with social learning. To see this, consider a naïve learner in the third period, defined as one who fails to account for the redundancy of previous signals, treating past actions as independent instead. In other words, a naïve learner assumes that previous agents have not taken into account any prior information, forming prices based solely on their respective private signals. Agent 3's posterior is then given by:

$$\begin{aligned}\widetilde{p}_2 = \widetilde{E}[D|s_2, I_2] &= \left(1 - \frac{\sigma^{-2}}{(\sigma^{-2}+\sigma^{-2}+\sigma^{-2})}\right) \times \widetilde{E}[D|I_2] + \frac{\sigma^{-2}}{(\sigma^{-2}+\sigma^{-2}+\sigma^{-2})} \times s_2 \\ &= \frac{2}{3} \times \left(\frac{1}{2} \times p_0 + \frac{1}{2} \times p_1\right) + \frac{1}{3} \times s_2 = \frac{1}{3} \times p_0 + \frac{1}{3} \times p_1 + \frac{1}{3} \times s_2\end{aligned} \tag{1.8}$$

The above equation states that our naïve learner would use a wrong prior given by the precision-weighted average of the previous agents' *posteriors* as opposed to their

[3] Note that the last result holds only in the case where agents act sequentially and there is only one agent per period. If, instead, there are multiple agents in a given period, say k of them, who are unable to observe each other's actions, the prior of the subsequent set of agents will not be equal to any of those agents' posterior beliefs or the average thereof. This is because the former would imply failure to absorb the private information of the remaining k-1 agents from period n-1, while the latter would lead to overweighting of commonly observed signals relative to the private signals from period n-1. Nevertheless, the result that the posterior of a given agent equals the weighted average of all available signals still holds.

signals: $\widetilde{E}[D|s_2, I_2] = \frac{\sigma^{-2}}{(\sigma^{-2}+\sigma^{-2})} \times (p_0 + p_1)$. Plugging in the expressions for p_0 and p_1 from equations (1.2) and (1.3) above, we obtain:

$$\widetilde{p}_2 = \frac{1}{2} \times s_0 + \frac{1}{6} \times s_1 + \frac{1}{3} \times s_2 \tag{1.9}$$

The last equation shows that naïve updating leads agents to assign wrong weights on prior signals. In particular, by overlooking the influence that signals further in the past have had on more recent actions, agents end up overweighting stale news. In contrast to Bayesian learners, naïve agents treat signals coming from early actions as distinct sources of information. This mistake gives rise to multiple channels of influence from early news: the direct channel arising from the placement of an explicit weight on past signals and the indirect one that emerges from their effect on intermediate observations. For a general $n > 2$, the price looks as follows:

$$\widetilde{p}_n = \frac{1}{1 \times 2} \times s_0 + \frac{1}{2 \times 3} \times s_1 + ... + \frac{1}{n \times (n+1)} \times s_{n-1} + \frac{1}{n+1} \times s_n \tag{1.10}$$

Comparing equations (1.6) and (1.10) we see that naïve learning is an issue of relative over- and under-weighting of the signals coming from previous periods: notice that both Bayesian and naïve learners assign the same weight to their private information. It therefore arises even if agents are more confident in their own signals, as long as they learn from past data to some extent. It can further be noted that weighting mistakes would be present regardless of whether the signal precisions are equal across agents or not: naïve learning implies over-weighting of old signals relative to optimal weights even with heterogeneity in signal precisions.

The key distinctions between a Bayesian and naïve approach to learning outlined in equations (1.4)-(1.10) guide my empirical analysis going forward. Specifically, the core of the paper seeks to benchmark the way that prices in the housing market influence each other against the two learning models described above. Note that deviations from Bayesian learning occur once early observable information gets embedded into intermediate actions. Bayesian learning implies that the effect of a given signal should not increase with the number of subsequent uses, rather we should expect it to decline with the arrival of more news as each individual piece of information receives a progressively lower weight. To see this, we can fix a price from a given period $k \geq 0$ and compute its covariance with prices from subsequent periods $n \geq k+1$ under the rational and the naïve models. We can then compare the evolution of covariance functions as the number of intermediate observations grows. For the rational model, we have:

$$Cov(p_n, p_k) = \frac{1}{n+1} \times \sigma^2 \tag{1.11}$$

The above expression shows that, as we increase n or the amount of interim prices observed by agent n but not by agent k, the covariance between p_n and p_k monotonically decreases. This is intuitive since optimal learning implies that agent n will assign proportionally lower weights to the information embedded in p_k as he observes more and more recent news. On the other hand, the covariance of the same two prices in the naïve model would be as follows:

$$Cov(\widetilde{p}_n, \widetilde{p}_k) = \sum_{0 \leq i < k} \left(\frac{1}{(i+1) \times (i+2)} \right)^2 \times \sigma^2 + \frac{1}{(k+1)^2 \times (k+2)} \times \sigma^2 \tag{1.12}$$

Note that the covariance in the naïve case is no longer decreasing with n. In particular, in this simple setting covariances do not depend on n and, as a result, prices in all subsequent periods will comove with $\widetilde{p}_k$ by the same amount[4]. To better explain the empirical results of Section 1.4 below, I develop a dynamic model with an evolving state of demand and add an additional commonly observed public signal which introduces correlation in the signals: the details of the model are presented in Section A.3 of the Appendix.

In the rest of the paper, I make use of an ideally-suited setting for analysing the comovement in house prices with the arrival of intermediate observations that potentially contain overlapping signals. Namely, starting from March 2012, the UK Land Registry has been publishing monthly housing transaction data on a regular basis on the twentieth working day of the subsequent month. Consequently, on this date sellers receive an information shock due to the release of house price data from the previous month. Prior to March 2012, the data was available for purchase under contract and there was no such a sharp and regular discontinuity in the information set of sellers. Figure A1 gives a graphical representation of the environment. Suppose, for instance, that the twentieth working day of March of a given year is March 28th: this is the date when the February transactions data is made publicly available. Sellers who list their properties after this date can thus make use of the latest set of price data to inform their decisions. Sellers who have listed their properties just a few days before, however, are not able to observe the data on February transactions and thus cannot infer any private signals. Comparing the correlation of February transactions with properties listed just before and just after the publication date, therefore, gives us an idea of the effect of pricing by comparables in the housing market. Any incremental effect on quotes posted in the post-publication period shows evidence that sellers use information on newly published prices to learn about the current state of demand. The results of this exercise are shown in

[4]It is important to note that throughout my empirical analysis, I consistently compare the effect of prices from a given period on subsequent listings based on the amount of intermediate information by fixing k and varying n. In particular, if we were to fix n and vary k, i.e., analyse the effect of prices from different periods on the same quote, the covariance implications would be different. To see this, note that equation (1.11) shows that under the rational model the covariance of p_n with different past prices p_k does not depend on the amount of intermediate observations. On the other hand, under the naïve model it can be shown from (1.12) that the same covariance is decreasing in k, i.e., increasing with the number of intermediate observations.

Section 1.4.1 and can be interpreted as the direct effect of pricing by comparables: they, however, fall short of explaining whether sellers incorporate new information in an optimal or naïve way. For this reason, after having established the baseline effect, in Section 1.4.2 I look at the way that the comovement of quotes with a given price evolves through chains of influence from subsequent listings. Specifically, using data on property listings and their timing, in addition to the price paid data, we can get a good approximation of each agent's information set at the time of setting the quote. We can then compare the covariance between a given price p and a subsequent quote q based on the number of intermediate quotes that are observable by q which may or may not contain information also embedded in p. Note that the regularity in the price publishing dates provides a good setting for estimating covariances of prices with subsequent quotes by taking into account the amount of interim information that sellers possess. To the extent that the listings posted on the two sides of the price publication date do not differ in a systematic way, comparing the evolution of influence from recent prices through sequences of listings around this date allows us to benchmark the estimated covariance coefficients to the Bayesian and naïve models described above. This enables us to determine if sellers are able to correctly extract the real news from a given price or if they instead end up overweighting commonly contained signals due to their failure to understand the duplication of information.

Throughout the analysis, I investigate the impact of transaction prices on listings of properties with similar characteristics in order to avoid any selection on observables. Moreover, I minimise concerns regarding the evolution of fundamentals by looking at a very tight window of listings posted in the four weeks surrounding the publication date. Similarly, I compare the effect of prices of sold properties that have at least one comparable listing before and one after the publication date to make sure that the results are not driven by systematic differences in the independent variable. In the next section, I provide a review of some of the existing literature that relates to this paper in order to outline its main contributions.

1.2 Previous Literature

The present paper relates to two broad strands of literature. First, it provides empirical evidence that complements the large body of theoretical literature on social learning beginning with the models of herd behaviour and informational cascades by Banerjee (1992) and Bikhchandani et al. (1992) in a setting with rational agents. Both papers show that when agents move sequentially and everyone observes all prior actions, using past observations to learn the information other agents might have had can lead to so-called herd behaviour where Bayesian agents stop listening to their own signals and follow everyone else. This in turns makes each agent's action less informative about their own signal and thus less useful to others. Banerjee

(1992) demonstrates that the welfare implications of this type of behaviour can be significant to the extent that agents might gain by constraining information sharing. The type of positive feedback effects present in this setting implies that outcomes can be very different across game repetitions and that this might lead to excess volatility in asset markets. Bikhchandani et al. (1992) further go on to show that this type of cascades are fragile in the sense that they can seemingly break down in a drastic manner with the arrival of a small amount of information or a slight possibility of a value change. They also demonstrate that the gradual release of public information once a cascade has started can reverse this and eventually lead to individuals settling into the correct cascade. The above papers inspect only herding effects that result in social settings with rational inference. Convergence on the wrong action with fully rational agents is, nonetheless, rare, it occurs primarily in cases where agents are not confident in their beliefs and, as Ho (1993) and Smith and Sorensen (2000) show, it arises in situations with coarse action or signal spaces. One of the early papers that study bounded rationality in social-learning environments is DeMarzo et al. (2003) who introduce the concept of "persuasion bias" defined as the failure to adjust for possible repetition of information coming either from one source over time or multiple sources connected through a network. In their paper they emphasise that the key issue causing this type of behaviour is the intractability of the path that led all prior individuals to form their beliefs. Theoretical papers that most tightly relate to the present article are Eyster and Rabin (2010) and Eyster and Rabin (2014) who study observational learning in rich-informaton settings with naïve agents. Specifically, Eyster and Rabin (2010) describe a form of so-called "inferential naïvety" whereby players learning from the observable actions of others fail to account for the influence of early actions on interim players' choices and, instead, treat all observations as purely driven by each player's private information. Just like in the simple model presented in Section 1.1 above, agents in their model move sequentially after receiving a private signal and observing the full history of past actions. They demonstrate that this type of behaviour can lead agents to converge to the wrong beliefs with full confidence to the point that they are made worse off by being able to observe the actions of previous movers. Perhaps most crucially for the subsequent tests, Eyster and Rabin (2014) prove that rational learning implies that in environments where agents share common observations, they should either never imitate more than one predecessor or rather engage in anti-imitating behaviour as well.

In terms of empirical literature, a closely-related paper that studies naïve learning is Murfin and Pratt (2019) who look at the market for corporate loans. They exploit the date on which a given loan is reported in Refinitiv's Dealscan database to identify the effect of new additions to the dataset on the pricing of subsequent loans. They find strong evidence of comparables pricing in credit markets and naïve inference whereby the effect of a given comparable increases by three to five percentage points in the presence of redundant channels of influence, up from a baseline effect of about 6-10%. The benefit of the present paper is that it makes use of a more cleanly defined shock to the information set of agents to identify the direct effect of pricing by comparables. Moreover, while Murfin and Pratt (2019) study the behaviour of investment bank professionals, I primarily look at households operating in the residential housing market who might be less sophisticated and, consequently, more prone to influence and to committing pricing errors. Furthermore, the purchase of a home most often is the biggest financial decision that households make which emphasises the importance of any pricing mistakes. Less related to the present study, numerous other papers analyse the use of the comparables pricing method in corporate finance[5]. Papers that study herding behaviour in wider financial markets are, among others: Lakonishok et al. (1992), Grinblatt et al. (1995), Hong et al. (2005) and Dasgupta et al. (2011) who analyse institutional herding among money managers; Alevy et al. (2007), Ivković and Weisbenner (2007) and Wang and Wang (2018) who look at portfolio choices of retail and professional investors; Hong et al. (2004) and Brown et al. (2008) who study stock market participation among neighbours; Fracassi (2017) who looks at peer-effects among corporate managers; Bailey et al. (2018) who study social network effects on individuals' housing decisions.

Perhaps most relevant to the present work is Glaeser and Nathanson (2017)'s research on suboptimal learning behaviour in the housing market that looks at homebuyers who extrapolate from past transaction data by assuming that past prices are pure manifestation of contemporaneous demand. They develop and calibrate a model of house prices and demonstrate that it matches fairly well the short-term autocorrelation, as well as the medium-term reversal and excess volatility of house prices observed in the data. Most interestingly, they find that bubble-like features are most severe when buyers have decent amount of data about past prices but limited information about fundamentals. Although I similarly look at naïve inference in the housing market, a key distinction between my paper and the one by Glaeser and Nathanson (2017) is that I study the implications of pricing biases on the part of homesellers. Furthermore, I provide more detailed micro-evidence on the pricing patterns that result from naïve learning by employing a rich dataset of house prices and characteristics. Specifically, the ability to observe a good proxy for the information set of prospective sellers allows me to identify chains of influence and obtain empirical estimates for the indirect effects of past prices on future listings that arise under naïve learning. I subsequently use these estimates to calibrate the structural parameters and show the effect of various shocks to agents' information sets on aggregate pricing dynamics.

[5]See, for instance, Baker and Ruback (1999), Bhojraj and Lee (2002) and Liu et al. (2002) for the study of the performance of this approach in equity valuation, Kim and Ritter (1999) and Purnanandam and Swaminathan (2004) for evidence on the use of the comparable firms multiples approach in initial public offerings, and DeAngelo (1990) and Kaplan and Ruback (1995) on its use in the market for corporate control.

The paper also relates to the broader literature on housing markets trying to explain the behaviour of market participants and aggregate market dynamics[6]. Piazzesi and Schneider (2009) present a model where a small number of irrationally optimistic individuals can have a large price impact without the need to obtain a large market share. Head et al. (2014) develop and calibrate a dynamic search model that generates close to half of the serial correlation in house price growth. Burnside et al. (2016) propose a model with heterogeneous beliefs and social interactions to study the boom-bust cycles prevalent in housing markets. Anenberg (2016) presents a micro-search model where sellers facing information frictions update their beliefs about house values with the arrival of buyers. His model is able to match many of the micro features present in the data and can explain half of the short-term persistence in aggregate price dynamics. Guren (2018) proposes a mechanism that amplifies frictions through strategic complementarity, i.e., the willingness of sellers to set listing prices close to the cross-sectional average in order to optimise the trade-off between selling price and time on the market. He shows that this mechanism causes sluggish price adjustment by sellers and can magnify momentum by a factor of two to three.

Finally, this paper touches on the literature on extrapolative expectations and behavioural biases. Fuster et al. (2010) propose a dynamic model where agents form expectations that overestimate the persistence of economic shocks. Similarly, Barberis et al. (2015) study a consumption asset-pricing model where only a group of agents form beliefs by extrapolating from past returns. Both papers find that the model fits the data on aggregate economic and financial variables well. In a similar vein, Kuchler and Zafar (2019) use survey data to show that individuals extrapolate from personal experience when forming beliefs about aggregate outcomes such as house price changes and unemployment levels.

1.3 Data and Summary Statistics

In this section I describe the data I use for the empirical analysis. The data on house prices comes from the Price Paid dataset published by the HM Land Registry. This data contains information on transactions of residential properties in England and Wales starting from 1995 to the present. Apart from some exemptions[7] all transactions of residential properties that have been sold for full market value are recorded and made publicly available by the UK Land Registry. The Price Paid dataset pro-

[6]For a survey of the literature on the microstructure of housing markets, see Han and Strange (2015). For a review of the theoretical and empirical literature on house price dynamics, see Cho (1996) and Glaeser and Nathanson (2015).

[7]Transactions that are excluded from the Price Paid dataset include commercial transactions, property transactions that have not been lodged with the HM Land Registry and properties sold below market value. For more details on the property sales not included in the dataset the reader can visit the HM Land Registry website: https://landregistry.data.gov.uk.

vides information on the date of the transaction[8], the transfer price, the full address of the property, as well as some additional characteristics about the property such as: the age of the property, i.e., whether the property is a new construction or an existing building; the duration of the lease (freehold or leasehold)[9]; and the property type categorised as either a detached, semi-detached, terraced house, or a flat.

Data on listed properties and listing prices comes from the Zoopla Property data[10]. Zoopla is the second largest provider of property data for consumers and property professionals in the UK, having access to over 27,000,000 residential property records and 15 years of price data. The full dataset available for research purposes contains over 5,000,000 records of properties listed for sale and over 3,000,000 records of properties advertised for rent. Zoopla's website is one of the most commonly used in the UK for listing properties for sale, second only to Rightmove but expanding in market coverage. The data mainly covers the period between 2009 through 2018 for properties located in Great Britain, with partial coverage from 2005. The key variables for my empirical work are the quoted prices along with the dates at which these have been updated. The data also gives information about the date on which the property has been initially listed and the date on which it has been withdrawn from the market. The above information is crucial to my empirical analysis as the goal of investigating the impact of newly available prices on new listings requires me to have a precise idea of the moment in time when listing prices are set/updated and the information set of sellers at the time. In addition, this dataset contains other property characteristics such as property location, property type[11], whether the property has been categorised as residential or commercial[12], number of bedrooms, number of reception rooms, number of bathrooms, number of floors and whether the property is listed for sale or for rent[13].

The final piece of data I use in my empirical analysis is the Domestic Energy Performance Certificates dataset from the Ministry of Housing, Communities and Local Government. Before 2008, the Energy Performance Certificates (EPC) for domestic

[8]This is the completion date of the sale as stated on the transfer deed.

[9]Note that first registration of leases for seven years or less are not recorded in the dataset.

[10]The access to the dataset has been provided by the University of Glasgow - Urban Big Data Centre. Access to the dataset for research purposes can be obtained directly through the Urban Big Data Centre. The data has been collected by Zoopla. Zoopla Limited, © 2019. Zoopla Limited. Economic and Social Research Council. Zoopla Property Data, 2019 [data collection]. University of Glasgow - Urban Big Data Centre.

[11]Property types include: barn conversion, block of flats, bungalow, business park, chalet, château, cottage, country house, detached bungalow, detached house, end-terrace house, equestrian property, farm, farm house, finca, flat, hotel/guest house, houseboat, industrial, land, leisure/hospitality, light industrial, link-detached house, lodge, longère, maisonette, mews house, mobile/park home, office, parking/garage, pub/bar, restaurant/cafe, retail premises, riad, semi-detached bungalow, semi-detached house, studio, terraced bungalow, terraced house, town house, unknown, villa and warehouse. For my analysis, I focus on the following property types: detached house, end-terrace house, flat, link-detached house, maisonette, mews house, semi-detached house, terraced house, town house, studio and villa.

[12]I keep only properties categorised as residential.

[13]I exclude properties listed for rent from my sample.

properties could be lodged on a voluntary basis. From 2008 onwards, however, it has become mandatory for accredited energy assessors to lodge the energy certificates. Consequently, the data coverage drastically improves around that time, as does my ability to match these with the Price Paid and Zoopla data. More specifically, the matching rate goes from a little over 50 percent in 1995 to over 90 percent around 2008. The dataset contains information on the address, property type, total floor area, number of storeys, number of rooms, floor level and height, along with many indicators of energy efficiency and quality of glazed surfaces.

Figure A2 displays heat maps of the spatial coverage of the data across England and Wales. Figures A2a and A2b show, for every year, the total number of transactions and listings, respectively, by local authority district. Comparing the two sets of maps, it can be noted that they display similar patterns and thus the listings data closely matches the true sales activity in the UK housing market. However, it can be seen that the Zoopla sample mainly covers the period between 2010 and 2017, with very few observations in 2009 and 2018. Figures A17a and A17b in the Appendix confirm this by showing the ratio of listings-to-transactions and the fraction of transactions whose listing information could be found in the Zoopla data across areas as a way of demonstrating the relative Zoopla coverage. As can be seen, between twenty and eighty percent of transactions are matched to their respective listings in the Zoopla dataset across most regions in the period between 2010 and 2017 with the coverage peaking between 2011 and 2016. Nonetheless, it is reassuring to know that the data is well-dispersed across space and time as this reduces the probability that the results presented later in the paper are driven by a small subsample unrepresentative of the aggregate dynamics of the UK housing market.

As I seek to investigate the effect of using the comparables pricing method in the housing market, most of my empirical work requires me to match listings with recent transactions of properties with similar characteristics. The goal is to replicate the natural approach that a seller would take when deciding at what price to list their property. For this purpose, I match listings to recently sold houses based on four criteria: (1) the property location measured using the first half of the postcode[14]; (2) a rural/urban indicator from the 2011 Census classification of Output Areas; (3) property type divided in four categories, these beeing a detached house, semi-detached house, terraced house and a flat, and; (4) number of rooms in the property[15,16].

[14]Postcodes in the UK are formed of five to seven alphanumeric characters and are typically split into two parts: the outward code and the inward code. In my work, I compare properties that have the same outward code which corresponds to properties that belong to the same subdistrict.

[15]The number of rooms variable of choice comes from the EPC dataset and it includes any living room, sitting room, dining room, bedroom, study and similar, a non-separated conservatory with an internal quality door and a kitchen/diner with a discrete sitting area. Excluded from the count are rooms used solely as a kitchen, utility room, bathroom, cloakroom, en-suite accommodation and similar, any rooms not having a window and any hallway, stairs or landing.

[16]I group into one room category properties having between six and ten rooms. Similarly, all

Table A1 provides summary statistics for the sample of listings and transactions that have at least one match and, therefore, form part of the empirical analysis. The main sample covers the period after March 2012, the date when the Land Registry began publishing monthly Price Paid data on a regular basis. However, the data before March 2012 is used in some of the robustness checks and thus I separately present summary statistics for this part of the sample for comparison. I remove observations where the listing or transaction price is below £10,000 or above £25,000,000 to make sure that outliers do not drive the results. I also eliminate properties that have more than twenty rooms as well as observations with no rooms. For the final sample, I end up with 1,983,528 listings and 2,521,505 transactions deemed comparable post March 2012; prior to March 2012, there are 1,007,942 such listings matched with 986,287 recent transactions. Looking at the price statistics, we can observe that quoted prices tend to be larger than transaction prices: post March 2012 the average price at which a property is listed equals £268,402, while the average price paid for a property is £256,734. In the earlier part of the sample both are slightly lower at £233,497 and £220,134, respectively, which is natural given that real-estate prices normally exhibit a positive trend. The data also confirms the positive skewness in house prices with the median listing and transaction prices being significantly lower than the average at £194,950 and £189,995, respectively, in the sample from 2012 onwards. The above results are consistent with findings from the previous literature (Merlo and Ortalo-Magné, 2004; Carrillo, 2012; Han and Strange, 2016; Guren, 2018). Second, I divide the data based on time-invariant property characteristics in order to show that the sample is well-balanced both across sets (listings and transactions) as well as across sample periods. In particular, about 15% to 19% of the properties in my sample are detached houses, 28% to 29% semi-detached houses, 31% to 34% terraced houses and the remaining 20% to 24% are flats. The average property has between four and five rooms and this is consistent across sample periods.

As part of my analysis focuses on ways that any potential mistakes made by sellers when setting quotes could be rectified by buyers at the selling stage, I also attempt to match listings to their respective ex-post transactions. To achieve this, I first match the data from Zoopla with the Price Paid data by property address; I then keep only the matches for which the transactions occurs at least four weeks[17] and no more than five years after the property has been listed on the market; I finally eliminate cases where the sale price is more than 50% above or below the final quote for that listing. This procedure leaves me with a sample of 2,086,462 listings

properties with more than ten rooms are also considered comparable to each other.

[17]Discussions with real-estate agents and Zoopla information suggests that, due to the lengthy conveyancing process, it on average takes about six weeks to complete a freehold sale and eight to ten weeks a leasehold one, but that this can go down to as little as a couple of weeks. As the Price Paid data contains the date when the sale has been completed, I take a conservative approach and remove occurrences with less than four weeks between listing and completion.

matched to their transactions for the period between 2009 and 2018. Figure A3 displays the distributions of the price differential and time on the market (TOM) for the set of matched properties. The mean and median values of the two distributions are represented by the green and blue vertical lines, respectively. Looking at Figure A3a, it is notable that the distribution of the percentage difference between the listed price and sale price exhibits a large spike at zero, namely, over 10% of the matched transactions occur at the ask price, consistent with the findings in Merlo and Ortalo-Magné (2004) and Guren (2018), among others. This suggests that the process of determining the listing price is very important given that, although buyers can negotiate the final price with sellers, this one often ends up being equal or very close to the quoted price. Unsurprisingly, we see that the price discount distribution is very asymmetric around zero with most of the properties being sold below the listed price and only about 12% being sold at a premium. The average and median properties in the sample sell 4.68% and 3.83% below listed price, respectively[18]. With regard to TOM, Figure A3b shows that properties sell within 28 weeks on average, with the median property selling within 21 weeks of listing[19]. It is also reassuring to see that 99% of the properties in the sample sell within no more than two years which confirms the matching quality. In Figure A18 of the Appendix I plot the time-series of the average and median price discount and TOM. We can note that there is a positive correlation between discount and time spent on the market although it seems that TOM is less sensitive to market conditions.

Table A2 provides summary statistics for the set of listings matched to their respective transactions in the restricted sample used in the later regressions, i.e., listings post March 2012 that have at least one comparable transaction in the prior month. Contrasted to the full sample of listings and transactions in Table A1, the matched sample is pretty similar across all characteristics, although it contains around 3% more houses, and consequently larger properties, at the expense of flats compared to the post 2012 sample in Table A1. The mean and median price discount equal -3.78% and -3.13%, with the average and median TOM being 26 and 21 weeks, respectively. Coupled with the results in Figure A18 in the Appendix, we can conclude that the majority of the observations in the test sample come from periods of hot housing markets. Thus, it would be interesting to see how agents behave in response to new information in times of moderate to good market conditions and contrast this with findings of previous papers that focus mostly on times of depressed housing markets (Anenberg, 2016).

To provide further evidence on the effectiveness of looking for comparable transactions that match across location, property type, number of rooms and time, I next

[18]The average and median discount with respect to the final quote equal -3.06% and -2.73%, respectively, which suggests that most price changes are likely to be downward revisions.

[19]It is important to bear in mind that this is the time difference between listing and sale completion. Taking into account the average time it takes to finalise a sale, we can conclude that the average (median) seller finds a buyer in about 21 (14) weeks.

show the fraction of the variation in prices that is explained by various characteristics. Figure A4 displays the R-squared obtained by regressing prices on the above fixed effects, separately for listings and transactions. I sequentially increase the number of fixed effects in the regressions in order to discern the incremental improvement in explanatory power. Comparing Figures A4a and A4b, we can note that the explanatory power of the various property characteristics and time effects is very similar for listing and transaction prices. Starting from the bottom, the date of the listing/transaction measured in months explains about 1% of the price variation - this is not surprising as we compare houses of very different types, size and location across the entire England. The combination of house innate characteristics such as type and number of rooms has a considerably larger explanatory power of about 12-13%. Location is by far the most important determinant of house prices, explaining close to 38% of the variation in prices alone, close to 45% of this variation when coupled with time effects and about 55% when combined with property type effects. This is in line with the well-known fact that location is the key feature driving property values. As we increase the number of fixed effects the R-squared gradually increases, reaching over 85% for the full set of characteristics used in the matching exercise. This result gives validity to the comparables search method I employ in my empirical work by re-affirming the assertion that the pairs of listing and transaction prices I use in the regressions are indeed largely driven by common variables. Consequently, any incremental effect of recent transaction prices on listings in the post-publication period that I find in the data would arise mainly on the account of changes to the informaton set of sellers. Figure A19 in the Appendix shows the explanatory power that the same set of characteristics have for the variation in the absolute and percentage price discount for the sample of transactions matched to their respective listings. We can note that the R-squared is considerably lower across most specifications and, in particular, time-invariant house effects explain just 25% of the variation in the dollar differential and less than 13% of the variation in the percentage price discount. Location is a much less important factor for price discounts explaining less than 10% of the variation. On the other hand, in contrast to the variation in price levels, the variation in price differences is much more significantly driven by time effects, re-asserting the conclusion that the price discount is an aggregate feature of the housing market that evolves similarly across properties of different types and location. The combination of all fixed effects achieves an R-squared of 78% and 68% for the pound and percentage price difference, respectively.

Before I proceed to present the results of my empirical analysis, I briefly investigate any potential selection biases that we might have to be aware of. First of all, it is important to note that the final sample used in the regressions below considers only transactions that serve as comparables to at least one listing posted before and one after the date when the price data is released. In this way, we can avoid potential concerns regarding systematic differences in the independent variable between the sets of treated and untreated listings. The results in the next section should therefore be interpreted as the incremental effect that the same set of prices have on subsequent listings following the Price Paid data publication dates. Accordingly, the only reason why the effect might be different comes from the discontinuity in the information set of sellers around these dates.

In Figure A5 I provide evidence on the similarity across the sets of treated and untreated listings. Namely, I regress listed prices on a dummy for the signed number of days between the listing and the closest price data publication date, by adding the usual fixed effects used in the matching process[20]. The figure shows that there is little variation in the prices at which properties are listed, controlling for house characteristics. On all days but two we cannot reject the null hypothesis that prices insignificantly differ relative to those of properties listed on the publishing date. Even on the two days where this difference is significant, it is never larger than £2,000. Furthermore, in untabulated analysis I regress listing prices on a dummy for treated, that is, I compare the price levels of listings occurring before and after the publication date[21]. The results show that there is no significant trend in listing prices around publishing dates: the coefficient on treated suggests that listing prices in the post publication periods are about £300 pounds larger, however, the p-value is above 0.15.

Finally, I investigate any potential selection into treatment by running a density test. Specifically, I test for a possible discontinuity in the density of observations in the days around price publishing dates. Examining the Zoopla data in more detail, however, shows that listings exhibit a strong pattern in terms of week days, with a lot of activity in the middle of the week (Tuesday through Friday) and significantly less listings being posted on weekends and Mondays. For this reason, I first regress the count of observations per date on days of the week dummies and conduct a McCrary test (McCrary, 2008) using the residuals from this regression. The results of the density test are shown in Figure A6. As is evident from the picture, there does not seem to be manipulation of the running variable around price data publication dates. The shape of the density function is fairly smooth without exhibiting a jump on the treatment day. This is confirmed by the p-value of the test which is equal to 0.638, well above the significance threshold.

Equipped with the above reassuring evidence, I now proceed to the next section where I present the empirical results of my study.

[20]Figure A20 in the Appendix plots the results of the same test for the sample period before March 2012 used in some of the robustness checks below.

[21]Specifically, I run the following regression: $q_i = \alpha + \beta \times Treated_i + FE + \varepsilon_i$, where the fixed-effects correspond to the characteristics the matching is based on, i.e., location, property type, number of rooms and month-year, and $Treated_i$ is a dummy that equals zero for listings that occur in the two weeks before the new transaction data is published and one for those that are posted in the two weeks after.

1.4 Results

The main goal of the paper is to investigate if sellers in the housing market are able to extract and use information from past prices in the optimal way or perhaps, due to the complexity of the chains of inter-influence among recent comparables, they are prone to double counting repeated information at the expense of real news. For this reason, I first provide some evidence that the comparables pricing approach is indeed used in this market by exploiting the shock to sellers' information sets that occurs on each Price Paid data publication date. I then present results on the indirect effect of past prices on future listings that arises due to repeated use of this approach by a sequence of sellers. Finally, I address the question of whether any mistakes made due to suboptimal learning in such an imperfect environment are corrected at the selling stage by looking at the sample of listings matched to their respective subsequent transactions.

1.4.1 Evidence of Pricing by Comparables

Before analysing how agents process newly released information when setting house prices we need to make sure that past prices of similar properties significantly affect their decisions. The housing market in the UK provides a natural experiment for testing this hypothesis, namely, whether agents behave differently after they have been exposed to the most recent set of transactions in their market of interest. Recall from Section 1.1 and Figure A1 that, starting from March 2012, the Land Registry publishes monthly transaction data for the previous month on the twentieth working day of each month. On these dates, sellers receive a shock to their information set. Specifically, in the days leading to the publication date, sellers, real-estate agents and other property professionals have access to the most recent prices only if they have been directly involved in the transaction or if they have access to other sources of private information. After the publication date, everybody can potentially observe the full set of transactions that have taken place in the previous month. In other words, individuals who list their properties before the twentieth working day of the month may not directly observe the prices at which similar properties have been sold in the past month, while those who do so after the publishing date will have access to this information. Notice that while sellers may not be aware of the release of information, recent prices are usually immediately incorporated in the statistics available on common property platforms such as Zoopla. This implies that the seller becomes inadvertently a user of the newly released data as long as he is guided by the information on these platforms. The sellers' lack of knowledge regarding the publishing dates makes the discontinuity in the information set less sharp but alleviates the concern that sellers strategically select when to list which further explains the results in Section 1.3. Together with the fact that sellers might have access to private information about recent transactions, this consideration implies that the findings in this section represent a lower bound for the true effect of newly released information on prices.

I start by comparing the effect of transaction prices from the past month on listings around the publication date. In particular, I match each listing price by the date on which it has been posted to its closest price publication date. In this way, listing prices posted in the roughly two weeks before the closest publication date do not observe the newest set of data and are thus untreated. Those posted in the two weeks after the most recent publication date are, by contrast, able to observe the latest set of pricing data and are therefore treated. The matching is done by following a natural approach mirroring that of a seller, i.e., by looking at prices at which properties comparable in location (measured by the first half of the postcode and by an indicator from the 2011 Census rural-urban classification of Output Areas), property type (flat, detached, semi-detached or a terraced house) and number of rooms have sold in the past month[22].

Table A3 displays the results of the following regressions:

$$log(q_i^{pre}) = \alpha^{pre} + \beta^{pre} \times log(p_j) + Controls + \varepsilon_i^{pre} \tag{1.13}$$

$$log(q_i^{post}) = \alpha^{post} + \beta^{post} \times log(p_j) + Controls + \varepsilon_i^{post} \tag{1.14}$$

$$log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i \tag{1.15}$$

where q_i is the initial listed price for property i, p_j is a transaction price for a comparable property j that has been or will be published on the closest publication date and $Treated_i$ is a dummy that turns on if the listing has been posted after the price publication date of the given month. The data runs from March 2012 to May 2018 as this is the sample period during which the Land Registry has been publishing price data on a regular monthly basis. I keep only comparable prices that have at least one treated and one untreated match to make sure that the set of prices affecting listings before and after publication dates is similar. Each listing has an average of 5.67 and a median of 4 comparable prices. Transaction prices on the other hand have an average of 3.09 treated and 3.03 untreated comparable listings, while the median number is 2 across both sets. In column (1) I regress only the set of untreated quotes on the transaction prices of the previous month, while in column (2) I repeat the procedure for the set of treated listings. In both regressions, I control for the distance in days between the date on which the comparable transaction took place and the date of the subsequent listing and for the interaction between this distance and the transaction price in order to account for any trends in housing prices. The difference in the coefficients on the price variable then gives an indication of

[22]The same approach of matching comparable properties is adopted throughout the rest of the analysis.

the extent to which past transactions affect future listing price decisions once they become publicly available. Specifically, comparing the two coefficients allows us to isolate the correlation between past prices and future quotes that arises due to the evolution of common fundamentals from the effect of deliberately using information contained in past prices to learn about the state of the housing market and inform future decisions. The first regression shows that the baseline effect of past transaction prices on listed prices in the following month is 84%. The magnitude of this coefficient confirms the well-established fact that prices in the residential property market exhibit high persistence. The coefficient from the second regression, however, suggests that the mere fact of being able to directly observe the latest set of transaction price data increases this effect by additional 0.45%. This incremental effect is statistically significant at the one percent level with an F-statistic of 7.6052. Columns (3)-(6) provide additional evidence of this result by running the regression specified in equation (1.15) on the full sample of treated and untreated listings. Column (3) controls for the time distance and the potential differences in the way that prices affect future listings across different distances, similarly to the regressions in the first two columns; column (4) also adds month-year fixed effects to account for the average level of listed prices across different periods; column (5) introduces a different control for time distance, namely, it allows for non-linear effects of prices on future listings by interacting the price with dummies for time distance measured in weeks, and finally; in column (6) I add transaction ID fixed effects, in addition to the previous controls, which allows me to account for common unobservables across listings matched to the same transaction. The results are robust across all specifications, namely, the incremental effect of prices on future quotes set after the data becomes publicly available remains at 0.45% and statistically significant. Even when we compare the effect by controlling for the average level of quotes matched to the same transaction, to alleviate concerns that some transactions might be matched to disproportionately more treated or untreated listings, the effect retains both its statistical and economic significance at 0.34%.

Tables A13-A17 in the Appendix provide additional evidence of the direct effect of newly published transactions: Table A13 refines the sample by varying the interval of time around publication dates in which quotes are considered and by limiting the number of comparables in order to make sure that the results are not driven by a small number of listings with too many comparable transactions; Table A14 shows robustness to the timing of publications, i.e., it controls for the week day of the publication and for whether this has occurred at the end of the month or the beginning of the subsequent month[23]; Table A15 focuses on existing properties only and investigates if the effect varies across properties in different price ranges; Table A16 adds real-estate agent fixed effects to make sure that the comparables pricing effect is not

[23]Notice that the twentieth working day of the month can sometimes fall at the beginning of the subsequent month if there are a lot of holidays, for instance.

absorbed by varying business practices across agents. All these refinements largely confirm the economic and statistical significance of the effect of news release on future listings. Finally, Table A17 shows that the results are robust to the inclusion of reference prices for listings, hereby defined as the price at which the current owner purchased the property, which have been shown to significantly affect listing price decisions (Andersen et al., 2019).

The results so far confirm the use of the comparables method in the residential housing market and give sense of the magnitude of its influence on future quotes. One might, however, still be worried that the set of sellers who choose to list their properties in the days following a publication date is different from the set of sellers who do so in the days before or that properties are systematically different along some unobservable dimension. To alleviate this concern, I next include in the sample all price updates in addition to the original quote for each listing and the dates at which these have occurred. This allows me to control for any potential unobservable differences in the immutable characteristics of sellers or their properties by adding listing ID fixed effects. Table A4 displays the results of this exercise which follows the regression specified in equation (1.15). In columns (1) and (2) I rerun the regression specifications from columns (4) and (6) of Table A3, respectively, without listing ID fixed effects in order to show that the treatment effect is similar in this extended sample: being able to directly observe the latest set of transaction data increases their effect on subsequent listings by about 0.43% and 0.30% once we add transaction ID fixed effects. Columns (3)-(6) introduce listing ID fixed effects to the regressions - in this way, the effect of the treatment is estimated solely by using the set of listings that have had at least one price change. As before, I control for the time distance in days between the transaction and the subsequent quote update and for its interaction with price in column (3); I add month-year fixed effects in column (4); in column (5) I replace the usual time distance control with dummies for time distance measured in weeks, and; in column (6) I include transaction ID fixed effects in addition to month-year and listing ID fixed effects. The magnitude of the coefficients of interest naturally decreases to about 0.08% as most of the variation is captured by the listing ID fixed effects, however, they remain statistically significant. Table A24 in the Appendix displays the results of the same regression on the restricted sample that includes only listings with more than one quote available. The incremental effect of transactions on quotes after the publication date is this time even larger at about 0.11% to 0.33% depending on the specification, with listing ID fixed effects included.

For robustness and to explore the heterogeneity in the response to different types of news, Tables A18-A21 provide separately the effect on quotes of positive, negative, large positive and large negative price shocks[24], respectively, on the full set

[24]To determine the sign and size of the shock, I run a hedonic regression of transaction prices and split the sample based on the sign and magnitude of the residuals.

of quotes that includes updates: sellers seem to respond more strongly to negative shocks and more so when these are large. Table A22 demonstrates that the incremental effect of newly published prices is stronger in periods and regions with low sales volume, while Table A23 similarly shows that this effect is stronger when the sales activity is below average at the local level.

The results presented above imply that the monthly publication of transaction prices is a salient feature of the UK residential housing market that significantly affects sellers' behaviour. To provide further evidence that the effect of prices around publication dates is indeed systematic and not coincidental, I next conduct a few robustness tests that are meant to rule out alternative hypotheses. For instance, one might think that we would observe a similar pattern in listing price behaviour if house prices were very persistent and had a trend, even if publication dates did not matter. Although the time controls should account for this possibility, to reject this hypothesis with more confidence, I employ two strategies: first, I conduct a difference-in-differences analysis whereby I take advantage of the large sample of data available and compare the effect of transactions on listings around publication dates prior to March 2012 and thereafter; second, I show that the effect is no longer present if I shift the publication date to a few days before or after the actual one. The results of these tests are presented in Tables A5 and A6. The sample I use for the analysis includes quote updates, however, for robustness I provide the results of the same analysis when only the original quote for each listing is included in Tables A25 and A26 of the Appendix.

Figure A7 illustrates the idea behind the first of these tests through an example. Although the Price Paid data was available to purchase under a licence from the Land Registry prior to March 2012, this was done at the discretion of the real estate agencies and other property data providers. This means that firms could get access to the data at varying dates that would most likely not always coincide with the twentieth working day of each month. As a result, there should not be a significant increase in the effect of past prices on future listings around the hypothetical publishing dates in the period before March 2012. The regressions in Table A5 make use of this change in setting by comparing the effect of prices on listings around publication dates before and after March 2012 via a difference-in-differences approach. Take, for instance, the twentieth working days of July 2011 and July 2013: these fell on July 28th and July 26th, respectively. Transaction data from June of the same year were made publicly available in 2013 but not in 2011. If we compare the effect of June prices on listings before and after July 26th 2013 to that on listings before and after July 28th 2011, we would be able to eliminate any systematic variation in quotes across different periods of a given month, provided that this does not drastically change in the years after 2012. The results of the following regression are presented in Table A5:

$$
\begin{aligned}
log(q_i) =& \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Post March 2012}_i + \beta_2 \times log(p_j) \times \textit{Treated}_i + \\
& \beta_3 \times log(p_j) \times \textit{Treated}_i \times \textit{Post March 2012}_i + \gamma_1 \textit{Post March 2012}_i + \gamma_2 \textit{Treated}_i + \\
& \gamma_3 \times \textit{Treated}_i \times \textit{Post March 2012}_i + \textit{Controls} + \varepsilon_i
\end{aligned}
\tag{1.16}
$$

where q_i is the listing price for property i, p_j is the price at which a comparable property has been transacted in the previous month, $\textit{Treated}_i$ is a dummy that turns on for listings posted on or after the twentieth working day of the month and $\textit{Post March 2012}_i$ is a dummy that turns on starting from March 2012. As usual, column (1) includes controls for the time distance between the listing and the comparable transaction and its interaction with price; column (2) adds month-year fixed effects; column (3) replaces the linear time distance control with dummies for time distance between the quote and the price measured in weeks, and; column (4) adds transaction ID fixed effects. It is interesting to see that the coefficient of interest in row 1 almost doubles compared to Table A3 and is now close to 0.75%. Coupled with the second row coefficients that display the effect of prices on listings after hypothetical publication dates before March 2012, we can conclude that the net effect of recently published prices on future listings is close to 0.35%. It is important to emphasise that the correlation between recent transaction prices and quotes decreases after the twentieth working day of the month before March 2012 as evidenced by the negative coefficients in the second row. This is not surprising as the "treated" listings naturally come after the "untreated" ones and we thus should expect that quotes closer to recent transactions have more correlated fundamentals than those further in the future. This result further strengthens the conclusion that publication dates provide a salient enrichment of the information set of sellers that they incorporate into their listing behaviour. In other words, the results presented in Table A3 above can be thought of as the lower bound for the direct effect of past prices on future listings that arises due to comparables pricing. The third row coefficients show that prices are generally more correlated in the period post March 2012 which indicates that sellers might now have a more frequent access to new data than before. Anecdotal evidence suggests that real-estate agencies used to purchase new Price Paid data less regularly such as every few months. The shift to monthly updates then represents an important increase in the frequency at which they would revise their price forecasts and client advice. Finally, when we compare the effect within listings matched to the same transaction by adding transaction ID fixed effects in column (4), we see that the effect of prices on listings does not significantly change around the hypothetical publishing dates before March 2012, however, it does significantly increase post March 2012 by additional 0.4%.

Moving on to Table A6, I now conduct a second type of robustness checks. This time, I limit the analysis to the sample period starting from March 2012 and I vary the publishing dates by seven days back and forth from the actual ones. I then look at the difference in the effect of prices from the previous month in the two weeks around these placebo publication dates using the regression specified in equation (1.15). The first four columns show the results when I consider as treated the listings that are posted at most seven days before the actual publication date, while in the remaining four columns I consider as untreated the listings that occur in the first week after the actual publication date. I add the usual controls for time distance and its interaction with price as well as month-year and transaction ID fixed effects to make the tests comparable to those in Tables A3 and A4. Looking at row one of Table A6, we can note that the price effect does not significantly change around each of these two sets of placebo publication days. In particular, although listings are still strongly correlated with transaction prices from the previous month, this correlation does not increase in the week following the hypothetical publishing dates. This finding corroborates our previous conclusion that days on which new Price Paid data is made publicly available by the Land Registry do matter and they significantly affect the behaviour of sellers.

I have so far provided evidence on the baseline effect of recent prices on subsequent quotes for properties up for sale. One might be led to think that this effect seems too small to be of any economic significance. It is worth remembering, however, that: (a) this is the effect of one single transaction, while most prospective sellers would look at multiple similar properties before making a pricing decision: to the extent that these comparables are driven by correlated signals their common component can be much more heavily over-weighted by future sellers; (b) I am very conservative in my comparables search strategy by matching listed to transacted properties only if they are identical across location, type, number of rooms, and if they occur in the roughly two weeks around the publication date, i.e., up to two and a half months following the sale; (c) the above estimates can be considered a lower bound for the true effect of comparables pricing due to the fact that the discontinuity around publication dates might not be perfectly sharp, but also due to the results in Table A5 where we see that the incremental effect of prices on future quotes after publishing dates is negative before March 2012. For all these reasons, the actual impact of past prices that results from the use of the comparables method is probably considerably larger. Moreover, I have yet to examine the way that the price effect changes with the number of interim channels of influence. To get a better idea of the total effect and its evolution, therefore, in the next section I study the manner in which the direct influence from recent prices gets amplified through the sequential use of past observable data on listings and transactions. I then use the obtained estimates to investigate the long-run effects of any mistakes on aggregate prices via a simple model of learning in Section 1.5.

1.4.2 Indirect Effect of Past Prices Through Intermediate Channels of Influence

In this section, I explore the way that sellers in the housing market process information they receive from comparables when the information sets across these might not necessarily be independent. More specifically, as transaction and listing prices become available, new sellers use them in order to learn private information about the state of housing demand. Suboptimal pricing behaviour can arise, however, if sellers do not appropriately account for the potential duplication of information: if everyone else takes the same approach, then the most recent prices incorporate information that is also embedded in older ones. Optimal learning requires sellers to distinguish between the new signal contained in most recent data and the part that has been influenced by previous prices or other commonly observed information. This issue is exacerbated by the fact that the ability to appropriately extract all the different pieces of information that drive recent prices entails knowledge of the full structure of information flows.

For this reason, I next investigate the importance of indirect effects whereby past directly observable prices potentially affect future seller behaviour also through other intermediate comparables. The evolution of the effect that prices have on future listings as the number of intermediate channels of influence grows can be benchmarked against the Bayesian and naïve learning models presented in Section 1.1 in the main body of the paper and Section A.3 in the Appendix in order to gain understanding about the way agents process information.

For the first set of tests, I focus on listings occurring in the second week after the publication of the latest transaction data that have at least one match within the set of sold properties. I then check the number of listings that are comparable to this pair in the week before and the week after the publication date. Figure A8 depicts the four possible cases that can arise. Specifically, going from left to right, the matched pair might have: (a) no comparable listings posted in any of the two weeks surrounding the Price Paid data publication date; (b) comparable listings only in the week before but not the week after; (c) comparable listings only in the week after but not the week before, and; (d) at least one similar listing in both weeks. Note that, although listings in the week before the data is published do not directly observe the recent transaction prices, they might still be correlated due to commonly observed signals and fundamentals. Sellers in the week after, on the other hand, are able to directly observe the latest transaction data and so they have a second channel of influence. Looking at listings posted in the second week following the publication date, therefore, allows us to test if agents are able to disentangle the different pieces of information embedded in a new observation and thus avoid double-counting redundant news. In particular, Section A.3 in the Appendix shows that, conditional on agents directly observing a given transaction and its price, Bayesian updating implies that its effect on future quotes should be monotonically decreasing with the number of intermediate links, i.e., the covariance between the two is expected to decline with the arrival of new information as agents optimally place lower weight on each individual signal. Conversely, if agents are unaware of or unable to discern the different channels of influence, then we might see the effect of that same transaction increase with the number of intermediaries relative to the Bayesian case.

Table A7 displays the results of this analysis. Across the four columns, I divide the sample of listings posted in week two post publication into four groups corresponding to the cases in Figure A8 above and I regress the quotes on the prices of comparable transactions which have just been made available:

$$log(q_i^s) = \alpha^s + \beta^s \times log(p_j^s) + Controls + \varepsilon_i^s \tag{1.17}$$

where q_i^s is the listed price, p_j^s is the price at which a comparable property has been transacted the month before and s is an index that captures whether the subsamples of quote-price pairs have no comparable listings, at least one comparable listing in the the week before, the week after or in both weeks around the publication date. In all regressions, I control for the number of comparable matches in each week and for the time distance between the price and the quote in question. Going from column (1) to column (4), we can see that the price coefficient is monotonically increasing. Column (1) gives the effect of recently published prices on listings in the second week following the publication date for the case where there are no comparable listings in the two weeks surrounding it. This effect is about 81.66% which can be thought of as the sum of the correlation that arises due to common signals and fundamentals and the direct effect of price j on quote i that results from the use of the comparables method. The following three columns then indicate the incremental effect that comes from the existence of additional links between the two. From column (2), we can infer that having a comparable listing in the week before the price data is published already increses this effect by additional 1.62%. That is, although the matched listing of week -1 does not directly observe price j, the fact that it is highly correlated to it because of commonly observed news makes agent i overweight this common signal when using both to inform his decisions. Column (3) shows an additional increase in the effect of 0.74% if the intermediate comparable is instead in the week post publication. Intuitively, the incremental effect here arises because, on top of being driven by common news, the intermediate listing is now able to directly observe transaction j and thus its price also explicitly embeds the signals contained in price j. A seller who uses both transaction j and the intermediate listing of the week before as two independent pieces of information is likely to overweight repeated news such as the private signal coming from transaction j and the common signal. Finally, column (4) shows that having all possible channels of influence present raises the coefficient on price j to 84.79%, 0.77% larger than in column (3) and a striking 3.13% larger than the baseline effect in column (1). This result further underscores our previous assertion that the direct effect estimated in Section 1.4.1 above is likely to be the lower bound of the overall impact that past prices have on future ones. Table A27 in the Appendix confirms these findings by presenting the results from a single regression on the full set of week two listings by interacting the effect of price with dummies for whether the quote-price pair have comparable listings in any of the two weeks surrounding the publication date. As evidenced by the coefficients in row (4) of this table, prices have on average more than 2.32% larger effect on future listings when multiple channels of influence are present. The coefficients in rows (2), (3) and (4) are statistically different from each other at the 5% level or higher as shown by the p-values obtained by running linear hypothesis tests. Table A28 in the Appendix provides evidence that the coefficients do not exhibit the same pattern in the sample before March 2012 and the difference between them often lacks statistical significance. Tables A29-A35 show that the results are largely robust after controlling for the reference price, in response to both positive and negative price shocks, also when these are large, and both in periods and regions with low as well as high sales activity and in times of high and low relative volume at the local level.

Before proceeding to an alternative experiment, I examine more closely the way that the effect from past prices evolves with the number of intermediate comparables in a given week. Table A8 presents the results of the following regression:

$$\begin{aligned} log(q_i) =& \alpha + \beta_0 \times log(p_j) + \sum_k \beta_k \times log(p_j) \times k\ Comps\ in\ week\ n_i + \\ & \sum_k \gamma_k \times k\ Comps\ in\ week\ n_i + Controls + \varepsilon_i \end{aligned} \tag{1.18}$$

where q_i is the listed price for property i posted in week two after the publication date, p_j is the transaction price for a comparable property j sold in the previous month and *k Comps in week* n_i is a dummy that turns on when quote i has k comparable listings in week n, n being either the week before or the week following the publication date. The first three columns show how the price effect changes as the number of comparable quotes in the week before the publication date increases. Going from row (2) to row (5), we can see that the coefficient is larger when there are multiple intermediate quotes available; the difference between the effect on quotes having one comparable and that on quotes having more than 3 is close to 1% and significant in all specifications but (1) which does not control for the number of comparables in the other weeks. The last three columns show the effect as the number of comparables in the week post publication increases. This time the effect on quotes in week two is much larger and it monotonically grows with the number of intermediate quotes as these directly observe the price and thus create additional redundant links between the two. The difference between the price effect on quotes with only one comparable and those with more than 3 is between 1.12-2.22% and is significant at the 1% level across all specifications.

To further corroborate the above conclusions, I now propose a second approach to examining the way that baseline effects get amplified by naïve learners over a sequence of listing prices. Figure A9 gives a visual representation of the chain of interactions between a given transaction and subsequent comparable listings. The selling price is depicted in blue, listings matched to it that occur prior to its publication date are depicted in light green and those occurring after its publication date are in dark green. Since sellers act sequentially in the housing market, observing a sequence of past prices requires them to disentangle between various sources of information that drive recent actions. Recall from the first set of tests and Section A.3 in the Appendix that rationality and full knowledge of the links across observations implies that agents should be able to extract the private signal coming from every new observation and avoid double counting information already embedded in prior actions. In other words, fully rational agents would not be disproportionately affected by past news based solely on the number of intermediate observations; we would thus expect to see the effect of a given transaction monotonically subside with the increase in the number of new comparables as a growing information set implies that each individual component gets a proportionally lower weight. Naïve agents, on the other hand, might fall in the trap of treating newly observed quotes as plain revelations of independent information about demand; as the number of in-between links increases, naïve learners would therefore keep overweighting the information embedded in early prices relative to the Bayesian framework. To test this hypothesis, I next compare the effect of a given transaction on listings occurring in the month around its publishing date by order of match, i.e., for every quote, I control for the number of comparable quotes that happen before it and for whether the quote in question occurs before or after the publication date.

Table A9 presents the results of this analysis. Specifically, I run the following regression:

$$\begin{aligned} log(q_i) =& \alpha + \beta_1 \times log(p_j) + \sum_{k=2}^{10} \beta_k^{pre} \times log(p_j) \times Comp\ Order\ k\ Pre_i + \\ & \sum_{k=2}^{10} \gamma_k^{pre} Comp\ Order\ k\ Pre_i + \sum_{k=1}^{10} \beta_k^{post} \times log(p_j) \times Comp\ Order\ k\ Post_i + \\ & \sum_{k=1}^{10} \gamma_k^{post} Comp\ Order\ k\ Post_i + Controls + \varepsilon_i \end{aligned} \tag{1.19}$$

where q_i is the listing price, p_j is the price at which a similar property has been sold in the previous month, *Comp Order k Pre*$_i$ is a dummy that turns on if the listing occurs in the period before the publication date and it is the k-th chronological match to transaction j, and *Comp Order k Post*$_i$ is a dummy that equals one if the listing occurs in the period after the price data is published and it is the k-th chronological post-publication match to transaction j. Due to its length, only the coefficients of interest are presented in Table A9 while the full table can be found in the Appendix (Table A36). Column (1) contains the baseline regression, column (2) adds the controls for distance in days between the listing and the matched transaction and the interaction with price, while column (3) also includes listing month-year fixed effects. I limit the number of comparables in the pre- and post-publication period to twenty as there are very few transactions with more than twenty comparables which makes the estimates very noisy. I estimate separate coefficients per order of match for the first nine matches while the tenth coefficient groups all matched listings of order ten or higher. To better visualise the results, I also plot the price coefficients along with their 95% confidence bounds from the specification in column (3) in Figure A10 below, while the coefficients for the other two specifications are graphically depicted in Figure A21 of the Appendix. The results are virtually unchanged across all specifications: the effect of past prices tends to slightly increase up until the seventh match in the period before publication, however, the magnitude of the incremental effect is not very large peaking at around 0.7% and losing statistical significance thereafter. The positive correlation can be attributed to the fact that fundamentals in the housing market are persistent which means that, although listings in the pre-publication period do not directly observe the price data from the previous month, they tend to co-move due to this underlying persistence and commonly observed public information. This is a key feature of the model presented in Section 1.5. Moreover, the slight upward trend suggests that sellers are unable to properly isolate the private signals from recent listings which in turn leads to placing too much weight on stale news. Moving forward to the tenth or higher-order matched listings, however, the incremental effect starts to wane, dropping down to below 0.04% as new information begins to dominate. What is interesting is that this trends gets completely turned around in the post-publication period: once transaction prices become publicly available, their influence on future listings sees a large jump and a significant upward trend as the order of match increases. More specifically, the incremental effect on the first match post-publication goes back up to around 0.15-0.25% relative to the first pre-publication match; it then steadily climbs to a striking 1.8% at the 8th match where it starts to level off. The coefficients for the post-publication period are both economically and statistically more significant that those in the pre-publication period. Although the effect on the early listings might be rationalised by claiming that sellers now learn about the private signals embedded in recently published prices, it would be very difficult to justify the upward trend as the number of intermediate comparables increases. To put it differently, once the news becomes public, we should see an immediate jump in the coefficient as the new information gets embedded into prices which should then remain flat for all future listings or even exhibit a downward trend with the arrival of new information from intermediate quotes. The fact that the effect is gradually increasing thus provides further proof that agents in the housing market have trouble discerning different drivers of past actions; instead, they treat new observations as independent of previous commonly observed ones. For additional evidence, Table A37 in the Appendix presents results on the linear trend in influence of past prices as the number of intermediate comparables grows. The results show that there is no significant change in the comovement between quotes and prices in the pre-publication period, however, the effect increases by about 0.07%-0.08% with each additional comparable in the post-publication period. Finally, Table A38 and Figure A22 in the Appendix provide evidence that the comovement patterns are very different before March 2012: in this sample, the influence from past prices remains largely flat as the number of interim quotes increases and we can even see that the incremental effect relative to the earliest match becomes negative for matches of order 10 or higher across both pre- and post-publication periods. The inability to use information contained in past prices, or the absence of regularity in the frequency of its arrival, means that sellers in the subsequent month have less correlated information sets before March 2012 which in turn implies that their pricing decisions will be less affected by the same set of past transactions.

For the final set of tests in this section, I focus on listings which have seen price updates, i.e., I search for listings for which I have at least two available prices posted on two different dates. This allows me to test for any amplification effects that arise due to the repeated use of comparables by taking into account any unobservable property and owner characteristics. Specifically, I now analyse the impact that transaction prices published just before a listing has been posted have on its subsequent quote updates. This enables me to investigate propagation effects within a given listing. To better see this, Figure A11 depicts potential ways that information available prior to the very first time a property has been listed could have an increasing influence on later price changes. A seller who has his property on the market for a while could still make use of new sources of information such as newly published listings. If this is true news uncorrelated with past signals, then the correlation between old prices and subsequent quote updates should mechanically decrease. However, if the new listings utilise data that is also observable to our seller from the very beginning, then he would again be faced with the challenging task of distinguishing between what is truly new information and stale news.

Figure A23 in the Appendix shows a histogram of the number of price changes per listing. In my sample, there are around 520,000 listings with one or more price changes that I am able to match with at least one prior transaction. However, most listings have very few price updates, with the vast majority having only one price change (358,939) and only 4,494 listings having five or more price updates. For this reason, I combine all price changes of order four or higher into one category in the regressions. Figure A24 in the Appendix displays a histogram of the number of days between the date the first listing price was set and the subsequent price changes. I limit the analysis to changes that have occurred within up to two years after the property was first listed. It can be seen that most of the price changes occur within two to three months of listing, with the modal number of days between the marketed date and the price change date being 28 (10,666 observations). However, sellers frequently update prices on the day following the listing date (7,529 observations). Nonetheless, it is also often the case that price changes occur even after 200 days have passed since the property was first listed (74,299 observations).

Table A10 presents the results of the following regression:

$$
\begin{aligned}
log(q_i^n) = &\alpha + \beta_1 \times log(p_j) + \sum_{n=2}^{5} \beta_n \times log(p_j) \times \textit{Update Number } n_i + \\
&\sum_{n=2}^{5} \gamma_n \textit{Update Number } n_i + \textit{Controls} + \varepsilon_i^n
\end{aligned}
\tag{1.20}
$$

where q_i^n is the n-th quote for listing i, p_j is the price at which a matched property was sold in the month before property i was initially listed and *Update Number* n_i is a dummy that turns on when quote i is the n-th change to the listing price for property i. The first column shows the results of the baseline regression, column (2) adds the time distance control and the interaction of price and time distance, column (3) adds month-year fixed effects and column (4) adds month-year and listing ID fixed effects. The coefficients and their 95% confidence bounds for the specification in column (3) are displayed in Figure A12 for visual inspection; Figure A25 in the Appendix shows the coefficients for the other three specifications. The results from the baseline regression in column (1) show that the incremental effect on the first price update of transaction prices published just before the property was listed on the market is about 0.97%. This effect gradually increases for subsequent price changes to reach 2.44% for updates of order four or higher. Moving on to column (2) we see that controlling for the fact that the relationship between the listing price and the independent variable naturally decreases with the passage of time makes the magnitude of the effect even larger. Specifically, the effect on the first price update now goes up to 1.32% while that on the fourth update is striking 3.59% larger than on the initially set price. Adding month-year fixed effects in column (3) leaves the results largely unchanged which suggests that the result is not driven by a mere trend in prices. Aggregate market dynamics do not absorb the increasing effect that past prices have on quote changes of higher order, i.e., controlling for average price levels confirms that listing prices are more heavily influenced by early observable information, which are likely to be heterogeneous across sellers who update their quotes in the same month[25]. Finally, the addition of listing ID fixed effects in col-

[25]Note that these findings are also consistent with sellers in the housing market exhibiting confir-

umn (4) slightly reduces the magnitude of the coefficients, however, they retain their statistical and economic significance across all specifications. Note that the present findings do not necessarily contradict the well-known fact that most price updates tend to be downward changes. For example, even among price downgrades, these results show that sellers who happened to observe a lot of positive news at the moment of listing tend to reduce their quotes by less relative to others, and vice versa.

The above results provide remarkable evidence of the notable influence that past information exerts on future seller behaviour both in the short as well as the medium term, taking into account the fact that a considerable fraction of price updates in the data occur even after six months of listing. The overreaction of sellers to stale information due to redundant channels of influence could potentially be an explanation to the documented stickiness in their pricing behaviour (Merlo et al., 2015).

I have hitherto provided convincing evidence of the challenges that agents in the housing market face due to the complex connections among sequentially-moving actors. I have described how this intricate environment coupled with sub-perfect knowledge about its structure can lead sellers to place disproportionate weight on stale news at the expense of truly new information by failing to account for commonly observed drivers of recent actions. Having said that, however, I have so far only considered one side of the housing market. In particular, it would be interesting to know if the potential mistakes that sellers make when trying to learn the state of demand are partially or fully corrected at the selling stage. In the next section I therefore look at the sample of listings matched to their respective transactions in order to answer this question.

1.4.3 Interaction with Buyers

For the final set of results, I analyse the relationship between varying degrees of influence from past prices among sellers and the ex-post discount that they are faced with at the transaction stage. Specifically, if sellers mis-estimate the state of demand in the housing market, then this should eventually be somewhat corrected by buyers making offers that are further from the listed price the more this one does not coincide with current fundamentals. To provide some indicative evidence of this, I now restrict my attention to the listings in the sample that I have matched to subsequent transactions. For each listing, I compute the price differential in percent between the first quote and the transaction price for that property[26]. I then split the sample of matched listings into five buckets corresponding to the five quantiles of the price discount distribution and run the following regression per bucket:

$$log(q_i^k) = \alpha^k + \beta^k \times log(p_j) + Controls + \varepsilon_i^k \tag{1.21}$$

mation bias.

[26] I do the same analysis with respect to the last quote available for a given listing, when there are multiple price changes, without any significant change in the results.

where q_i^k is the first quoted price for listing i which is in quantile k of the price discount distribution and p_j is the transaction price for a similar property which has been published in the month before the listing was first posted. I include the usual controls for the time distance between past prices and matched quotes and its interaction with price, as well as listing month-year fixed effects in order to absorb any aggregate pricing dynamics. The sample used in these tests contains 1,067,282 listings in the post March 2012 period that are paired with their corresponding subsequent transactions. Dividing the sample based on the price discount distribution leads to the following five buckets: (1) properties sold at a price that is more than 7.5% lower than the listed price; (2) properties sold at a discount of between 7.5% and 4.2% to listed price; (3) properties sold at a discount of between 4.2% and 2.2%; (4) properties sold at the listed price or a discount of up to 2.2%, and; (5) properties sold at a premium to quoted price. This division restates the strong skewness in the distribution of price discount whereby we observe that over 74% of the properties are sold at a discount, with 12% of properties selling at the listed price and only 14% being sold at a premium, consistent with previous findings.

Table A11 displays the results of the regression in equation (1.21). For a visual representation, I plot the coefficients per bucket along with their 95% confidence bounds in Figure A13. The dashed horizontal line represents the price effect using the full sample of matched listings. The results suggest a strong U-shaped relationship between the effect of recent transactions on listings and the ex-post price discount. In particular, sellers who end up selling their properties at a very large discount or at a premium to listed price tend to be the ones who were more heavily influenced by past prices. The difference in the price coefficients between the extremes (buckets 1 and 5) and the middle buckets, which contain properties sold closer to listed price, shows an 8% increase in the influence from recently observed data for those cases where the final price deviates the most from the quoted one. This piece of evidence suggests that making ″wrong″ inference about demand by over-weighting stale information is indeed somewhat corrected in the final stage of the selling process. Interestingly, the effect goes both ways, i.e., not only do we see overpriced houses being sold at large discounts, but also properties that have been under-valued by sellers see their final transaction prices increase the most due to strong buyer competition.

Unlike most other financial markets, the market for residential housing clears along two dimensions (Yavas and Yang, 1995; Chen and Rosenthal, 1996; Carrillo, 2012). Specifically, mis-valued properties might take longer to sell if sellers are not willing to accept a significant discount. The relationship between making wrong inference and time on the market is, however, likely not monotone as is the case with the absolute price differential. In particular, while an overpriced property would take longer to sell, an underpriced one should sell very quickly due to high demand. As a result, I now investigate the relationship between the effect of past prices and time on the market by dividing the sample in two subcategories: properties sold at a discount, and properties sold at a premium to listed price. I run the regression specified in equation (1.21) by splitting the properties into five buckets based on the quantiles of the distribution of time spent on the market. Table A12 displays the results of this exercise: in Panel A I look at properties sold at a discount to listed price (793,203 unique observations), while Panel B considers properties sold at the listed price or above (274,079 distinct listings). Note that the time on the market distribution is quite different for properties sold below the quoted price, compared to those sold at or above it. In particular, while the average property sold at a discount spends about 28 weeks on the market with the median property selling in 22 weeks, the average and median properties sold at premium spend 22.5 and 17 weeks on the market, respectively, which corresponds to approximately five weeks faster selling time for the second sample. It is a well known fact that volume and prices are highly correlated in the housing market, i.e., houses sell more quickly and at higher prices in hot markets, relative to cold markets (Genesove and Mayer, 2001; Glaeser and Nathanson, 2015). This suggests that the time coverage of properties sold at a discount compared to those sold at a premium is probably quite different. As a result, it is crucial to introduce time effects to the regression that would capture the average price level in the housing market in a given month. In this way, I compare relative sensitivity to past prices across TOM buckets after partialling out any aggregate temporal variation of house prices.

Figure A14 plots the effect of recent prices on listings by quantile of the TOM distribution: Figure A14a shows the coefficients for the sample of listings sold at a discount, while Figure A14b provides the coefficients for the listings sold at a premium. It can be observed that the effect is contrasting across the two samples. Namely, for the sample of properties sold below the quoted price, there is a positive, albeit delicate connection between the amount of time spent on the market and the degree of correlation with past price data. It seems that listed prices for properties that take longer to sell are more heavily affected by recent comparable transactions, however, the statistical significance of this effect is close to zero. In particular, we cannot rule out the hypothesis that the correlation with recent transactions is equal to the sample average across all TOM buckets. The results are quite different in Panel B of Table A12: it is now evident that listed prices for properties that take the least time to sell have been most influenced by past news, with the difference in the effect being 7-9% larger for properties with the shortest TOM compared to the rest of the sample. On the other hand, we cannot reject the hypothesis that there is no significant variation in the way that sellers in the top four quantiles of the TOM distribution make inference from past available data. The findings above suggest that the housing market is at least to some extent efficient at correcting mistakes made by sellers due to wrong inference[27]. Furthermore, the analysis conducted in this section hints at a potential

[27]Note, however, that it is impossible to determine whether the final transaction prices are fully

asymmetry in the way that the market achieves this rectification. Specifically, while overpricing is typically corrected through sellers accepting discounts to their ask price, undervaluation gets amended through both the price channel (buyers willing to pay a price above the quoted one) and the time channel (significantly underpriced properties take less time to sell).

The empirical evidence on the inference biases of sellers provided in this section indicates that stale information might have long-lasting effects on their behaviour due to their inability to disentangle between redundant and new signals. In the penultimate section, I therefore investigate the economic impact that pricing mistakes might have on aggregate house market dynamics in the long run.

1.5 Economic Magnitude of Learning Mistakes

The aim of this section is to provide some indication of what the effect of individual pricing mistakes that result from naïve learning as demonstrated in Section 1.4.2 might be on the aggregate picture, once we consider that the majority of house market agents are likely subject to the same learning biases. This can further help us gain understanding of some recent market phenomena, such as the impact of the Brexit referendum vote and various stamp duty holidays on housing market dynamics. For this purpose, I sketch a simple model whose goal is to capture some of the key features of real-estate markets and the information structure established in the rest of the text.

Let us assume that the log of house prices are governed by a fundamental δ_t which follows an AR(1) process with persistence parameter ρ and mean a:

$$\delta_t = a + \rho\delta_{t-1} + \varepsilon_t \text{ , } \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2) \tag{1.22}$$

The above assumption is for simplicity and is meant to capture, in reduced-form, the excess demand that prospective sellers face. This implies that prices are determined based on the conditional expectation of δ_t at time t:

$$p_{i,t} = \mathbb{E}_{i,t}[\delta_t] \tag{1.23}$$

where $p_{i,t}$ is the transaction price for a property sold at time t. Agents do not observe the realisation of the fundamental δ_t and, as a result, estimate its value from available information. In particular, the informational structure is characterised by the presence of public and private signals. In every period, prices are determined after the observation of a private signal $s_{i,t}$ about the fundamental value:

$$s_{i,t} = \delta_t + \eta_{i,t} \text{ , } \eta_{i,t} \overset{iid}{\sim} \mathcal{N}(0, \sigma_\eta^2) \tag{1.24}$$

corrected without observing the fundamental. Moreover, to the extent that buyers make inference from observable data, they may also be subject to similar biases as sellers.

The noise terms are independent and identically distributed across individuals and time. If there are multiple transactions occurring in a given period, they are all formed based on a different private signal. There is also one publicly observable signal s_t arriving every k periods[28]:

$$s_t = \delta_t + u_t \, , \, u_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_u^2) \tag{1.25}$$

The public signal noise is also identically and independently distributed across time. This signal represents any public information that agents might take advantage of to make inference about housing demand and prices[29]. Finally, agents also observe the full history of past transactions which they also use to extract the private signals that agents in previous periods have received. At every period t, therefore, the information sets of agents consist of the full history of past prices, the history of public signals and their own private signals. They thus form conditional expectations of δ_t and set prices accordingly, as follows:

$$p_{i,t} = \mathbb{E}[\delta_t | s_{i,t}, s^t, p^{t-1}] \tag{1.26}$$

where $s^t = \{s_0, s_{0+k}, s_{0+2k}, ...\}, p^{t-1} = \{p_0, p_1, ..., p_{t-1}\}$, denote the full history of public signals and transaction prices, respectively, that agents at time t observe, while $s_{i,t}$ is agent i's private signal. Given the model described in equations (1.22)-(1.26) above, we can trace the learning process of agents who act sequentially. Specifically, I solve for the posterior beliefs, and consequently prices, for both the Bayesian and the naïve case in order to compare the house price dynamics that these generate. Section A.3.2 in the Appendix provides detail on the procedure of forming posterior beliefs and the recursion that can be used to update beliefs given the new signals from a given period. It can be shown that a Bayesian learner i in period t would form prices as follows:

$$\begin{aligned} p_{i,t} = \delta^i_{t|t} = & w_t s_{i,t} + (1-w_t)[a + \widetilde{w}_{t-1}\rho\bar{s}_{t-1}] + \\ & (1-w_t)(1-\widetilde{w}_{t-1})[\rho a + \widetilde{w}_{t-2}\rho^2\bar{s}_{t-2}] + ... + \\ & (1-w_t)(1-\widetilde{w}_{t-1})(1-\widetilde{w}_{t-2})...(1-\widetilde{w}_1)[\rho^{t-1}a + \widetilde{w}_0\rho^t\bar{s}_0] + \\ & (1-w_t)(1-\widetilde{w}_{t-1})(1-\widetilde{w}_{t-2})...(1-\widetilde{w}_1)(1-\widetilde{w}_0)\rho^t s_0 \end{aligned} \tag{1.27}$$

where $p_{i,t}$ is the price set by a Bayesian agent i in period t, $\delta^i_{t|t}$ is his conditional expectation of the fundamental value δ_t given all information available at time t, $s_{i,t}$ is agent i's own private signal, s_0 is the public signal arriving in the first period,

[28]Here I assume that there is a single public signal arriving at $t = 0$ for simplicity of exposition, however, I vary the frequency of public signal arrival in the simulation exercise.

[29]As explained in Section A.3 of the Appendix, the public signal might involve a housing price index published at regular frequencies. Alternatively, it can also be interpreted as representing local area characteristics or amenities visible to everyone.

$\bar{s}_{t-k}$, $\forall k \leq t$, is the precision-weighted average[30] of the private signals across all n agents in period $t-k$, and w_t and $\widetilde{w}_{t-k}$, $\forall k \leq t$ are weights that the agents in period t assign to all available signals. These are determined based on the signals' relative precisions with regard to the current state of the underlying as explained in Section A.3.2 of the Appendix. Similarly, we can show that naïve learners would form beliefs in a slightly different way:

$$\begin{aligned}\widetilde{p}_{i,t} = \widetilde{\delta}^i_{t|t} =& w_t s_{i,t} + (1-w_t)[a + \widetilde{w}_{t-1}\rho\bar{\widetilde{p}}_{t-1}] + \\ & (1-w_t)(1-\widetilde{w}_{t-1})[\rho a + \widetilde{w}_{t-2}\rho^2\bar{\widetilde{p}}_{t-2}] + ... + \\ & (1-w_t)(1-\widetilde{w}_{t-1})(1-\widetilde{w}_{t-2})...(1-\widetilde{w}_1)[\rho^{t-1}a + \widetilde{w}_0\rho^t\bar{p}_0] + \\ & (1-w_t)(1-\widetilde{w}_{t-1})(1-\widetilde{w}_{t-2})...(1-\widetilde{w}_1)(1-\widetilde{w}_0)\rho^t s_0\end{aligned} \tag{1.28}$$

where $\widetilde{p}_{i,t}$ is the price set by a naïve agent i in period t, $\widetilde{\delta}^i_{t|t}$ is his conditional expectation of the fundamental value δ_t given all information available at time t, w_t and $\widetilde{w}_{t-k}$, $\forall k \leq t$ are the same weights as defined in the Bayesian case and $\bar{\widetilde{p}}_{t-k}$, $\forall k \leq t$ is the average price across all n agents in period $t-k$ weighted by the relative private signal precisions. Comparing equations (1.27) and (1.28), we can note that the difference between Bayesian and naïve agents is that naïve learners treat all past prices as independent signals, i.e., they fail to account for the fact that past agents have similarly set prices by looking at the actions of yet earlier agents. They, therefore, assign the same weights as the Bayesian agents but directly to the observed prices as opposed to the properly extracted signals. This would lead them to overweight stale news at the expense of more recent information since old news have already been accounted for in recent prices.

To test the magnitude of the effect of naïve learning, I simulate a market with the above characteristics and compare the impact of various shocks on prices under the Bayesian and the naïve framework. For this, I first calibrate the model parameters, specifically the various signal precisions, using the empirical estimates from Section 1.4.2 above. The details of the calibration procedure are presented in Section A.3.1 in the Appendix. The parameters that govern the underlying process are estimated by running a monthly regression of log aggregate house prices on aggregate income. The predicted values of this regression are then used to fit an AR(1) process that yields estimates for a, ρ and σ^2_ε. Figure A15 depicts the impact of a shock to the public signal. I plot the response of naïve prices in pink, that of Bayesian prices in blue and the underlying in green. The shock is normalised to correspond to a £10,000 increase in prices which is about 5% for an average house priced at £200,000. In the first row figures, I hold the frequency of public signal arrival fixed by assuming that a new commonly observed signal arrives every period and I vary the number of agents: going from left to right, I plot impulse responses for the cases with one, five

[30] As the precisions of all private signals are assumed to be the same, $\bar{s}_t$ is an equally-weighted average for all t.

and ten transactions per month. Looking at the first figure which corresponds to the case with only one agent per period, we can see that naïve prices take longer to recover relative to the Bayesian case although the difference is not very large. This difference gets amplified, however, once we increase the number of agents per period. Figure A15b shows that, in the case with five transactions per month, after 20 years the effect on Bayesian prices has been eliminated while naïve prices are still about £700 above fundamental value which corresponds to 7% of the initial shock. The effect is even more striking in Figure A15c which assumes that there are ten agents per period: in this case, while rational prices have essentially converged to the truth in about 12 years, naïve prices are still higher by more than £2,500; even after 20 years more than 10% of the shock persists. Two opposing effects generate these results. First, note that the increase in the number of observations per period leads to Bayesian agents learning faster. Second, naïve agents are, on the contrary, harmed by the availability of more data since it takes them more time to converge as n increases. This implies that increase in the amount of information about past prices is not necessarily beneficial in environments where agents are prone to making wrong inference by double-counting commonly contained signals. Looking at the second row of Figure A15, I now keep the number of agents constant at ten per month and vary the frequency at which new public signals are released, from every month in Figure A15d to every twelve months in Figure A15f. The results here are less surprising, namely, it takes both types of agents longer to converge when public news is more sparse. However, naïve learners are relatively more affected by this since they always overweight old news, and therefore the shock, and even more so when information arrives less frequently. This causes naïve prices to still be more than £3,500 above fundamentals, 35% of the initial shock, even after 20 years.

We can, so far, conclude that naïve learning leads agents to overreact to public shocks relative to Bayesian learners. This cannot be generalised, however, to any type of shock. In particular, Figure A16 displays impulse responses to a £10,000 shock to the underlying demand. The specifications across figures are the same as in Figure A15, namely I do the same comparative statics by varying the number of agents and the frequency at which public news gets released. We can observe a striking difference in the response of naïve and Bayesian learners relative to the previous example. Specifically, naïve agents underreact to the shock when this one represents true changes in demand. This is intuitive as the real shock to δ_t gets suppressed by stale information coming from previous observations. Note that the effect is even more significant due to the high persistence in the fundamental since this implies that shocks to demand can take a very long time to recover from. As a result, in the worst case scenario of Figure A16f, with ten agents per period and public signals arriving every twelve months, naïve prices are still about £4,000 away from true fundamentals in the long run.

The above results suggest that pricing mistakes arising from naïve learning have an important economic impact which can be very long-lasting and cause changes in pricing patterns that are unrelated to true fundamentals. Moreover, the above findings shed light on some real-world dynamics of housing markets. In particular, they could potentially explain why real-estate prices in the UK were largely unaffected in the wake of the Brexit referendum vote even though this event pointed toward a significant drop in future housing demand. On the other hand, the naïve learning model would clarify the effectiveness of stamp duty tax holidays and similar policy measures as this type of salient public information is predicted to have a positive impact on housing market activity and prices for an extended period of time even after their end.

1.6 Conclusions

In this paper, I provide evidence on the learning behaviour of sellers in the market for residential housing. I show that valuation by comparables is a commonly used pricing method in the housing market which makes prices sensitive to the quality and quantity of past observations that gets released. Crucially, I demonstrate that the failure to fully grasp the structure of information flows leads sellers to overweight signals coming from old data as these get repeatedly embedded in subsequent observations. Finally, I present and simulate a model to show how the excessive sensitivity to stale and common news can cause prices to exhibit large swings that might be unrelated to fundamentals.

Although I use the housing market as an ideal laboratory for analysing naïve inference and its implications, the findings of this paper extend more generally to other markets where this pricing method is regularly employed, but also even more broadly to any social setting where economic agents use past observations to inform their decisions (e.g., leisure choices, political opinions, financial decisions). As long as the structure of the network through which information disseminates is not perfectly known to the individuals that form part of it, making inference by approximating past actions as being pure revelations of the private signals of previous actors would lead agents to place disproportionately high importance on early signals relative to more recent ones. Moreover, the findings above suggest that this behaviour can give rise to situations where agents over-react to noise and under-react to true changes in fundamentals. Crucially, the degree of (over)under-reaction is increasing with the availability of data on past actions and decreasing with the frequency at which agents receive new (public or private) signals about fundamentals. The results of this paper, therefore, give support to policies that facilitate access to reliable information about economic fundamentals. However, perhaps counter-intuively, they predict that the improvement in the ability to observe the actions of other individuals might actually contribute to pricing mistakes.

2. Revealed Expectations and Learning Biases: Evidence from the Mutual Fund Industry

FRANCESCO NICOLAI AND SIMONA RISTESKA[1]

How do investors form their return expectations? Do they take all available information into account? Does personal experience play a crucial role in the formation of expectations? We attempt to answer these questions by looking at mutual fund managers' stock return expectations as revealed by their portfolio holdings. We exploit the fact that, under a large class of models, the optimal portfolio rule has a similar functional form; using a three dimensional panel consisting of the portfolio holdings of mutual fund managers over a period of thirty-five years, we are able to extract a measure of subjective expected returns for every manager in our panel by exploiting the variation across stocks over time between and within managers. To see this, consider a mean-variance investor for whom the vector of physical expected returns is given by the following formula:

$$\mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f\mathbf{1}] = \gamma_{i,t}\Sigma_t \boldsymbol{w^*_{i,t}} \tag{2.1}$$

where $\mathbb{E}_{i,t}[\cdot]$ is the conditional expectation operator taken under investor i's information set at time t, $\boldsymbol{r_{t+1}} - r_f\mathbf{1}$ is a vector of excess returns, $\gamma_{i,t}$ is the coefficient of relative risk aversion of manager i at time t, Σ_t is the conditional covariance matrix of stock returns and $\boldsymbol{w^*_{i,t}}$ is the time t vector of optimal portfolio weights of investor i. The above expression for expected excess returns is obtained by inverting the first-order condition of a mean-variance investor provided that we have a good measure of conditional covariances Σ_t. In these regards we follow Merton (1980) and argue that investors' disagreement should mainly regard expected returns and not variances and covariances. We show that empirically this is a good approximation. In Section 2.2.1 we show that - as long as we correctly interpret the manager-specific time-varying parameter $\gamma_{i,t}$ - many optimal portfolio models give rise to a subjective expected return similar to (2.1); whenever that is not the case, we can saturate the model with fixed effects in order to split the total demand into a mean-variance component and a hedging component; to isolate the effect of risk aversion from the effect of subjective expected returns we resort to the very general principle that,

[1]We benefited from helpful comments from Ulf Axelson, Nicholas Barberis, Pasquale Della Corte, Daniel Ferreira, Boyan Jovanovic, Christian Julliard, Samuli Knupfer, Ralph Koijen, Avner Langut, Dong Lou, Ian Martin, Igor Makarov, Cameron Peng, Asaf Razin, Andrew Redleaf, Andrea Tamoni, Michela Verardo and the participants at the LSE seminar, the 2019 Yale Whitebox Conference, the 2019 Belgrade Young Economists Conference. Any errors or omissions are the responsibility of the authors.

given the cross-section of assets the manager invests in, risk aversion is a manager-specific quantity, while expected returns are at the same time asset-specific. The information contained in the cross section of holdings, therefore, greatly reduces the issue of separating the variation due to the manager's preferences from the one due to beliefs.

We start by providing evidence in Section 2.3 that more than 50% of the variation in expected returns is explained by a common time-varying factor and we are able to explain almost 90% of the variation with manager-time and stock-time components. This suggests that a saturated regression will likely allow us to isolate the idiosyncratic part of expected returns affected by manager-stock-time specific effects; we focus on this component to explore the extent to which managers' beliefs are affected by experience. In particular, we investigate whether fund managers put more emphasis on past stock returns that they have *personally* experienced over their investment career. To begin, we consider the effect of the simple average of past observed returns on portfolio holdings decisions. Having experienced a one standard deviation higher average return on a given stock causes the manager to inflate his expected excess return by between 10.3 and 15.1 basis points (after partialling out the effect of common stock and manager characteristics). This effect is both statistically and economically significant and it is almost an order of magnitude larger than that of other commonly used predictors. Nonetheless, the effect of average experienced returns masks important heterogeneity in the influence of past returns observed at different points in time: when we move on to examining the particular shape of the learning curve we find evidence of a differential impact. We start by providing non-parametric results that do not require taking a stance on the precise functional form that investors use to weight past experienced returns. Mutual fund managers in our sample are subject to the so-called serial-position effect: the tendency to predominantly remember the initial and the last observations in a series. More precisely, managers' investment decisions and beliefs are particularly affected by the returns they have experienced early on during their stock-specific experience and those they have experienced most recently. In other words, professional investors seem to exhibit the *primacy* and *recency* bias.

As one would expect, the effect is stronger for single-managed funds and decays fast as the number of managers in a team increases: the effect of recently experienced returns on managers in a single managed fund is twice as large compared to managers working with at least one other professional; the effect of early returns is an order of magnitude larger.

We also show that the differential effect of taxes on capital gains and losses cannot explain these findings since the effect of early career experience is still present even when the manager switches to a different fund. At most, tax considerations can explain 20% of the estimated influence of past returns on portfolio choices and expected returns.

Armed with the reduced-form evidence, we provide a tentative parametric estimation of the managers' learning function. In particular, the results in the reduced-form estimation seem to suggest a non-monotonic learning function. For this reason we adopt a variation of the parametrisation of the learning function in Malmendier and Nagel (2016) that allows for a variety of decreasing and increasing, convex and concave, monotone and non-monotone learning weights. We find that fund managers on average do indeed place a disproportionate weight on personal past experience and that this biases the expected returns recovered from their stock holdings, after having adjusted for risk and risk aversion. When we allow for time-varying weights on past stock returns, we show that mutual fund managers tend to place excessive weight on returns experienced at the beginning of their careers and in the most recent quarters compared to those in the middle period, suggesting that both early-career and recent experience seem to be important determinants of the investment behaviour of a large class of professional investors. For instance, a manager with the median stock-specific experience of 9 quarters assigns around 1.84 times larger weight on the return experienced in the most recent quarter compared to the benchmark of 1/9, while the weight on the first experienced return is 3.13 times larger than the benchmark. We thus reconcile two conflicting strands of the literature: similarly to Malmendier and Nagel (2011) and Malmendier and Nagel (2016), we confirm that investors do overweight their personal experience and manifest a *recency bias*, but - at the same time - we show that professional investors also place a disproportionately large weight on returns that have been experienced in the early part of their investing career, similarly to the findings of Kaustia and Knüpfer (2008) and Hirshleifer et al. (2021). When looking at co-managed funds, we show that a large fraction of the effect of early experience washes out while the effect of recently experienced returns persists; this might be due to the fact that, while there is large heterogeneity in early experience, recently experienced returns are mostly shared among managers within a team.

Finally, in the last part of the paper we focus on risk preferences. Notice from equation (2.1) that, while risk aversion varies at the manager-time level, beliefs vary at the manager-time-stock level. This lets us separate *variation* in adjusted portfolio holdings that is due to the managers' risk appetite from differences in beliefs, but does not inform us regarding their *level*. Once we make some minimal assumptions to pin down their level, we show that individual expected returns tend to be quite biased and that preferences display significant heterogeneity across individuals and time. Moreover, on average, mutual fund managers display an Arrow (1965)-Pratt (1964) coefficient of relative risk aversion between 0.915 and 1.283.

The rest of the paper is organised as follows: Section 2.1 provides an overview of recent literature. We proceed by showing that most of the literature relies on evidence from surveys obtained from non-professional investors or, when not affected by these concerns, on a relatively limited amount of data. We argue that the present paper tries to solve the aforementioned issues. Section 2.2 describes how we can separate the variation in expected returns from the variation in risk aversion or other factors in a wide class of models. Section 2.3 gives details of the data used in our empirical work and provides some summary statistics. Section 2.4 provides the non-parametric results of our analysis, while Section 2.5 describes and show the results of our parametric approach. In Section 2.6, we tackle the question of the level of risk aversion of investment professionals. Finally, Section 2.7 provides concluding remarks.

2.1 Previous Literature

The issue of whether economic agents learn with experience has been explored to some extent by the existing literature. Evidence from the literature in psychology and economics shows that personal experience exerts a larger influence on behaviour compared to other shared sources of information[2], especially very recent and very early experience. These two phenomena are usually referred to as the *recency* and the *primacy* effect and they generate what is known to researchers in psychology as the U-shaped serial-position curve[3].

Diving deeper into the field of finance there is growing evidence that personal experience affects financial behaviour. Kaustia and Knüpfer (2008) and Chiang et al. (2011) show that the likelihood of participating in subsequent IPOs is affected by returns experienced in previous offerings. Choi et al. (2009) provide evidence that investors with high return or low volatility on their 401(k) savings tend to invest a larger fraction of their wealth. Using data from the Survey of Consumer Finances from 1960 to 2007, Malmendier and Nagel (2011) find that individuals born before the 1920s who have experienced the lackluster stock market returns during the Great Depression report higher risk aversion, lower expected returns and are less likely to invest in the stock market. Those that happened to experience lower bond market returns tend to reduce their bond holdings. They also find that returns experienced in the previous year contribute four to six times more to future investment decisions than those experienced thirty years ago. In a similar vein, Malmendier and Nagel (2016) analyse the effect of life-time experience on inflation

[2]For early evidence on the concept of reinforcement learning, see the seminal study by Thorndike (1898). A large body of theoretical and empirical literature studies the role of personal experience in learning, see, for instance, Tversky and Kahneman (1973) for a discussion of the availability bias, Fazio et al. (1978) for experimental evidence on the differential processing of information that results from direct versus indirect experience, Roth and Erev (1995) and Erev and Roth (1998) for experimental data and theory regarding learning in sequential games, Camerer and Ho (1999) for a combined model of reinforcement and belief-based learning, Simonsohn et al. (2008) for experimental analysis of the effect of personal experience in a game theory context.

[3]The psychology literature on these topics goes beyond the scope of this paper. Among others, see Nipher (1878), Ebbinghaus (1913) and Murdock (1962) for evidence on the serial-position effect; for evidence on the primacy effect, see Asch (1946); the recency effect is explored by Deese and Kaufman (1957). See Murdock (1974) for a survey.

expectations using the Reuters/Michigan Survey of Consumers; they show that the effect is stronger for younger respondents, and has a direct effect on their borrowing and savings decisions. Malmendier et al. (2021) analyse the effect of experienced inflation on members of the FOMC board and find similar results. Greenwood and Nagel (2009) investigate the effect of experience on mutual fund managers during the dot-com bubble of the late 1990s. The authors use age as a proxy for experience and show that younger managers were investing more in technology stocks compared to similar older managers and displayed a more pronounced trend-chasing behavior. Chernenko et al. (2016) study the effect of experience on a panel of mutual funds holdings of MBS during the 2003-2007 mortgage boom and show that less experienced managers had larger positions in these securities, especially those backed by subprime mortgages; moreover they show that personal experience outside of the fund had an effect on portfolio choice behaviour. Andonov and Rauh (2020) analyse the effect of experienced returns on a cross-section of U.S. Pension Fund managers, showing a significant effect of past experience on the expected returns that these investors report in annual target asset allocations; in particular, earlier experiences have a stronger effect on investment behaviour. Giglio et al. (2021) look at retail investors' portfolio allocations and match them to beliefs elicited from surveys. They find that stated beliefs have a low explanatory power for the timing of trades, however, they are able to predict the direction and size of those trades that do occur. Finally, there is evidence that experienced risk affects financial behaviour: Knüpfer et al. (2017) show that experienced labour market distress affects portfolio choices, while Lochstoer and Muir (forthcoming) find that individuals have extrapolative beliefs about market volatility.

While the contribution of the above papers is substantial, we argue that most of them are affected by one or more of the following issues: reliance on evidence obtained from surveys where agents report their subjective expected returns, focus on non-professional investors who spend limited time investing and, usually, invest relatively small amounts, and reliance on limited time-series or cross-sections implying that it is harder to perform statistical inference.

Regarding the first issue, the task of recovering investors' expectations is a particularly tricky one. It is well known at least since Harrison and Kreps (1979) that asset prices reveal only risk-neutral expectations of market participants; a way to circumvent this problem is, therefore, to focus attention on expectations elicited from surveys. Most of these measures seem to display high correlations as Greenwood and Shleifer (2014) point out. However Cochrane (2017) argues that there is no guarantee that people report their "true-measure unconditional mean" in surveys. In these regards, Adam et al. (2021) provide evidence that surveyed expected returns are inconsistent with risk-neutral expected returns, ambiguity averse/robust expected returns or any other risk-adjusted returns[4]. However, nothing guarantees that the

[4]Appendix B.3 shows that our framework can also deal with this type of preferences.

reported expected returns are exactly representative of the mathematical physical expectation of investors. Consider for instance a survey respondent that interprets the question as asking "what is the most likely return" instead of "what is the expected return". In that case, the respondent will provide a measure of the modal return rather than its average taken across states of the world. Although the previous example may seem far-fetched, Martin (2017) shows that - for a log investor who holds the market - the physical distribution of returns is asymmetric and, for instance, at the height of the crisis, while the expected return on the S&P 500 was above 20% per year, the author recovers a probability of almost 20% of a 20% decline in the index. Large probability masses far from the mean imply large discrepancies between modal, median and average returns. Beliefs reflected in portfolio choices are more informative and represent the primary object of interest, given that it is ultimately changes in demand and supply that determine the variation in prices. Malmendier and Nagel (2011), Andonov and Rauh (2020) and Giglio et al. (2021) show that portfolio choices are consistent with stated beliefs, but the explanatory power is only partial, while - by construction - our beliefs are fully consistent with trading behaviour.

Regarding the second issue, we argue that there are reasons to believe that sophisticated professional investors might behave differently compared to households and, for this reason, we focus our attention on mutual fund managers; they also routinely follow the stock market and therefore there might be reasons to expect them to be less prone to biases or memory issues. While this seems to be true in the case of IPO subscriptions (Chiang et al., 2011), we show that our investors display large biases even though we cannot provide a direct comparison to households. It should also be noted that, to the extent that these financial intermediaries represent a large fraction of total stock market activity, their beliefs will be an important driver of stock price movements.

Finally, concerning the third issue, many of the papers dealing with institutional investors focus on specific events (e.g., Greenwood and Nagel (2009) or Chernenko et al. (2016)) or rely on limited time series data (e.g., Andonov and Rauh (2020)). The aim of the present paper is to be more general and explore whether the effect of experienced returns is common across periods and stocks and represents a permanent trait of professional investors' behaviour.

2.2 Methodology

In this section we provide a detailed description of our empirical strategy. We first explain how we obtain a measure of expected returns given portfolio holdings. We argue that in a wide set of models - including a mean-variance benchmark - we are able to separate the effect of risk and risk aversion from the effect of return beliefs by using the cross-section of manager holdings. We then describe the way we deal with the issue of estimating covariance matrices and, finally, our plan for identifying risk aversion.

2.2.1 Recovering Subjective Expected Returns

Portfolio choices reveal information about future stock return expectations: this is the main insight of Sharpe (1974)'s *indirect approach* to mean-variance optimisation whereby beliefs about expected returns are inferred from portfolio holdings, rather than the other way around[5]. Consider the problem of investor i who maximises his value function by choosing his portfolio allocations into a risk-free and N risky assets:

$$\max_{\{\boldsymbol{w_{i,t}},\ldots\}} J_{i,t}(W_{i,t}) \tag{2.2}$$

where $J_{i,t}(.)$ is the value function of the investor evaluated at his current wealth $W_{i,t}$. When returns follow a geometric Brownian motion, the law of motion for wealth is:

$$\frac{dW_{i,t}}{W_{i,t}} = r_f dt + \boldsymbol{w}'_{\boldsymbol{i,t}}(\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})dt - \Delta C_{i,t}dt + \boldsymbol{w}'_{\boldsymbol{i,t}}\Sigma_t^{\frac{1}{2}}\boldsymbol{dZ_t} \tag{2.3}$$

where r_f is the instantaneous risk-free rate (or the instantaneous rate of return of any other *reference* asset with respect to which excess returns are computed), $\boldsymbol{\mu_{i,t}}$ is an $N \times 1$ vector of stock return drifts as perceived by investor i, $\boldsymbol{w_{i,t}}$ is an $N \times 1$ vector of stock portfolio weights, $\Sigma_t^{\frac{1}{2}}$ is an $N \times N$ matrix of instantaneous loadings on the Brownian motion processes $\boldsymbol{Z_t}$, $\Delta C_{i,t}$ is the (net) outflow of resources[6], and $\mathbf{1}$ is an $N \times 1$ vector of ones.

The investor chooses his optimal portfolio by selecting $\boldsymbol{w_{i,t}}$. Notice that we are deliberately vague about other potential choice variables, i.e., our analysis follows solely from the optimality conditions for the portfolio holdings and the fact that current wealth is the only state variable. Standard dynamic optimisation arguments (Back, 2017) give the following optimality condition:

$$\boldsymbol{w^*_{i,t}} = -\frac{J_{W_{i,t}}}{W_{i,t}J_{W_{i,t}W_{i,t}}}\Sigma_t^{-1}(\boldsymbol{\mu_{i,t}} - r_f\mathbf{1}) \tag{2.4}$$

where $J_{W_{i,t}}$ and $J_{W_{i,t}W_{i,t}}$ are the first and second derivatives of the value function with

[5] Black and Litterman (1992) start from the same insight to obtain portfolio holdings that combine the manager's views with average realised returns in a consistent way; Cohen et al. (2008) and Shumway et al. (2011) use a similar approach to extract a measure of beliefs from portfolios holdings. The former paper measures the *best ideas* of mutual funds as the investment positions for which the authors can extract the largest expected returns, while the latter analyses the rationality implications of extracted beliefs.

[6] For a standard consumption maximisation problem we can interpret $\Delta C_{i,t} = \frac{C_{i,t} - Y_{i,t}}{W_{i,t}}$, i.e., the instantaneous flow of consumption $C_{i,t}$ net of the income flow $Y_{i,t}$, expressed as a fraction of wealth $W_{i,t}$. In this setting $\Delta C_{i,t}$ can be loosely interpreted as the net outflow of money the mutual fund manager is subject to in each period because of redemption/creation of new fund shares. Because of Markovianity we have that $\Delta C_{i,t} = \Delta C(W_{i,t})$.

respect to current wealth and therefore $-\frac{W_{i,t}J_{W_{i,t}W_{i,t}}}{J_{W_{i,t}}}$ is the Arrow (1965)-Pratt (1964) coefficient of instantaneous relative risk aversion measuring the curvature of the value function with respect to wealth, which we denote $\gamma_{i,t} \equiv -\frac{W_{i,t}J_{W_{i,t}W_{i,t}}}{J_{W_{i,t}}}$. Notice that equation (2.4) is a generalisation of the optimal demand employed by Koijen and Yogo (2019)[7]. We can invert the optimality condition (2.4) in order to get an expression for expected excess returns as a function of optimal holdings and Σ_t. In particular, we have that:

$$\boldsymbol{\mu_{i,t}} - r_f\mathbf{1} = \gamma_{i,t}\Sigma_t \boldsymbol{w^*_{i,t}} \tag{2.5}$$

If we had information about the level of the investor's risk aversion $\gamma_{i,t}$ and the covariance matrix Σ_t, we could obtain an exact measure of his subjective expectations of future one-period ahead excess returns $\boldsymbol{\mu_{i,t}} - r_f\mathbf{1}$. We follow Merton (1980) in arguing that investors should share beliefs regarding Σ_t; we later provide evidence in support of this assumption. To isolate the effect of $\gamma_{i,t}$, let us consider each element of the vector of excess returns $\boldsymbol{\mu_{i,t}} - r_f\mathbf{1}$. At each point in time t, for each stock j, each manager i forms a measure of expected excess return which we can denote by $(\boldsymbol{\mu_{i,t}} - r_f\mathbf{1})_j$[8]. By simply keeping track of the subscripts one can realise that there is variation in expected returns across managers, stocks and time, i.e., along the three dimensions i, j, t. On the other hand, the coefficient of relative risk aversion $\gamma_{i,t}$ varies only at the i-t level, implying that the cross-section of holdings for manager i at time t gives us enough information to isolate the variation in beliefs from the variation in risk aversion which acts as a level shifter on the demand for risky assets[9]. When instantaneous returns are normally distributed and wealth is the only state variable, any utility function (and therefore any value function $J_{i,t}(W_{i,t})$) gives rise to a demand as the one in (2.4). We can extend this approach to a wide class of models where there is an $L \times 1$ vector of Markovian state variables $\boldsymbol{X_t}$ with the following law of motion:

$$\boldsymbol{dX_t} = \phi(\boldsymbol{X_t})dt + \Gamma(\boldsymbol{X_t})\boldsymbol{dZ_t} \tag{2.6}$$

Standard dynamic optimisation arguments imply that, in that case, the optimal demand will be:

$$\boldsymbol{w^*_{i,t}} = -\frac{J_{W_{i,t}}}{W_{i,t}J_{W_{i,t}W_{i,t}}}\Sigma_t^{-1}\left((\boldsymbol{\mu_{i,t}} - r_f\mathbf{1}) - \sum_{l=1}^{L}\frac{J_{W_{i,t}X_{l,t}}}{J_{W_{i,t}}}\boldsymbol{K_{l,t}}\right) \tag{2.7}$$

[7] The optimal demand in equation (7) of Koijen and Yogo (2019) is equivalent to our specification whenever $-\frac{W_{i,t}J_{W_{i,t}W_{i,t}}}{J_{W_{i,t}}} = 1$, i.e., investors have logarithmic utility. It is easy to incorporate short sale constraints in our setting as we show in Appendix B.3.

[8] $(\boldsymbol{\mu_{i,t}} - r_f\mathbf{1})_j$ is the $j-$th element of the vector of expected excess returns for manager i, time t, i.e., $\boldsymbol{\mu_{i,t}} - r_f\mathbf{1} = [(\boldsymbol{\mu_{i,t}} - r_f\mathbf{1})_1, ..., (\boldsymbol{\mu_{i,t}} - r_f\mathbf{1})_j, ..., (\boldsymbol{\mu_{i,t}} - r_f\mathbf{1})_N]'$.

[9] For the reader who is familiar with the textbook mean-variance optimisation, this is analogous to the fact that the selection of the tangency portfolio does not depend on the investor's risk aversion which merely influences the relative proportion of wealth invested in the risk-free and risky assets.

where $\frac{J_{W_{i,t}X_{l,t}}}{J_{W_{i,t}}} = \frac{\partial \log J_{W_{i,t}}}{\partial X_{l,t}}$ measures the semi-elasticity of the marginal utility of wealth $J_{W_{i,t}}$ with respect to the Markovian state variable $X_{l,t}$, and $\boldsymbol{K_{l,t}} = \Sigma_t^{\frac{1}{2}} \boldsymbol{\Gamma_{l,t}}$ represents the vector of instantaneous covariances between returns and the state variable $X_{l,t}$. Let us denote the hedging demand $\boldsymbol{H_{i,t}} \equiv \sum_{l=1}^{L} \frac{J_{W_{i,t}X_{l,t}}}{J_{W_{i,t}}} \boldsymbol{K_{l,t}}$. There are many settings in which we can still disentangle variation in beliefs from variation in hedging demands[10]. First, we might consider the possibility that the mutual fund is facing borrowing constraints. We show in the Appendix that in this case the expected return can be recovered from:

$$(\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})_j = \gamma_{i,t} \left(\Sigma_t \boldsymbol{w^*_{i,t}} \right)_j + H_{i,t} \tag{2.8}$$

Similarly, suppose mutual funds managers are ranked according to a common summary statistic (e.g. alpha over a benchmark). The expected excess return can then be approximated by:

$$(\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})_j = \gamma_{i,t} \left(\Sigma_t \boldsymbol{w^*_{i,t}} \right)_j + H_{j,t} \tag{2.9}$$

The previous examples show that, by saturating the regressions with the proper fixed effects, we are able to use the cross-section of assets of a particular investor to separate the effect of changes in beliefs (which vary at the i, j, t level) from the effect of changes in risk aversion (varying at the i, t level) and hedging demand (as long as this varies at a coarser level).

As a caveat, notice that the only situation where we would be unable to separate changes in the hedging demand from changes in beliefs is if the hedging demand varied at the stock-manager-time level (i.e., if we had $H_{i,j,t}$). This would undermine any attempt to recover variation in beliefs from variation in portfolio holdings; however, the results in the paper would not lose their relevance. First of all, as shown by Moreira and Muir (2019) in the case of time-varying expected returns and volatilities, optimal portfolios can be closely approximated by an affine transformation of the standard mean-variance portfolio. Second, even if expected excess returns cannot be separated from hedging demands, it is not easy to conceive of a mechanism where past experience has a large impact on hedging demands. Third, even if this were the case, we could still interpret all the results in terms of scaled demands ($\Sigma_t \boldsymbol{w^*_{i,t}}$) as opposed to beliefs. Asset prices are ultimately determined by investors' holdings and the variation thereof; it would be nice to know whether the effect on investors' demands goes through expected returns ($\boldsymbol{\mu_{i,t}} - r_f \mathbf{1}$), risk aversion ($\gamma_{i,t}$) or hedging demands ($\boldsymbol{H}$), but ultimately what matters is the fact that part of the variation in the cross-section and the time-series of assets holdings is due to the returns that the agent has experienced. Having said that, in what follows, we impose the

[10]For more details, see Appendix B.3 where we analyse the case of borrowing and short selling constraints, concerns about model misspecification and the issue of benchmarking.

previously discussed restrictions in order to disentangle the different mechanisms. We, therefore, assume that the issue of hedging demands can be solved by saturating the regression with the appropriate levels of fixed effects. In the following two sections, we tackle the two remaining problems, namely, the estimation of the conditional covariance matrix and level of risk aversion.

2.2.2 Estimating the Covariance Matrix

As can be seen in the previous section, in order to construct a measure of one-period ahead expected excess returns, we need to have a measure of the conditional covariance matrices. In this paper we rely on an argument set forth by Merton (1980), which states that, in principle, all investors should agree on Σ_t since it can be very precisely estimated by using increasingly more granular data. In practice it is unavoidable to take a stance on how to estimate the conditional covariance matrix. To make sure that our results do not depend on the chosen estimator for Σ_t, we decide to take three different approaches for this exercise:

1. As a first measure, we compute the sample covariance matrix of stock returns:

$$\hat{\Sigma}_t^{d,1} = \frac{1}{t-1}(R_t - \bar{\boldsymbol{r}}_{\boldsymbol{t}}\mathbf{1}')(R_t - \bar{\boldsymbol{r}}_{\boldsymbol{t}}\mathbf{1}')'$$

   where $R_t = [\boldsymbol{r}_{\mathbf{1},t}, ..., \boldsymbol{r}_{j,t}, ..., \boldsymbol{r}_{N,t}]'$ is an $N \times t$ matrix that contains past realised returns as rows, $\bar{\boldsymbol{r}}_{\boldsymbol{t}}$ is an $N \times 1$ vector that collects sample average returns computed at time t, and $\mathbf{1}$ is a $t \times 1$ vector of ones. We estimate $\hat{\Sigma}_t^{d,1}$ from a one-year rolling window of daily returns[11] and we scale it by $K = \frac{\text{nb. obs.}}{\text{nb. quarters}} = 63.07$ days to obtain our first estimator as $\hat{\Sigma}_t^1 = K \times \hat{\Sigma}_t^{d,1}$. It is well known that it is extremely hard to estimate correlations between stocks and correlations close to unity in absolute value tend to give extreme long-short portfolios. For this reason we resort to the next two measures of the sample covariance matrix;

2. Our second estimate makes use of a Bayesian Stein Shrinkage estimator. We follow Touloumis (2015) and compute the daily covariance matrix $\hat{\Sigma}_t^{d,2}$ as a weighted-average of the sample covariance matrix $\hat{\Sigma}_t^{d,1}$ and a target matrix Σ_t^{target} which imposes zero correlations across stocks:

$$\hat{\Sigma}_t^{d,2} = \lambda\hat{\Sigma}_t^{d,1} + (1-\lambda)\Sigma_t^{target}$$

   where Σ_t^{target} is a diagonal matrix where the elements on the diagonal are the

[11] The reader might be worried about the fact that we estimate expected returns employing covariance matrices that rely on past return data, to subsequently regress on past realised returns. However, notice that the same covariance estimates are shared in the cross-section of managers, which is not true for past experienced returns. Furthermore, our estimates of covariance matrices employ only one year of data while the average manager has more than three years of experience with a given stock.

sample estimated variances, namely $\Sigma_t^{target} = \hat{\Sigma}_t^{d,1} * I_N$ where $*$ denotes the Hadamard product and I_N is a $N \times N$ identity matrix where N is the number of stocks. The estimator of quarterly covariances is then: $\hat{\Sigma}_t^2 = K \times \hat{\Sigma}_t^{d,2}$;

3. In our third and final approach, we again apply a similar Bayesian Stein Shrinkage Estimator:

$$\hat{\Sigma}_t^{d,3} = \lambda \hat{\Sigma}_t^{d,1} + (1-\lambda)\tilde{\Sigma}_t^{target}$$

Following Ledoit and Wolf (2004), $\tilde{\Sigma}_t^{target}$ is a diagonal matrix with the constant average daily sample variance on the diagonal, namely $\tilde{\Sigma}_t^{target} = \frac{tr(\hat{\Sigma}_t^{d,1})}{N} I_N$, where $tr(\hat{\Sigma}_t^{d,1})$ is the trace of the covariance matrix, and I_N is a $N \times N$ identity matrix where N is the number of stocks. The estimator is then: $\hat{\Sigma}_t^3 = K \times \hat{\Sigma}_t^{d,3}$.

More details on the construction of $\hat{\Sigma}_t^2$ and $\hat{\Sigma}_t^3$ and the optimal choice of λ are provided in Appendix B.4. We show in the rest of the paper that the way we compute the covariance matrices is not very relevant for our results. This should be expected given that, as long as managers' estimates of covariances are very similar in the cross-section, up to the first order, the covariance matrix behaves like a stock-time fixed effect and therefore will be absorbed by those in the saturated regressions.

2.2.3 Recovering Risk Aversion

Having discussed the identification of hedging demands and the way we estimate covariance matrices, we now turn to the issue of risk aversion. Let us first disregard any hedging demand for simplicity. The portfolio choice in this case takes the form of (2.4). Note that while we can separate changes in beliefs from changes in $\gamma_{i,t}$, the investor's risk aversion, we are unable to determine their levels. As a simple example, notice that $\tilde{\gamma}_{i,t} = 2 \times \gamma_{i,t}$ and $\tilde{\boldsymbol{\mu}}_{\boldsymbol{i,t}} - r_f\mathbf{1} = 2 \times (\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f\mathbf{1})$ would yield the exact same portfolio choice as that implied by $\gamma_{i,t}$ and $\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f\mathbf{1}$. In Section 2.6, we impose a plausible restriction on the *level* of subjective expected returns and risk aversion, namely, that fund managers expectations are formed in such a way to minimise the difference with ex-post realised returns[12]. Start from the following identities:

$$\begin{aligned}
\boldsymbol{r_{t+1}} - r_f\mathbf{1} &= \mathbb{E}_t[\boldsymbol{r_{t+1}} - r_f\mathbf{1}] + \boldsymbol{\epsilon_{t+1}} && (2.10)\\
&= (\boldsymbol{\mu_{i,t}} - r_f\mathbf{1}) + \boldsymbol{\epsilon_{i,t+1}} && (2.11)\\
&= \frac{\gamma_{i,t}}{\gamma_{i,t}}(\boldsymbol{\mu_{i,t}} - r_f\mathbf{1}) + \boldsymbol{\epsilon_{i,t+1}} && (2.12)\\
&= \gamma_{i,t}(\Sigma_t \boldsymbol{w^*_{i,t}}) + \boldsymbol{\epsilon_{i,t+1}} && (2.13)
\end{aligned}$$

[12]Conditional expectations are the best predictor in a mean square sense, i.e., given the information set $\mathcal{F}_t$ and the random variable y_{t+1}, the conditional expectation $\mathbb{E}[y_{t+1}|\mathcal{F}_t]$ minimises $\mathbb{E}[(y_{t+1} - f_t)^2]$ over all the $\mathcal{F}_t$-measurable functions f_t.

The first line of the above expression is a definition for $\boldsymbol{\epsilon_{t+1}}$: realised returns have to be equal to expected returns plus an orthogonal prediction error. In the second line, we assume that the subjective expectation $(\boldsymbol{\mu_{i,t}} - r_f\mathbf{1})$ and the error $\boldsymbol{\epsilon_{i,t+1}}$ made by the investor are orthogonal. This can be interpreted as a requirement that the expected return is consistent with the law of iterated expectations [13]. The third line multiplies and divides this expectation by the investor's risk aversion $\gamma_{i,t}$. In the empirical counterpart, this will require that the instantaneous relative risk aversion is known to the manager at time t. Finally, we use equation (2.5) to rewrite (2.12) as (2.13). We can, therefore, pin down the level of risk aversion $\gamma_{i,t}$ by running multiple regressions across managers and/or time of stock realised returns on scaled portfolio weights. For instance, if we think that risk aversion is a manager-specific quantity we can run the following regression:

$$r_{j,t+1} - r_f = \alpha_i + \beta_i(\Sigma_t \boldsymbol{w^*_{i,t}})_j + \varepsilon_{i,j,t+1} \tag{2.14}$$

where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \boldsymbol{w^*_{i,t}})_j$ is the demand for stock j for manager i, at time t scaled by the conditional covariance matrix Σ_t. The estimate for α_i will then be a measure of the bias or residual hedging demand. If $\alpha_i = 0$, i.e., the bias or hedging demand is not statistically different from zero, we would be able to interpret the estimate for β_i as the average coefficient of relative risk aversion of manager i, that is $\beta_i = \gamma_i$. It is important to notice that, while it might be interesting to pin down the *levels* of risk aversion and beliefs of each manager, the identification of the learning parameters comes from differential *changes* in beliefs in the cross-section of stocks held, hence it is not affected by our choice of the risk aversion parameter.

2.3 Data and Summary Statistics

In this section we describe the data that we use in the empirical analysis. Data on mutual funds and mutual fund managers' information are obtained from the Center for Research on Security Prices (CRSP)[14] Mutual Fund database. Given that we aim to conduct our analysis at the fund manager level, as opposed to the fund level, we need to construct a dataset of managers' careers. To do this, we first obtain a list of

[13]To see this remember that, according to our notation, the expected excess return of manager i using his information set at time t is $\mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f\mathbf{1}] = \boldsymbol{\mu_{i,t}} - r_f\mathbf{1}$. We can therefore rewrite (2.11) as $\boldsymbol{r_{t+1}} - r_f\mathbf{1} = \mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f\mathbf{1}] + (\boldsymbol{r_{t+1}} - r_f\mathbf{1} - \mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f\mathbf{1}])$. If the law of iterated expectations applies under manager i's expectation, i.e., if $\mathbb{E}_i[\mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f\mathbf{1}]] = \mathbb{E}_i[\boldsymbol{r_{t+1}} - r_f\mathbf{1}]$, it is easy to show that:

- $\mathbb{E}_i[(\boldsymbol{r_{t+1}} - r_f\mathbf{1} - \mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f\mathbf{1}])] = \mathbf{0}$, i.e., there is no unconditional bias,
- $\mathbb{E}_i[\mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f\mathbf{1}](\boldsymbol{r_{t+1}} - r_f\mathbf{1} - \mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f\mathbf{1}])'] = 0_{N\times N}$, i.e., the perceived expected return and the error are uncorrelated.

[14]University of Chicago. Center for Research in Security Prices, I. (1960)

the managers that at any point in time are managing at least one equity fund. We then split each occurrence of multiple managers managing a fund at the same time into separate observations. We also disregard all the cases in which no manager name is available and all the observations where we have words such as "team", "group", "partners" or others that do not allow us to infer who was managing the fund. The most challenging part, however, is to account for cases in which a typo in the fund manager's name causes CRSP to treat the same manager as two different individuals. As an illustration, an individual named John Smith could, for example, appear as "John Smith", "J. Smith", "J Smith" or just "Smith". In order to tackle this issue, we first match names into pairs using a string matching algorithm. We match similar names using three different string distances: the cosine, Jaccard and Jaro-Wrinkler metrics, and we apply rather large distance-specific thresholds that allow us to keep the names which are sufficiently close. We subsequently proceed by manually checking the matched results which amount to more than 15,000 pairs of matched names. Out of these pairs, our manual exercise leaves us with roughly 20% of real matches which suggests that we are quite flexible with the distance thresholds. It is important to stress that, although our manual check might contain some errors, i.e., false positive matches and/or false match rejections, so long as these mistakes are random they only introduce noise in our estimates and cause no bias. More details on the process are provided in Appendix B.4. After matching the names we assign a unique index to each manager in order to build their careers. This exercise leaves us with 3,214 unique managers in our sample. We next match the above managerial data with CRSP mutual fund data based on the first and last date when a manager has been managing a given fund. We remove index funds, fixed-income funds and funds which mainly own foreign equities following Evans (2010), Benos et al. (2010) and Kacperczyk et al. (2006)[15]. We then match the fund information with mutual fund holdings data from the Thomson-Reuters Institutional Holdings database, using Russ Wermer's MFLinks tables. We finally merge the above data with CRSP data on stock returns and risk-free rates and Compustat-Capital IQ[16] data on firm fundamentals. Since we have monthly mutual fund and return data while holdings data are only available on a quarterly basis, we compute quarterly stock returns from the CRSP monthly data and proceed by merging with Compustat quarterly data. The final dataset comprises of over 13 millions observations for 3,214 distinct managers in the period 1980-2015[17]. Table B1 provides descriptive statistics. The first panel reports summary statistics regarding average and median past returns experienced by managers. As one should expect, past experienced returns tend to be right skewed with mean average returns that are larger than mean me-

[15]Details on the removed funds can be found in Appendix B.4.

[16]Standard & Poor's Compustat Services, I. (1962)

[17]The number of observations includes a sizeble fraction of holdings that have zero weights but are included because they are part of the manager investment universe. The investment universe is constructed similarly to Koijen and Yogo (2019).

dian returns (2.4% and 1.4%, respectively). While the standard deviation of average experienced returns is similar to the one of median experienced returns (10% and 11% respectively), counterintuitively, the latter seem to be more dispersed, implying that negative experienced returns tend to be right skewed (so that the median is smaller than the average) and positive experienced returns tend to be left skewed (so that the median is larger than the average).

The second panel of Table B1 regards expected returns, which are computed as explained in Section 2.2.1. In the rest of the paper we provide six measures of expected excess returns which we denote (1)-(6). The first issue regards the inclusion of zero weights[18]. Measures (1)-(3) include only positive weights, while measures (4)-(6) do include the zero weights[19]. Measures (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, measures (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and measures (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$. It is clear from the table that the measures are quite similar in terms of summary statistics. All the measures have an average expected excess return of about 1% per quarter and a median expected excess return of about 0.6%. It should also be noted that, while we have about 12.7 million data points if we consider the zero weights, the number of observations drops to about 5.4 million once we remove the zeros. Figure B1 sheds light on the sources of variation in beliefs. We provide a decomposition of the variation in expected excess returns according to measure (1) by regressing it against various fixed effects. Manager and stock fixed effects explain a small fraction of excess returns (11.63% and 14.20%, respectively), while time fixed effects explain more than half (55.73%) of the variation. This suggests that manager and stock immutable characteristics are relatively less important than aggregate time-varying factors in the formation of expectations. When we separately include manager, stock and time fixed effects the explanatory power rises to almost seventy percent (68.21%). If we allow for interactions between fixed effects, we can explain up to almost ninety percent (89.43%) of the variation in expected excess returns when we include manager-time and stock-time fixed effects. From this decomposition we learn that the largest part of the changes in expected returns is due to time-varying factors, then stock-specific characteristics and, finally, factors related to the manager. The addition of manager-time and stock-time fixed effects will remove the greatest majority of the variation in expected excess returns and will, thus, ensure that the results are driven by idiosyncratic variation in expected returns unexplained by systematic factors. This gives more credibility to our identification strategy.

Finally, we consider the data related to the managers' careers which can be anal-

[18]Similarly to the present paper, Koijen and Yogo (2019) discuss how the analysis might be affected by including or excluding zero weights.

[19]It might be important to know whether zero weights arise by choice or because the manager cannot short sell stocks that would otherwise appear with negative weights. Appendix B.3 shows how the optimal choice of a manager is affected by short selling constraints and how to deal with them when trying to recover beliefs.

ysed with the help of the last panel of Table B1 and Figures B2 and B3. The upper panel of Figure B2 provides information regarding the experience of the managers in the sample. We plot the number of managers by the first time they appear in the sample, which we call the starting date of the fund manager and denote it by $t_{i,0}$. The sample extends from 1980 to 2015 and covers a period of 35 years. Notice, however, that there are fewer managers who start their career in the first ten years compared to the rest of the sample. This can be attributed to low data coverage during the 1980s. Most of the managers in our sample begin their career in the late 1990s. We can observe, however, a wide range of manager starting dates up until the last sample year. We then proceed to construct a tenure variable which measures how many quarters have passed since the start of the manager's career, i.e., for a given manager i and date t, $\text{tenure}_{i,t} = t - t_{i,0}$[20]. The lower panel of Figure B2 displays the number of managers with a given level of accumulated tenure over the sample period, i.e., the empirical distribution of $(t - t_{i,0})$ for all i, t. Most of the managers in our sample are relatively young and inexperienced, but again, there is quite a large variation in tenure as well, ranging from less than a year up to some managers that are present in the whole sample (i.e., for a period of 35 years). Note that, by construction, the number of observations with a given level of accumulated tenure should be decreasing as, for example, a manager who has 5 quarters of accumulated tenure must also have accumulated 4 quarters of experience previously. In practice, this could be violated for two reasons: the first reason is that mutual funds were required to report holdings at a semi-annual level up until 2003 and only later regulators enforced quarterly reporting, as a result, some funds used to report holdings on a quarterly basis while others did so only on a semi-annual basis prior to 2003; second, there might be some missing data in our sample which means that we might be able to observe a given manager's career and holdings in a particular quarter but not in the previous one. The bottom panel of Table B1 shows that the average tenure is of 26.9 quarters (almost 7 years), but because of the positive skewness manifested in Figure B2, the median tenure is of only 22 quarters (5.5 years). We then proceed to the main object of interest of the paper, which is the relationship between each manager and stock. Figure B3 describes the relationship between fund managers and individual stock holdings. The first panel displays the date when a given stock-manager pair has first appeared in our sample which we call the starting date. For each manager i and stock j we can denote the starting date as $t_{i,j,0}$. Unsurprisingly, the largest number of such initiations have occurred in the late nineties and early 2000s, i.e., when the number of managers in our sample significantly increases. There is, however, large variation in the stock-manager starting dates which we exploit as part of our identification strategy. To see this, the second histogram depicts the length of the personal experience of a given manager with a

[20]Notice that for each manager we disregard the first quarter of experience, i.e., $t_{i,0}$, when computing the statistic.

given stock, i.e., for each manager i, stock j and date t, $\text{experience}_{i,j,t} = t - t_{i,j,0}$. It is clear from the histogram that there is a large variation in experience. The third panel of Table B1 shows that it ranges from 1 to 139 quarters, with a standard deviation of about 12.9 quarters. The standard deviation is of similar magnitude compared to the average (about 13.2 quarters) and the median experience (9 quarters). The main hypothesis of the paper is that this variation in stock-specific experience is associated with a variation in expected returns across managers. Finally, we can look at the maximal experience achieved for each stock-manager pair, in the bottom panel of Figure B3 and Table B1[21]. While the average maximal experience and its standard deviation are similar to the above (13.9 and 12 quarters respectively), the median maximal experience is larger (11 quarters, compared to 9 quarters of experience).

In the next section, we present the reduced-form results of our empirical analysis.

2.4 Reduced-form Results

The main hypothesis of the paper is that past experienced returns affect expected future returns. Moreover, if that is the case, we would like to further explore whether certain periods carry more relevance than others. In what follows, we show that differential stock-specific experience across managers indeed matters in the formation of expectations and, in particular, differences in the first and the most recent few quarters of experience play the most crucial role.

The empirical specification in this section relies on the following argument: we conjecture that the manager will try to estimate future returns by looking at the returns he has experienced over his career. A manager i with $T_{i,j,t}$ quarters of experience with a given stock j at time t might use the average experienced return as a sufficient statistic when forming expectations, i.e., his expected return for that stock can be represented as:

$$\mathbb{E}_{i,t}[r_{j,t+1}] = \beta \bar{r}_{i,j,t} = \beta \left(\frac{1}{T_{i,j,t}} \sum_{k=1}^{T_{i,j,t}} r_{j,t+1-k} \right) \tag{2.15}$$

where $\bar{r}_{i,j,t}$ denotes the equal-weighted average of stock j returns observed over the investor's experience horizon. Notice that the variation in the length of past experience $T_{i,j,t}$ allows us to exploit the cross-section of managers holding a given stock j as our source of differential treatment[22]. The coefficient β captures the average effect that past observed returns have on expectations formation, while the implicit constant weight $\omega_k = \omega = \frac{1}{T_{i,j,t}}$ means that all past observations are equally-weighted.

[21]For each manager i and stock j, the maximal experience is defined as $\text{max. experience}_{i,j} = \max_t\{\text{experience}_{i,j,t}\}$.

[22]On the other hand, the variation in the length of past experience $T_{i,j,t}$ for a given manager i at time t across different stocks is what helps us in disentangling preferences from expected returns.

This choice implies that investors attach equal importance to all observations, however, as the length of experience grows every observation receives a progressively lower weight. Note that this approach does not restrict managers from incorporating other sources of information in their estimation. This can be easily taken into account by saturating the regression with the proper controls. To reiterate, fixed effects, for instance, would account for the information or characteristics that all managers, or all stocks in the portfolio of a given manager, have in common; the coefficient on the average experienced return would thus provide a measure of the incremental effect of experience[23]. We, therefore, show in Table B2 the results of the following regression:

$$\mu_{i,j,t} - r_f = \beta \bar{r}_{i,j,t} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{2.16}$$

where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t}$ is the previously defined equal-weighted average experienced return[24], $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To better disentangle the effect of experience we focus on the subsample of single-managed funds[25]. The results in the table confirm our main hypothesis: having experienced an increase of one standard deviation in average quarterly return leads to an increase in the expected excess return of between 0.103% and 0.151%; the results are both economically and statistically large and display very minor variation across specifications. This validates our intuition that the estimation method for the covariance matrix is not very consequential. Similarly, the inclusion of the zero weights has no effect on our main findings, even though the drop in R-squared shows that the zeros are indeed informative and cannot be fully explained by the fixed effects alone. The within R-squared shows that the average experienced returns explain between 0.6% and 0.9% of the variation in expected returns. While this might seem low, it is in fact in line with the findings of Koijen and Yogo (2019) that observable characteristics explain a small part of the variation in investors' demands which is mostly explained by latent factors. Table B11 in Appendix B.6 reports the results of a similar regression with manager-time and stock fixed effects, and a number of time-varying stock characteristics, namely, profitability, investment, book-to-market ratio, market equity, and dividend-price ratio. The findings are similar in magnitude and statistically significant, and show that the effect of experienced returns is almost an order of magnitude larger than that of other known characteristics, confirming again the findings of Koijen and Yogo (2019) that standard predictors have a hard

[23]Notice that this implies that managers could very well use all past realised returns when they form expectations and this would be absorbed by the stock-time fixed effects. In particular, β would then measure the relative over-weighting of experienced returns.

[24]All the regressions in the paper use standardised explanatory variables for ease of interpretation.

[25]Section 2.4.1 analyses the case of co-managed funds, showing indeed that most of the effect washes out when we aggregate across managers.

time explaining portfolio choices[26].

So far, we have assumed that the effect of experience is homogeneous. Alternatively, we could allow for more flexible weights in order to investigate whether certain periods matter more than others. Consider the following modified weight: $\omega_k = \frac{\delta_k}{T_{i,j,t}}$, such that $\frac{1}{T_{i,j,t}} \sum_{k=1}^{T_{i,j,t}} \delta_k = 1$. Namely, the manager estimates future returns from the *weighted* average of past experienced returns:

$$\mathbb{E}_{i,t}[r_{j,t+1}] = \beta \sum_{k=1}^{T_{i,j,t}} \frac{\delta_k}{T_{i,j,t}} r_{j,t+1-k} = \sum_{k=1}^{T_{i,j,t}} \beta \delta_k \frac{r_{j,t+1-k}}{T_{i,j,t}} = \sum_{k=1}^{T_{i,j,t}} \tilde{\beta}_k \tilde{r}_{j,t+1-k} \tag{2.17}$$

The weighting term δ_k is a number centred around one measuring the relative over- or under-weighting of a given past observation. If $\delta_k < 1$, then returns observed k-periods ago are under-weighted, while if $\delta_k > 1$ they are over-weighted relative to the previous benchmark. The last equality in equation (2.17) shows that if we rewrite $\tilde{\beta}_k = \beta\delta_k$ and $\tilde{r}_{j,t+1-k} = \frac{r_{j,t+1-k}}{T_{i,j,t}}$, then we can run a regression on experience-adjusted returns and obtain:

$$\beta = \frac{1}{T_{i,j,t}} \sum_{k=1}^{T_{i,j,t}} \tilde{\beta}_k, \quad \delta_k = \frac{\tilde{\beta}_k}{\beta} \tag{2.18}$$

that is, the average effect of past experience can be obtained as the average of the k coefficients $\tilde{\beta}_k$, while the relative weight assigned to the k-periods ago return is given as the ratio of the coefficient on the k-th term and the equal-weighted average of all coefficients.

In practice, this approach breaks down if we have to deal with varying experience lengths $T_{i,j,t}$, as the number of regressors would change together with $T_{i,j,t}$. For this reason, we group past returns into buckets as a means of fixing the number of regressors. In our first such specification we divide the stock-specific experience of the manager into five non-overlapping buckets of equal length, $|\Delta T^q_{i,j,t}|$, with $q = \{1, 2, 3, 4, 5\}$[27]. Table B3 reports the results of the following regression:

$$\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_q \bar{r}_{i,j,t\in\Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{2.19}$$

for $Q = 5$ and where $\bar{r}_{i,j,t\in\Delta T^q_{i,j,t}}$, $q \in \{1, 2, 3, 4, 5\}$, is the average return in the q-th bucket. Table B4 reports the results for ten non-overlapping buckets of equal length,

[26]We do not report results for median experienced returns which are virtually identical.

[27]To cast this specification in terms of the previously discussed model, let us denote each bucket by $\Delta T^q_{i,j,t}$ and its length by $|\Delta T^q_{i,j,t}|$. We then have that $\delta_k = \beta_q \frac{T_{i,j,t}}{|\Delta T^q_{i,j,t}|}$, where for each time index k in bucket $\Delta T^q_{i,j,t}$ we assign a common effect β_q and take the average return $\bar{r}_{i,j,t\in\Delta T^q_{i,j,t}} = \sum_{k\in\Delta T^q_{i,j,t}} \frac{r_{j,t+1-k}}{|\Delta T^q_{i,j,t}|}$. Notice that $\frac{T_{i,j,t}}{|\Delta T^q_{i,j,t}|} \approx 5$, where the approximation derives from the fact that we have to split ties when the experience length is not a multiple of five.

i.e., the specification in equation (2.19) for $Q = 10$. In both cases we focus on the subsample of single-managed funds. To better visualise the results, the estimated coefficients of a regression with five buckets are reported in the upper panel of Figure B4, while the bottom panel reports the results for ten buckets. The picture immediately reveals that the effect of past experienced returns is clearly neither constant nor monotone. Consider, for instance, our first model of expected returns with $Q = 5$ for which we show results in column (1) of Table B3: a one standard deviation increase in experienced average quarterly return in the most recent or in the earliest period of holding the stock increases the expected return by roughly 0.25% ($\beta_1 = 0.276$ and $\beta_5 = 0.238$); on the other hand, the effect of an increase of one standard deviation midway through the manager's experience has an effect lower by almost an order of magnitude ($\beta_3 = 0.041$). Figure B4 confirms that the effect of experienced returns is "U-shaped" regardless of whether we include the zero weights and independently from the estimator for the covariance matrix used. The lower panel of the figure reports the results for $Q = 10$, painting almost an identical picture. The coefficients for ten buckets are similar in magnitude to those for the regression with five buckets and follow the same "U-shaped" pattern. We report in Appendix B.6 the results for various other specifications: Tables B12 and B13 report the results of the previous models with stock fixed effects and the previously mentioned controls, while Tables B14 and B15 describe the results for a model with three non-overlapping equal-sized buckets; finally Tables B16, B17 and B18, B19 report the results for three non-overlapping buckets of unequal length (with stock-time fixed effects or stock fixed effects and varying controls), where the first and last buckets consist of four and eight quarters, respectively. All these specifications confirm the previously discussed results: experienced returns are important in determining expected returns and most of the impact comes from most recent and earliest stock-specific observations. This is evidence in favour of the so-called serial-position effect, concept well studied among researchers in psychology (Murdock, 1974). Moreover, our findings reconcile two apparently distinct phenomena observed in previous research: on the one hand, Malmendier and Nagel (2011) show that economic agents are principally affected by recent experience, while on the other hand Kaustia and Knüpfer (2008), Hirshleifer et al. (2021) and Hoffmann et al. (2017) report evidence in favour of the *primacy effect* or *first impression bias*. We show that both effects are present in mutual fund managers and that they need to be separately considered.

2.4.1 Co-managed Funds

So far we have focused our attention on single-managed funds, but one might be interested to know whether the above findings are, in fact, weaker when managers work in teams. In this section we check the impact of the number of managers within a team on the effect of experience. Our hypothesis is that personal stock-specific experiences should partly offset each other within a team, so long as the managers that form part of the team have followed different career paths. To explore this hypothesis, we run the following regression:

$$\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_{q,n} \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{2.20}$$

We split managers into subsamples based on the number of co-managers they work with, i.e., $n_{i,t} \in$ {1, 2, 3, 4 or more} signifies that the manager works in a team of one, two, three or four or more people. We thus obtain a different set of coefficients $\beta_{q,n}$ for each combination of buckets and size of the management team. We report in Table B5 the results of this exercise for $Q = 5$ buckets; the results for the specification with 10 buckets are reported in Table B20 in Appendix B.6. To better visualise the results, Figure B5 displays the coefficients $\beta_{q,n}$. The two plots on the left-hand-side of the figure show the results for measure (1) while the right-hand-side plots display the coefficients for measure (4). The first row reports the results for $Q = 5$ and the bottom row for $Q = 10$ buckets. As one can see in Table B5, the coefficient on the most recently experienced returns for single-managed funds is more than twice as large as the same for funds managed by two managers; the difference in coefficients is even larger for the earliest bucket of returns, more specifically, the effect of returns observed at the beginning of a stock-specific experience is more than ten times greater for single-managed funds compared to funds managed by at least two people. The effect on managers working in teams of three or more is orders of magnitude lower, while still statistically significant for recent experienced returns. On the other hand, the effect of early returns loses significance. The above is visually confirmed by the plots in Figure B5 showing a rather steep decrease in the coefficients on the earliest bucket of returns across teams of different sizes, especially when going from a single-managed fund to a fund managed by two professionals. The findings are equally pronounced for the specification with ten buckets.

The results seem to suggest that a considerable part of personal experience washes out in the cross-section of managers working in the same team, and more so the further we go in the past since managers are more likely to change teams over a longer period of time. On the other hand, recent returns affect all co-managers in a similar way as they have presumably gone through the same recent experience, having been working for the same fund. This could justify the difference in spreads observed between buckets at different horizons, especially if we compare single-managed funds with those managed by two individuals.

2.4.2 Taxes

In what follows, we investigate the impact of taxes on managers' investment decisions and the potential explanation the tax regime might have thereof. More specifically, we examine whether tax considerations can absorb the effect that past experience has on portfolio weights and expectations formation. The differential treatment of short-term and long-term capital gains in terms of their taxation, together with the possibility to offset capital gains with capital losses, suggests that mutual funds will try to defer the realisation of gains and accelerate the realisation of losses. This implies that it is optimal from the point of view of minimising the tax bill for mutual funds to hold on to assets that performed well in the past and sell assets that had subpar performances[28].

This, in turn, means that the previous results could be simply driven by tax considerations. One way to solve the problem is to model the optimal selling decision in the spirit of Barclay et al. (1998) or Sialm and Zhang (2020) and check if the effect of experienced returns survives after we have taken tax considerations into account. However, in what follows we take a reduced-form approach and make use of the large amount of data on managers who have managed different funds in their career. In particular, we focus on the subsample of manager-stock pairs where the manager had positive holdings of the stock in the past while managing a different mutual fund compared to the one that he is currently working for. In this setting, tax considerations should be muted given that capital gain overhangs cannot be transferred from one fund to another.

Table B6 reports the results of a regression of expected returns on five buckets of past experience for only those managers that have changed funds, while Table B7 reports the results when we split the previous experience in ten buckets. The results are then summarised in Figure B6 where the upper panel reports the results for five buckets and the lower panel for ten buckets. While the number of observations is greatly reduced (from about 800,000 to slightly more than 110,000 observations if we do not include zero weights, and from about 2 million to approximately 225,000 if we do), the economic and statistical significance of the coefficients is virtually unchanged confirming the previous findings: experienced returns have a sizeable influence on expected excess returns, with the majority of the effect coming from the extreme buckets. If, for instance, we consider measure (1) we notice that the coefficient on the most recent bucket goes from 0.276 to 0.224, while on the earliest one from 0.238 to 0.199. We infer, therefore, that no more than 20% of the effect might be due to tax considerations and we confirm both the *recency* and the *first impression bias*.

Having presented the reduced-form results of our analysis, we now develop a simple three-parameters model of learning and proceed with its estimation.

[28]Bergstresser and Poterba (2002) show that inflows to mutual funds, and therefore managers' compensation, are affected by the amount of unrealised capital gains, implying that there might be a tension between postponing capital gains indefinitely to provide better after-tax returns for current investors and attracting new investors. Barclay et al. (1998) explicitly tackle this question, showing that indeed managers tend to realise gains early to attract new investors.

2.5 Parametric Estimation

The reduced-form evidence of the previous section teaches us that: experience matters, i.e., average experienced returns are an important determinant of expected returns and; the effect of experience is neither constant nor monotone, in particular, earliest and most recent experience matter the most. However, as shown in Section 2.4, estimating the shape of the weighting function requires us to drop a sizeable amount of observations and potentially lose significant information. For this reason we now posit a functional form for the learning weights and try to estimate its parameters. As Figures B4, B5 and B6 show, we need to allow for non-monotone weights if we want to accurately fit the data. Similarly to Section 2.4, we assume that the manager uses a weighted average of experienced returns in order to predict future returns. Recall the model in equation (2.17) where the weights $\frac{\delta_{i,j,t,k}}{T_{i,j,t}}$ capture the differential effect of returns experienced at different points in time. In this section we directly model these weights as follows:

$$\omega_{i,j,t,k} = \frac{\delta_{i,j,t,k}}{T_{i,j,t}} = \frac{(T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}} \tag{2.21}$$

The functional form in equation (2.21) is similar to the one used by Malmendier and Nagel (2011) and Malmendier and Nagel (2016)[29]. The weighting function used in these papers depends only on $T_{i,j,t} - k$ and, as such, it confounds two separate effects: the *first impression bias* and the *recency bias*. On the other hand, our weighting function has the advantage of disentangling between these effects: the term $T_{i,j,t} - k$ measures the distance between the return observed at time $t + 1 - k$ and the beginning of a stock-specific experience, hence capturing the *first impression bias*, while k measures the distance from the current date t, thus capturing the *recency bias*. Figure B7 shows how flexible the parsimonious parametrisation introduced in equation (2.21) is. We plot in blue the weighting function for a manager with $T_{ij,t} = 50$ quarters of experience for all the combinations of $\{\lambda_1, \lambda_2\} \in \{-0.1, 0, 0.1\} \times \{-0.1, 0, 0.1\}$[30] and compare it to the black dashed line representing the benchmark $\frac{1}{T_{i,j,t}}$ where the manager equally weights each observation that forms part of his experience. The first parameter, λ_1, governs the strength of the *first impression bias*: when it is negative, the manager is overweighting early experiences relative to the benchmark scenario. The second parameter, λ_2, controls the strength of the *recency bias*: when the sign of λ_2 is negative the manager overweights recent observations relative to the benchmark, and vice versa. As one can see from the examples in Figure B7, using only two parameters we are able to capture a variety of shapes including linear, convex or concave, increasing or decreasing, monotone

[29]Our weighting scheme collapses to the one used by Malmendier and Nagel (2011) when $\lambda_2 = 0$.

[30]Figure B12 in Appendix B.6 plots the weighting function for $\{\lambda_1, \lambda_2\} \in \{-2, 0, 2\} \times \{-2, 0, 2\}$.

or non-monotone weighting schemes arising from the interplay of the *recency* and *first impression bias*. Given the evidence from the reduced-form regressions we expect λ_1 and λ_2 to be negative, implying that the managers are subject to both effects. Similarly to the model in equation (2.19) we include manager-time and stock-time fixed effects to get rid of potentially time-varying unobservable characteristics shared across stocks and managers, respectively. Table B8 reports the NLS estimates of the following regression[31]:

$$\mu_{i,j,t} - r_f = \beta \left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{i,j,t+1-k} \right) + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{2.22}$$

$$\omega_{i,j,t,k} = \frac{(T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}}$$

Consistent with the reduced-form evidence, both λ_1 and λ_2 are negative and statistically significant across all specifications. The magnitude of the effects is illustrated in Figure B8 where we plot the weighting functions at median and average experience of $T_{i,j,t} = 9, 13$ quarters, respectively, using the empirically estimated values for λ_1 and λ_2 under model (1). It is evident that the weighting function is always convex and non-monotone, implying that managers overweight the most recent and the earliest returns observed; for instance, a manager with an experience of nine quarters will assign a weight of 0.204 (0.347) to the most recent (earliest) observation, which is 1.84 (3.13) times the benchmark of 1/9. On the contrary, he will only assign a weight of 0.043 to the middle observation which is 0.39 times the benchmark weight. The results display a slight asymmetry with λ_1 being always larger in magnitude than λ_2 implying that the *recency bias* is marginally weaker compared to the *first impression bias*. This is, however, not a robust feature of the data: Table B21 in Appendix B.6 shows that λ_1 and λ_2 are almost identical once we include only manager-time and stock fixed effects, indicating that a large fraction of the recency bias might be captured by stock-time fixed effects as we should expect. Pinning down the actual magnitude of the two biases is extremely difficult given that we have to get rid of a large fraction of the variation in expected returns to achieve identification. Finally, the parameter β in Table B8 measures the average impact of past experience on expected excess returns: the estimates range between 0.139 and 0.207. This is about 4 basis points larger than the baseline results in Table B2 where we do not allow for varying weights[32]. We therefore confirm that once we take into account the possibility that recent and early returns might have a differential impact,

[31] Appendix B.5 provides more details on the estimation procedure.

[32] Note that all the results presented refer to standardised variables. In the case of the results in this section we estimate β and then scale its value by the standard deviation of $\left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{j,t+1-k}\right)$. This is to avoid directly scaling the weighted average which would affect the computation of the gradient of the right hand side of equation (2.22) needed to obtain standard errors.

we find an incremental effect of experience on expected returns.

2.6 Risk Aversion

As explained in Section 2.2.3, our methodology allows us to examine in more detail the preferences of investors. Recall equations (2.10)-(2.13); if we assume that subjective expected returns obey the law of iterated expectations, we are able to extract the risk aversion of managers by exploiting the cross-section of individual stock holdings. Running regressions of realised excess returns on scaled demands, as shown in equation (2.14), we can obtain an estimate for the risk aversion parameter γ and the bias (or residual hedging demand). We start this section by providing evidence from pooled regressions and then proceed to show results pertaining to the distribution of γ_i obtained from multiple regressions. Table B9 reports the results of the following pooled regression:

$$r_{j,t+1} - r_f = \alpha + \gamma(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1} \tag{2.23}$$

where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \boldsymbol{w}^*_{i,t})_j$ is the scaled demand for stock j of manager i, at time t. If we assume that preferences are constant across managers and time, we obtain a risk aversion coefficient close to unity (between 0.915 and 1.283 across specifications) for our *representative investor*. While the estimate is low compared to other measures obtained from equity returns (Mehra and Prescott, 1985; Kocherlakota, 1996), it is consistent with measures derived from labour choices (Chetty, 2006) and option prices (Martin, 2017). Our *representative investor* displays a quite large and statistically significant bias (or residual hedging demand) of about 1% per quarter.

The pooled results in Table B9 mask a sizeable amount of variability across managers. For this reason, we proceed by estimating separate regressions, one for each manager in the sample:

$$r_{j,t+1} - r_f = \alpha_i + \gamma_i(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1} \tag{2.24}$$

Given that there seems to be limited difference resulting from the choice of the covariance matrix Σ_t, we report the results using the sample covariance $\hat{\Sigma}^1_t$. Table B10 reports summary statistics for elicited risk aversion and bias referring to measure (1)[33]. We obtain a median (average) relative risk aversion of 1.117 (1.236), in line with the pooled results; however, there is a wide dispersion in the estimates with a standard deviation of 5.850. The estimates display positive skewness and are leptokurtic. When we allow for variation in preferences across managers, the bias is reduced on average: the mean bias is only 0.7% and the median bias is 1%

[33] The results for measure (4) can be found in Appendix B.6.

per quarter. Figure B9 displays histograms of the distribution of α_i and γ_i after we have removed outliers. Unfortunately our methodology does not prevent us from obtaining negative values for γ_i whenever the cross-section of revealed beliefs is negatively correlated with realised returns. Most of the mass, however, seems to fall in the positive value region.

We then proceed to exploit the variation of preferences across managers and analyse whether tenure affects risk aversion and bias. Figure B10 displays the bias and the risk aversion as a function of tenure for measures (1) and (4). It is hard to detect a specific pattern in either of the measures; longer tenures seem to be dominated by noise, given that they make use of fewer estimations by construction. Finally, Figure B11 reports the results by date: also in this case it is hard to detect any conclusive evidence. Unfortunately, our measures of risk aversion cannot be used to predict or explain future returns given that they have been obtained from them: by construction they represent the best linear predictor of $r_{j,t+1} - r_f$ given the information contained in $(\Sigma_t \boldsymbol{w}^*_{i,t})_j$.

2.7 Conclusions

This paper contributes to the literature on the effect of personal experience on learning and expected returns by analysing a large sample of more than 3,000 professional investors (mutual fund managers) that have been tracked throughout their careers in the 35 years period between 1980 and 2015. Section 2.2.1 shows that in a variety of cases it is possible to invert the portfolio demands of our investors to obtain their subjective expected returns by using the identifying assumption that, while beliefs vary at the stock-investor-time level, risk aversion varies at the investor-time level, i.e., risk aversion is constant in the cross-section of stock holdings of a given manager. Similarly, we are able to account for many cases in which demands display a hedging component by saturating the regressions with fixed effects. Indeed, as we show in Section 2.3, almost ninety percent of recovered expected returns can be explained by manager-time and stock-time fixed effects. We then provide reduced-form evidence that professional investors overweight experienced returns compared to other information shared across stocks and individuals: having experienced a one standard deviation increase in quarterly returns on average leads to an increased expected return of about 10-15 basis points per quarter. Various reduced-form specifications in Section 2.4 and the parametric estimation in Section 2.5 confirm that the effect of experienced returns is neither constant nor monotone. We show that investors exhibit *recency* and *first impression bias*: an investor with a stock-specific experience of nine quarters overweights the most recently observed quarterly returns by 1.84 times and the earliest experienced returns by 3.13 times relative to the constant weight benchmark. These results are most apparent for managers working alone, as opposed to in a team of two or more, suggesting that a significant fraction, though not the entirety, of the effect of personal experience cancels out once aggregated. By looking at managers who have switched funds, we eliminate the possibility that these findings are purely driven by tax considerations: more than 80% of the effect remains unexplained by tax concerns. We finally turn to the issue of estimating risk aversion and find that a *representative investor* displays a coefficient of relative risk aversion around unity. The paper also finds that individual investors exhibit biases when forming expectations. Finally, when we look at more disaggregated measures, we find that there is a large heterogeneity in biases and risk aversion across time and investors. The results in the paper can inform theorists willing to model the preferences and the learning behaviour of professional investors in a way that is consistent with the evidence obtained from portfolio holdings. Consistent with theory, more than half of the variation in expected excess returns can be explained by a common time varying component. However, an incremental forty percent is due to investor-specific and stock-specific time-varying effects, hinting at the possibility of time variation in preferences and stock-specific factors shared across investors. Finally, if interested in modelling the idiosyncratic part of expected returns, one should pay particular attention to behavioural factors which play a prominent role as shown by the evidence provided in this paper.

3. Living on the Edge: the Salience of Property Taxes in the UK Housing Market

FRANCESCO NICOLAI, MARCO PELOSI AND SIMONA RISTESKA[1]

A standard tenet of economic theory is that the statutory incidence of taxes is irrelevant for their economic incidence[2]. It should also be the case that whether a tax is paid at the moment of transaction or later is irrelevant for its incidence, as long as we take into account the time value of money and the riskiness of the cash flows. By looking at the UK residential property market, this paper shows that this is not the case and that deferred taxes have a markedly lower incidence compared to taxes paid at the time of decision-making.

Together with France, the United Kingdom is one of the few countries receiving a sizeable fraction of revenues from property taxes, amounting to about 4.3% of GDP or more than £84 billion in 2016 (European Commission (2018)). The two main taxes levied on domestic properties are the Stamp Duty Land Tax and the council tax. The former is a tax levied on the transaction value of land and any buildings and structures thereon. The fact that its statutory incidence falls on the buyer, who is required to pay the tax liability to the HM Revenue and Customs within very few weeks from the completion of the transaction, and the fact that the tax represents a lump sum ranging between 1% and 7% of the property value are features that make the stamp duty tax particularly salient at the moment of purchase. The latter, which is the focus of the present paper, is a tax levied by the local government on a yearly basis. The council tax is levied on the resident, as opposed to the house owner, and is based on the property value in 1991. While the council tax is extremely salient at the moment when it needs to be paid, we show that this is not the case at the moment when properties are purchased even though, in present value terms, it is similar to or even larger than the stamp duty tax. By using the geographical discontinuity at the border of different local authorities in the London area, we are able to estimate the incidence of the council tax on property prices and contrast it with the incidence of the stamp duty tax estimated, among others, by Best and Kleven (2018). The London area is particularly suitable for the estimation because of the sharp nature of the

[1]We are grateful to Vicente Cuñat, Daniel Ferreira, Dirk Jenter, Christian Julliard, Daniel Paravisini, Andrea Tamoni, Michela Verardo and the participants at the LSE PhD seminar for the useful comments on the paper. We thank Vittorio Raoul Tavolaro for invaluable research assistance. The paper contains HM Land Registry data © Crown copyright and database right 2019. The data is licensed under the Open Government Licence v3.0.1. We thank the University of Glasgow - Urban Big Data Centre for providing Zoopla property data. Zoopla Limited, © 2019. Zoopla Limited. Economic and Social Research Council. Zoopla Property Data, 2019 [data collection]. University of Glasgow - Urban Big Data Centre. Any errors or omissions are the responsibility of the authors.

[2]Kotlikoff and Summers (1987) provide a detailed review of classical theory on tax incidence.

council borders and the large dispersion in council tax rates across Boroughs. For instance, Figure C1 depicts a road that is at the border of the Borough of Westminster and the Borough of Kensington and Chelsea. As can be seen from the picture, the houses on both sides of the street are otherwise identical except for the fact that they pay quite different council tax amounts: the ones on the left pay £2,279 per year in council tax while those on the right pay £1,421 per year. If we discount the future payments as a perpetuity at a rate of 4%, similar to the mortgage rates observed in sample, we obtain that the difference between the two present values amounts to £21,450 (about \$28,000). The tax differentials become even more significant once we consider the fact that many London Boroughs share services, such as waste management, and that many other amenities, such as access to parks, schooling and religious facilities, are not strictly limited to residents of a given Borough. In Section 3.3 we show that the price of similar properties on opposite sides of a border does not adjust for differentials in council tax amounts. By employing a variety of estimators, we establish that the council tax incidence is never statistically negative. We then proceed in Section 3.4 to set up a model where downpayment-constrained households purchase a house and pay two sets of taxes: a lump sum stamp duty tax levied at the moment of purchase and a periodic council tax. We move on to perform a Bayesian analysis in Section 3.4.1 where we provide a posterior range for council tax incidence using priors that are economically motivated. In all these estimates, the incidence of council tax on property prices is too low relative to existing estimates of the incidence of other property taxes, even after accounting for time value of money and the fact that discount rates might be larger because of borrowing constraints. These findings can be rationalised in a model where agents neglect taxes that are levied in the future. We show in Section 3.4.2 that then a trade-off between the two types of taxes arises: the stamp duty tax is distortionary because agents are liquidity constrained; on the other hand, the council tax leads agents to over-consume the housing good and, therefore, distorts their consumption choices by reducing available income. As a result, we demonstrate that the Government can optimally tune the two taxes to minimise distortions for a given level of revenue.

The present paper adds on to the burgeoning literature on behavioural public finance and the salience of taxes (or the lack thereof). Chetty et al. (2009) is the first paper to empirically estimate how differences in salience can alter the behaviour of economic agents. They intervene in a grocery store in order to modify the salience of sales taxes and show that the incidence on buyers is largely reduced when taxes are made fully salient. In a second experiment they compare the effect of excises taxes, which are included in posted prices, and sales taxes, which are not explicitly included, on alcohol demand and again show that tax salience plays an important role in consumer behaviour. The setting in the present paper is quite similar to the second experiment in Chetty et al. (2009), given that the stamp duty tax is paid upfront while the council tax is deferred and thus less salient. For policy reasons, however, the question of property taxes is of greater importance because of the large amounts of money involved and the fact that it is very difficult for agents to learn since buying a new property is typically a once-in-a-lifetime event. Following Chetty et al. (2009), other papers have also explored the question of tax salience, for instance, Feldman and Ruffle (2015) and Feldman et al. (2018) have replicated the findings in laboratory experiments, while Finkelstein (2009) similarly shows that the introduction of electronic toll payments raises toll expenditures. Taubinsky and Rees-Jones (2018) further explore the topic by showing that there is large variation in the way agents react to tax salience and investigate policy implications. The present paper is also akin to Allcott (2011) who demonstrates that a similar bias is present in the automobile market, namely, car buyers fail to correctly price in the future energy cost at the time of purchase. As in Allcott (2011), our conclusions also rely on the choice of an appropriate discount factor. We show in Section 3.4.1 that the bias persists even after allowing for large discount rates. In a similar vein, using Norwegian data, Agarwal and Karapetyan (2016) explore the effect of non-salient debt features on households' purchasing decisions and show that they do not fully factor in the added cost. The authors show that the mispricing was eliminated once these features became fully salient. Finally, the paper extends the literature on property taxes; among others, we use the results of Besley et al. (2014) and, in particular, Best and Kleven (2018) to compare our estimates of the council tax incidence with their stamp duty incidence estimates in order to highlight the lack of salience of the former.

The rest of the paper is organised as follows: Section 3.1 describes the data and the institutional setting; Section 3.2 gives evidence of the geographical distribution of council taxes and points out that this can significantly bias our estimates if not appropriately taken care of, before proceeding with the details of our identification strategies; Section 3.3 presents the empirical estimates of the council tax incidence; Section 3.4 develops a stylised model to help interpret the findings and shows that the estimated incidence is too low to be consistent with fully-salient taxes, before exploring some policy implications; and finally, Section 3.5 summarises and concludes the paper.

3.1 Data

To estimate the incidence of council taxes we need access to data on property characteristics and house prices, as well as council taxes paid. Price paid data on house transactions are readily available from the HM Land Registry website. This dataset contains information about all residential properties transacted in England and Wales from 1995 that have been sold for full market value[3]. The dataset comprises of the

[3]Data excluded from the dataset include commercial transactions, property transactions that have not been lodged in with HM Land Registry and transactions made below market value. For

price paid, the transaction date and, most importantly, the address of the house which allows us to pinpoint the exact location of every property. Additionally, the data provide us with information on the property type, which can be one of five possible categories (a detached, semi-detached, or terraced house, a flat/maisonette and other), the age of the property (classified into old or new to distinguish between newly built properties and already established buildings) and the duration of tenure, i.e., whether the property is under a freehold or leasehold[4].

Since we would ideally like to compare properties that are as similar to each other as possible, we need more information on property characteristics. For this purpose we make use of two additional datasets: the Zoopla Property data and Domestic Energy Performance Certificates. The Zoopla Property data[5] has been collected by Zoopla, one of the UK's leading providers of property data for consumers and property professionals, giving free access to information on 27,000,000 property records, up to 1,000,000 property listings and 15 years of sold prices data. The dataset covers the period between 1st January 2010 and 31st March 2019 for properties located in Great Britain (England, Wales, Scotland). The dataset contains details on characteristics such as property location, property type[6], whether the property has been categorised as residential or commercial[7], number of bedrooms, number of floors, number of bathrooms, number of receptions and whether the property is listed for sale or for rent[8]. In addition, we also have access to the asking price for both rents and sales, however, we use the more accurate transaction price from the HM Land Registry dataset. The second source of house characteristics comes from the Ministry of Housing, Communities and Local Government. On their website, one can access the Energy Performance Certificates (EPC) for domestic and non-domestic buildings. For domestic properties, before 2008 certificates could be lodged on a voluntary basis. From 2008 onwards, however, it has become mandatory for accredited energy assessors to lodge the energy certificates. Consequently, the data coverage drastically improves around that time, as does our ability to match these data with the price paid data. More specifically, the matching rate jumps from

more details on the property sales not included in the dataset the reader can visit the HM Land Registry website: https://www.gov.uk/guidance/about-the-price-paid-data.

[4]Note that leases of seven years or less are not recorded in the dataset.

[5]The access to the dataset has been kindly provided by the University of Glasgow - Urban Big Data Centre. Access to the dataset for research purposes can be obtained directly through the Urban Big Data Centre. The data has been collected by Zoopla. Zoopla Limited, © 2019. Zoopla Limited. Economic and Social Research Council. Zoopla Property Data, 2019 [data collection]. University of Glasgow - Urban Big Data Centre.

[6]Property types include: barn conversion, block of flats, bungalow, business park, chalet, château, cottage, country house, detached bungalow, detached house, end terrace house, equestrian property, farm, farm house, finca, flat, hotel/guest house, houseboat, industrial, land, leisure/hospitality, light industrial, link-detached house, lodge, longère, maisonette, mews house, mobile/park home, office, parking/garage, pub/bar, restaurant/cafe, retail premises, riad, semi-detached bungalow, semi-detached house, studio, terraced bungalow, terraced house, town house, unknown, villa and warehouse.

[7]We keep only properties categorised as residential.

[8]For the time being we only keep properties listed for sale.

about 50 percent to over 90 percent around 2008. The dataset contains information on the location, property type, total floor area, number of storeys, number of rooms, floor level and height, along with many indicators of energy efficiency and quality of glazed surfaces. The final piece of data needed to conduct our analysis is related to council tax data; in the following section we describe in more detail the functioning of this property tax and the relevant data.

3.1.1 Council Tax

The taxation of properties in the United Kingdom is peculiar compared to other OECD countries, representing a rather large source of both central Government and local authorities' revenues. The three main taxes levied on properties are the council tax, business rates and stamp duty taxes. Council taxes are levied on each occupier, rather than on the owner, of domestic properties. The tax is one of the few levies in Great Britain being both set and collected by local authorities (Boroughs in the case of London) and it represents one of their major sources of revenue (around one-third of total revenue), the other sources being commercial property taxes (business rates) and transfers from the central Government. The tax is based on a classification in eight bands (A-H) based on the value of the property as established by the Valuation Office in 1991; newly built properties are assigned to a band, after having their current value converted into the value of an equivalent property in 1991. Each London Borough is responsible for setting the annual tax amount to be paid by a property in band D every year; the amount to be paid by other bands is automatically set as a ratio to the amount for band D[9]. Bands C and D represent the largest fraction of dwellings (about 50 percent of the total), but there is variation across Boroughs with central properties being skewed towards higher valued bands compared to properties in outer Boroughs. Figure C2 shows the time series of the council tax payable per band per Borough. Each panel in the figure depicts the amount payable by different bands showing that, by construction, the tax moves in locksteps across bands. More interestingly, it should be noted that there is a wide dispersion in amounts payable across Boroughs, even though the ranking across different local authorities remains almost constant with the only exception being the Borough of Hammersmith and Fulham where taxes have been slashed starting from the late 2000s. After a marked increase in council tax rates in the early 2000s, the freeze mandated by the central Goverment after the 2008 financial crisis is visible in the time series; since 2011, taxes can be raised only by a centrally set amount unless a local referendum allows the authority to do so. We show in Section 3.2.1 that the geographical distribution of council tax rates is not random and could severely bias any estimate of incidence, given that central (and pricier) Boroughs tend to set lower council tax rates. This is

[9]The ratios are constant across Boroughs and are as follows: band A 6/9, band B 7/9, band C 8/9, band D 1, band E 10/9, band F 13/9, band G 15/9, band H 2.

mainly because central Boroughs tend to have larger fraction of properties in higher bands; for instance, the Borough of Kensington and Chelsea raises more than fifty percent of its revenues from bands G and H, while Barking and Dagenham raise less than five percent from these bands.

We obtain information on council tax band assignment from the website of the Valuation Office Agency, which provides data on the full address and the council tax band for each property in Great Britain. The average amount to be paid in each London Borough by each band in the period 1999-2018 is obtained from the London Datastore managed by the Greater London Authority.

In the following section, we provide some descriptive statistics of the data we have mentioned so far.

3.1.2 Descriptive Statistics

Figure C3 shows the distribution of transaction prices for domestic properties in London, truncated to exclude extremely high property prices which are, however, included in the analysis. The data consists of 889,925 observations in the period between 1999 and 2018 for which property characteristics and council tax information is available. We confirm that the distribution is highly skewed with the average and median property values being £366,528 and £250,000, respectively. It is immediately obvious that there is a large degree of bunching in prices, as noted for instance in Best and Kleven (2018). The bunching mainly happens just before stamp duty notches, which allows Best and Kleven (2018) to estimate the local incidence of this tax. Figure C4, for instance, shows the large extent of bunching at the threshold of £250,000 (upper panel) and £500,000 (lower panel) where the stamp duty tax jumps from 1% to 3% and from 3% to 4%, respectively. Best and Kleven (2018) estimate a rather large incidence of stamp duty tax on property prices and argue in favour of evidence of rather strict borrowing constraints; we use their estimates to inform our analysis of the incidence of the council tax, allowing us to disentangle how much of the incidence is due to borrowing constraints (or the lack thereof) and how much is attributable to pure time discount. Figure C5 shows the distribution of house prices per band. The vertical red lines depict the median price within each band. As one should expect, higher bands tend to have houses with higher average prices although there is a large dispersion within bands. This is because prices have increased a lot over the past twenty years, especially for more central and higher-banded properties. This makes it essential that we compare only transactions occurring in close periods. Moreover, one should notice that the number of properties belonging to bands C and D dominates the rest, as previously mentioned. In Figures C6, C7, C8, C9 and C10 we show that there is a wide dispersion of transaction prices based on house characteristics such as property type, number of rooms, property age and duration. There is a disproportionate amount of flats in our sample, which we see as an advantage in our estimation, as flats are much more likely to be similar to each other relative to other property types. Detached houses are most expensive, with a median price of £525, 000, followed by semi-detached houses (£319, 950) and terraced houses (£270, 000), and finally, flats are the cheapest category (£195, 000). Naturally, the house price is increasing in the number of rooms with the median value of each additional room being about £40, 000 in the full sample. Newly-built properties represent a minority in our sample and trade at a small discount relative to established buildings. This is due to the geographic distribution of the housing stock in London where older properties tend to be in the more sought-after central areas. However, there is some heterogeneity when we look at the year of construction: properties built before 1949 sold at a median of £287, 000 close to those built after 2003 (£275, 000), while properties built in the period 1950-1982 and 1983-2002 sold at lower prices (£215, 000 and £200, 000, respectively). This pattern can be explained both by differences in type and location across groups. Finally, it can be noted that properties under a freehold ownership have a higher median price (£305, 000) compared to leasehold properties (£195, 000).

After having described the data, we proceed to the discussion of our empirical strategy in the next section.

3.2 Empirical Strategy

3.2.1 Evidence of Selection

The main issue that arises when estimating the incidence of council taxes is the fact that the cross-sectional distribution of council tax amounts across Boroughs is very strongly correlated with other characteristics that affect house prices. To see this, Figure C11 shows a map of the distribution of Band D council tax amounts payable for each London Borough along with the respective distribution of house prices. Panel C11a shows the distribution of council taxes in 2000, where taxes increase moving from yellow to red; Panel C11b the distribution of house prices in the same year, where prices increase moving from light blue to brown. Panel C11c shows the distribution of council taxes in 2018, while panel C11d the distribution of house prices in the same year. It is visually striking that councils with lower taxes tend to have higher house prices. For instance, the City of Westminster had the lowest Band D council tax in 2000 (£375.17) and the second highest average house price (£357, 925), after the Borough of Kensington and Chelsea (£726, 908) which had the fourth lowest council tax (£623.38). In 2018 the same holds true, with the City of Westminster having the lowest council tax (£710.50) and the second highest average price (£1, 612, 231), after Kensington and Chelsea (£3, 040, 547) which had the fifth lowest council tax (£1, 139.41). In general, it is clear from the map that Boroughs that lie further from the centre tend to have higher council taxes and lower prices, while the more central Boroughs tend to exhibit the opposite pattern. To confirm the intuition obtained from Figure C11, we can run a naïve regression of house prices on house characteristics and council tax payable without controlling for the geographical location of the property, i.e.:

$$p_{idbt} = \beta\tau_{dbt} + \delta_{bt} + \zeta' x_{idbt} + \varepsilon_{idbt} \tag{3.1}$$

where p_{idbt} is the price of house i in Borough d, band b at time t; τ_{dbt} is the council tax amount for a house in Borough d, band b at time t; δ_{bt} are year-band fixed effects; and x_{idbt} are controls which include the property size measured in squared meters, number of rooms, property type, age, duration and month which controls for seasonality in the housing market (Ngai and Tenreyro, 2014). Table C1 reports the results of regression (3.1); the first column provides the baseline result where month and year-band fixed effects are included in order to remove the mechanical correlation between increasing property prices and taxes over time and the fact that moving from band A to band H goes hand in hand with higher house prices. If we took this evidence at face value, we would conclude that the incidence of council tax is extremely large and statistically significant with a point estimate of -231.2. To give intuition, using a discount rate of $r = 4\%$ (similar to the risk-free rate observed in sample) this would roughly imply that an extra £1 in present value of taxes would lead to a drop in prices of $r \times \beta = 4\% \times 231.2 = £9.25$. It is obvious that this figure is only the artefact of the negative correlation between the value of properties and the average tax within councils as observed in Figure C11. Extremely negative coefficients are obtained in columns (2), (3) and (4) where we control for the property size, number of rooms, property type, whether the property is newly-built and whether it is a leasehold. The smallest of these coefficients in absolute value, i.e., -228.7 in column (3), would imply an incidence of $r \times \beta = 4\% \times 228.7 = £9.15$ which is still unreasonably high. Table C2 shows similar estimates when we include all the variables available as controls. To further corroborate the negative correlation between property prices and council taxes due to geographical selection, we provide the results of the following two-step estimation. First, we regress house prices on characteristics to obtain hedonic residuals:

$$p_{idbt} = \zeta' x_{idbt} + \varepsilon_{idbt} \tag{3.2}$$

For each Borough, band, year, we compute the median residual price ε^{med}_{dbt} and proceed to regress it on council tax amount payable including year-band fixed effects:

$$\varepsilon^{med}_{dbt} = \beta\tau_{dbt} + \delta_{bt} + \eta_{dbt} \tag{3.3}$$

The results are reported in Table C3. The vector of predictors x_{idbt} in the first-stage hedonic regression includes: month fixed effects in column (1); month, property size, number of rooms in column (2); month, property size, number of rooms and property type in column (3); and month, property size, number of rooms, property type and indicators for whether the property is newly-built and a leashold in column (4). Similarly, Table C4 reports results when the dependent variable in the second stage is the average hedonic residual $\bar{\varepsilon}_{dbt}$ per Borough, band, year, i.e.:

$$\bar{\varepsilon}_{dbt} = \beta \tau_{dbt} + \delta_{bt} + \eta_{dbt} \tag{3.4}$$

Both tables confirm the previous finding that Boroughs with higher house values tend to impose lower council tax bills: the coefficients are negative and statistically significant, ranging from -183.6 to -368.4.

The results provided so far imply that special care needs to be taken before using the geographical variation in council taxes to estimate their incidence on house prices. For this reason in our identification strategy we compare only houses that lie extremely close, i.e., no more than 500 meters and mainly closer than 200 meters, to the border between two adjacent Boroughs in order to disentangle the actual incidence of the tax from the geographical distribution of taxes across Boroughs. Throughout the rest of the paper, the reader should bear in mind that the geographical distribution of council taxes entails that any estimated incidence is, at most, an upper bound for the *true* incidence. This is because, if buyers value certain characteristics upon purchasing a house, these should be capitalised in the house price which, in this case, acts almost like a sufficient statistic for their value; the results of Figure C11 and Tables C1-C4 signal that houses with more highly valued characteristics (and higher prices) tend to be located in Boroughs with lower taxes, thus inflating any estimate of tax incidence. A second and more subtle reason why we can only estimate an upper bound for the incidence has to do with our identification strategy. By comparing similar dwellings on opposite sides of a border, we implicitly assume that the buyer always has an outside option during the price bargaining process. As a result, the buyer would be much more elastic than an otherwise identical buyer involved in the purchase of a house located in the heart of a Borough where there is no outside option in terms of council tax. We show in Section 3.4 that the seller bears the full incidence of the tax at the border, while that is not necessarily the case at an interior point. In general, even in the absence of perfect substitutes across council borders, it is reasonable to conjecture that the incidence is still much larger at the border compared to the council centre, where the agent would have to move long distance in order to pay a different council tax rate.

In the next section we describe the identification strategies that allow us to estimate the incidence of council taxes as precisely as possible given the present setting, bearing in mind that any attempt is likely to result in an over-estimation of the *true* incidence.

3.2.2 Identification Strategies

We use two different identification strategies to measure an upper bound of the incidence of council tax on property prices: grid regressions and a matching algorithm.

Grid Regressions

The first strategy compares houses that lie in close proximity by dividing the area of London in a grid and assigning a fixed effect to each square in the grid. By doing so, we are de-facto comparing two houses that are otherwise identical but lie on opposite sides of a given border between two Boroughs. Figure C12 graphically depicts our first approach. Panel C12a shows a grid of squares with equal sizes superposed on a map of London. Panel C12b shows a more detailed picture of the Boroughs in the centre[10]. We then proceed to select the squares that have two houses that: are sold in the same year, are in the same council tax band and lie on opposite sides of the border; Panel C12b displays in blue examples of such squares. It can be noticed that we discard observations for which the border is located on the Thames River bank. To avoid relying on an arbitrary division, we use three different grids, namely one grid divides the area in 50×50 squares, another divides it in 100×100 squares and, finally, the last grid is a 150×150 one. These squares have an approximate size of 800 meters, 400 meters and 250 meters, respectively. While the maximal possible distance between houses can be inferred as $\sqrt{2} \times$ square side length, we choose to remove observations that are more than 500 meters far from the border. Figure C13 shows the distribution of distances to the border for our different specifications. As mentioned, no house lies more than 500 meters away from the border, and most of the observations are about 200 meters away from the closest border. As we proceed to refine our grids by subdividing into a larger number of squares, we can see that we lose observations in the 200 meters-500 meters range; this reduces our power significantly, but ensures that we compare houses that are indeed in very close proximity.

Our strategy consists of running within square regressions whereby we compare houses that are sold in the same year and in the same council tax band, specifically:

$$p_{ibgdt} = \beta\tau_{bdt} + \delta_{bgt} + \zeta' x_{ibgdt} + \varepsilon_{ibgdt} \tag{3.5}$$

where p_{ibgdt} is the price of house i, in council tax band b, grid square g, Borough d, and year t; τ_{bdt} is the council tax amount for band b, Borough d in year t; and x_{ibgdt} are house-specific controls. The presence of the band-grid square-year fixed effects δ_{bgt} guarantees that the regression compares houses that are in the same square, same council tax band and are sold in the same year, implying that our identification

[10] The three main Boroughs depicted in the picture are, starting from left, Hammersmith and Fulham, Kensington and Chelsea and the City of Westminster.

assumption is that they systematically differ only due to the amount of council tax paid, after partialling out the effect of house characteristics that we add to increase our precision. It should be noticed that, as mentioned above, *better* Boroughs, i.e., Boroughs with higher average prices, tend to have lower council taxes, implying that - if we leave some hidden characteristic out of our regression - the estimate of β is most likely going to overstate the *true* incidence. To give an example, while highly unlikely given the sharp nature of the borders, one could argue that there is a name tag value of living in certain Boroughs over others, for instance, a house in Westminster commands a premium over a similar house on the other side of the border in Brent. The fact that Westminster has a lower tax compared to Brent implies that we would overestimate the incidence of the tax because of the name tag value of living in Westminster. In general, to reverse this bias and claim that the *true* incidence might be higher than the one we estimate, the reader should think of some hidden characteristic that systematically causes people to prefer living in a Borough with worse amenities compared to a Borough with better ones.

The following section presents our second identification strategy which relies on a matching estimator rather than grid squares fixed effects.

Matching Estimator

Our second identification approach consists of pairwise matching of houses on opposite sides of a given border. To find the closest match, we need to define a distance: in what follows, we rely on a Euclidean distance and a distance based on a linear model. Under the first one, we restrict the possible matches to be: no more than 500 meters away from each other, sold in the same year, in the same council tax band, and to both be either old or newly-built and freehold or leasehold properties. For each property we then choose the closest match as the one minimising the Euclidean distance $d(i,j) = \sqrt{\sum_{k=1}^{K}(x_{ik} - x_{jk})^2}$, where i is the original property, j indexes the possible matches on the other side of the border, x_{ik} are house i characteristics and x_{jk} are house j characteristics. We then run within-pair regressions:

$$p_{ibdt} = \beta\tau_{bdt} + \delta_{ij} + \zeta' x_{ibdt} + \varepsilon_{ibdt} \tag{3.6}$$

where δ_{ij} are ij-pair dummies and x_{ibdt} are house i-specific features. The second choice of distance is based on a linear pricing model:

$$p_{it} = \alpha + \beta' x_{it} + \varepsilon_{it} \tag{3.7}$$

where x_{it} similarly contains house-specific characteristics as above. We then compute the model-predicted price $\hat{p}_{it} = \hat{\alpha} + \hat{\beta}' x_{it}$. As before, we restrict the pairing to houses sold in the same year, band, old/new and leasehold/freehold categories and no further than 500 meters from each other. For each property i we pick the closest match j as the one that minimises the following distance: $d(i,j) = |\hat{p}_{it} - \hat{p}_{jt}|$. To estimate the incidence, we run within pair-regressions as in equation (3.6) where the δ_{ij} dummies are determined according to the new matching algorithm. As in Section 3.2.2 the identification is valid as long as the only systematic difference within pairs is the amount of council tax. As previously explained, any other omitted variable would most likely lead us to estimate an upper bound for the incidence, given the geographical distribution of council taxes.

3.3 Results

3.3.1 Grid Estimator

Table C5 presents the results of the grid regressions described in Section 3.2.2 where we use a 50×50 grid and include band-grid square ID-year fixed effects to compare the effect of council taxes on properties in the same band, sold in the same year, located in the same grid square but on opposite sides of a border as in equation (3.5). The controls we include are as follows: column (1) uses month fixed effects to control for housing market seasonality; column (2) adds number of rooms fixed effects and controls for property size; column (3) also adds property type fixed effects, and; column (4) includes an indicator for newly-built and leasehold properties. These are our default specifications throughout the rest of the paper. In all columns the coefficient on council taxes is statistically indistinguishable from zero and always with the wrong sign. The lack of significance cannot be attributed to lack of statistical power in the regressions given that other control variables are always strongly statistically significant. For instance, the effect of one additional squared metre ranges between £4,537 and £4,627, newly-built properties command a premium of about £33,400 and freehold properties sell for £76,000 more relative to leaseholds. The same conclusion can be drawn from Table C6 where we expand the regressions to include all available house price predictors, showing that even relatively minor characteristics such as the number of lighting outlets or the presence of fireplaces in the property have a significant effect on prices.

Table C7 displays the grid regression results for grids of different sizes: column (1) uses a grid that divides the London area into 50×50 squares, column (2) 100×100, and column (3) 150×150. This might help to alleviate concerns that grids made of large squares might be comparing houses that are rather distant from each other. The specification is otherwise same as the one in column (4) of Table C5. The coefficient on council tax remains statistically insignificant and the point estimate varies from positive to negative across columns: this is precisely what we should expect when a regressor has no effect on the outcome variable and simply reacts to the noise in the sample. The fact that the R-squared is very high (between 77% and 83%) and that all other coefficients are precisely estimated confirms our previous finding that the incidence of the council tax is indistinguishable from zero. In Table C8 we augment the regressions by adding all additional house characteristics: the coefficient on council tax ranges from -11.8 to 75.4 and is never statistically lower than zero.

To make sure that the confounding effect of the stamp duty notches does not play a role in our estimation results, Table C9 presents the results of the grid regressions when we remove the two main stamp duty notches at £250,000 and £500,000. Column (1) excludes only the first notch, column (2) the second, and column (3) removes both. The results are virtually unchanged, with the incidence still being statistically insignificant, small in magnitude, and always displaying the wrong sign. As previously mentioned, the large R-squared and the fact that the remaining coefficients are precisely estimated guarantees that this is not due to lack of power.

Finally, Tables C10 and C11 provide estimates of council tax incidence using a similar two-step approach as in Tables C3 and C4, i.e., by first obtaining residual hedonic prices as follows:

$$p_{ibdgt} = \zeta' x_{ibdgt} + \varepsilon_{ibdgt} \tag{3.8}$$

and subsequently regressing the median or average hedonic residuals for each Borough, band, grid square and year on council tax amounts:

$$\varepsilon^{med}_{bdgt} = \beta \tau_{bdt} + \delta_{bgt} + \eta_{bdgt} \tag{3.9}$$

$$\bar{\varepsilon}_{bdgt} = \beta \tau_{bdt} + \delta_{bgt} + \eta_{bdgt} \tag{3.10}$$

where δ_{bgt} are band-grid square-year fixed effects included to ensure that we compare values of houses in the same council tax band, sold in the same year and located in the same square of the grid. As usual, we restrict the analysis to grid squares with at least two houses located on different sides of a border and present the four standard specifications. The results confirm the previous finding: both the median and average hedonic residuals are not decreasing in the council tax amount paid, suggesting that the incidence of this tax on house prices is not different from zero.

In the following section we supplement the evidence by presenting results using our second identification strategy.

3.3.2 Matching Estimator

Tables C12, C13 and C14 show the results of our second estimation approach where we explicitly match similar dwellings on opposite sides of a border as described in Section 3.2.2. As previously mentioned, all the results are obtained using housing pairs on opposite sides of a border no more than 500 metres apart, sold in the same year, in the same council tax band and which are both either old or newly-built and leasehold or freehold properties. Table C12 displays the results where closest pairs have been determined by minimising the Euclidean distance $d(i,j) = \sqrt{\sum_{k=1}^{K}(x_{ik} - x_{jk})^2}$, where the vectors x_i and x_j consist of property size and number of rooms in columns (1) and (2), and also energy cost in columns (3) and (4). All the variables are standardised to be comparable. This procedure leads to 57,612 and 57,323 observations of property pairs with 71,578 and 71,656 unique transactions in columns (1)-(2) and (3)-(4), respectively[11]. After having obtained the pairs, we run the regression specified in equation (3.6). The presence of δ_{ij} pair fixed effects amounts to regressing the difference in prices of matched houses on the difference in council tax paid, controlling for other property characteristics along which the matched properties may differ. Consistent with the results obtained with the grid estimator, none of the coefficients on council tax is statistically significantly negative. As pointed out before, this result is not attributable to lack of statistical power: for instance, the coefficient on size is highly statistically significant and has the same order of magnitude as the ones obtained with the earlier estimator[12]. Table C13 confirms these findings under the linear matching algorithm where pairs are chosen by minimising the distance $d(i,j) = |\hat{p}_{it} - \hat{p}_{jt}|$, where the predicted prices $\hat{p}_{it}$ and $\hat{p}_{jt}$ are obtained from a linear model as in equation (3.7). As before, columns (1) and (2) match properties based on size and number of rooms, while columns (3) and (4) add energy cost. Finally, Table C14 presents the last set of results for the linear model where we allow each property to be paired with more than one similar property on the other side of the border, as long as the absolute difference in predicted prices is less than 30% of the largest predicted price, namely: $|\hat{p}_{it} - \hat{p}_{jt}| < 0.3 \times \max\{\hat{p}_{it}, \hat{p}_{jt}\}$. While the point estimates range between -5.24 and -8.19, none of the coefficients is statistically different from zero as in all previous specifications. We shed more light on the interpretation of these and the previous results in Section 3.4.1.

The empirical findings above demonstrate that council tax differences never significantly explain house price differences. Moreover, while absence of evidence, namely the fact that agents seem to be insensitive to taxes that are postponed to the future, does not directly imply evidence of absence, many point estimates are positive and hence with the wrong sign. Bearing these estimates in mind, in the next section we develop a simple model that allows us to propose a plausible explanation for the above results. We subsequently calibrate the model using a Bayesian approach informed by all of the above estimates and briefly discuss policy implications.

[11]Notice that any given transaction can be the closest match for more than one property. In order to take care of this redundancy we cluster standard errors at the transaction ID level.

[12]Notice that, compared to the default specifications used in Tables C1, C5, C7 and C9, the indicators for newly-built and leasehold properties have been dropped given that properties are constrained to be identical along these dimensions.

3.4 Model

In what follows, we present a simple multi-period model of housing-consumption choice in order to calibrate the above results. We begin with the optimisation problem of an agent who chooses at time $t = 0$ an infinite stream of consumption $\{c_t\}_{t=0}^{\infty}$ and a composite housing good h:

$$\max_{\{\{c_t,d_t\}_{t=0}^{\infty},h,\mathbb{1}_{\{A\}},\mathbb{1}_{\{B\}}\}} \quad U(\{c_t\}_{t=0}^{\infty},h) = c_0 + \sum_{t=1}^{\infty}\beta^t u(c_t) + \sum_{t=0}^{\infty}\beta^t \log(h) \tag{3.11}$$

$$s.t. \quad c_0 + h(p_{A0}\mathbb{1}_{\{A\}} + p_{B0}\mathbb{1}_{\{B\}} + \tau_S) \leq w_0 + d_0 \tag{3.12}$$

$$c_t + h(\tau_{At}\mathbb{1}_{\{A\}} + \tau_{Bt}\mathbb{1}_{\{B\}}) + d_{t-1}(1+r) \leq w_t + d_t \quad t = 1,2,3,... \tag{3.13}$$

$$d_t \leq \alpha h(p_{At}\mathbb{1}_{\{A\}} + p_{Bt}\mathbb{1}_{\{B\}}) \quad t = 0,1,2... \tag{3.14}$$

For simplicity, the utility of the agent is chosen to be time-separable and separable in consumption and housing. The utility function is quasi-linear in c_0 in order to get rid of income effects, as is standard practice in the public finance literature. For tractability and to separate the effects of stamp duty and council tax, the agent purchases the housing good only once at $t = 0$. There are two Boroughs, A and B, with exogenously chosen and potentially different council tax rates. We assume that there is equal supply of housing in both Boroughs[13]. Equation (3.12) is the first-period budget constraint: the agent spends his initial endowment w_0 on consumption c_0 and the after-tax cost of his housing demand h. When he buys a house, the agent pays the pre-tax price $p_{i0}, i = A, B$, and, in addition, he also needs to pay the stamp duty tax τ_S hereby assumed to be proportional to the quality-adjusted level of housing demand. If his total demand exceeds his initial endowment, the agent can borrow additional funds d_0 for one period at the risk-free rate. The budget constraints for all subsequent periods are identical and given by equation (3.13): from time $t = 1$ onwards, the agent spends his endowment w_t on his optimal consumption choice c_t and to pay the council tax $\tau_{it}, i = A, B$, that corresponds to the Borough where he has chosen to locate at time $t = 0$. He also needs to repay his short-term debt from the previous period inclusive of interest $d_{t-1}(1+r)$, and is allowed to borrow again at the same terms in order to balance his budget constraint. Finally, the last constraint in equation (3.14) is the financing constraint: the agent cannot borrow more than a fraction α of the pre-tax cost of his housing demand. This can potentially generate very large incidence for the stamp duty tax since the lump sum nature of this tax tightens the leverage constraint. The Lagrangian for the

[13]This assumption is crucial and de-facto eliminates the potential for a differential elasticity of supply with respect to council taxes at the border. We consider this assumption quite reasonable given that the greatest majority of the housing stock in London has been constructed well before the introduction of this tax in the early 90s as shown in Figures C8 and C9.

above problem can be written as:

$$\begin{aligned}\mathcal{L} = \; & U(\{c_t\}_{t=0}^{\infty}, h) - \lambda_0(c_0 + h(p_{B0} + \tau_S) - w_0 - d_0) \\ & - \sum_{t=1}^{\infty} \lambda_t(c_t + h\tau_{Bt} + d_{t-1}(1+r) - w_t - d_t) - \sum_{t=0}^{\infty} \mu_t(d_t - \alpha h p_{Bt}) \\ & - h\mathbb{1}_{\{A\}} \left[\lambda_0(p_{A0} - p_{B0}) + \sum_{t=1}^{\infty} \lambda_t(\tau_{At} - \tau_{Bt}) - \alpha \sum_{t=0}^{\infty} \mu_t(p_{At} - p_{Bt}) \right]\end{aligned} \tag{3.15}$$

where we use the fact that $\mathbb{1}_{\{B\}} = 1 - \mathbb{1}_{\{A\}}$. Notice that the Lagrangian is monotone in the choice of Borough $\mathbb{1}_{\{A\}}$, therefore, the choice of where to locate can be separated from the consumption and housing-quality choices. The agent chooses to live in Borough A if:

$$p_{A0} - p_{B0} \leq - \sum_{t=1}^{\infty} \frac{\lambda_t}{\lambda_0}(\tau_{At} - \tau_{Bt}) + \alpha \sum_{t=0}^{\infty} \frac{\mu_t}{\lambda_0}(p_{At} - p_{Bt}) \tag{3.16}$$

i.e., if the price differential between the same-quality house in Boroughs A and B more than compensates for the present value of the difference in future council tax payments and the collateral value of the house. In equilibrium, markets clear if equation (3.16) holds with equality which, from now onwards, we assume to be the case. Assuming that the agent is indifferent between living in Boroughs A and B, we proceed by suppressing the Borough subscripts and denote the price of the house as p_t and the council tax as τ_t. The first-order conditions for an interior solution are:

$$1 = \lambda_0 \tag{3.17}$$

$$\beta^t u'(c_t) = \lambda_t \quad \forall t = 1, 2, 3, ... \tag{3.18}$$

$$-\lambda_t + \lambda_{t+1}(1+r) + \mu_t = 0 \quad \forall t = 0, 1, 2, ... \tag{3.19}$$

$$\frac{h^{-1}}{(1-\beta)} = \lambda_0(p_0 - \alpha\frac{\mu_0}{\lambda_0}p_0 + \tau_S) + \sum_{t=0}^{\infty} \lambda_{t+1}\tau_{t+1} - \sum_{t=0}^{\infty} \lambda_{t+2}\frac{\mu_{t+1}}{\lambda_{t+2}}\alpha p_{t+1} \tag{3.20}$$

Combining the first-order conditions for consumption and for the optimal debt choice, we obtain the following Euler equation:

$$\frac{\lambda_{t+1}}{\lambda_t} = \beta\frac{u'(c_{t+1})}{u'(c_t)} = \frac{1}{1 + r + \frac{\mu_t}{\lambda_{t+1}}} \tag{3.21}$$

The above Euler equation implies that the agent's discount factor is equal to the inverse of the risk-free rate and a liquidity premium $\frac{\mu_t}{\lambda_{t+1}}$, arising from the fact that the house has some collateral value. In order to simplify the exposition, we assume that in equilibrium the liquidity premium is constant and equal to $\frac{\mu_t}{\lambda_{t+1}} = k$, that house prices grow at a constant rate g, i.e., $p_{it} = p_{i0}(1+g)^t$, and council tax amounts grow at a constant rate $\tilde{g}$, i.e., $\tau_{it} = \tau_{i1}(1+\tilde{g})^{t-1}$. Re-arranging equations (3.16), (3.20) and (3.21), we obtain the final no-arbitrage condition and housing demand:

$$(p_{A0} - p_{B0})\left(1 - \frac{\alpha k}{r+k-g}\right) = -(\tau_{A1} - \tau_{B1})\frac{1}{r+k-\tilde{g}} \tag{3.22}$$

$$\frac{h^{-1}}{(1-\beta)} = p_0\left(1 - \frac{\alpha k}{r+k-g}\right) + \tau_S + \frac{\tau_1}{r+k-\tilde{g}} \tag{3.23}$$

The first equation is the equilibrium condition of how house prices should behave across Boroughs: the house price differential, after having taken into account the collateral value $\frac{\alpha k}{r+k-g}$, needs to match (the negative of) the present value of the council tax differential. The second equation states that the agent's marginal utility of housing is equal to the house price inclusive of (the present value of) all taxes and collateral value. It is important to note that the no-arbitrage condition (3.22) in general gives a different incidence compared to the one obtained from the housing demand (3.23). This is because the former holds only at the border between two Boroughs where the outside option, i.e., the option to buy an otherwise identical house on the other side of the border, implies that the supply bears the whole burden of the tax. In particular, from equation (3.22) we obtain an incidence of:

$$\frac{dp_0}{d\tau_1} = -\frac{1}{r+k-\tilde{g}} \times \frac{r+k-g}{r+(1-\alpha)k-g} \tag{3.24}$$

On the other hand, for both houses on the border as well as houses in the middle of a given Borough we can define the optimal demand from equation (3.23) as $D(p_0, \tau_1, \tau_S) = h^*(p_0, \tau_1, \tau_S)$. Equating with the optimal supply, $S(p_0) = D(p_0, \tau_1, \tau_S)$, and after total differentiation we obtain the standard formula for the incidence:

$$\frac{dp_0}{d\tau_1} = -\frac{\frac{\partial D}{\partial \tau_1}}{\frac{\partial D}{\partial p_0} - \frac{\partial S}{\partial p_0}} = -\frac{1}{r+k-\tilde{g}} \times \frac{1}{\frac{r+(1-\alpha)k-g}{r+k-g} + \tilde{\eta}_S} \tag{3.25}$$

where $\tilde{\eta}_S = \frac{\partial S}{\partial p_0}\frac{p_0}{S}\frac{p_0\left(1-\frac{\alpha k}{r+k-g}\right)+\tau_S+\frac{\tau_1}{r+k-\tilde{g}}}{p_0} = \eta_S\frac{p_0\left(1-\frac{\alpha k}{r+k-g}\right)+\tau_S+\frac{\tau_1}{r+k-\tilde{g}}}{p_0}$ is a slightly modified version of the supply elasticity η_S that takes into account the price inclusive of taxes and collateral value. In general, we have that:

$$\frac{1}{\frac{r+(1-\alpha)k-g}{r+k-g} + \tilde{\eta}_S} \leq \frac{r+k-g}{r+(1-\alpha)k-g} \tag{3.26}$$

implying that the incidence at the border between Boroughs is an upper bound for the *true* council tax incidence as long as the modified elasticity of supply is non-negative, i.e., $\tilde{\eta}_S \geq 0$. Notice that the modified elasticity of supply $\tilde{\eta}_S$ is positive as long as the true elasticity of supply η_S is positive.

3.4.1 Calibration

The model in the previous section allows us to better interpret the empirical results of Section 3.3. By using equations (3.22), (3.23) and (3.24) we get[14]:

$$\frac{dp_0}{d\tau_1} = \frac{dp_0}{d\tau_S} \times \frac{1}{r + k - \tilde{g}} \tag{3.27}$$

i.e., the incidence of the council tax can be interpreted as the present value of the sum of the incidence of the stamp duty tax discounted at the liquidity-adjusted cost of capital $r + k$ with growth rate $\tilde{g}$. In what follows we use the results in Tables C5 - C14 and provide further direction on how to interpret them. We treat each estimate as a separate model m. Conditional on the model being true and given a common prior distribution $p(\beta_\tau | m) = p(\beta_\tau)$ about the true incidence of council tax and the likelihood function of the data $p(y|\beta_\tau, m)$ we can use Bayes' rule to express the posterior distribution for the incidence under each model m as:

$$p(\beta_\tau | y, m) = \frac{p(y|\beta_\tau, m) \times p(\beta_\tau)}{\int p(y|\beta_\tau, m) \times p(\beta_\tau) d\beta_\tau} \tag{3.28}$$

We then proceed to obtain the model-averaged posterior distribution as:

$$p(\beta_\tau | y) = \sum_m p(\beta_\tau | y, m) p(m|y) \tag{3.29}$$

The computational burden of equation (3.29) is significant, therefore, we proceed with the simplifying assumptions described in Appendix C.4. We always start from a normally-distributed prior $\beta_\tau \sim \mathcal{N}(b_\tau, \sigma_\tau^2)$ and likelihood function which leads to a normal posterior. As detailed in Appendix C.4 the mean of the prior is chosen by calibrating the parameters g, $\tilde{g}$, r and α based on historical data and matching the stamp duty incidence to results in Best and Kleven (2018). For robustness we also vary the precision of the prior and provide results for five different specifications: $p(\beta_\tau) = \mathcal{N}(-150, 50^2), \mathcal{N}(-100, 50^2), \mathcal{N}(-50, 50^2), \mathcal{N}(-150, 75^2), \mathcal{N}(-50, 25^2)$.

Figure C15 plots the model-averaged density of the posterior distribution for the council tax incidence. Panel (a) displays the posterior density for a constant standard deviation of the prior of 50, while (b) for a standard deviation equal to half the prior mean. It can be noted that the shape of the posterior is similar across specifications and that it displays a significant shift of mass toward zero. Table C15 provides the quantiles, the mode and the mean of the posterior distribution of the incidence. The median posterior incidence ranges between -22.87 and -2.17, well below the median implied by the model calibration which has informed the prior.

[14]This assumes that $\tilde{\eta}_S = 0$, i.e., that the supply of housing is fixed in the short term.

The last column reports the ratio between the two, giving the implied attenuation bias displayed by agents. Given the model parameters the price reaction to council taxes is between 4% and 37% of what the price reaction to the stamp duty tax would imply from agents who fully perceive the tax.

The results above become striking once coupled with the extent to which house buyers react to stamp duty taxes. When buyers are liquidity-constrained, their effective discount rates become large and, therefore, one might be tempted to attribute the previous evidence solely to extreme discounting of future cash flows. If we are willing to take this view, we would have to assume discount rates ranging between 23.4% and 231.9% in order to fit the posterior estimates of the council tax incidence. Moreover, it should be noted that every estimate of the council tax incidence is conditioned on an estimate of the stamp duty incidence, i.e., the discount rate is not a free parameter in the calibration. To put it differently, changing the discount rate to match a reasonable incidence for the council tax would lead to an incidence of the stamp duty tax that is inconsistent with current estimates in the literature. The fact that the incidence of the stamp duty is large but not extreme implies that the liquidity premium cannot be the only source of the low council tax incidence. Third, in our estimation we use relatively concentrated priors around the model-informed incidence; had we allowed the likelihood to dominate by assigning diffuse priors, we would have obtained much lower estimates compared to the conservative ones provided so far. One way to explain these findings is by hypothesising that, when buying their properties, agents discount tax payments that happen in the future disproportionately compared to those that occur concurrently with the purchase. It is difficult to argue that this might be due to uncertainty associated with council tax payments given that differences in council tax amounts across Boroughs are very smooth and predictable as shown in Figure C2. This leaves us with another plausible alternative explanation: agents fail to fully internalise the difference in council tax payments across Boroughs upon purchasing a property, either because this is much less salient compared to the stamp duty tax[15], or because they fail to appreciate the magnitude of its present value[16]. Notice also that the results so far suggest that there is somebody who does not take the council tax differentials into account in a fully-rational way, but this does not need to be the house buyer: our previous analysis goes through even if the buyer is fully aware of the tax and hopes to shift its incidence onto the subsequent buyer, or the renter in the case of buy-to-let property

[15]It is also possible that the tax is fully salient to agents but, due to mental accounting, they fail to integrate its present value into the house price they are willing to pay. Other explanations could be related to search costs and cognitive costs.

[16]For a property in band D worth, say, £300,000, the stamp duty tax in 2018 would amount to £9,000. If the buyer could choose whether to buy the property in the Borough of Camden or the Borough of Westminster, the difference in council tax would amount to about £778 in 2018 which, in present value using a discount rate of 4%, would be equal to £19,450, more than twice the value of the stamp duty tax.

transactions[17].

Motivated by these findings, we explore some policy implications in the following section.

3.4.2 Implications for Tax Policy

Given the results in the previous section, it seems reasonable to argue that agents fail to fully perceive deferred taxes. As a result, we propose a modified version of the model above that allows for non-fully salient taxes. We extend our analysis to properties that are potentially far from the border and, therefore, allow the elasticity of supply η_S to be non-zero. Recall that the incidence estimates coming from the border in Section 3.3 are an upper bound for the incidence in the middle of Boroughs. For simplicity, let us assume we are in an equilibrium where the leverage constraint (3.14) is binding, i.e., $d_t = \alpha h p_t$. If we multiply each of the constraints (3.12) and (3.13) by $\frac{1}{(1+r+k)^t}$ and add them together, we obtain the following consolidated budget constraint:

$$c_0 + \frac{c_1}{(1+r+k)} + \frac{c_2}{(1+r+k)^2} + ... + \tilde{p}h = w_0 + \frac{w_1}{(1+r+k)} + ... = I \tag{3.30}$$

where $\tilde{p} = p_0\left(1 - \frac{\alpha k}{r+k-g}\right) + \tau_S + \frac{\tau_1}{r+k-\tilde{g}}$ is the tax-inclusive house price. For simplicity of exposition, define $p = p_0\left(1 - \frac{\alpha k}{r+k-g}\right)$ and $\tau = \frac{\tau_1}{r+k-\tilde{g}}$, so that we can rewrite $\tilde{p} = p + \tau_S + \tau$. Following Chetty et al. (2009), Farhi and Gabaix (2020) and Goldin (2015), we assume that the agent misperceives taxes with attenuation factor γ, i.e., he solves the following maximisation problem:

$$\max_{\{\{c_t\}_{t=0}^{\infty},h\}} U(\{c_t\}_{t=0}^{\infty},h) = c_0 + \log(h) + \sum_{t=1}^{\infty} \beta^t \left(u(c_t) + \log(h)\right) \tag{3.31}$$

s.t.

$$c_0 + \frac{c_1}{(1+r+k)} + \frac{c_2}{(1+r+k)^2} + ... + \tilde{p}_\gamma h = w_0 + \frac{w_1}{(1+r+k)} + ... = I \tag{3.32}$$

where the perceived house price is:

$$\tilde{p}_\gamma = p + \tau_S + \gamma\tau, \quad \gamma \in [0,1] \tag{3.33}$$

Recall from the previous section that the attenuation factor for the council tax implied by the data ranges between 0.04 and 0.37. Notice that while the agent perceives the above budget constraint, he has to satisfy the actual budget constraint (3.30) given by the rational model. As pointed out in Reck (2016), it is crucial to

[17] Note that we largely interpret the results as evidence of overpricing. Another possibility is that the properties on the low council tax side of borders are relatively underpriced and it is, therefore, sellers who fail to incorporate the tax discount into their ask price.

decide what choice variable bears the burden of adjustment. Given our assumption about the quasi-linear utility function in first-period consumption c_0, it is natural to let c_0 be the shock absorber. This choice amounts to assuming the following train of events: 1) the agent misperceives the council tax he will have to pay going forward and, as a result, buys "too much" quality-adjusted housing; 2) following this, he realises that the actual amount of taxes he will have to pay is beyond his budget; 3) consequently, the agent adjusts his consumption in the first period keeping everything else constant. Denoting the observed demands as $\hat{c_0}, \hat{c_t}, \hat{h}$, and the optimal demands absent any behavioural frictions as c_0^*, c_t^*, h^*, we have the following first-order conditions:

$$\hat{c_t} = [u']^{-1}\left(\frac{1}{(\beta(1+r+k))^t}\right) = c_t^* \tag{3.34}$$

$$\hat{h} = [(1-\beta)\tilde{p}_\gamma]^{-1} \neq [(1-\beta)\tilde{p}]^{-1} = h^* \tag{3.35}$$

$$\hat{c_0} = I - \sum_{t=1}^{\infty} \frac{\hat{c_t}}{(1+r+k)^t} - \hat{h}\tilde{p} \neq c_0^* \tag{3.36}$$

As previously mentioned, the optimality condition for future consumption remains as before. However, equation (3.35) shows that the agent demands "too much" housing due to the fact that the perceived price $\tilde{p}_\gamma$ is lower than the true price $\tilde{p}$, as long as $\gamma < 1$. As a result, because of quasi-linearity in the utility function, $\hat{c_0}$ adjusts to absorb the reduction in available income. The previous discussion highlights the fact that misperception of the house price affects both consumption and housing demand, albeit in opposite directions. This implies that a benevolent social planner needs to carefully balance the two distortions when setting the optimal tax policy. To see this more formally, let us adopt the approach of Goldin (2015) and assume that the Government chooses the optimal (property) tax combination in order to raise a fixed amount of revenue and maximise the utility of the buyer[18]. For convenience, define the present value of council tax revenue from the Government's point of view, discounted at the risk-free rate, as $\tilde{\tau} = \frac{\tau_1}{r-\tilde{g}}$. The total revenue raised from a given buyer is:

$$R = (\tau_S + \tilde{\tau})h = \left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)h \tag{3.37}$$

The second equality of the above equation shows that the Government discounts the revenue raised through council taxes at a lower rate than agents due to the presence of borrowing constraints. The Government can twick the two taxes to maintain

[18] In what follows, we abstract from analysing the effect on the utility of the seller.

revenue-neutrality. In particular, a revenue-neutral tax change is such that:

$$\left[h+\left(\tau_S+\tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial \tau_S}\right]\Delta\tau_S=-\left[\frac{r+k-\tilde{g}}{r-\tilde{g}}h+\left(\tau_S+\tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial \tau}\right]\Delta\tau \tag{3.38}$$

This implies that the change in stamp duty per unit change in council tax needed to maintain revenue-neutrality is:

$$\frac{\Delta\tau_S}{\Delta\tau}=-\frac{\frac{r+k-\tilde{g}}{r-\tilde{g}}h+\left(\tau_S+\tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial \tau}}{h+\left(\tau_S+\tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial \tau_S}}=-\frac{\frac{r+k-\tilde{g}}{r-\tilde{g}}h+\left(\tau_S+\tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\theta_\tau\frac{\partial h}{\partial p}}{h+\left(\tau_S+\tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\theta_{\tau_S}\frac{\partial h}{\partial p}} \tag{3.39}$$

where $\theta_{\tau_S}=\frac{\frac{\partial h}{\partial \tau_S}}{\frac{\partial h}{\partial p}}$ and $\theta_\tau=\frac{\frac{\partial h}{\partial \tau}}{\frac{\partial h}{\partial p}}$ tell us how responsive the demand is with respect to taxes relative to pre-tax prices. From equations (3.33) and (3.35) we infer that $\theta_{\tau_S}=1$ and $\theta_\tau=\gamma$ in our model. The indirect utility function for an inattentive agent is:

$$V(p,\tau_S,\tau)=I-\sum_{t=1}^{\infty}\frac{\hat{c}_t}{(1+r+k)^t}-\hat{h}(p+\tau_S+\tau)+\sum_{t=1}^{\infty}\beta^t u(\hat{c}_t)+\frac{\log(\hat{h})}{(1-\beta)} \tag{3.40}$$

where $\hat{c}_t=[u']^{-1}\left(\frac{1}{(\beta(1+r+k))^t}\right)$ and $\hat{h}=\hat{h}(p,\tau_S,\tau)=[(1-\beta)(p+\tau_S+\gamma\tau)]^{-1}$ from the agent's first-order conditions. Differentiate the indirect utility function above to obtain:

$$\frac{dV}{d\tau}=-\hat{h}\left(\frac{dp}{d\tau}+\frac{\partial\tau_S}{\partial\tau}+1\right)+\left[\frac{\partial U}{\partial h}-(p+\tau_S+\tau)\right]\left[\frac{dp}{d\tau}+\theta_{\tau_S}\frac{\partial\tau_S}{\partial\tau}+\theta_\tau\right]\frac{\partial\hat{h}}{\partial p} \tag{3.41}$$

where $\frac{dp}{d\tau}=\frac{\partial p}{\partial\tau}+\frac{\partial p}{\partial\tau_S}\frac{\partial\tau_S}{\partial\tau}$ is the total incidence of the council tax after having taken into account the shift in stamp duty to guarantee revenue neutrality. As in Goldin (2015), the change in welfare can be decomposed into four components: the first part, i.e., $-\hat{h}\left(\frac{dp}{d\tau}+\frac{\partial\tau_S}{\partial\tau}+1\right)$ measures the direct welfare effect of a tax shift due to the alleviation of the borrowing constraint; the second part, i.e., $\left[\frac{\partial U}{\partial h}-(p+\tau_S+\tau)\right]$ is the behavioural wedge and it represents the difference between perceived and actual prices; the third component, i.e., $\left[\frac{dp}{d\tau}+\theta_{\tau_S}\frac{\partial\tau_S}{\partial\tau}+\theta_\tau\right]$ is equal to the change in prices as perceived by the agent; and the fourth component, i.e., $\frac{\partial\hat{h}}{\partial p}$ is the impact of a change in prices on demand for housing. With no bias, i.e., when $\gamma=1$ the perceived price is equal to the actual price and the envelope theorem ensures that the second component above is equal to zero. As a consequence, the optimal tax policy depends on the sign of the first term[19]. If this is positive, it is optimal for the government to set $\tau_S=0$, if negative, $\tau_S=R$. It is easy to show that when $\gamma=1$

[19] Notice that $\frac{\partial\tau_S}{\partial\tau}<-1$ because $r+k-\tilde{g}>r-\tilde{g}$, $\theta_\tau<\theta_{\tau_S}$ and $\frac{\partial h}{\partial p}<0$. The above assumes that $\frac{r+k-\tilde{g}}{r-\tilde{g}}h+\left(\tau_S+\tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial\tau}>0$ and $h+\left(\tau_S+\tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right)\frac{\partial h}{\partial\tau_S}>0$, i.e., the Government is on the upward sloping part of the Laffer curve. The term $\frac{\partial p}{\partial\tau}+\frac{\partial p}{\partial\tau_S}\frac{\partial\tau_S}{\partial\tau}$ is usually positive since agents react less to a decrease in council tax relative to a revenue-neutral increase in the stamp duty.

this term is unambiguously positive as long as $\eta_S > 0$. The Government should then choose a zero stamp duty tax in order to alleviate the agent's liquidity constraint. In the presence of biases, however, there is a trade-off between the two inefficiencies: 1) the liquidity constraint and differences in salience make increasing the stamp duty tax less efficient than raising the council tax; 2) on the other hand, raising the council tax causes a shift in demand away from c_0 which in our example is the shock absorber. In the extreme case when there are no liquidity constraints, it is optimal to impose no council tax. Otherwise, the problem of the social planner amounts to choosing the optimal combination of stamp duty and council tax to jointly solve the following two equations:

$$\hat{h}\left(\tau_S + \tau\frac{r+k-\tilde{g}}{r-\tilde{g}}\right) = R \tag{3.42}$$

$$\frac{dV}{d\tau} = 0 \tag{3.43}$$

Figure C16 reports the optimal mix of taxes computed for a house worth £430,000 which is the median value of properties in band D in 2017. The property pays a stamp duty of £11,500 and we assume that it pays a yearly council tax of £1,419.73, the in-sample median amount in the corresponding band and year. The upper panel shows how the optimal combination varies as a function of the discount rate $r + k$, while the bottom panel varies the attenuation parameter γ. The figures confirm the above intuition. From Figure C16a we can see that when the liquidity premium is zero, the optimal policy is to levy only the stamp duty tax. For a small liquidity premium there is an optimal mix that includes positive amounts of both taxes, however, the borrowing constraints become dominant fairly quickly and make it optimal to set a stamp duty of zero. Figure C16b, on the other hand, focuses on the effect of salience. Even when the council tax is entirely non-salient, i.e., $\gamma = 0$, it is still optimal to raise a little over 20% of revenue through it. As the tax becomes more salient, its distortionary effect on c_0 decreases, therefore, its proportion should increase, up to the point where it becomes the only form of taxation for γ greater than 0.25. It should be noted, however, that this assumes that tax policy changes do not affect any of the parameters. In practice, changing the tax mix can change the inattention parameter γ.

3.5 Conclusions

This paper studies the incidence of property taxes in the UK housing market. By using a geographical discontinuity approach, exploiting the considerable difference in council tax rates across London Boroughs, we show that agents significantly underreact to council taxes. Our empirical estimates of council tax incidence on house prices is never significantly negative and this lack of significance cannot be attributed to lack of power. This is in sharp contrast to the large stamp duty incidence estimated by Best and Kleven (2018) and suggests that agents do not pay sufficient attention to taxes deferred to the future, or possibly points to evidence of very large search frictions or other cognitive costs. In Section 3.4.2, we touch upon the policy implications of our findings, however, one should be aware of issues arising when manipulating tax rates given that there is no guarantee that changes in policies are not followed by changes in tax salience and therefore behaviour. The analysis in this paper relies on data from the residential property market, however, it can also be extended to other domains of tax policy. One general take-away from the present work is that transaction taxes, such as the stamp duty tax, have a large incidence on transaction prices while deferred taxes, such as the council tax, have a lower effect on prices but potentially higher impact on consumption choices. This implies that the optimal mix of taxes may be some combination of the two. The analysis can be extended, for instance, to financial securities where the fact that a transaction tax might be very distortionary does not imply that it is optimal to raise revenues only through capital gains[20] or dividend taxes.

The findings in the paper keep open the question of the nature of the channels through which inattentive households correct their mistakes and adjust their consumption policies, once neglected taxes materialise. Access to disaggregated expenditure data could help shed light on this matter: this can be done by analysing differences in consumption responses at the border between Boroughs, which we should expect to arise whenever agents fail to optimally account for tax differences and are forced to adjust their expenditures ex-post to meet their budget constraints.

[20]While the capital gains tax is a transaction tax, the fact that it is borne by the seller of the asset suggests that agents could still underreact to it as it is a deferred tax and, therefore less salient compared to a tax charged at the moment of purchase like the stamp duty tax.

Bibliography

Adam, Klaus, Dmitry Matveev, and Stefan Nagel, "Do Survey Expectations of Stock Returns Reflect Risk Adjustments?," *Journal of Monetary Economics*, 2021, *117*, 723–740.

Agarwal, Sumit and Artashes Karapetyan, "Salience and Mispricing: Homebuyers' Housing Decisions," *Norges Bank Working Paper 21*, 2016.

Alevy, Jonathan E., Michael S. Haigh, and John A. List, "Information cascades: Evidence from a Field Experiment with Financial Market Professionals," *The Journal of Finance*, 2007, *62* (1), 151–180.

Allcott, Hunt, "Consumers' Perceptions and Misperceptions of Energy Costs," *American Economic Review*, May 2011, *101* (3), 98–104.

Andersen, Steffen, Cristian Badarinza, Lu Liu, Julie Marx, and Tarun Ramadorai, "Reference Dependence in the Housing Market," CEPR Discussion Papers 14147, C.E.P.R. Discussion Papers 2019.

Anderson, Evan W., Lars Peter Hansen, and Thomas J. Sargent, "A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection," *Journal of the European Economic Association*, 2003, *1* (1), 68–123.

Andonov, Aleksandar and Joshua D. Rauh, "The Return Expectations of Institutional Investors," *Working Paper*, 2020.

Anenberg, Elliot, "Information Frictions and Housing Market Dynamics," *International Economic Review*, 2016, *57* (4), 1449–1479.

Arrow, Kenneth J., "The Theory of Risk Aversion," in "Aspects of The Theory of Risk Bearing" 1965.

Asch, Solomon E., "Forming Impressions of Personality.," *The Journal of Abnormal and Social Psychology*, 1946, *41* (3), 258.

Back, Kerry E., *Asset Pricing and Portfolio Choice Theory*, Oxford University Press, 2017.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel, "The Economic Effects of Social Networks: Evidence from the Housing Market," *Journal of Political Economy*, 2018, *126* (6), 2224–2276.

Baker, Malcolm and Richard Ruback, "Estimating Industry Multiples," *Harvard University*, 1999.

Banerjee, Abhijit V., "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 1992, *107* (3), 797–817.

Barber, Brad M., Xing Huang, and Terrance Odean, "Which Factors Matter to Investors? Evidence from Mutual Fund Flows," *The Review of Financial Studies*, 2016, *29* (10), 2600–2642.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, "X-CAPM: An Extrapolative Capital Asset Pricing Model," *Journal of Financial Economics*, 2015, *115* (1), 1–24.

Barclay, Michael J., Neil D. Pearson, and Michael S. Weisbach, "Open-end Mutual Funds and Capital-gains Taxes," *Journal of Financial Economics*, 1998, *49* (1), 3 – 43.

Benos, Evangelos, Marek Jochec, and Victor Nyekel, "Can Mutual Funds Time Risk Factors?," *The Quarterly Review of Economics and Finance*, 2010, *50* (4), 509 – 514.

Bergstresser, Daniel and James Poterba, "Do After-tax Returns Affect Mutual Fund Inflows?," *Journal of Financial Economics*, 2002, *63* (3), 381 – 414.

Berk, Jonathan B. and Jules H. van Binsbergen, "Assessing Asset Pricing Models Using Revealed Preference," *Journal of Financial Economics*, 2016, *119* (1), 1–23.

Besley, Timothy, Neil Meads, and Paolo Surico, "The Incidence of Transaction Taxes: Evidence from a Stamp Duty Holiday," *Journal of Public Economics*, 2014, *119* (C), 61–70.

Best, Michael and Henrik Jacobsen Kleven, "Housing Market Responses to Transaction Taxes: Evidence From Notches and Stimulus in the U.K," *Review of Economic Studies*, 2018, *85* (1), 157–193.

Bhojraj, Sanjeev and Charles MC Lee, "Who Is My Peer? A Valuation-based Approach to the Selection of Comparable Firms," *Journal of Accounting Research*, 2002, *40* (2), 407–439.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change in Informational Cascades," *Journal of Political Economy*, 1992, *100* (5), 992–1026.

Black, Fisher and Robert Litterman, "Global Portfolio Optimization," *Financial Analysts Journal*, 1992, *48* (05), 28–43.

Bracke, Philippe and Silvana Tenreyro, "History Dependence in the Housing Market," 2020.

Brown, Jeffrey R., Zoran Ivković, Paul A. Smith, and Scott Weisbenner, "Neighbors Matter: Causal Community Effects and Stock Market Participation," *The Journal of Finance*, 2008, *63* (3), 1509–1531.

Brunnermeier, Markus K. and Christian Julliard, "Money Illusion and Housing Frenzies," *The Review of Financial Studies*, 2008, *21* (1), 135–180.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo, "Understanding Booms and Busts in Housing Markets," *Journal of Political Economy*, 2016, *124* (4), 1088–1147.

Camerer, Colin and Teck Hua Ho, "Experience-weighted Attraction Learning in Normal Form Games," *Econometrica*, 1999, *67* (4), 827–874.

Carrillo, Paul E., "An Empirical Stationary Equilibrium Search Model of the Housing Market," *International Economic Review*, 2012, *53* (1), 203–234.

Chen, Yongmin and Robert W. Rosenthal, "Asking Prices as Commitment Devices," *International Economic Review*, 1996, pp. 129–155.

Chernenko, Sergey, Samuel G. Hanson, and Adi Sunderam, "Who Neglects Risk? Investor Experience and the Credit Boom," *Journal of Financial Economics*, 2016, *122* (2), 248–269.

Chetty, Raj, "A New Method of Estimating Risk Aversion," *American Economic Review*, 2006, *96* (5), 1821–1834.

_ , Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *American Economic Review*, September 2009, *99* (4), 1145–1177.

Chiang, Yao-Min, David A. Hirshleifer, Yiming Qian, and Ann E. Sherman, "Do Investors Learn from Experience? Evidence from Frequent IPO Investors," *The Review of Financial Studies*, 2011, *24* (5), 1560–1589.

Cho, Man, "House Price Dynamics: A Survey of Theoretical and Empirical Issues," *Journal of Housing Research*, 1996, pp. 145–172.

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick, "Reinforcement Learning and Savings Behavior," *The Journal of Finance*, 2009, *64* (6), 2515–2534.

Cochrane, John H., "Macro-Finance," *Review of Finance*, 2017, *21* (3), 945–985.

Cohen, Randolph B., Christopher Polk, and Bernhard Silli, "Best Ideas," LSE Research Online Documents on Economics 24471, London School of Economics and Political Science, LSE Library 2008.

Cvitanic, Jaksa and Ioannis Karatzas, "Convex Duality in Constrained Portfolio Optimization," *The Annals of Applied Probability*, 1992, *2* (4), 767–818.

Dasgupta, Amil, Andrea Prat, and Michela Verardo, "The Price Impact of Institutional Herding," *The Review of Financial Studies*, 2011, *24* (3), 892–925.

Davidson, Russell and James G. MacKinnon, "Artificial Regressions," *A Companion to Theoretical Econometrics*, 2001, pp. 16–37.

Davis, Morris and Erwan Quintin, "On the Nature of Self-Assessed House Prices," *Real Estate Economics*, 2017, *45* (3), 628–649.

DeAngelo, Linda Elizabeth, "Equity Valuation and Corporate Control," *Accounting Review*, 1990, pp. 93–112.

Deese, James and Roger A. Kaufman, "Serial Effects in Recall of Unorganized and Sequentially Organized Verbal Material.," *Journal of Experimental Psychology*, 1957, *54* (3), 180.

DeMarzo, Peter M., Dimitri Vayanos, and Jeffrey Zwiebel, "Persuasion Bias, Social Influence, and Unidimensional Opinions," *The Quarterly Journal of Economics*, 2003, *118* (3), 909–968.

Ebbinghaus, Hermann, "Memory: A Contribution to Experimental Psychology," 1913, p. 142.

Erev, Ido and Alvin E. Roth, "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, 1998, pp. 848–881.

European Commission, "Taxation Trends in the European Union," *DG Taxation and Customs Union*, 2018.

Evans, Richard B., "Mutual Fund Incubation," *The Journal of Finance*, 2010, *65* (4), 1581–1611.

Eyster, Erik and Matthew Rabin, "Naïve Herding in Rich-Information Settings," *American Economic Journal: Microeconomics*, 2010, *2* (4), 221–243.

_ and _ , "Extensive Imitation is Irrational and Harmful," *The Quarterly Journal of Economics*, 2014, *129* (4), 1861–1898.

Fama, Eugene F. and Kenneth R. French, "A Five-factor Asset Pricing Model," *Journal of Financial Economics*, 2015, *116* (1), 1–22.

Farhi, Emmanuel and Xavier Gabaix, "Optimal Taxation with Behavioral Agents," *American Economic Review*, 2020, *110* (1), 298–336.

Fazio, Russell, Mark Zanna, and Joel Cooper, "Direct Experience and Attitude-Behavior Consistency: An Information Processing Analysis," *Personality and Social Psychology Bulletin*, 1978, *4*, 48–51.

Feldman, Naomi E. and Bradley J. Ruffle, "The Impact of Including, Adding, and Subtracting a Tax on Demand," *American Economic Journal: Economic Policy*, February 2015, *7* (1), 95–118.

__ , Jacob Goldin, and Tatiana Homonoff, "Raising the Stakes: Experimental Evidence on the Endogeneity of Taxpayer Mistakes," *National Tax Journal*, June 2018, *71* (2), 201–230.

Finkelstein, Amy, "E-ztax: Tax Salience and Tax Rates," *The Quarterly Journal of Economics*, 2009, *124* (3), 969–1010.

Fracassi, Cesare, "Corporate Finance Policies and Social Networks," *Management Science*, 2017, *63* (8), 2420–2438.

Fuster, Andreas, David Laibson, and Brock Mendel, "Natural Expectations and Macroeconomic Fluctuations," *Journal of Economic Perspectives*, 2010, *24* (4), 67–84.

Genesove, David and Christopher Mayer, "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *The Quarterly Journal of Economics*, 2001, *116* (4), 1233–1260.

Giacoletti, Marco and Christopher A. Parsons, "Peak-Bust Rental Spreads," *Available at SSRN 3482198*, 2019.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, "Five Facts About Beliefs and Portfolios," *American Economic Review*, 2021, *111* (5), 1481–1522.

Glaeser, Edward L. and Charles G. Nathanson, "Housing Bubbles," in "Handbook of Regional and Urban Economics," Vol. 5, Elsevier, 2015, pp. 701–751.

__ and __ , "An Extrapolative Model of House Price Dynamics," *Journal of Financial Economics*, 2017, *126* (1), 147–170.

Goldin, Jacob, "Optimal Tax Salience," *Journal of Public Economics*, 2015, *131* (C), 115–123.

Greenwood, Robin and Andrei Shleifer, "Expectations of Returns and Expected Returns," *The Review of Financial Studies*, 2014, *27* (3), 714–746.

__ and Stefan Nagel, "Inexperienced Investors and Bubbles," *Journal of Financial Economics*, 2009, *93* (2), 239–258.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, "Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior," *American Economic Review*, 1995, pp. 1088–1105.

Guren, Adam M., "House Price Momentum and Strategic Complementarity," *Journal of Political Economy*, 2018, *126* (3), 1172–1218.

Han, Lu and William C. Strange, "The Microstructure of Housing Markets," in Gilles Duranton, J. V. Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, Vol. 5 of *Handbook of Regional and Urban Economics*, Elsevier, 2015, chapter 0, pp. 813–886.

_ and _ , "What Is the Role of the Asking Price for a House?," *Journal of Urban Economics*, 2016, *93*, 115–130.

Harrison, J. Michael and David M. Kreps, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 1979, *20* (3), 381–408.

Head, Allen, Huw Lloyd-Ellis, and Hongfei Sun, "Search, Liquidity, and the Dynamics of House Prices and Construction," *American Economic Review*, 2014, *104* (4), 1172–1210.

Hirshleifer, David A., Ben Lourie, Thomas Ruchti, and Phong Truong, "First Impressions and Analyst Forecast Bias," *Review of Finance*, 2021, *25* (2), 325–364.

Ho, Lee In, "On the Convergence of Informational Cascades," *Journal of Economic Theory*, 1993, *61* (2), 395–411.

Hoffmann, Arvid O. I., Zwetelina Iliewa, and Lena Jaroszek, "Wall Street Crosses Memory Lane: How Witnessed Returns Affect Professionals' Expected Returns," in "Paris December 2017 Finance Meeting EUROFIDAI-AFFI" 2017.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, "Social Interaction and Stock-market Participation," *The Journal of Finance*, 2004, *59* (1), 137–163.

_ , _ , and _ , "Thy Neighbor's Portfolio: Word-of-Mouth Effects in the Holdings and Trades of Money Managers," *The Journal of Finance*, 2005, *60* (6), 2801–2824.

Ivković, Zoran and Scott Weisbenner, "Information Diffusion Effects in Individual Investors' Common Stock Purchases: Covet Thy Neighbors' Investment Choices," *The Review of Financial Studies*, 2007, *20* (4), 1327–1357.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, "Unobserved Actions of Mutual Funds," *The Review of Financial Studies*, 2006, *21* (6), 2379–2416.

Kaplan, Steven N. and Richard S. Ruback, "The Valuation of Cash Flow Forecasts: An Empirical Analysis," *The Journal of Finance*, 1995, *50* (4), 1059–1093.

Kaustia, Markku and Samuli Knüpfer, "Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions," *The Journal of Finance*, 2008, *63* (6), 2679–2702.

Kim, Moonchul and Jay Ritter, "Valuing IPOs," *Journal of Financial Economics*, 1999, *53* (3), 409–437.

Knüpfer, Samuli, Elias Rantapuska, and Matti Sarvimäki, "Formative Experiences and Portfolio Choice: Evidence from the Finnish Great Depression," *The Journal of Finance*, 2017, *72* (1), 133–166.

Kocherlakota, Narayana, "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, 1996, *34* (1), 42–71.

Koijen, Ralph S.J. and Motohiro Yogo, "A Demand System Approach to Asset Pricing," *Journal of Political Economy*, 2019, *127* (4), 1475–1515.

Kotlikoff, Laurence and Lawrence Summers, "Tax Incidence," in A. J. Auerbach and M. Feldstein, eds., *Handbook of Public Economics*, 1 ed., Vol. 2, Elsevier, 1987, chapter 16, pp. 1043–1092.

Kuchler, Theresa and Basit Zafar, "Personal Experiences and Expectations about Aggregate Outcomes," *The Journal of Finance*, 2019, *74* (5), 2491–2542.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, "The Impact of Institutional Trading on Stock Prices," *Journal of Financial Economics*, 1992, *32* (1), 23–43.

Ledoit, Olivier and Michael Wolf, "Honey, I Shrunk the Sample Covariance Matrix," *The Journal of Portfolio Management*, 2004, *30*, 110–119.

Liu, Jing, Doron Nissim, and Jacob Thomas, "Equity Valuation Using Multiples," *Journal of Accounting Research*, 2002, *40* (1), 135–172.

Lochstoer, Lars A. and Tyler Muir, "Volatility Expectations and Returns," *The Journal of Finance*, forthcoming.

Maenhout, Pascal J., "Robust Portfolio Rules and Asset Pricing," *The Review of Financial Studies*, 2004, *17* (4), 951–983.

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *The Quarterly Journal of Economics*, 2011, *126* (1), 373–416.

_ and _ , "Learning from Inflation Experiences," *The Quarterly Journal of Economics*, 2016, *131* (1), 53–87.

_ , _ , and Zhen Yan, "The Making of Hawks and Doves," *Journal of Monetary Economics*, 2021, *117*, 19–42.

Martin, Ian W. R., "What Is the Expected Return on the Market?," *The Quarterly Journal of Economics*, 2017, *132* (1), 367–433.

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

Mehra, Rajnish and Edward C. Prescott, "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 1985, *15* (2), 145–161.

Merlo, Antonio and François Ortalo-Magné, "Bargaining Over Residential Real Estate: Evidence from England," *Journal of Urban Economics*, 2004, *56* (2), 192–216.

_ , _ , and John Rust, "The Home Selling Problem: Theory and Evidence," *International Economic Review*, 2015, *56* (2), 457–484.

Merton, Robert C., "On Estimating the Expected Return on the Market : An Exploratory Investigation," *Journal of Financial Economics*, 1980, *8* (4), 323–361.

Moreira, Alan and Tyler Muir, "Should Long-Term Investors Time Volatility?," *Journal of Financial Economics*, 2019, *131* (3), 507 – 527.

Murdock, Bennet B. Jr., "The Serial Position Effect of Free Recall.," *Journal of Experimental Psychology*, 1962, *64* (5), 482.

_ , *Human Memory: Theory and Data.*, Lawrence Erlbaum, 1974.

Murfin, Justin and Ryan Pratt, "Comparables Pricing," *The Review of Financial Studies*, 2019, *32* (2), 688–737.

Ngai, L. Rachel and Silvana Tenreyro, "Hot and Cold Seasons in the Housing Market," *American Economic Review*, 2014, *104* (12), 3991–4026.

Nipher, Francis E., "On the Distribution of Errors in Numbers Written from Memory," *Transactions of the Academy of Science of St. Louis*, 1878, *3*, 10–1.

Piazzesi, Monika and Martin Schneider, "Momentum Traders in the Housing Market: Survey Evidence and a Search Model," *American Economic Review*, 2009, *99* (2), 406–11.

Pratt, John W., "Risk Aversion in the Small and in the Large," *Econometrica*, 1964, *32*, 122–136.

Purnanandam, Amiyatosh K. and Bhaskaran Swaminathan, "Are IPOs Really Underpriced?," *The Review of Financial Studies*, 2004, *17* (3), 811–848.

Reck, Daniel, "Taxes and Mistakes: What's in a Sufficient Statistic?," *Working Paper*, April 2016.

Roth, Alvin E. and Ido Erev, "Learning in Extensive-form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, 1995, *8* (1), 164–212.

Sharpe, William F., "Imputing Expected Security Returns from Portfolio Composition," *Journal of Financial and Quantitative Analysis*, 1974, *9* (03), 463–472.

Shumway, Tyler, Maciej J. Szefler, and Kathy Yuan, "The Information Content of Revealed Beliefs in Portfolio Holdings," LSE Research Online Documents on Economics, London School of Economics and Political Science, LSE Library 2011.

Sialm, Clemens and Hanjiang Zhang, "Tax-Efficient Asset Management: Evidence from Equity Mutual Funds," *The Journal of Finance*, 2020, *75* (2), 735–777.

Simonsohn, Uri, Niklas Karlsson, George Loewenstein, and Dan Ariely, "The Tree of Experience in the Forest of Information: Overweighing Experienced Relative to Observed Information," *Games and Economic Behavior*, 2008, *62* (1), 263–286.

Smith, Lones and Peter Sorensen, "Pathological Outcomes of Observational Learning," *Econometrica*, 2000, *68* (2), 371–398.

Standard & Poor's Compustat Services, I., "Compustat," 1962.

Taubinsky, Dmitry and Alex Rees-Jones, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment," *Review of Economic Studies*, 2018, *85* (4), 2462–2496.

Tepla, Lucie, "Optimal Portfolio Policies with Borrowing and Shortsale Constraints," *Journal of Economic Dynamics and Control*, 2000, *24* (11-12), 1623–1639.

Thorndike, Edward L., "Animal Intelligence: An Experimental Study of the Associative Processes in Animals.," *The Psychological Review: Monograph Supplements*, 1898, *2* (4), i.

Touloumis, Anestis, "Nonparametric Stein-type Shrinkage Covariance Matrix Estimators in High-dimensional Settings," *Computational Statistics & Data Analysis*, 2015, *83* (C), 251–261.

Tversky, Amos and Daniel Kahneman, "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology*, 1973, *5* (2), 207–232.

University of Chicago. Center for Research in Security Prices, I., "CRSP databases," 1960.

van Binsbergen, Jules H., Michael W. Brandt, and Ralph S.J. Koijen, "Optimal Decentralized Investment Management," *The Journal of Finance*, 2008, *63* (4), 1849–1895.

Wang, Guocheng and Yanyi Wang, "Herding, Social Network and Volatility," *Economic Modelling*, 2018, *68*, 74–81.

Xu, Gan-Lin and Steven E. Shreve, "A Duality Method for Optimal Consumption and Investment Under Short- Selling Prohibition. I. General Market Coefficients," *The Annals of Applied Probability*, 1992, *2* (1), 87–112.

Yavas, Abdullah and Shiawee Yang, "The Strategic Role of Listing Price in Marketing Real Estate: Theory and Evidence," *Real Estate Economics*, 1995, *23* (3), 347–368.

A. Appendix to Learning from Past Prices: Evidence from the UK Housing Market

A.1 Tables

Table A1: Listings and Comparable Transactions - Summary Statistics

The table presents summary statistics for the set of listings and recent comparable transactions that have at least one match. Summary statistics are presented separately for the sample of data before March 2012 and post March 2012. Nb. of Observations refers to the total number of unique listings and transactions, respectively. Listing price is the first quote at which a property has been listed, while transaction price is the final agreed price between the buyer and the seller. Property type refers to the built-form of the property which can be one of four possible categories: detached, semi-detached, terraced house or a flat. Number of rooms refers to the total number of habitable rooms in the property. I report the following statistics on the distribution of prices and number of rooms: Mean is the average value, Min is the lowest value, P25, Median and P75 are the 25-th, 50-th and 75-th percentile of the distributions, respectively, and Max is the highest value observed in the sample. For property type, I report the fraction of observations that are of a given type.

	Listings		Transactions	
	Pre March 2012	Post March 2012	Pre March 2012	Post March 2012
Nb. of Observations	1,007,942	1,983,528	986,287	2,521,505
Listing/Transaction Price				
Mean	£233,497	£268,402	£220,134	£256,734
Min	£10,500	£10,500	£10,300	£10,018
P25	£125,000	£129,995	£119,995	£125,000
Median	£178,500	£194,950	£170,000	£189,995
P75	£265,000	£310,000	£250,000	£300,000
Max	£17,500,000	£25,000,000	£19,250,000	£18,500,000
Property type (%)				
Detached	16.00	15.23	17.34	18.67
Semi-detached	28.11	28.84	27.84	29.12
Terraced	31.66	34.15	31.27	31.72
Flat	24.23	21.77	23.55	20.49
Number of rooms				
Mean	4.47	4.53	4.49	4.59
Min	1	1	1	1
P25	3	3	3	4
Median	4	4	4	5
P75	5	5	5	5
Max	18	19	18	19

Table A2: Listings Matched to Respective Transactions - Summary Statistics

The table presents summary statistics for the set of listings that have been matched to their respective ex-post transactions in the period from March 2012. Listing price is the first quote at which a property has been listed, while transaction price is the final agreed price between the buyer and the seller. Price discount is the percentage difference between the initial listed price and the final transaction price. TOM is time on the market measured in weeks. Property type refers to the built-form of the property which can be one of four possible categories: detached, semi-detached, terraced house or a flat. Number of rooms refers to the total number of habitable rooms in the property. I report the following statistics on the distribution of prices, price discounts, time on the market and number of rooms: Mean is the average value, Min is the lowest value, P25, Median and P75 are the 25-th, 50-th and 75-th percentile of the distributions, respectively, and Max is the highest value observed in the sample. For property type, I report the fraction of observations that are of a given type. Nb. of Observations refers to the total number of transactions in the matched sample.

Listing Price		*Transaction Price*	
Mean	£255,038	Mean	£245,569
Min	£12,000	Min	£11,000
P25	£134,950	P25	£127,000
Median	£190,000	Median	£186,500
P75	£299,950	P75	£290,000
Max	£15,000,000	Max	£16,200,000
Price Discount (%)		*TOM (weeks)*	
Mean	-3.78	Mean	26.48
Min	-63.52	Min	4.00
P25	-6.41	P25	14.43
Median	-3.13	Median	20.57
P75	0.00	P75	31.43
Max	1.16	Max	256.86
Number of rooms		*Property type (%)*	
Mean	4.62	Detached	15.14
Min	1	Semi-detached	32.16
P25	4	Terraced	35.28
Median	5	Flat	17.42
P75	5		
Max	16		
Nb. of Observations		1,067,282	

Table A3: Effect of Transaction Prices on Quotes Around Publishing Dates

The table presents the results of the following regression: $log(q_i) = \alpha + \beta \times log(p_j) + Controls + \varepsilon_i$, where q_i is the listed price for property i and p_j is the transaction price for a comparable property j sold in the previous month. Columns (1) and (2) present the results of running separate regressions for the set of untreated and treated listings, where treated listings are those that are able to directly observe the most recent price data. Columns (3)-(6) combine the two samples in a single regression of the following form: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where $Treated_i$ is a dummy that turns on when the listing price has been set in the period following the price data publication date. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (5). Column (5) instead includes time distance (measured in weeks) dummies and their interaction with log price. Fixed-effects included are: listing month-year dummies in columns (4)-(6), and; transaction ID dummies in column (6). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	Untreated	Treated	Full Sample			
	(1)	(2)	(3)	(4)	(5)	(6)
Price × Treated			0.0045***	0.0045***	0.0039**	0.0034***
			(0.0015)	(0.0015)	(0.0015)	(0.0011)
Price	0.8401***	0.8446***	0.8401***	0.8402***	0.8367***	
	(0.0026)	(0.0036)	(0.0023)	(0.0023)	(0.0030)	
Treated			-0.0550***	-0.0548***	-0.0459**	-0.0387***
			(0.0185)	(0.0185)	(0.0183)	(0.0134)
Controls						
Price x Time distance	Yes	Yes	Yes	Yes	No	Yes
Price x Time distance dummies	No	No	No	No	Yes	No
Fixed-Effects						
Month-year	No	No	No	Yes	Yes	Yes
Transaction ID	No	No	No	No	No	Yes
Observations	3,698,564	3,768,386	7,466,950	7,466,950	7,466,950	7,466,950
R^2	0.7028	0.7056	0.7043	0.7050	0.7050	0.8689
Within R^2	–	–	–	0.7032	0.7032	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A4: Effect of Transaction Prices on Quotes - Including Listing Price Updates

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set/updated in the period following the most recent price data publication date. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (5). Column (5) instead includes time distance (measured in weeks) dummies and their interaction with log price. Listing ID fixed effects are included in specifications (3)-(6). Additional fixed-effects include: listing month-year dummies in all columns but (3) and; transaction ID dummies in columns (2) and (6). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
Price × Treated	0.0043***	0.0030***	0.0008***	0.0008***	0.0008***	0.0025***
	(0.0012)	(0.0009)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Price	0.8418***		-0.0012***	−0.0001		
	(0.0020)		(0.0002)	(0.0001)		
Treated	-0.0521***	-0.0342***	-0.0135***	-0.0080***	-0.0080***	-0.0331***
	(0.0150)	(0.0103)	(0.0027)	(0.0024)	(0.0024)	(0.0023)
Controls						
Price × Time distance	Yes	Yes	Yes	Yes	No	Yes
Price × Time distance dummies	No	No	No	No	Yes	No
Fixed-Effects						
Listing ID	No	No	Yes	Yes	Yes	Yes
Month-year	Yes	Yes	No	Yes	Yes	Yes
Transaction ID	No	Yes	No	No	No	Yes
Observations	11,410,244	11,410,244	11,410,244	11,410,244	11,410,244	11,410,244
R^2	0.7080	0.8695	0.9985	0.9989	0.9989	0.9994
Within R^2	0.7053	0.0000	0.0011	0.0002	0.0002	0.0156

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A5: Effect of Transaction Prices on Quotes - Before vs After March 2012

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Post March 2012}_i + \beta_2 \times log(p_j) \times \textit{Treated}_i + \beta_3 \times log(p_j) \times \textit{Treated}_i \times \textit{Post March 2012}_i + \gamma_1 \textit{Post March 2012}_i + \gamma_2 \textit{Treated}_i + \gamma_3 \times \textit{Treated}_i \times \textit{Post March 2012}_i + \textit{Controls} + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month, *Post March 2012*$_i$ is a dummy that equals one for listings published starting from March 2012 and *Treated*$_i$ is a dummy that turns on when the listing price has been set/updated in the period following the most recent price data publication date. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (3). Column (3) instead includes time distance (measured in weeks) dummies and their interaction with log price. Fixed-effects included are: listing month-year dummies in all columns but (1) and; transaction ID dummies in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price × Treated × Post March 2012	0.0075***	0.0072***	0.0072***	0.0040**
	(0.0021)	(0.0021)	(0.0021)	(0.0018)
Price × Treated	-0.0041**	-0.0038*	-0.0043**	-0.0010
	(0.0020)	(0.0020)	(0.0020)	(0.0016)
Price × Post March 2012	0.0513***	0.0508***	0.0508***	
	(0.0020)	(0.0020)	(0.0020)	
Price	0.7889***	0.7893***	0.7875***	
	(0.0024)	(0.0024)	(0.0028)	
Treated	0.0459*	0.0417*	0.0492**	0.0102
	(0.0236)	(0.0235)	(0.0235)	(0.0190)
Post March 2012	-0.6188***	-0.6381***	-0.6382***	
	(0.0241)	(0.0247)	(0.0248)	
Treated × Post March 2012	-0.0878***	-0.0839***	-0.0838***	-0.0443**
	(0.0246)	(0.0245)	(0.0245)	(0.0216)
Controls				
Price × Time distance	Yes	Yes	No	Yes
Price × Time distance dummies	No	No	Yes	No
Fixed-Effects				
Month-year	No	Yes	Yes	Yes
Transaction ID	No	No	No	Yes
Observations	16,367,900	16,367,900	16,367,900	16,367,900
R^2	0.6805	0.6814	0.6814	0.8565
Within R^2	–	0.6782	0.6782	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A6: Effect of Transaction Prices on Quotes Around Placebo Publishing Dates

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set/updated in the week before (first four columns) or one week after (last four columns) the closest price data publication date. Only listing prices from the two weeks surrounding the placebo publishing dates are considered. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (3) and (7). Columns (3) and (7) instead include time distance (measured in weeks) dummies and their interaction with log price. Additional fixed-effects include: listing month-year dummies in all columns but (1) and (5) and; transaction ID dummies in columns (4) and (8). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	7 days before				7 days after			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Price × Treated	0.0014	0.0012	0.0006	0.0000	0.0016	0.0012	0.0007	-0.0007
	(0.0016)	(0.0016)	(0.0016)	(0.0014)	(0.0016)	(0.0016)	(0.0016)	(0.0013)
Price	0.8457***	0.8451***	0.8428***		0.8436***	0.8435***	0.8429***	
	(0.0026)	(0.0026)	(0.0028)		(0.0035)	(0.0035)	(0.0033)	
Treated	-0.0208	-0.0192	-0.0106	0.0008	-0.0222	-0.0164	-0.0100	0.0086
	(0.0194)	(0.0194)	(0.0194)	(0.0162)	(0.0191)	(0.0190)	(0.0191)	(0.0160)
Controls								
Price × Time distance	Yes	Yes	No	Yes	Yes	Yes	No	Yes
Price × Time distance dummies	No	No	Yes	No	No	No	Yes	No
Fixed-Effects								
Month-year	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Transaction ID	No	No	No	Yes	No	No	No	Yes
Observations	4,569,583	4,569,583	4,569,583	4,569,583	4,705,129	4,705,129	4,705,129	4,705,129
R^2	0.7061	0.7069	0.7069	0.8785	0.7104	0.7113	0.7113	0.8811
Within R^2	–	0.7038	0.7038	0.0000	–	0.7085	0.7085	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A7: Indirect Price Effects Through Intermediate Listings

The table presents the results of the following regression: $log(q_i^s) = \alpha^s + \beta^s \times log(p_j^s) + Controls + \varepsilon_i^s$, where q_i^s is the listed price for property i and p_j^s is the transaction price for a comparable property j sold in the previous month. The sample includes quotes that have been set/updated during the second week following the most recent price data publication date. Each column refers to a different subsample s of quote-price pairs: column (1) considers quotes with no comparable listings in the previous two weeks; column (2) quotes with comparable listings only in the week before the publication date; column (3) quotes with comparable listings only in the week after the publication date, and; column (4) quotes with comparable listings in both weeks. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns, as well as controls for the number of comparable quotes in the current and each of the two previous weeks. Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	No prior comps	Comps in wk -1	Comps in wk +1	Comps in all wks
	(1)	(2)	(3)	(4)
Price	0.8166***	0.8328***	0.8402***	0.8479***
	(0.0094)	(0.0103)	(0.0084)	(0.0054)
Controls				
Price x Time distance	Yes	Yes	Yes	Yes
Nb. of comps per week	Yes	Yes	Yes	Yes
Observations	384,735	332,801	486,689	1,728,033
R^2	0.6927	0.6947	0.7092	0.7110
Adjusted R^2	0.6927	0.6947	0.7092	0.7110

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A8: Indirect Price Effects by Number of Intermediate Comparable Listings

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \sum_k \beta_k \times log(p_j) \times k\ Comps\ in\ week\ n_i + \sum_k \gamma_k \times k\ Comps\ in\ week\ n_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and k *Comps in week* n_i is a dummy that turns on when quote i has k comparable listings in week n. The sample includes quotes that have been set/updated during the second week following the most recent price data publication date. Columns (1)-(3) investigate the indirect effect of past transactions on quotes based on the number of comparable listings in the week before the price data publication date, while columns (4)-(6) consider the indirect price effect based on the number of comparable listings in the first week after the publication date. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns, as well as controls for the number of comparable quotes in weeks +1, +2 and -1, +2 and their interaction with log price in columns (2)-(3) and (5)-(6), respectively. Columns (3) and (6) also include listing month-year fixed effects. Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	$n=-1$			$n=+1$		
	(1)	(2)	(3)	(4)	(5)	(6)
Price	0.8330***	0.8367***	0.8358***	0.8277***	0.8333***	0.8323***
	(0.0041)	(0.0043)	(0.0043)	(0.0041)	(0.0043)	(0.0043)
Price × 1 comp in week n (1)	0.0080***	0.0090***	0.0089***	0.0121***	0.0137***	0.0135***
	(0.0024)	(0.0024)	(0.0024)	(0.0025)	(0.0025)	(0.0025)
Price × 2 comps in week n	0.0137***	0.0157***	0.0157***	0.0153***	0.0187***	0.0186***
	(0.0029)	(0.0030)	(0.0030)	(0.0028)	(0.0029)	(0.0029)
Price × 3 comps in week n	0.0123***	0.0154***	0.0146***	0.0199***	0.0252***	0.0248***
	(0.0036)	(0.0038)	(0.0038)	(0.0034)	(0.0036)	(0.0036)
Price × > 3 comps in week n (2)	0.0132***	0.0199***	0.0194***	0.0233***	0.0359***	0.0350***
	(0.0031)	(0.0037)	(0.0037)	(0.0029)	(0.0035)	(0.0035)
1 comp in week n	-0.0990***	-0.1117***	-0.1122***	-0.1441***	-0.1636***	-0.1627***
	(0.0290)	(0.0294)	(0.0293)	(0.0297)	(0.0299)	(0.0299)
2 comps in week n	-0.1678***	-0.1930***	-0.1922***	-0.1835***	-0.2253***	-0.2256***
	(0.0350)	(0.0361)	(0.0360)	(0.0341)	(0.0350)	(0.0349)
3 comps in week n	-0.1460***	-0.1847***	-0.1764***	-0.2358***	-0.3009***	-0.2970***
	(0.0440)	(0.0459)	(0.0458)	(0.0414)	(0.0431)	(0.0429)
> 3 comps in week n	-0.1490***	-0.2354***	-0.2275***	-0.2695***	-0.4238***	-0.4131***
	(0.0380)	(0.0447)	(0.0446)	(0.0349)	(0.0418)	(0.0417)
(2)−(1)	0.0052	0.0109***	0.0105***	0.0112***	0.0222***	0.0215***
p-value	(0.1193)	(0.0029)	(0.0046)	(0.0002)	(0.0000)	(0.0000)
Controls						
Price x Time distance	Yes	Yes	Yes	Yes	Yes	Yes
Price x Nb. of comps per week	No	Yes	Yes	No	Yes	Yes
Fixed-Effects						
Month-year	No	No	Yes	No	No	Yes
Observations	2,932,258	2,932,258	2,932,258	2,932,258	2,932,258	2,932,258
R^2	0.7067	0.7068	0.7077	0.7068	0.7069	0.7078
Within R^2	–	–	0.7050	–	–	0.7051

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A9: Chain Effects of Prices on Quotes by Order of Match

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_1 \times log(p_j) + \sum_{k=2}^{10} \beta_k^{pre} \times log(p_j) \times Comp\ Order\ k\ Pre_i + \sum_{k=2}^{10} \gamma_k^{pre} Comp\ Order\ k\ Pre_i + \sum_{k=1}^{10} \beta_k^{post} \times log(p_j) \times Comp\ Order\ k\ Post_i + \sum_{k=1}^{10} \gamma_k^{post} Comp\ Order\ k\ Post_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and *Comp Order k Pre (Post)*$_i$ is a dummy that turns on when quote i is the k-th sequential match to transaction j in the period before (after) the price data publication date. The sample includes listings in the one-month period surrounding the publication date that have a comparable transaction which has at least one treated and one untreated match. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1). Column (3) also includes listing month-year fixed effects. Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)
Price	0.8388***	0.8404***	0.8401***
	(0.0011)	(0.0020)	(0.0020)
Price × 2nd Untreated	0.0010	0.0011	0.0011
	(0.0016)	(0.0016)	(0.0016)
Price × 3rd Untreated	0.0043**	0.0045**	0.0044**
	(0.0020)	(0.0020)	(0.0020)
Price × 4th Untreated	0.0054**	0.0057**	0.0056**
	(0.0023)	(0.0023)	(0.0023)
Price × 5th Untreated	0.0060**	0.0063**	0.0062**
	(0.0027)	(0.0027)	(0.0027)
Price × 6th Untreated	0.0039	0.0042	0.0041
	(0.0033)	(0.0033)	(0.0033)
Price × 7th Untreated	0.0068*	0.0071*	0.0070*
	(0.0039)	(0.0039)	(0.0039)
Price × 8th Untreated	0.0021	0.0024	0.0025
	(0.0044)	(0.0044)	(0.0044)
Price × 9th Untreated	0.0036	0.0040	0.0041
	(0.0055)	(0.0055)	(0.0055)
Price × 10th or more Untreated	0.0000	0.0004	0.0003
	(0.0038)	(0.0038)	(0.0038)
Price × 1st Treated	0.0015	0.0023	0.0025
	(0.0014)	(0.0017)	(0.0017)
Price × 2nd Treated	0.0039**	0.0048**	0.0050***
	(0.0016)	(0.0019)	(0.0019)
Price × 3rd Treated	0.0054***	0.0064***	0.0065***
	(0.0019)	(0.0022)	(0.0022)
Price × 4th Treated	0.0106***	0.0116***	0.0116***
	(0.0022)	(0.0025)	(0.0025)
Price × 5th Treated	0.0082***	0.0093***	0.0092***
	(0.0026)	(0.0029)	(0.0029)

Continued on next page

Table A9 – *Continued from previous page*

	(1)	(2)	(3)
Price × 6th Treated	0.0098***	0.0109***	0.0107***
	(0.0031)	(0.0033)	(0.0033)
Price × 7th Treated	0.0117***	0.0129***	0.0126***
	(0.0037)	(0.0039)	(0.0039)
Price × 8th Treated	0.0177***	0.0188***	0.0186***
	(0.0044)	(0.0046)	(0.0046)
Price × 9th Treated	0.0099*	0.0110**	0.0108**
	(0.0051)	(0.0052)	(0.0052)
Price × 10th or more Treated	0.0101***	0.0113***	0.0112***
	(0.0036)	(0.0039)	(0.0039)
Controls			
Price x Time distance	No	Yes	Yes
Fixed-Effects			
Month-year	No	No	Yes
Observations	11,292,009	11,292,009	11,292,009
R^2	0.7081	0.7081	0.7089
Within R^2	–	–	0.7063

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A10: Effect of Prices on Quote Updates

The table presents the results of the following regression: $log(q_i^n) = \alpha + \beta_1 \times log(p_j) + \sum_{n=2}^{5} \beta_n \times log(p_j) \times Update\ Number\ n_i + \sum_{n=2}^{5} \gamma_n Update\ Number\ n_i + Controls + \varepsilon_i^n$, where q_i^n is the n-th listed price update for property i, p_j is the transaction price for a comparable property j sold in the month before property i was initially listed and $Update\ Number\ n_i$ is a dummy that turns on when q_i^n is the n-th consecutive quote update for property i. The sample includes listings in the post March 2012 period that have at least one price change and a comparable transaction that has been published just before the listing was first posted. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1). Column (3) also includes listing month-year fixed effects and column (4) includes listing ID as well as month-year fixed effects. Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price	0.8320***	0.8345***	0.8310***	-0.0105***
	(0.0010)	(0.0012)	(0.0013)	(0.0002)
Price × 1st Change	0.0097***	0.0132***	0.0129***	0.0070***
	(0.0002)	(0.0009)	(0.0009)	(0.0002)
Price × 2nd Change	0.0160***	0.0224***	0.0217***	0.0131***
	(0.0014)	(0.0021)	(0.0021)	(0.0004)
Price × 3rd Change	0.0194***	0.0282***	0.0272***	0.0184***
	(0.0028)	(0.0036)	(0.0036)	(0.0007)
Price × ≥ 4th Change	0.0244***	0.0359***	0.0349***	0.0225***
	(0.0074)	(0.0080)	(0.0079)	(0.0014)
1st Price Change	-0.1560***	-0.2037***	-0.1990***	-0.1224***
	(0.0019)	(0.0105)	(0.0105)	(0.0025)
2nd Price Change	-0.2542***	-0.3407***	-0.3306***	-0.2299***
	(0.0164)	(0.0250)	(0.0250)	(0.0049)
3rd Price Change	-0.3099***	-0.4294***	-0.4146***	-0.3204***
	(0.0340)	(0.0429)	(0.0428)	(0.0091)
≥ 4th Price Change	-0.3793***	-0.5362***	-0.5226***	-0.3972***
	(0.0894)	(0.0959)	(0.0950)	(0.0171)
Controls				
Price x Time distance	No	Yes	Yes	Yes
Fixed-Effects				
Month-year	No	No	Yes	Yes
Listing ID	No	No	No	Yes
Observations	5,868,384	5,868,384	5,868,384	5,868,384
R^2	0.7160	0.7161	0.7170	0.9981
Within R^2	–	–	0.7101	0.2670

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A11: Relation between Effect of Past Prices and Future Price Discount

The table presents the results of the following regression: $log(q_i^k) = \alpha^k + \beta^k \times log(p_j) + Controls + \varepsilon_i^k$, where q_i^k is the first quoted price for listing i which is in quantile k of the price discount distribution and p_j is the transaction price for a similar property which has been published in the month before the listing was originally posted. The sample includes listings in the sample period starting from March 2012 that have been matched to their respective ex-post transactions. The first five columns present the coefficients on recent transaction prices per quantile of price discount: column (1) considers listings sold at the largest price discount while column (5) listings that were sold at a premium to quoted price. The final column includes all matched listings to give an idea of the average price effect. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price as well as listing month-year fixed effects are included in all regressions. Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)	(5)	Full Sample
Price	0.8671***	0.8198***	0.7851***	0.7921***	0.8737***	0.8254***
	(0.0053)	(0.0053)	(0.0054)	(0.0049)	(0.0073)	(0.0027)
Controls						
Price x Time distance	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-Effects						
Month-year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	956,899	1,000,010	985,789	1,411,376	813,810	5,167,884
R^2	0.7392	0.7209	0.7084	0.6936	0.7041	0.7153
Within R^2	0.7332	0.7183	0.7072	0.6926	0.6984	0.7144

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A12: Relation between Effect of Past Prices and Time on the Market

The table presents the results of the following regression: $log(q_i^k) = \alpha^k + \beta^k \times log(p_j) + Controls + \varepsilon_i^k$, where q_i^k is the first quoted price for listing i which is in quantile k of the time-on-the-market (TOM) distribution and p_j is the transaction price for a similar property which has been published in the month before the listing was originally posted. The sample includes listings in the sample period starting from March 2012 that have been matched to their respective ex-post transactions. Panel A considers properties that sold at a discount to listed price, while Panel B focuses on properties sold at a premium. The first five columns present the coefficients on recent transaction prices per TOM quantile: column (1) looks at listings that took the least time to sell, while column (5) listings that had the longest duration. The final column includes all matched listings to give an idea of the average price effect. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price as well as listing month-year fixed effects are included in all regressions. Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

Panel A: Properties Sold at Discount						
	(1)	(2)	(3)	(4)	(5)	Full Sample
Price	0.8216***	0.8120***	0.8139***	0.8245***	0.8306***	0.8211***
	(0.0065)	(0.0055)	(0.0063)	(0.0058)	(0.0062)	(0.0030)
Controls						
Price x Time distance	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-Effects						
Month-year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	736,353	905,252	680,331	752,650	638,851	3,713,437
R^2	0.7230	0.7232	0.7273	0.7276	0.7258	0.7242
Within R^2	0.7220	0.7222	0.7253	0.7243	0.7165	0.7223

Panel B: Properties Sold at Premium						
	(1)	(2)	(3)	(4)	(5)	Full Sample
Price	0.9042***	0.8113***	0.8152***	0.8157***	0.8303***	0.8431***
	(0.0115)	(0.0117)	(0.0109)	(0.0107)	(0.0115)	(0.0054)
Controls						
Price x Time distance	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-Effects						
Month-year	Yes	Yes	Yes	Yes	Yes	Yes
Observations	314,764	253,091	291,296	317,690	277,606	1,454,447
R^2	0.7080	0.7002	0.6939	0.6782	0.6814	0.6933
Within R^2	0.7026	0.6940	0.6887	0.6739	0.6781	0.6895

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

A.2 Figures

Figure A1: Timeline of Price Data Publication

The figure presents the natural experiment that generates shocks to the information set of sellers: beginning in March 2012, the Land Registry publishes regular monthly data on housing transactions on the twentieth working day of the subsequent month. For example, transaction prices, depicted in blue, from February are published on the twentieth working day of March which is 28th March in this case. The property listings published at the beginning of March and before the publication date, depicted in light green, do not observe the data on February transactions, while those published after this date, depicted in dark green, may observe February price data and therefore can use this to make inference about demand.

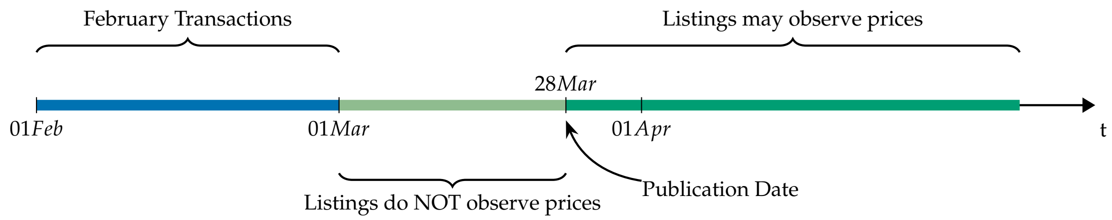
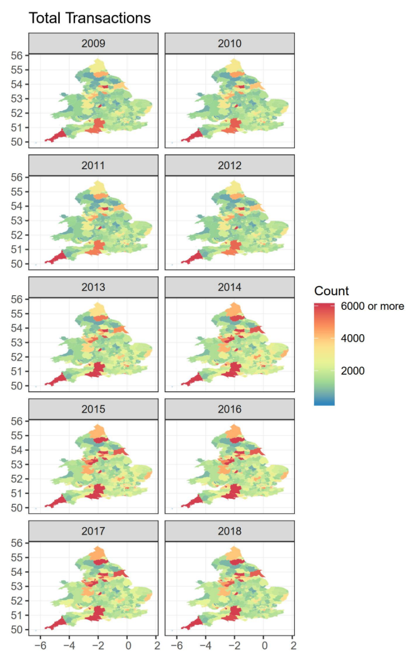
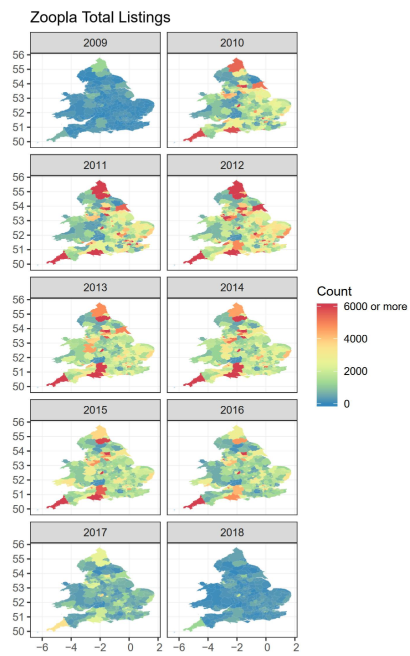


Figure A2: Geographic Coverage

The figure plots heat maps of the geographic coverage of the transaction and listing data between 2009-2018 by year across England and Wales, computed as the total number of observations by local authority district. Figure A2a displays the total number of transactions, while figure A2b the total number of unique listings in the sample.



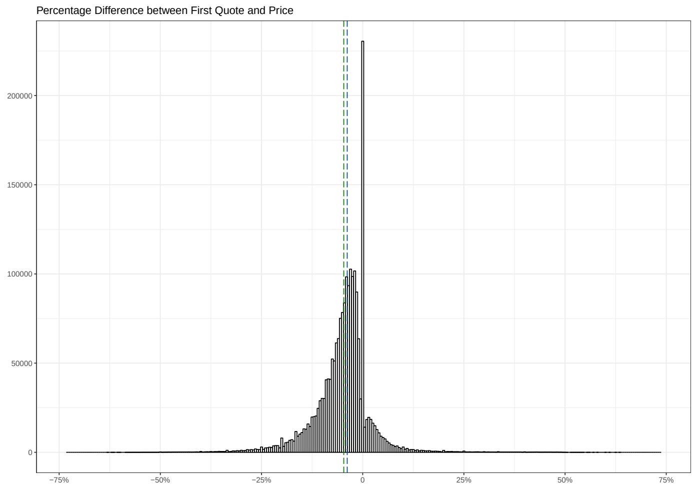
(a) Transactions



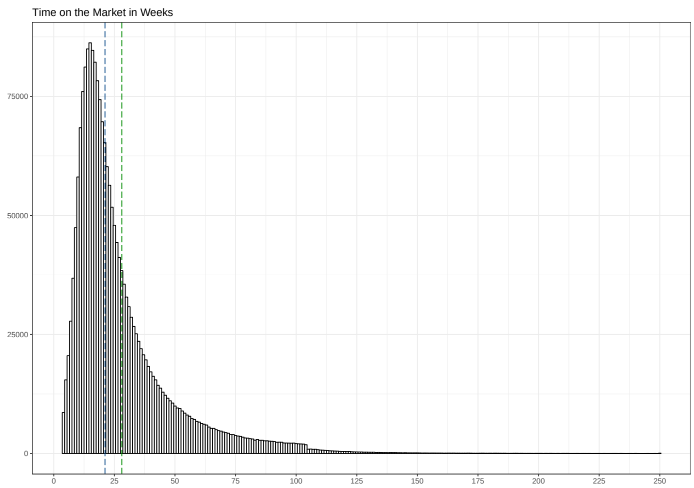
(b) Listings

Figure A3: Price Discount and Time on the Market for Matched Listings

The figure displays the distributions of price discount and time on the market (TOM) for the set of property listings that were matched to their respective ex-post transactions in the sample from 2009 to 2018. Figure A3a plots the histogram of the percentage difference between the first listed price and the final transaction price, while Figure A3b shows the histogram of time on the market measured as the number of weeks from listing to sale completion. The mean and median values of the two distributions are represented by the green and blue vertical lines, respectively.



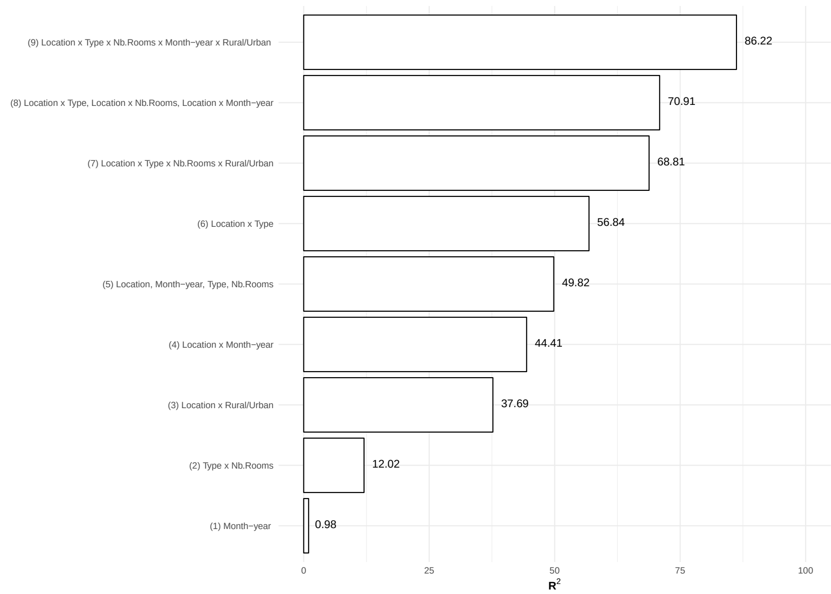
(a) Discount with respect to First Quote



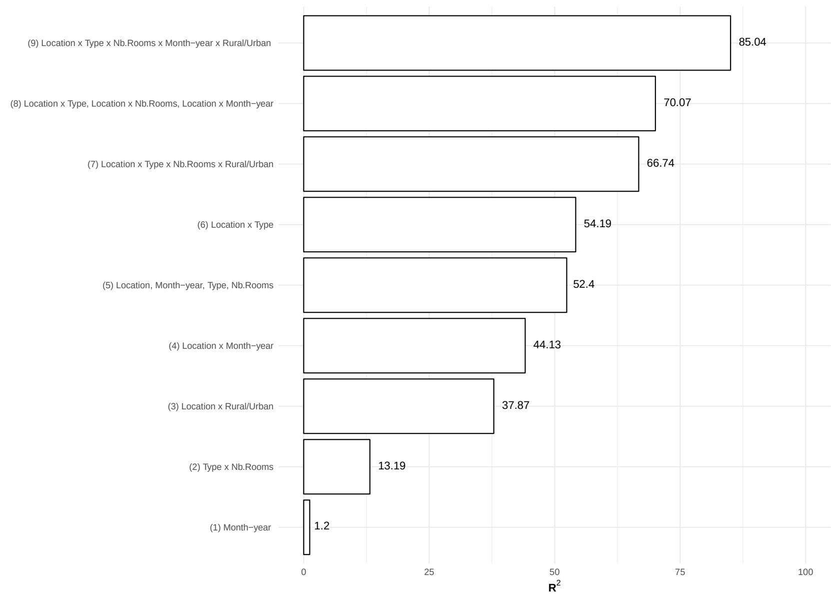
(b) Time on the Market (weeks)

Figure A4: Fraction of Explained Variation in Prices

The figure displays the percentage of the variation in the price data that is explained by observable characteristics, measured as the R-squared from a regression of prices on various fixed effects. Figure A4a shows the variation explained in the listing data and Figure A4b in the transaction data. Fixed effects included are: month-year of the listing or transaction; property type (detached, semi-detached, terraced house or a flat); number of rooms in the property, where properties with between 6 and 10 rooms are placed in one bucket and properties with more than 10 rooms in another; location, measured as the address outcode, and; a rural/urban area indicator from the 2011 Census classification of Output Areas.



(a) Listing Prices



(b) Transaction Prices

Figure A5: Variation In Quotes Around Publishing Dates

The figure plots the results from a regression of listing prices on dummies for the signed number of days between the listing date and the closest price data publication date for the sample after March 2012. The regression is specified as follows: $q_i = \alpha + \sum_{\Delta=-15}^{15} \gamma_\Delta \Delta_i + FE + \varepsilon_i$, where the fixed-effects correspond to the characteristics the matching is based on, i.e., location, property type, number of rooms and month-year, and Δ_i is a dummy for the signed difference in days between the date on which a listing is posted and the closest publication date. The baseline coefficient is the one for listings posted exactly on the publication date. The vertical lines represent the 95% confidence bounds for the point estimates.

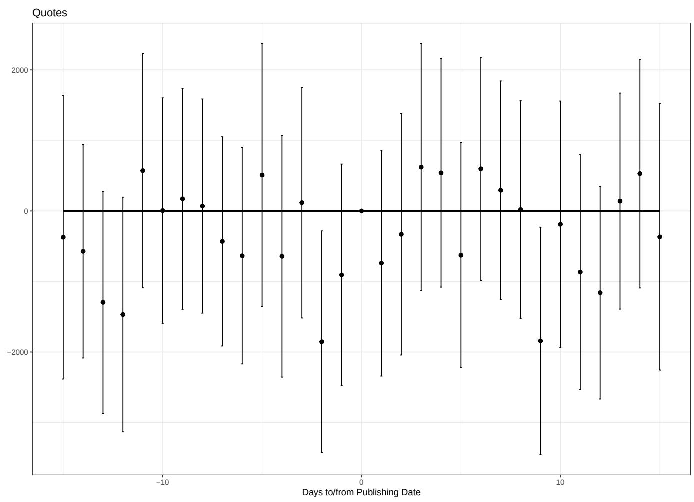


Figure A6: Density of Listing Observations Around Publishing Dates

The figure displays the smoothed density of the number of listing observations per day around price data publication dates for the sample after March 2012, where I fit two polynomials on each side of the publication date. The total daily count is first regressed on day-of-the-week dummies and the residuals of this regression are used for the density test.

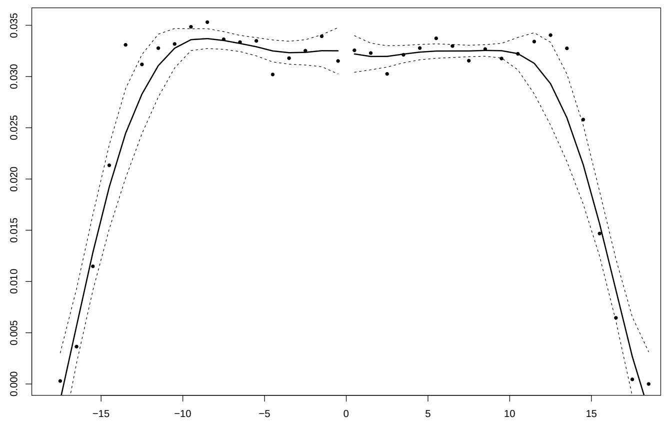


Figure A7: Difference-in-differences Analysis: Pre vs Post March 2012

The figure presents the change to the institutional setting that occurred in March 2012: beginning in March 2012, the Land Registry publishes regular monthly data on transactions on the twentieth working day of the subsequent month. The second figure shows that transaction prices, depicted in blue, from June 2013, are published on the twentieth working day of July which is 26th July in this case. Property listings posted at the beginning of June and before the publication date, depicted in light green, do not observe the June transactions data, while those posted after this date, depicted in dark green, may observe and therefore use use this data to make inference about market demand. The first figure shows that in the period before March 2012, there was no such regular shock to the sellers' information set. For example, June transactions data was not made publicly available on the twentieth working day of July 2011, 28th July. Listings that occurred throughout the months of July and August 2011, therefore, are depicted in light green as they might not observe the June transactions data or, at least, not have this information arrive at the same regular intervals.

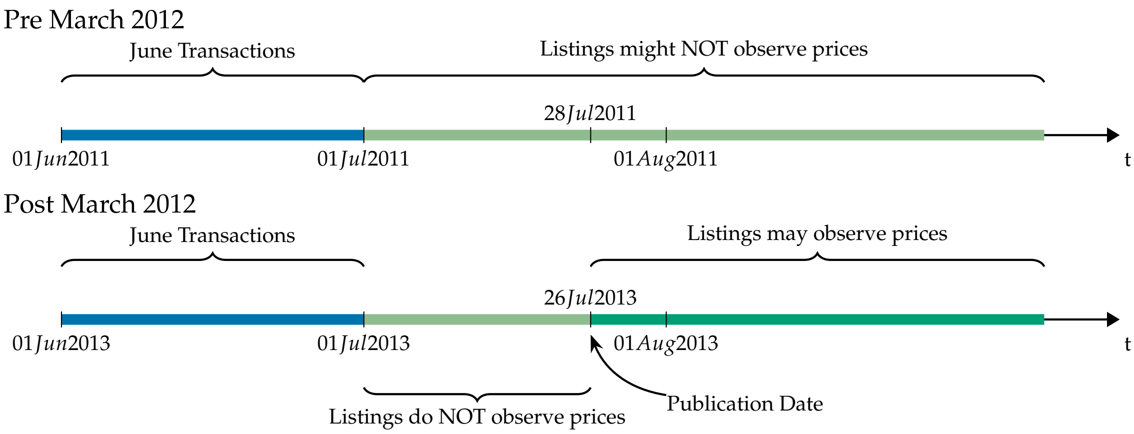


Figure A8: Experiment I: Indirect Price Effects Through Intermediate Listings

The figure provides an example of the possible channels through which a given price observation might have an influence on subsequent sellers. The blue circles represent transaction prices from a given month, the light green ones (Q-1) are listings that were posted in the week before the price data becomes available, while the dark green circles, Q+1 and Q+2, are listings posted in the first or second week following the publication date, respectively. Focusing on the listings posted in week two after publication and their links to the transaction prices from the previous month, I show the four possible cases that can arise. Going from left to right, there may be: (a) no comparable listing posted in any of the two weeks surrounding the Price Paid data publication date; (b) comparable listings only in the week before but not the week after; (c) comparable listings only in the week after but not the week before, and; (d) at least one similar listing in both weeks.

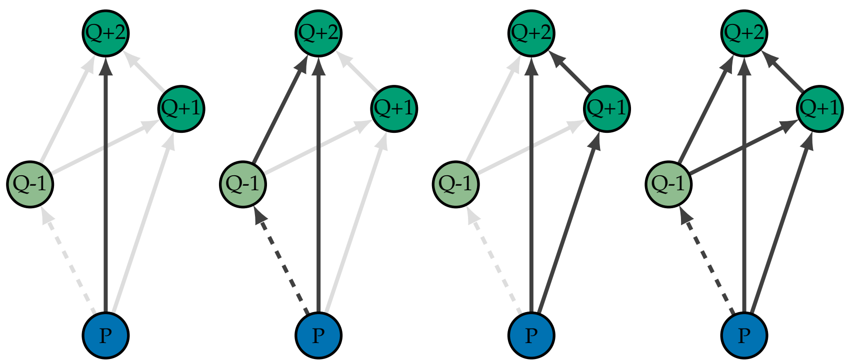


Figure A9: Experiment II: Chain Effects of Prices on Quotes by Order of Match

The figure shows how a given price observation can have an increasing number of channels of indirect influence on future listings as the number of interim comparables grows. The blue circle repesents a given transaction price, the light green circles are listings posted before its publication date and the dark green circles are listings posted after. The listings are indexed in order to capture their chronological arrival in the market and provide an idea of the information set of subsequent sellers.

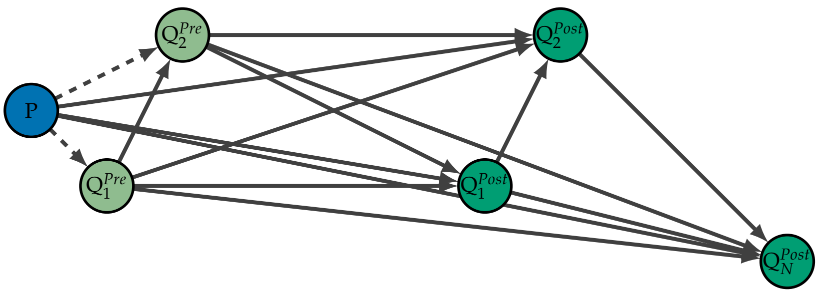


Figure A10: Effect of Transaction Prices by Order of Match

The figure plots the price coefficients from the following regression along with their 95% confidence bounds: $log(q_i) = \alpha + \beta_1 \times log(p_j) + \sum_{k=2}^{10} \beta_k^{pre} \times log(p_j) \times Comp\ Order\ k\ Pre_i + \sum_{k=2}^{10} \gamma_k^{pre} Comp\ Order\ k\ Pre_i + \sum_{k=1}^{10} \beta_k^{post} \times log(p_j) \times Comp\ Order\ k\ Post_i + \sum_{k=1}^{10} \gamma_k^{post} Comp\ Order\ k\ Post_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and *Comp Order k Pre (Post)*$_i$ is a dummy that turns on when quote i is the k-th sequential match to transaction j in the period before (after) the price data publication date. The sample includes listings in the one-month period surrounding the publication date that have a comparable transaction which has at least one treated and one untreated match. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price, as well as listing month-year fixed effects are included in this specification.

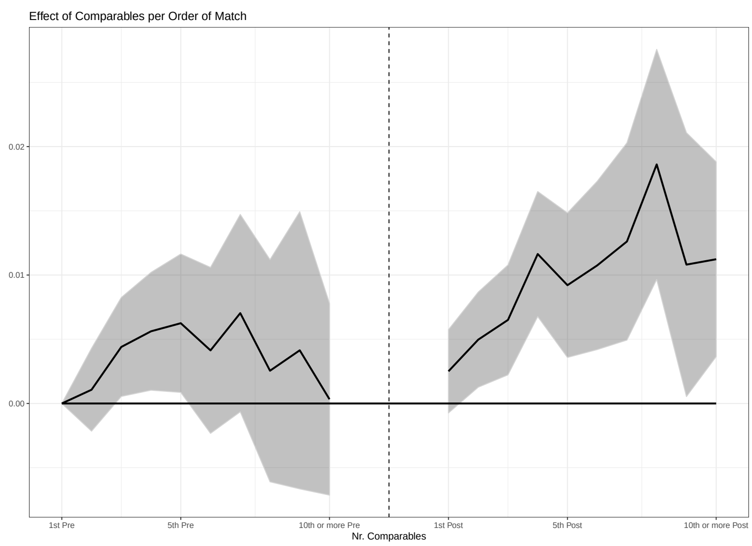


Figure A11: Experiment III: Effect of Prices on Quote Updates

The figure shows the way that a given price observation can exercise an increasing influence on a listing via its effect on other observable listings. The blue circle depicts a transaction from the previous month, the green circles are the ordered quote changes for a listing that has been first posted following the release of the transaction data and the pink circles represent other listings posted while the property of interest is still on the market.

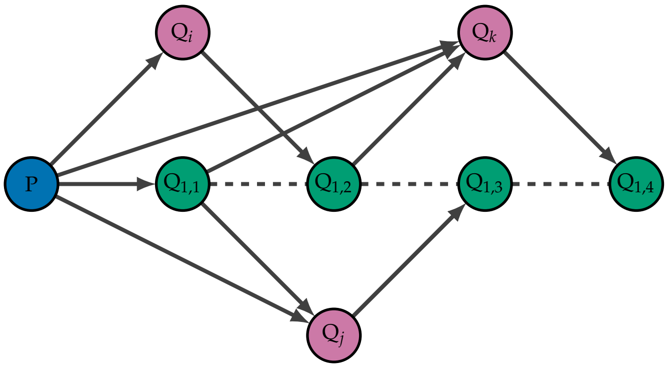


Figure A12: Effect of Transaction Prices on Quote Updates

The figure displays the price coefficients from the following regression along with their 95% confidence bounds: $log(q_i^n) = \alpha + \beta_1 \times log(p_j) + \sum_{n=2}^{5} \beta_n \times log(p_j) \times Update\ Number\ n_i + \sum_{n=2}^{5} \gamma_n Update\ Number\ n_i + Controls + \varepsilon_i^n$, where q_i^n is the n-th listed price update for property i, p_j is the transaction price for a comparable property j sold in the month before property i was initially listed and $Update\ Number\ n_i$ is a dummy that turns on when q_i^n is the n-th consecutive quote update for property i. The sample includes listings in the post March 2012 period that have at least one price change and a comparable transaction that has been published just before the listing was first posted. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price as well as month-year fixed effects are included in this specification.

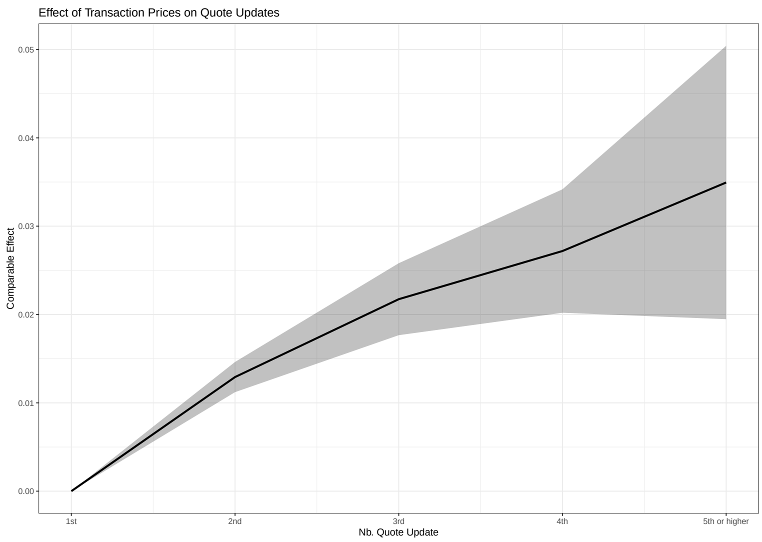


Figure A13: Effect of Past Prices by Price Discount

The figure plots the price coefficients of the following regressions along with their 95% confidence bounds: $log(q_i^k) = \alpha^k + \beta^k \times log(p_j) + Controls + \varepsilon_i^k$, where q_i^k is the first quoted price for listing i which is in quantile k of the price discount distribution and p_j is the transaction price for a similar property which has been published in the month before the listing was originally posted. The sample includes listings in the sample period starting from March 2012 that have been matched to their respective ex-post transactions. Each coefficient comes from a regression of listings from a different quantile of the price discount distribution: the first coefficient is based on listings sold at the largest price discount while the last one on listings that were sold at a premium to quoted price. The horizontal line represents the average price effect across the full sample. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price, as well as listing month-year fixed effects are included in the regressions.

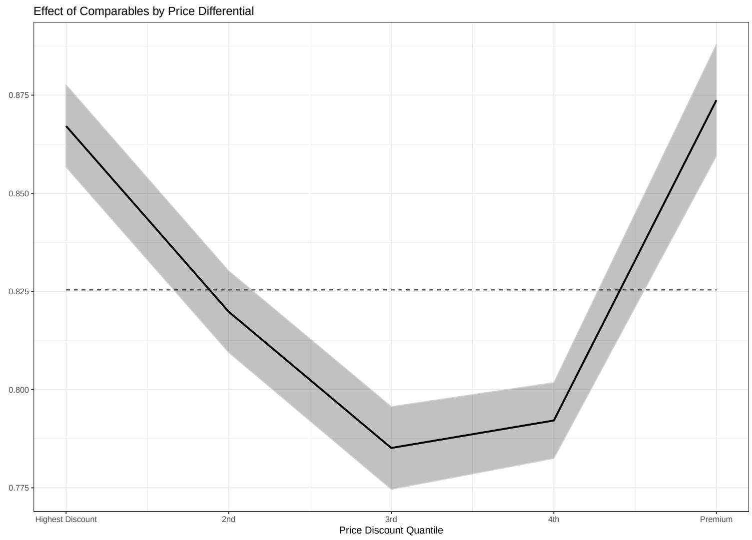
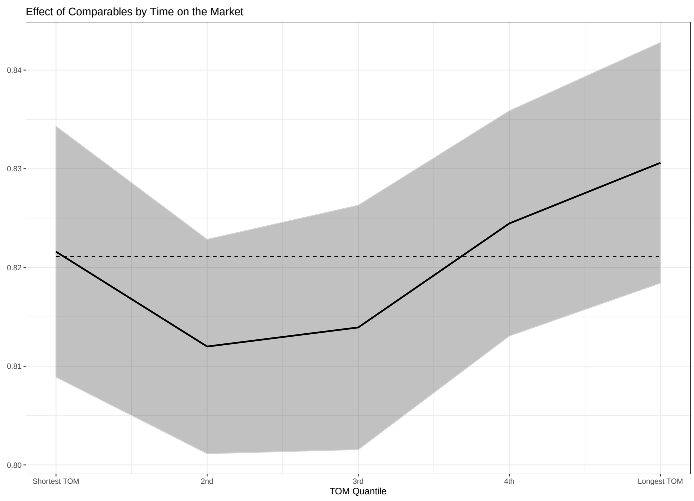
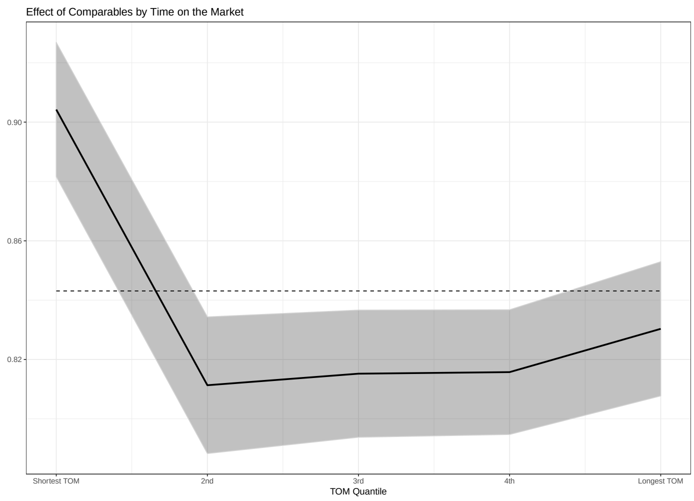


Figure A14: Effect of Past Prices by Time on the Market

The figure plots the price coefficients of the following regressions along with their 95% confidence bounds: $log(q_i^k) = \alpha^k + \beta^k \times log(p_j) + Controls + \varepsilon_i^k$, where q_i^k is the first quoted price for listing i which is in quantile k of the TOM distribution and p_j is the transaction price for a similar property which has been published in the month before the listing was originally posted. The sample includes listings after March 2012 that have been matched to their respective ex-post transactions. Figure A14a considers properties that sold at a discount to listed price, while Figure A14b properties sold at a premium. Each coefficient comes from a regression of listings from a different quantile of the TOM distribution: the first coefficient is based on listings that took the least time to sell while the last one on listings that had the longest duration. The horizontal line represents the average price effect across the full sample. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price, as well as listing month-year fixed effects are included in the regressions.



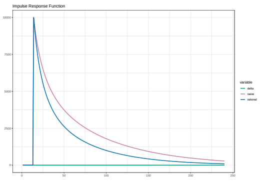
(a) Properties Sold at Discount



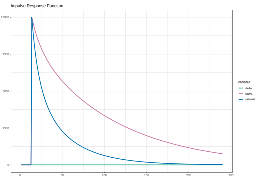
(b) Properties Sold at Premium

Figure A15: Impulse Response to a Shock to the Public Signal

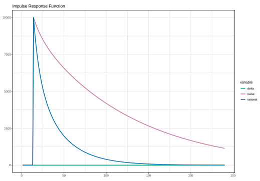
The figure plots impulse responses of prices to a shock to the public signal in period 13. The response of naïve prices is depicted in pink, that of rational prices is in blue and the underlying state of demand is plotted in green. The various figures vary the number of simultaneous price-setters in a given period, $n = 1, 5, 10$ and the frequency at which the public signal arrives, $k = 1, 6, 12$ months. The shock is standardised to correspond to a £10,000 increase in prices on impact.



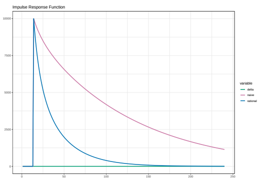
(a) $n = 1, k = 1$



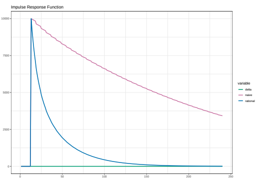
(b) $n = 5, k = 1$



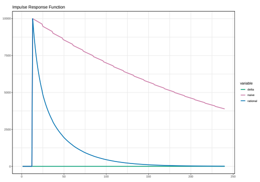
(c) $n = 10, k = 1$



(d) $n = 10, k = 1$



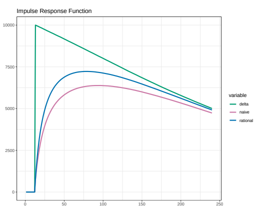
(e) $n = 10, k = 6$



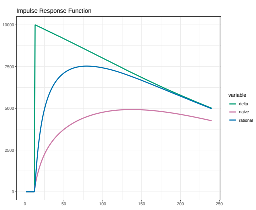
(f) $n = 10, k = 12$

Figure A16: Impulse Response to a Shock to the Underlying

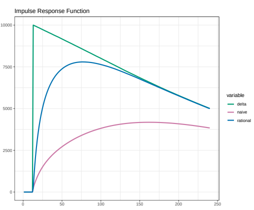
The figure plots impulse responses of prices to a shock to the underlying in period 13. The response of naïve prices is depicted in pink, that of rational prices is in blue and the underlying state of demand is plotted in green. The various figures vary the number of simultaneous price-setters in a given period, $n = 1, 5, 10$ and the frequency at which the public signal arrives, $k = 1, 6, 12$ months. The shock is standardised to correspond to a £10,000 increase in the underlying on impact.



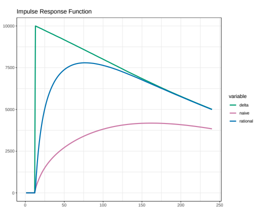
(a) $n = 1, k = 1$



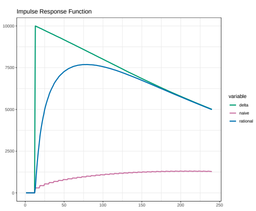
(b) $n = 5, k = 1$



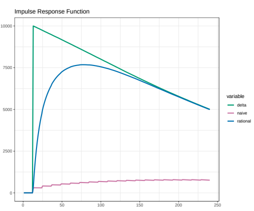
(c) $n = 10, k = 1$



(d) $n = 10, k = 1$



(e) $n = 10, k = 6$



(f) $n = 10, k = 12$

A.3 Structural Estimation

In this section I sketch a stylised model that builds on the results from Section 1.1 in the main body of the paper in order to gain intuition regarding the way that covariances between prices and subsequent quotes are expected to change as the number of intermediate comparables grows under the Bayesian and naïve learning models. I make the model more realistic compared to Section 1.1 by allowing the underlying state to change over time and for the presence of commonly observed signals. Subsequently, I use the estimated coefficients from the empirical study of indirect effects obtained in Section 1.4.2 to provide some evidence regarding the magnitude of the impact of pricing mistakes on aggregate market dynamics.

For simplicity, let us assume that the log of house prices are determined by the fundamental δ_t which follows an AR(1) process with persistence parameter ρ and mean a:

$$\delta_t = a + \rho\delta_{t-1} + \varepsilon_t \ , \ \varepsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2) \tag{A.1}$$

This can be thought of as a reduced-form way of modelling the demand that sellers face. As a result, prospective sellers set listing prices based on their expectation of δ_t:

$$p_{i,t} = \mathbb{E}_{i,t}[\delta_t] \tag{A.2}$$

$$q_{i,t} = \mathbb{E}_{i,t}[\delta_t] \tag{A.3}$$

where $q_{i,t}$ is the log quote set by agent i at time t and $p_{i,t}$ is a transaction price for a property sold at time t. Agents do not observe the realisation of δ_t and, therefore, try to estimate its value from available information. In particular, the informational structure is characterised by the presence of public and private signals. Each seller receives a private signal $s^q_{i,t}$ before choosing the quote:

$$s^q_{i,t} = \delta_t + \nu_{i,t} \ , \ \nu_{i,t} \overset{iid}{\sim} \mathcal{N}(0, \sigma_\nu^2) \tag{A.4}$$

The noise terms are independent and identically distributed across individuals and time. If there are multiple sellers in a given period, they all receive a different private signal. There is also a publicly observable signal s_t arriving every k periods:

$$s_t = \delta_t + u_t \ , \ u_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_u^2) \tag{A.5}$$

In this case too, the noise is identically and independently distributed across time. The public signal represents any public information that sellers might use to make inference about housing demand and prices, for instance a housing price index published at regular frequencies. Alternatively, we can interpret it as representing local area characteristics or amenities visible to everyone. Finally, sellers also observe the full history of past transactions and listings which they also use to extract the private signals that agents in previous periods have received. The signal contained in past prices has the same form as the other two signals but possibly a different precision:

$$s^p_{i,t} = \delta_t + \eta_{i,t}\ ,\ \eta_{i,t} \overset{iid}{\sim} \mathcal{N}(0,\, \sigma^2_\eta) \tag{A.6}$$

I make a distinction between the private signals embedded in prices and those from quotes to account for the fact that the final transaction price can be adjusted upon interaction with the buyer. In other words, the signal extracted from transaction prices contains additional information about demand and (idiosyncratic) buyer characteristics to the extent that buyers have some bargaining power. This approach of modelling private signals differently across quotes and prices can be seen as a reduced-form way of capturing housing market features common in the housing literature without resorting to more complicated search models.[1] At every point in time, the information sets of agents consist of the full history of past prices, the history of public signals and their own private signals. They thus form conditional expectations of δ_t and set quotes accordingly, as follows:

$$p_{i,t} = \mathbb{E}[\delta_t | s^p_{i,t}, s^t, p^{t-1}, q^{t-1}] \tag{A.7}$$

$$q_{i,t} = \mathbb{E}[\delta_t | s^q_{i,t}, s^t, p^{t-1}, q^{t-1}] \tag{A.8}$$

where $s^t = \{s_0, s_{0+k}, s_{0+2k}, ...\}, p^{t-1} = \{p_0, p_1, ..., p_{t-1}\}, q^{t-1} = \{q_0, q_1, ..., q_{t-1}\}$ denote the full history of public signals, transaction prices and listing prices, respectively, that agents at time t observe[2], while $s^p_{i,t}$ or $s^q_{i,t}$ is agent i's private signal.

A.3.1 Parameter Calibration

The goal of this exercise is to help us gain understanding about the way that the effect of a given price p on subsequent listings q evolves as the number of interim comparables increases under the Bayesian learning framework. Specifically, I use the setting of Figure A8 and Table A7 where I look at the effect of recent transaction prices on quotes posted in week two after the publication date. As in Table A7, there are four types of sellers depending on what is in their information set:

- Seller 1 observes the newly published price data but has no comparable listings in the two-week period around the publishing date;

[1] The specifics of the housing market microstructure and the bargaining process go beyond the scope of this paper. Here, I simply attempt to provide some evidence on the economic magnitude of the effect of naïve inference by sellers. Using a more involved search model would render the interpretation of the results more difficult without changing the big picture. See Han and Strange (2015) for a survey of the literature on the microstructure of housing markets.

[2] As price data are published on a monthly basis, agents observe the history of past prices up until the previous month. Quotes can, however, be observed at a higher frequency on property websites. Here I assume that sellers can observe all past listings up to the previous week.

- Seller 2 observes the newly published price data and has at least one comparable listing in the week before and no comparable in the week after the publishing date;
- Seller 3 observes the newly published price data and has no comparable listing in the week before and at least one comparable in the week after the publishing date, and;
- Seller 4 observes the newly published price data and has at least one comparable listing both in the week before and the week after the publishing date.

The object of interest for this analysis is the covariance of four types of quotes with the most recently published prices. For simplicity, I assume that the newly published comparable transaction price, denoted as p_0, has been determined based on agent 0's private signal $s^p_{0,0}$ and a public signal[3] that englobes the full history of past information s^0:

$$p_0 = \mathbb{E}[\delta_0 | s^p_{0,0}, s^0] = w^p_0 \times s^p_{0,0} + (1 - w^p_0) \times s^0 \tag{A.9}$$

where the weight that agent 0 assigns to his private signal is proportional to the signal precision, i.e., $w^p_0 = \frac{\sigma_\eta^{-2}}{\sigma_\eta^{-2} + \sigma_w^{-2}}$. The variance of the posterior belief of agent 0 $P_{0|0}$ is given by:

$$P_{0|0} = (w^p_0)^2 \times \sigma^2_\eta + (1 - w^p_0)^2 \times \sigma^2_w \tag{A.10}$$

One period later, at $t = 1$, prospective sellers determine listing prices using available information. Type 1 sellers observe p_0, their own private signal $s^q_{1,1}$, a new public signal s_1[4] and they also directly observe the same public signal s^0 that has already been accounted for by agent 0. Rational sellers understand that p_0 already incorporates the original public signal s^0 and thus avoid double-counting it. They form their posterior belief about δ_1, and hence the quote, as follows:

$$\begin{aligned} q_{1,1} &= \mathbb{E}[\delta_1 | s^q_{1,1}, s_1, p_0, s^0] = \mathbb{E}[\delta_1 | s^q_{1,1}, s_1, p_0] \\ &= w^q_1 \times \left[\frac{\sigma_v^{-2}}{\sigma_v^{-2} + \sigma_u^{-2}} \times s^q_{1,1} + \frac{\sigma_u^{-2}}{\sigma_v^{-2} + \sigma_u^{-2}} \times s_1 \right] + (1 - w^q_1) \times (a + \rho \times p_0) \end{aligned} \tag{A.11}$$

where $w^q_1 = P_{1|0} \times \left(P_{1|0} + \left(\frac{\sigma_v^{-2}}{\sigma_v^{-2} + \sigma_u^{-2}} \right)^2 \times \sigma^2_v + \left(\frac{\sigma_u^{-2}}{\sigma_v^{-2} + \sigma_u^{-2}} \right)^2 \times \sigma^2_u \right)^{-1}$ and $P_{1|0} = \rho^2 \times P_{0|0} + \sigma^2_\varepsilon$ is the variance of agent 1's prior belief about δ_1 given the available information up to time 0. Note that under the Bayesian learning framework agent 1 does not

[3]The public signal here can be interpreted as the prior belief of agent 0 based on his information set before receiving the private signal. For this reason, I allow this prior to have a different precision compared to the periodic public signals, as follows: $s^0 = \delta_0 + w_0$, $w_0 \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_w)$.

[4]For this exercise, I assume that a new public signal arrives every period, i.e., k=1.

assign an explicit weight on the initial public signal s^0, rather he treats p_0 as a sufficient statistic for all information up to $t = 0$, being aware that it already embeds s^0. Naïve sellers, however, fail to recognise this, believing that p_0 is solely determined based on agent 0's private signal, i.e., they believe $p_0 = \widetilde{\mathbb{E}}[\delta_0 | s^p_{0,0}]$. As a result, they treat the newly observed price as independent from the public signal, leading them to assign an explicit weight to the public signal when forming beliefs about the state of housing demand:

$$\begin{aligned}\widetilde{q}_{1,1} &= \widetilde{\mathbb{E}}[\delta_1 | s^q_{1,1}, s_1, p_0, s^0] \\ &= w^q_1 \times \left[\frac{\sigma_v^{-2}}{\sigma_v^{-2} + \sigma_u^{-2}} \times s^q_{1,1} + \frac{\sigma_u^{-2}}{\sigma_v^{-2} + \sigma_u^{-2}} \times s_1\right] + \\ &\quad (1 - w^q_1) \times [a + w^p_0 \times \rho p_0 + (1 - w^p_0) \times \rho s^0]\end{aligned} \tag{A.12}$$

where the weight w^q_1 is the same as in the Bayesian case. Notice that when forming his prior belief the naïve agent assigns weights to p_0 and s^0 that a Bayesian learner would assign to the correctly extracted signals $s^p_{0,0}$ and s^0, respectively. The problem with naïve learners is that they believe that past actions are purely driven by private signals, $p_0 = s^p_{0,0}$, when instead p_0 has effectively been determined using all available information at $t = 0$, as in equation (A.9). This leads naïve agents to overweight the commonly observed public signal relative to Bayesian learners, as they account for it both directly (through its explicit weight in the prior belief) and indirectly (through its effect on p_0).

Type 2 sellers observe p_0 and another listing, denoted by $q_{0,1}$, posted before the price data publication date[5]. They also observe the original public signal s^0, the public signal from period 1, s_1, and a new private signal, $s^q_{2,1}$. The Bayesian learner would extract only the private signals embedded in $q_{0,1}$ and p_0 and form beliefs as follows:

$$\begin{aligned}q_{2,1} &= \mathbb{E}[\delta_1 | s^q_{2,1}, q_{0,1}, s_1, p_0, s^0] = \mathbb{E}[\delta_1 | s^q_{2,1}, s^q_{0,1}, s_1, p_0] \\ &= w^q_2 \times \left[\frac{\sigma_v^{-2}}{2\sigma_v^{-2} + \sigma_u^{-2}} \times (s^q_{2,1} + s^q_{0,1}) + \frac{\sigma_u^{-2}}{2\sigma_v^{-2} + \sigma_u^{-2}} \times s_1\right] + (1 - w^q_2) \times [a + \rho \times p_0]\end{aligned} \tag{A.13}$$

where $w^q_2 = P_{1|0} \times \left(P_{1|0} + 2 \times \left(\frac{\sigma_v^{-2}}{2\sigma_v^{-2}+\sigma_u^{-2}}\right)^2 \times \sigma_v^2 + \left(\frac{\sigma_u^{-2}}{2\sigma_v^{-2}+\sigma_u^{-2}}\right)^2 \times \sigma_u^2\right)^{-1}$. Notice that both agent 2 and the interim agent forming quote $q_{0,1}$ observe s^0 and s_1. Since seller 2 knows their precisions, he can easily infer what the private signal embed-

[5] Note that the interim agent does not observe p_0 and therefore cannot learn agent 0's private signal. He thus sets his quote $q_{0,1}$ based on the original public signal s^0, the new public signal from period 1, s_1, and his own private signal, $s^q_{0,1}$, as follows: $q_{0,1} = \mathbb{E}[\delta_1 | s^q_{0,1}, s_1, s^0] = w^q_0 \times \left[\frac{\sigma_v^{-2}}{\sigma_v^{-2}+\sigma_u^{-2}} \times s^q_{0,1} + \frac{\sigma_u^{-2}}{\sigma_v^{-2}+\sigma_u^{-2}} \times s_1\right] + (1 - w^q_0) \times (a + \rho s^0)$, where $w^q_0 = (\rho^2 \times \sigma_w^2 + \sigma_\varepsilon^2) \times \left(\rho^2 \times \sigma_w^2 + \sigma_\varepsilon^2 + \left(\frac{\sigma_v^{-2}}{\sigma_v^{-2}+\sigma_u^{-2}}\right)^2 \times \sigma_v^2 + \left(\frac{\sigma_u^{-2}}{\sigma_v^{-2}+\sigma_u^{-2}}\right)^2 \times \sigma_u^2\right)^{-1}$.

ded in $q_{0,1}$ is and avoid double counting information. His posterior belief about δ_1 is thus equal to a weighted-average of the prior belief (for which p_0 is again a sufficient statistic) and the average of the newly obtained signals in period 1, i.e., $s^q_{2,1}$, $s^q_{0,1}$ and s_1, weighted by their precisions. The naïve sellers of type 2 make the same mistake as the type 1 naïve sellers, i.e., they incorrectly believe that $p_0 = \widetilde{\mathbb{E}}[\delta_0|s^p_{0,0}]$ and $q_{0,1} = \widetilde{\mathbb{E}}[\delta_1|s^q_{0,1}]$. As a result, the effect of naïve inference is now two-fold as the initial public signal has been embedded both in p_0 and $q_{0,1}$. Since the private signals received by agent 2 and the interim agent who sets $q_{0,1}$ are equally-precise, agent 2 sets his quote as follows:

$$\begin{aligned}\widetilde{q}_{2,1} &= \widetilde{\mathbb{E}}[\delta_1|s^q_{2,1}, q_{0,1}, s_1, p_0, s^0] \\ &= w^q_2 \times \left[\frac{\sigma_v^{-2}}{2\sigma_v^{-2}+\sigma_u^{-2}} \times (s^q_{2,1} + q_{0,1}) + \frac{\sigma_u^{-2}}{2\sigma_v^{-2}+\sigma_u^{-2}} \times s_1\right] + \\ &\quad (1-w^q_2) \times [a + w^p_0 \times \rho p_0 + (1-w^p_0) \times \rho s^0]\end{aligned} \tag{A.14}$$

Similarly to seller 1, the naïve type 2 seller assigns weights that would be correct if past prices and quotes were truly equal to the private signals of the preceding agents. As this is not the case, however, seller 2 ends up overweighting the common signal through two indirect channels: its influence on p_0 and that on $q_{0,1}$.

Moving on to type 3 sellers, recall that the only difference with type 2 sellers is that they observe listing $q_{1,1}$, instead of $q_{0,1}$, which is set after the publication date and, in turn, directly observes p_0 as well. Rational agents would recognise this and extract the private signal from $q_{1,1}$ in order to avoid double counting the public signals s^0 and s_1 as well as the private information coming from p_0. Accordingly, it follows that the weights they would assign are the same as for the rational type 2 sellers:

$$\begin{aligned}q_{3,1} &= \mathbb{E}[\delta_1|s^q_{3,1}, q_{1,1}, s_1, p_0, s^0] = \mathbb{E}[\delta_1|s^q_{3,1}, s^q_{1,1}, s_1, p_0] \\ &= w^q_2 \times \left[\frac{\sigma_v^{-2}}{2\sigma_v^{-2}+\sigma_u^{-2}} \times (s^q_{3,1} + s^q_{1,1}) + \frac{\sigma_u^{-2}}{2\sigma_v^{-2}+\sigma_u^{-2}} \times s_1\right] + (1-w^q_2) \times [a + \rho \times p_0]\end{aligned} \tag{A.15}$$

Similarly, naïve type 3 sellers assign the same weights as naïve type 2 sellers under the beliefs that $p_0 = \widetilde{\mathbb{E}}[\delta_0|s^p_{0,0}]$ and $q_{1,1} = \widetilde{\mathbb{E}}[\delta_1|s^q_{1,1}]$. As a result, they set the quote as follows:

$$\begin{aligned}\widetilde{q}_{3,1} &= \widetilde{\mathbb{E}}[\delta_1|s^q_{3,1}, \widetilde{q}_{1,1}, s_1, p_0, s^0] \\ &= w^q_2 \times \left[\frac{\sigma_v^{-2}}{2\sigma_v^{-2}+\sigma_u^{-2}} \times (s^q_{3,1} + \widetilde{q}_{1,1}) + \frac{\sigma_u^{-2}}{2\sigma_v^{-2}+\sigma_u^{-2}} \times s_1\right] + \\ &\quad (1-w^q_2) \times [a + w^p_0 \times \rho p_0 + (1-w^p_0) \times \rho s^0]\end{aligned} \tag{A.16}$$

An important distinction arises when comparing naïve agents of types 2 and 3. Namely, although they assign the exact same weights, the fact that $\widetilde{q}_{1,1}$, unlike $q_{0,1}$, is formed using information from p_0 could lead to different covariances with p_0 between type 2 and type 3 sellers. This is because agent 3 has two different channels of influence from the private signal in p_0, one through the direct effect of p_0 on $\widetilde{q}_{3,1}$ and another due to the indirect effect of p_0 through $\widetilde{q}_{1,1}$. In addition, there are now four implicit weights on the public signal: the direct effect, the indirect effect through p_0 and the indirect effect through $\widetilde{q}_{1,1}$, which can further be decomposed into its direct effect on $\widetilde{q}_{1,1}$ and the indirect effect on $\widetilde{q}_{1,1}$ through p_0.

Finally, sellers of type 4 have the richest information set: they observe p_0 as well as both $q_{0,1}$ and $q_{2,1}$, in addition to the private and public signals, $s^q_{4,1}$, s_1 and s^0. As usual, rational type 4 sellers extract and use only the private information from the intermediate quotes, leading to the following beliefs:

$$\begin{aligned} q_{4,1} &= \mathbb{E}[\delta_1 | s^q_{4,1}, q_{2,1}, q_{0,1}, s_1, p_0, s^0] = \mathbb{E}[\delta_1 | s^q_{4,1}, s^q_{2,1}, s^q_{0,1}, s_1, p_0] \\ &= w^q_4 \times \left[\frac{\sigma_v^{-2}}{3\sigma_v^{-2} + \sigma_u^{-2}} \times (s^q_{4,1} + s^q_{2,1} + s^q_{0,1}) + \frac{\sigma_u^{-2}}{3\sigma_v^{-2} + \sigma_u^{-2}} \times s_1 \right] + \\ &\quad (1 - w^q_4) \times [a + \rho \times p_0] \end{aligned} \tag{A.17}$$

where $w^q_4 = P_{1|0} \times \left(P_{1|0} + 3 \times \left(\frac{\sigma_v^{-2}}{3\sigma_v^{-2}+\sigma_u^{-2}} \right)^2 \times \sigma_v^2 + \left(\frac{\sigma_u^{-2}}{3\sigma_v^{-2}+\sigma_u^{-2}} \right)^2 \times \sigma_u^2 \right)^{-1}$. Naïve type 4 agents instead, believing that $p_0 = \widetilde{\mathbb{E}}[\delta_0 | s^p_{0,0}]$, $q_{0,1} = \widetilde{\mathbb{E}}[\delta_1 | s^q_{0,1}]$ and $q_{2,1} = \widetilde{\mathbb{E}}[\delta_1 | s^q_{2,1}]$, assign weights as follows:

$$\begin{aligned} \widetilde{q}_{4,1} &= \widetilde{\mathbb{E}}[\delta_1 | s^q_{4,1}, \widetilde{q}_{2,1}, q_{0,1}, s_1, p_0, s^0] \\ &= w^q_4 \times \left[\frac{\sigma_v^{-2}}{3\sigma_v^{-2} + \sigma_u^{-2}} \times (s^q_{4,1} + \widetilde{q}_{2,1} + q_{0,1}) + \frac{\sigma_u^{-2}}{3\sigma_v^{-2} + \sigma_u^{-2}} \times s_1 \right] + \\ &\quad (1 - w^q_4) \times [a + w^p_0 \times \rho p_0 + (1 - w^p_0) \times \rho s^0] \end{aligned} \tag{A.18}$$

The initial public signal s^0 influences type 4 sellers via six different channels, the private signal from p_0 affects $\widetilde{q}_{4,1}$ via two channels and the new public signal s_1 is also counted multiple times through its effect on $q_{0,1}$, $\widetilde{q}_{2,1}$ and the directly assigned weight.

The results above show how agents overweight repeated information at the expense of novel signals and this generates differences in the covariance between quotes and prices under the naïve model relative to the Bayesian case. To more clearly see the differences in comovement patterns under the two models, I compute the covariances of the four sets of quotes with p_0 for both the Bayesian and naïve agents. It can be shown that the covariances in the Bayesian case take the following forms:

$$Cov(q_{1,1}, p_0) = \rho Var(\delta_0) + (1 - w^q_1)(w^p_0)^2 \times \rho\sigma^2_\eta + (1 - w^q_1)(1 - w^p_0)^2 \times \rho\sigma^2_w \tag{A.19}$$

$$Cov(q_{2,1}, p_0) = \rho Var(\delta_0) + (1 - w^q_2)(w^p_0)^2 \times \rho\sigma^2_\eta + (1 - w^q_2)(1 - w^p_0)^2 \times \rho\sigma^2_w \tag{A.20}$$

$$Cov(q_{3,1}, p_0) = Cov(q_{2,1}, p_0) \tag{A.21}$$

$$Cov(q_{4,1}, p_0) = \rho Var(\delta_0) + (1 - w_4^q)(w_0^p)^2 \times \rho\sigma_\eta^2 + (1 - w_4^q)(1 - w_0^p)^2 \times \rho\sigma_w^2 \tag{A.22}$$

Note that the only difference in the covariance expressions above is in the weights $(1 - w_i^q)$ that multiply the two terms related to the respective inverse precisions of the private signal $s_{0,0}^p$ and the initial public signal s^0. As these weights are decreasing with i, it follows that Bayesian updating would imply that the covariances, and therefore the betas, should be monotonically decreasing with the increase in the number of intermediate comparables, regardless of the parameter values. In other words, as the information set of agents grows, they optimally assign a lower weight to each individual signal. Benchmarking the results of the empirical analysis from Section 1.4.2 against these predictions, we can reject the hypothesis that sellers in the housing market act in a Bayesian way. On the other hand, we can derive the same covariances for the naïve case and compare:

$$Cov(\widetilde{q}_{1,1}, p_0) = \rho Var(\delta_0) + (1 - w_1^q)(w_0^p)^3 \times \rho\sigma_\eta^2 + (1 - w_1^q)(1 - w_0^p)^2(1 + w_0^p) \times \rho\sigma_w^2 \tag{A.23}$$

$$\begin{aligned} Cov(\widetilde{q}_{2,1}, p_0) = \; & \rho Var(\delta_0) + (1 - w_2^q)(w_0^p)^3 \times \rho\sigma_\eta^2 + \\ & [w_2^q k_2(1 - w_0^q) + (1 - w_2^q)(1 - w_0^p)(1 + w_0^p)](1 - w_0^p) \times \rho\sigma_w^2 \end{aligned} \tag{A.24}$$

$$\begin{aligned} Cov(\widetilde{q}_{3,1}, p_0) = \; & \rho Var(\delta_0) + [w_2^q k_2(1 - w_1^q) + (1 - w_2^q)](w_0^p)^3 \times \rho\sigma_\eta^2 + \\ & [w_2^q k_2(1 - w_1^q) + (1 - w_2^q)](1 + w_0^p)(1 - w_0^p)^2 \times \rho\sigma_w^2 \end{aligned} \tag{A.25}$$

$$\begin{aligned} Cov(\widetilde{q}_{4,1}, p_0) = \; & \rho Var(\delta_0) + [w_4^q k_4(1 - w_2^q) + (1 - w_4^q)](w_0^p)^3 \times \rho\sigma_\eta^2 + \\ & [w_4^q k_4(1 - w_0^q)(1 + w_2^q k_2) + w_4^q k_4(1 - w_2^q)(1 - w_0^p)(1 + w_0^p) + \\ & (1 - w_4^q)(1 - w_0^p)(1 + w_0^p)](1 - w_0^p) \times \rho\sigma_w^2 \end{aligned} \tag{A.26}$$

where $k_2 = \frac{\sigma_v^{-2}}{2\sigma_v^{-2} + \sigma_u^{-2}}$ and $k_4 = \frac{\sigma_v^{-2}}{3\sigma_v^{-2} + \sigma_u^{-2}}$. The covariances may no longer be decreasing due to the overweighting of stale information embedded in p_0 relative to new signals coming from intermediate comparables. We can note that the original public signal is always more heavily weighted in the covariances between naïve quotes and prices relative to the rational case. This is because, on top of assigning the optimal explicit weight to it, naïve sellers also get an indirect influence through its effect on previous prices/quotes. On the other hand, the private signal coming from p_0 might be both under- or over-weighted across agents, depending on the relative precisions.

As the results in Section 1.4.2 cannot be reconciled with Bayesian updating, I postulate that sellers are subject to naïve learning and use these results in a calibration exercise. Specifically, I estimate the signal precisions described above, i.e., the precisions of the original and periodic public signals, σ_w^{-2} and σ_u^{-2}, and the two types of private signals, σ_η^{-2} and σ_v^{-2}, using equations (A.23)-(A.26) and the results from Table A7. The parameters that govern the underlying process are estimated by running a monthly regression of log aggregate house prices on aggregate income. The predicted values from the above regression are used to fit an AR(1) process that yields estimates for a, ρ and σ_ε^2. The calibrated parameters are then employed in simulations in order to evaluate the aggregate impact of naïve inference on house prices in the long run.

A.3.2 Estimating the Magnitude of the Effect of Information Shocks under the Bayesian and Naïve Filters

Given the model described in equations (A.1)-(A.8) above, we can trace the learning process of sellers who act sequentially. Assuming there are n agents per period, the first set of agents set prices using their own private signal and the public signal[6,7]. For simplicity, I here assume that there is a single public signal arriving at $t = 0$, however, I vary the frequency of public signal arrival in the simulations. I first describe the updating process for fully rational agents and subsequently specify how this differs from naïve updating. Let us denote the posterior belief of agents at time t by $\delta_{t|t}$, it then follows that:

$$p_0^i = \delta_{0|0}^i = \mathbb{E}[\delta_0 | s_0^i, s_0] = w_0 \times s_0^i + (1 - w_0) \times s_0 \tag{A.27}$$

where $w_0 = \frac{\sigma_\eta^{-2}}{\sigma_\eta^{-2} + \sigma_u^{-2}}$, as before. Denoting the variance of the posterior belief of each seller acting in period t by $P_{t|t}$, we have:

$$P_{0|0} = w_0^2 \times \sigma_\eta^2 + (1 - w_0)^2 \times \sigma_u^2 \tag{A.28}$$

The second set of agents in period $t = 1$ observe the same public signal s_0 and the n prices from the previous period, along with their own private signals. Unlike in the standard single-file example, their prior belief, therefore, is not simply equal to the posterior of any of the preceding agents, rather it is a function of the average private signal from the previous period. Denoting this prior belief by $\delta_{t|t-1}$, we obtain:

$$\delta_{1|0} = a + \widetilde{w}_0 \times \rho \bar{s}_0 + (1 - \widetilde{w}_0) \times \rho s_0 \tag{A.29}$$

where $\widetilde{w}_0 = \frac{n\sigma_\eta^{-2}}{\sigma_u^{-2} + n\sigma_\eta^{-2}}$ and $\bar{s}_0$ is the equally-weighted average of private signals at time 0. The variance of the prior belief of agents in period t is denoted as $P_{t|t-1}$. This

[6] I assume that same-period agents do not observe each other's actions and thus cannot use each other's signals to inform their decisions.

[7] For the simulation exercise, I focus solely on transaction prices and, therefore, disregard the quote setting procedure. As a result, the only relevant type of private signal is the one embedded in final prices, i.e., $s_{i,t}^p$. For ease of exposition, I hereafter denote this signal simply by s_t^i.

can be computed recursively as follows:

$$P_{1|0} = \rho^2 \widetilde{P}_{0|0} + \sigma_\varepsilon^2 \tag{A.30}$$

where $\widetilde{P}_{0|0} = \widetilde{w}_0^2 \times \frac{1}{n}\sigma_\eta^2 + (1-\widetilde{w}_0)^2 \times \sigma_u^2$ in order to adjust for the fact that there are n private signals coming from period $t-1$. Each period agent i forms his posterior belief and the price by mixing the above prior and his private signal:

$$p_1^i = \delta_{1|1}^i = w_1 \times s_1^i + (1-w_1) \times \delta_{1|0} \tag{A.31}$$

where $w_1 = P_{1|0} \times (P_{1|0} + \sigma_\eta^2)^{-1}$ is the Kalman gain. From here onward, we can define the recursion through which agents form and update their beliefs in a sequential way. The prior beliefs are computed by adjusting for the number of observations from the previous period:

$$\delta_{t|t-1} = a + \widetilde{w}_{t-1} \times \rho \bar{s}_{t-1} + (1-\widetilde{w}_{t-1}) \times \rho \delta_{t-1|t-2} \tag{A.32}$$

where $\widetilde{w}_{t-1} = P_{t-1|t-2} \times (P_{t-1|t-2} + \frac{1}{n}\sigma_\eta^2)^{-1}$. The variance of this prior can be computed as follows:

$$P_{t|t-1} = \rho^2 P_{t-1|t-1} + \sigma_\varepsilon^2 = \rho^2 [P_{t-1|t-2} - P_{t-1|t-2}^2 (P_{t-1|t-2} + \frac{1}{n}\sigma_\eta^2)^{-1}] + \sigma_\varepsilon^2 \tag{A.33}$$

Finally, agent i forms his posterior belief updated for his private signal and sets the price accordingly:

$$p_t^i = \delta_{t|t}^i = w_t \times s_t^i + (1-w_t) \times \delta_{t|t-1} \tag{A.34}$$

where $w_t = P_{t|t-1}(P_{t|t-1} + \sigma_\eta^2)^{-1}$. Plugging in the expressions for the prior beliefs recursively, we can outline the way that prices depend on all past signals:

$$\begin{aligned} p_t^i = \delta_{t|t}^i = & w_t s_t^i + (1-w_t)[a + \widetilde{w}_{t-1}\rho \bar{s}_{t-1}] + \\ & (1-w_t)(1-\widetilde{w}_{t-1})[\rho a + \widetilde{w}_{t-2}\rho^2 \bar{s}_{t-2}] + ... + \\ & (1-w_t)(1-\widetilde{w}_{t-1})(1-\widetilde{w}_{t-2})...(1-\widetilde{w}_1)[\rho^{t-1} a + \widetilde{w}_0 \rho^t \bar{s}_0] + \\ & (1-w_t)(1-\widetilde{w}_{t-1})(1-\widetilde{w}_{t-2})...(1-\widetilde{w}_1)(1-\widetilde{w}_0)\rho^t s_0 \end{aligned} \tag{A.35}$$

The difference between rational and naïve sellers is that naïve learners treat all past prices as independent signals by failing to account for the fact that previous sellers have similarly formed their beliefs by looking at yet earlier prices. They, therefore, assign the same weights as the rational agents but directly to the observed prices as opposed to the signals extracted:

$$\begin{aligned}\widetilde{p}_t^i = \widetilde{\delta}_{t|t}^i = & w_t s_t^i + (1-w_t)[a + \widetilde{w}_{t-1}\rho\bar{\bar{p}}_{t-1}] + \\ & (1-w_t)(1-\widetilde{w}_{t-1})[\rho a + \widetilde{w}_{t-2}\rho^2\bar{\bar{p}}_{t-2}] + ... + \\ & (1-w_t)(1-\widetilde{w}_{t-1})(1-\widetilde{w}_{t-2})...(1-\widetilde{w}_1)[\rho^{t-1}a + \widetilde{w}_0\rho^t\bar{p}_0] + \\ & (1-w_t)(1-\widetilde{w}_{t-1})(1-\widetilde{w}_{t-2})...(1-\widetilde{w}_1)(1-\widetilde{w}_0)\rho^t s_0\end{aligned} \tag{A.36}$$

This leads them to overweight old signals at the expense of more recent information since these have also been incorporated into the prices set by more recent sellers. To determine the magnitude of the effect of naïve learning given the estimates obtained in the empirical analysis, I simulate a market with the above characteristics and compare the impact of various information shocks on prices in the rational and naïve settings.

A.4 Additional Tables and Figures

Table A13: Effect of Transaction Prices on Quotes - Sample Refinements

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set in the period following the price data publication date. Columns (1)-(2) present the results for the sample of listings that excludes those posted exactly on publishing dates; columns (3)-(4) restrict the sample to listings posted in the two weeks around the publishing date and; columns (5)-(6) limit the number of comparables to no more than 30 per listing. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns. Fixed-effects included are: listing month-year dummies in all columns, and; transaction ID dummies in columns (2), (4) and (6). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	No quotes on pub dates		Within 7 days of pub date		Less than 30 comps	
	(1)	(2)	(3)	(4)	(5)	(6)
Price × Treated	0.0041***	0.0031***	0.0056***	0.0038**	0.0035**	0.0033***
	(0.0016)	(0.0011)	(0.0021)	(0.0018)	(0.0015)	(0.0011)
Price	0.8400***		0.8457***		0.8387***	
	(0.0023)		(0.0037)		(0.0023)	
Treated	-0.0501***	-0.0351**	-0.0637**	-0.0406*	-0.0423**	-0.0373***
	(0.0193)	(0.0138)	(0.0254)	(0.0210)	(0.0183)	(0.0132)
Controls						
Price × Time distance	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-Effects						
Month-year	Yes	Yes	Yes	Yes	Yes	Yes
Transaction ID	No	Yes	No	Yes	No	Yes
Observations	7,075,069	7,075,069	2,908,410	2,908,410	7,421,440	7,421,440
R^2	0.7052	0.8696	0.7058	0.8800	0.7056	0.8694
Within R^2	0.7033	0.0000	0.7036	0.0001	0.7038	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A14: Effect of Transaction Prices on Quotes - Sensitivity to Publication Days

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set in the period following the price data publication date. In columns (1)-(2) dummies for the day-of-the-week of the publishing date are also interacted with price p_j and $Treated_i$ (for brevity I report only the subset of relevant coefficients). Columns (3)-(4) restrict the sample to cases where the publishing date occurred at the end of the month, while columns (5)-(6) to cases where it fell at the beginning of the next month. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns. Fixed-effects included are: listing month-year dummies in all columns, and; transaction ID dummies in columns (2), (4) and (6). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	Pub day of week		Pub day at end of month		Pub day at beginning of month	
	(1)	(2)	(3)	(4)	(5)	(6)
Price × Treated	0.0057***	0.0034***	0.0040**	0.0033***	0.0107*	0.0057
	(0.0019)	(0.0011)	(0.0016)	(0.0011)	(0.0063)	(0.0045)
Price × Treated × Monday	-0.0019	0.0007*				
	(0.0034)	(0.0004)				
Price × Treated × Tuesday	-0.0070**	−0.0000				
	(0.0030)	(0.0003)				
Price × Treated × Wednesday	0.0040	-0.0004				
	(0.0030)	(0.0003)				
Price × Treated × Thursday	-0.0048*	0.0000				
	(0.0029)	(0.0003)				
Price	0.8402***		0.8411***		0.8225***	
	(0.0023)		(0.0023)		(0.0097)	
Treated	-0.0669***	-0.0383***	-0.0483**	-0.0365***	-0.1131	-0.0537
	(0.0233)	(0.0134)	(0.0190)	(0.0138)	(0.0789)	(0.0576)
Controls						
Price × Time distance	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-Effects						
Month-year	Yes	Yes	Yes	Yes	Yes	Yes
Transaction ID	No	Yes	No	Yes	No	Yes
Observations	7,466,950	7,466,950	6,959,170	6,959,170	507,780	507,780
R^2	0.7050	0.8689	0.7072	0.8701	0.6735	0.8525
Within R^2	0.7032	0.0000	0.7055	0.0000	0.6697	0.0002

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A15: Effect of Transaction Prices on Quotes - Existing Houses and by Price Range

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set in the period following the price data publication date. Columns (1)-(2) present the results for the sample of listings that excludes newly-built properties; columns (3)-(4) restrict the sample to quotes that are below median, and; columns (5)-(6) to quotes above median. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns. Fixed-effects included are: listing month-year dummies in all columns, and; transaction ID dummies in columns (2), (4) and (6). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	Existing houses only		Price below median		Price above median	
	(1)	(2)	(3)	(4)	(5)	(6)
Price × Treated	0.0045***	0.0029***	0.0046*	0.0040**	0.0081**	0.0028
	(0.0015)	(0.0011)	(0.0025)	(0.0017)	(0.0032)	(0.0020)
Price	0.8431***		0.5467***		0.5934***	
	(0.0023)		(0.0034)		(0.0048)	
Treated	-0.0545***	-0.0333**	-0.0547*	-0.0468**	-0.1000**	-0.0306
	(0.0186)	(0.0134)	(0.0294)	(0.0206)	(0.0407)	(0.0248)
Controls						
Price × Time distance	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-Effects						
Month-year	Yes	Yes	Yes	Yes	Yes	Yes
Transaction ID	No	Yes	No	Yes	No	Yes
Observations	7,225,115	7,225,115	3,919,114	3,919,114	3,089,563	3,089,563
R^2	0.7093	0.8719	0.3981	0.7039	0.4363	0.7309
Within R^2	0.7076	0.0000	0.3958	0.0000	0.4327	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A16: Effect of Transaction Prices on Quotes - Controlling for Listing Agent

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set in the period following the price data publication date. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (3). Column (3) instead includes time distance (measured in weeks) dummies and their interaction with log price. Fixed-effects included are: real-estate agent dummies in all columns; listing month-year dummies in columns (2)-(4), and; transaction ID dummies in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price × Treated	0.0037***	0.0042***	0.0039***	0.0023**
	(0.0013)	(0.0013)	(0.0013)	(0.0010)
Price	0.5581***	0.5455***	0.5418***	
	(0.0019)	(0.0019)	(0.0026)	
Treated	-0.0444***	-0.0524***	-0.0481***	-0.0240**
	(0.0154)	(0.0154)	(0.0153)	(0.0115)
Controls				
Price x Time distance	Yes	Yes	No	Yes
Price x Time distance dummies	No	No	Yes	No
Fixed-Effects				
Agent ID	Yes	Yes	Yes	Yes
Month-year	No	Yes	Yes	Yes
Transaction ID	No	No	No	Yes
Observations	7,443,824	7,443,824	7,443,824	7,443,824
R^2	0.7899	0.7925	0.7925	0.8970
Within R^2	0.3690	0.3540	0.3540	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A17: Effect of Transaction Prices on Quotes - Controlling for Reference Price

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set in the period following the price data publication date. Controls for the purchase price of the listed property are included in all columns. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (3). Column (3) instead includes time distance (measured in weeks) dummies and their interaction with log price. Additional fixed-effects include: listing month-year dummies in all columns but (1), and transaction ID dummies in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price × Treated	0.0035***	0.0040***	0.0036***	0.0030***
	(0.0013)	(0.0013)	(0.0013)	(0.0011)
Price	0.5173***	0.5108***	0.5089***	
	(0.0020)	(0.0020)	(0.0028)	
Treated	-0.0431***	-0.0500***	-0.0447***	-0.0343***
	(0.0163)	(0.0162)	(0.0161)	(0.0130)
Purchase Price	0.5009***	0.5081***	0.5081***	0.2276***
	(0.0014)	(0.0014)	(0.0014)	(0.0013)
Controls				
Price x Time distance	Yes	Yes	No	Yes
Price x Time distance dummies	No	No	Yes	No
Fixed-effects				
Month-year	No	Yes	Yes	Yes
Transaction ID	No	No	No	Yes
Observations	7,457,256	7,457,256	7,457,256	7,457,256
R^2	0.7752	0.7768	0.7768	0.8759
Within R^2	–	0.7755	0.7755	0.0544

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A18: Effect of Transaction Prices on Quotes - Positive Shocks

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set/updated in the period following the most recent price publication date. The sample of sold properties includes only those whose transaction price is above the predicted price obtained using a hedonic regression of the form: $log(p_j) = FE + \varepsilon_j$, where the fixed effects refer to the month-year of the transaction, location measured as the address outcode, property type (detached, semi-detached, terraced house or a flat) and number of rooms. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (3). Column (3) instead includes time distance (measured in weeks) dummies and their interaction with log price. Additional fixed-effects include: listing month-year dummies in all columns but (1), and transaction ID dummies in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price × Treated	0.0035**	0.0034**	0.0029*	0.0028***
	(0.0015)	(0.0015)	(0.0015)	(0.0010)
Price	0.9137***	0.9087***	0.9021***	
	(0.0025)	(0.0025)	(0.0032)	
Treated	-0.0410**	-0.0404**	-0.0339*	-0.0319**
	(0.0187)	(0.0186)	(0.0184)	(0.0127)
Controls				
Price × Time distance	Yes	Yes	No	Yes
Price × Time distance dummies	No	No	Yes	No
Fixed-Effects				
Month-year	No	Yes	Yes	Yes
Transaction ID	No	No	No	Yes
Observations	5,464,143	5,464,143	5,464,143	5,464,143
R^2	0.7696	0.7711	0.7711	0.8723
Within R^2	–	0.7638	0.7638	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A19: Effect of Transaction Prices on Quotes - Negative Shocks

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set/updated in the period following the most recent price publication date. The sample of sold properties includes only those whose transaction price is below the predicted price obtained using a hedonic regression of the form: $log(p_j) = FE + \varepsilon_j$, where the fixed effects refer to the month-year of the transaction, location measured as the address outcode, property type (detached, semi-detached, terraced house or a flat) and number of rooms. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (3). Column (3) instead includes time distance (measured in weeks) dummies and their interaction with log price. Additional fixed-effects include: listing month-year dummies in all columns but (1), and transaction ID dummies in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price × Treated	0.0049***	0.0054***	0.0047***	0.0035***
	(0.0015)	(0.0015)	(0.0015)	(0.0010)
Price	0.9188***	0.9205***	0.9231***	
	(0.0025)	(0.0025)	(0.0032)	
Treated	-0.0595***	-0.0642***	-0.0552***	-0.0404***
	(0.0185)	(0.0185)	(0.0183)	(0.0124)
Controls				
Price × Time distance	Yes	Yes	No	Yes
Price × Time distance dummies	No	No	Yes	No
Fixed-Effects				
Month-year	No	Yes	Yes	Yes
Transaction ID	No	No	No	Yes
Observations	5,946,101	5,946,101	5,946,101	5,946,101
R^2	0.7507	0.7517	0.7517	0.8665
Within R^2	–	0.7509	0.7509	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A20: Effect of Transaction Prices on Quotes - Large Positive Shocks

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set/updated in the period following the most recent price publication date. The sample of sold properties includes only those for which the log transaction price is by more than 0.20 greater than the predicted price obtained using a hedonic regression of the form: $log(p_j) = FE + \varepsilon_j$, where the fixed effects refer to the month-year of the transaction, location measured as the address outcode, property type (detached, semi-detached, terraced house or a flat) and number of rooms. This cutoff roughly corresponds to the top 20% of the residual price distribution. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (3). Column (3) instead includes time distance (measured in weeks) dummies and their interaction with log price. Additional fixed-effects include: listing month-year dummies in all columns but (1), and transaction ID dummies in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price × Treated	0.0041**	0.0036*	0.0026	0.0029**
	(0.0021)	(0.0021)	(0.0020)	(0.0013)
Price	0.9192***	0.9140***	0.9069***	
	(0.0036)	(0.0035)	(0.0044)	
Treated	-0.0481*	-0.0418	-0.0299	-0.0329**
	(0.0258)	(0.0257)	(0.0254)	(0.0161)
Controls				
Price × Time distance	Yes	Yes	No	Yes
Price × Time distance dummies	No	No	Yes	No
Fixed-Effects				
Month-year	No	Yes	Yes	Yes
Transaction ID	No	No	No	Yes
Observations	2,208,678	2,208,678	2,208,678	2,208,678
R^2	0.7813	0.7827	0.7828	0.8762
Within R^2	–	0.7748	0.7748	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A21: Effect of Transaction Prices on Quotes - Large Negative Shocks

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set/updated in the period following the most recent price publication date. The sample of sold properties includes only those for which the log transaction price is by more than 0.20 lower than the predicted price obtained using a hedonic regression of the form: $log(p_j) = FE + \varepsilon_j$, where the fixed effects refer to the month-year of the transaction, location measured as the address outcode, property type (detached, semi-detached, terraced house or a flat) and number of rooms. This cutoff roughly corresponds to the bottom 20% of the residual price distribution. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (3). Column (3) instead includes time distance (measured in weeks) dummies and their interaction with log price. Additional fixed-effects include: listing month-year dummies in all columns but (1), and transaction ID dummies in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price × Treated	0.0051**	0.0057***	0.0050**	0.0043***
	(0.0020)	(0.0020)	(0.0020)	(0.0013)
Price	0.9275***	0.9294***	0.9356***	
	(0.0034)	(0.0034)	(0.0045)	
Treated	-0.0622**	-0.0672***	-0.0571**	-0.0489***
	(0.0242)	(0.0241)	(0.0240)	(0.0154)
Controls				
Price × Time distance	Yes	Yes	No	Yes
Price × Time distance dummies	No	No	Yes	No
Fixed-Effects				
Month-year	No	Yes	Yes	Yes
Transaction ID	No	No	No	Yes
Observations	2,644,044	2,644,044	2,644,044	2,644,044
R^2	0.7630	0.7643	0.7643	0.8704
Within R^2	–	0.7619	0.7618	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A22: Effect of Transaction Prices on Quotes by Sales Volume

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set/updated in the period following the most recent price publication date. The sample is split based on sales activity: the first two columns show the effect on quotes of transaction prices in months and regions with sales volume below the sample average, while the last two columns show the results for the subsample with above-average sales volume. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns. Additional fixed-effects include: listing month-year dummies in all columns, and; transaction ID dummies in columns (2) and (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	Low Volume		High Volume	
	(1)	(2)	(3)	(4)
Price × Treated	0.0054**	0.0037**	0.0034**	0.0027***
	(0.0022)	(0.0015)	(0.0015)	(0.0010)
Price	0.8358***		0.8392***	
	(0.0036)		(0.0023)	
Treated	-0.0648**	-0.0417**	-0.0412**	-0.0312**
	(0.0267)	(0.0179)	(0.0179)	(0.0122)
Controls				
Price × Time distance	Yes	Yes	Yes	Yes
Fixed-effects				
Month-year	Yes	Yes	Yes	Yes
Transaction ID	No	Yes	No	Yes
Observations	3,461,461	3,461,461	7,948,783	7,948,783
R^2	0.7212	0.8781	0.6974	0.8631
Within R^2	0.7082	0.0000	0.6956	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A23: Effect of Transaction Prices on Quotes by Excess Volume at the Local Level

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set/updated in the period following the most recent price publication date. The sample is split based on relative sales activity: the first two columns show the effect on quotes of transaction prices in months with below-average sales volume for the local area, while the last two columns show the results for months with above-average sales volume. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns. Additional fixed-effects include: listing month-year dummies in all columns, and; transaction ID dummies in columns (2) and (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	Low Relative Volume		High Relative Volume	
	(1)	(2)	(3)	(4)
Price × Treated	0.0058***	0.0041***	0.0024	0.0021*
	(0.0017)	(0.0012)	(0.0018)	(0.0012)
Price	0.8216***		0.8341***	
	(0.0028)		(0.0027)	
Treated	-0.0712***	-0.0463***	-0.0282	-0.0240
	(0.0208)	(0.0145)	(0.0215)	(0.0148)
Controls				
Price × Time distance	Yes	Yes	Yes	Yes
Fixed-effects				
Month-year	Yes	Yes	Yes	Yes
Transaction ID	No	Yes	No	Yes
Observations	5,618,545	5,618,545	5,791,699	5,791,699
R^2	0.7081	0.8679	0.7106	0.8700
Within R^2	0.6730	0.0000	0.6945	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A24: Effect of Transaction Prices on Quotes - within Listing Price Updates

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set/updated in the period following the most recent price publication date. Only listings that have at least one treated and one untreated quote are included. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (5). Column (5) instead includes time distance (measured in weeks) dummies and their interaction with log price. Listing ID fixed effects are included in specifications (3)-(7). Additional fixed-effects include: listing month-year dummies in all columns but (3); transaction ID dummies in columns (2), (6) and (7), and; order of quote update dummies in column (7). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Price × Treated	0.0028**	0.0037***	0.0013***	0.0011***	0.0011***	0.0033***	0.0021***
	(0.0014)	(0.0012)	(0.0003)	(0.0002)	(0.0002)	(0.0003)	(0.0002)
Price	0.8353***		-0.0024***	0.0003	0.0003		
	(0.0026)		(0.0003)	(0.0003)	(0.0004)		
Treated	-0.0362**	-0.0466***	-0.0173***	-0.0104***	-0.0106***	-0.0400***	-0.0221***
	(0.0165)	(0.0139)	(0.0034)	(0.0030)	(0.0030)	(0.0037)	(0.0030)
Controls							
Price × Time distance	Yes	Yes	Yes	Yes	No	Yes	Yes
Price × Time distance dummies	No	No	No	No	Yes	No	No
Fixed-Effects							
Listing ID	No	No	Yes	Yes	Yes	Yes	Yes
Month-year	Yes	Yes	No	Yes	Yes	Yes	Yes
Transaction ID	No	Yes	No	No	No	Yes	Yes
Order of price update	No	No	No	No	No	No	Yes
Observations	3,817,123	3,817,123	3,817,123	3,817,123	3,817,123	3,817,123	3,817,123
R^2	0.7136	0.9203	0.9962	0.9971	0.9971	0.9989	0.9991
Within R^2	0.7061	0.0001	0.0006	0.0008	0.0009	0.0131	0.0011

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A25: Effect of Transaction Prices on First Quotes - Before vs After March 2012

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Post March 2012}_i + \beta_2 \times log(p_j) \times \textit{Treated}_i + \beta_3 \times log(p_j) \times \textit{Treated}_i \times \textit{Post March 2012}_i + \gamma_1 \textit{Post March 2012}_i + \gamma_2 \textit{Treated}_i + \gamma_3 \times \textit{Treated}_i \times \textit{Post March 2012}_i + \textit{Controls} + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month, *Post March 2012*$_i$ is a dummy that equals one for listings published starting from March 2012 and *Treated*$_i$ is a dummy that turns on when the listing price has been set in the period following the most recent price publication date. Only the initial quotes of listings are included. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (3). Column (3) instead includes time distance (measured in weeks) dummies and their interaction with log price. Fixed-effects included are: listing month-year dummies in all columns but (1) and; transaction ID dummies in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price × Treated × Post March 2012	0.0094***	0.0090***	0.0089***	0.0033
	(0.0027)	(0.0027)	(0.0027)	(0.0024)
Price × Treated	-0.0060**	-0.0056**	-0.0060**	0.0001
	(0.0025)	(0.0025)	(0.0025)	(0.0021)
Price × Post March 2012	0.0460***	0.0461***	0.0461***	
	(0.0023)	(0.0023)	(0.0023)	
Price	0.7917***	0.7917***	0.7901***	
	(0.0027)	(0.0027)	(0.0033)	
Treated	0.0667**	0.0625**	0.0687**	-0.0071
	(0.0302)	(0.0301)	(0.0301)	(0.0248)
Post March 2012	-0.5650***	-0.5862***	-0.5865***	
	(0.0270)	(0.0278)	(0.0278)	
Treated × Post March 2012	-0.1083***	-0.1033***	-0.1030***	-0.0314
	(0.0321)	(0.0320)	(0.0320)	(0.0282)
Controls				
Price × Time distance	Yes	Yes	No	Yes
Price × Time distance dummies	No	No	Yes	No
Fixed-Effects				
Month-year	No	Yes	Yes	Yes
Transaction ID	No	No	No	Yes
Observations	10,585,043	10,585,043	10,585,043	10,585,043
R^2	0.6773	0.6782	0.6782	0.8561
Within R^2	–	0.6756	0.6756	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A26: Effect of Transaction Prices on First Quotes Around Placebo Publishing Dates

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times Treated_i + \gamma Treated_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $Treated_i$ is a dummy that turns on when the listing price has been set in the week before (first four columns) or one week after (last four columns) the closest price publication date. Only the initial quotes of listings posted in the two weeks surrounding the placebo publishing dates are included. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (3) and (7). Columns (3) and (7) instead include time distance (measured in weeks) dummies and their interaction with log price. Additional fixed-effects include: listing month-year dummies in all columns but (1) and (5) and; transaction ID dummies in columns (4) and (8). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	7 days before				7 days after			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Price × Treated	0.0021	0.0019	0.0012	0.0012	0.0007	0.0004	-0.0000	-0.0013
	(0.0021)	(0.0021)	(0.0021)	(0.0018)	(0.0021)	(0.0021)	(0.0021)	(0.0018)
Price	0.8430***	0.8428***	0.8410***		0.8406***	0.8409***	0.8411***	
	(0.0031)	(0.0031)	(0.0035)		(0.0042)	(0.0042)	(0.0040)	
Treated	-0.0279	-0.0266	-0.0169	-0.0110	-0.0127	-0.0076	-0.0014	0.0139
	(0.0255)	(0.0254)	(0.0255)	(0.0212)	(0.0254)	(0.0253)	(0.0254)	(0.0214)
Controls								
Price × Time distance	Yes	Yes	No	Yes	Yes	Yes	No	Yes
Price × Time distance dummies	No	No	Yes	No	No	No	Yes	No
Fixed-Effects								
Month-year	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Transaction ID	No	No	No	Yes	No	No	No	Yes
Observations	2,831,845	2,831,845	2,831,845	2,831,845	2,900,284	2,900,284	2,900,284	2,900,284
R^2	0.7021	0.7029	0.7029	0.8785	0.7068	0.7077	0.7077	0.8817
Within R^2	–	0.7006	0.7005	0.0000	–	0.7055	0.7055	0.0000

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A27: Indirect Price Effects Through Intermediate Listings - Full Sample

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Comps in w-1}_i + \beta_2 \times log(p_j) \times \textit{Comps in w+1}_i + \beta_3 \times log(p_j) \times \textit{Comps in all weeks}_i + \gamma_1 \times \textit{Comps in w-1}_i + \gamma_2 \times \textit{Comps in w+1}_i + \gamma_3 \times \textit{Comps in all weeks}_i + \textit{Controls} + \varepsilon_i$, where q_i is the listed price for property i which has been set/updated during the second week following the most recent price data publication date, p_j is the transaction price for a comparable property j sold in the previous month, *Comps in w-1*$_i$ is a dummy that turns on if the listing-transaction pair have at least one other comparable match in the week before the price publication date but none in the week after, *Comps in w+1*$_i$ is a dummy for pairs that have at least one match in the week after but none in the week before, and *Comps in all weeks*$_i$ is a dummy for listing-transaction pairs with at least one match in each week. The table also reports the p-values of linear hypothesis tests of the difference in the price coefficients. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1), while controls for the number of comparable quotes in the current and each of the two previous weeks are included in columns (3)-(4). Listing month-year fixed effects are included in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price	0.8242***	0.8231***	0.8231***	0.8226***
	(0.0020)	(0.0044)	(0.0044)	(0.0043)
Price × Comps in week -1$^{(1)}$	0.0089***	0.0089***	0.0089***	0.0083***
	(0.0032)	(0.0032)	(0.0032)	(0.0032)
Price × Comps in week +1$^{(2)}$	0.0176***	0.0176***	0.0177***	0.0169***
	(0.0029)	(0.0029)	(0.0029)	(0.0029)
Price × Comps in all weeks$^{(3)}$	0.0239***	0.0240***	0.0238***	0.0232***
	(0.0025)	(0.0025)	(0.0025)	(0.0025)
Comps in week -1	-0.1140***	-0.1138***	-0.1151***	-0.1071***
	(0.0392)	(0.0392)	(0.0392)	(0.0391)
Comps in week +1	-0.2130***	-0.2130***	-0.2150***	-0.2067***
	(0.0356)	(0.0356)	(0.0356)	(0.0356)
Comps in all weeks	-0.2854***	-0.2858***	-0.2890***	-0.2837***
	(0.0303)	(0.0303)	(0.0303)	(0.0303)
(2)−(1)	0.0087***	0.0087***	0.0088***	0.0086***
p-value	(0.0082)	(0.0083)	(0.0084)	(0.0091)
(3)−(2)	0.0063**	0.0064**	0.0061**	0.0063**
p-value	(0.0154)	(0.0147)	(0.0179)	(0.0158)
(3)−(1)	0.0150***	0.0151***	0.0149***	0.0149***
p-value	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Controls				
Price x Time distance	No	Yes	Yes	Yes
Nb. of comps per week	No	No	Yes	Yes
Fixed-Effects				
Month-year	No	No	No	Yes
Observations	2,932,258	2,932,258	2,932,258	2,932,258
R^2	0.7067	0.7068	0.7068	0.7077
Within R^2	–	–	–	0.7050

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A28: Indirect Price Effects Through Intermediate Listings - Before March 2012

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Comps in w-1}_i + \beta_2 \times log(p_j) \times \textit{Comps in w+1}_i + \beta_3 \times log(p_j) \times \textit{Comps in all weeks}_i + \gamma_1 \times \textit{Comps in w-1}_i + \gamma_2 \times \textit{Comps in w+1}_i + \gamma_3 \times \textit{Comps in all weeks}_i + \textit{Controls} + \varepsilon_i$, where q_i is the listed price for property i which has been set/updated during the second week following the most recent (placebo) price data publication date, p_j is the transaction price for a comparable property j sold in the previous month, $\textit{Comps in w-1}_i$ is a dummy that turns on if the listing-transaction pair have at least one other comparable match in the week before the price publication date but none in the week after, $\textit{Comps in w+1}_i$ is a dummy for pairs that have at least one match in the week after but none in the week before, and $\textit{Comps in all weeks}_i$ is a dummy for listing-transaction pairs with at least one match in each week. The regressions are estimated using data from the sample before March 2012. The table also reports the p-values of linear hypothesis tests of the difference in the price coefficients. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1), while controls for the number of comparable quotes in the current and each of the two previous weeks are included in columns (3)-(4). Listing month-year fixed effects are included in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price	0.7656***	0.7510***	0.7504***	0.7548***
	(0.0037)	(0.0085)	(0.0085)	(0.0085)
Price × Comps in week -1$^{(1)}$	0.0257***	0.0257***	0.0258***	0.0229***
	(0.0064)	(0.0064)	(0.0064)	(0.0064)
Price × Comps in week +1$^{(2)}$	0.0303***	0.0303***	0.0306***	0.0267***
	(0.0056)	(0.0056)	(0.0056)	(0.0056)
Price × Comps in all weeks$^{(3)}$	0.0216***	0.0215***	0.0180***	0.0135***
	(0.0047)	(0.0047)	(0.0047)	(0.0047)
Comps in week -1	-0.3179***	-0.3178***	-0.3240***	-0.2814***
	(0.0762)	(0.0762)	(0.0762)	(0.0759)
Comps in week +1	-0.3700***	-0.3696***	-0.3813***	-0.3282***
	(0.0673)	(0.0673)	(0.0673)	(0.0671)
Comps in all weeks	-0.2494***	-0.2488***	-0.2345***	-0.1697***
	(0.0565)	(0.0565)	(0.0565)	(0.0565)
(2)−(1)	0.0046	0.0046	0.0048	0.0038
p-value	(0.4899)	(0.4922)	(0.4782)	(0.5784)
(3)−(2)	-0.0087*	-0.0088*	-0.0126**	-0.0132**
p-value	(0.0901)	(0.0893)	(0.0146)	(0.0102)
(3)−(1)	-0.0041	-0.0042	-0.0078	-0.0094
p-value	(0.4910)	(0.4859)	(0.1883)	(0.1107)
Controls				
Price x Time distance	No	Yes	Yes	Yes
Nb. of comps per week	No	No	Yes	Yes
Fixed-Effects				
Month-year	No	No	No	Yes
Observations	1,355,042	1,355,042	1,355,042	1,355,042
R^2	0.6014	0.6014	0.6024	0.6044
Within R^2	–	–	–	0.6018

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A29: Indirect Price Effects - Controlling for Reference Price

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Comps in w-1}_i + \beta_2 \times log(p_j) \times \textit{Comps in w+1}_i + \beta_3 \times log(p_j) \times \textit{Comps in all weeks}_i + \gamma_1 \times \textit{Comps in w-1}_i + \gamma_2 \times \textit{Comps in w+1}_i + \gamma_3 \times \textit{Comps in all weeks}_i + \textit{Controls} + \varepsilon_i$, where q_i is the listed price for property i which has been set/updated during the second week following the most recent price data publication date, p_j is the transaction price for a comparable property j sold in the previous month, *Comps in w-1*$_i$ is a dummy that turns on if the listing-transaction pair have at least one other comparable match in the week before the price publication date but none in the week after, *Comps in w+1*$_i$ is a dummy for pairs that have at least one match in the week after but none in the week before, and *Comps in all weeks*$_i$ is a dummy for listing-transaction pairs with at least one match in each week. Controls for the initial purchase price of the listed property are included in all columns, controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1), while controls for the number of comparable quotes in the current and each of the two previous weeks are included in columns (3)-(4). Listing month-year fixed effects are included in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price	0.5137***	0.5155***	0.5153***	0.5088***
	(0.0022)	(0.0038)	(0.0038)	(0.0037)
Price × Comps in week -1	0.0098***	0.0098***	0.0098***	0.0090***
	(0.0029)	(0.0029)	(0.0029)	(0.0029)
Price × Comps in week +1	0.0116***	0.0116***	0.0117***	0.0111***
	(0.0027)	(0.0027)	(0.0027)	(0.0027)
Price × Comps in all weeks	0.0155***	0.0155***	0.0157***	0.0150***
	(0.0023)	(0.0023)	(0.0023)	(0.0023)
Comps in week -1	-0.1217***	-0.1214***	-0.1205***	-0.1106***
	(0.0355)	(0.0355)	(0.0355)	(0.0353)
Comps in week +1	-0.1409***	-0.1409***	-0.1406***	-0.1330***
	(0.0326)	(0.0326)	(0.0326)	(0.0325)
Comps in all weeks	-0.1805***	-0.1805***	-0.1815***	-0.1762***
	(0.0277)	(0.0277)	(0.0277)	(0.0276)
Purchase Price	0.4884***	0.4884***	0.4886***	0.4941***
	(0.0020)	(0.0020)	(0.0020)	(0.0020)
Controls				
Price x Time distance	No	Yes	Yes	Yes
Nb. of comps per week	No	No	Yes	Yes
Fixed-effects				
Month-year	No	No	No	Yes
Observations	2,912,960	2,912,960	2,912,960	2,912,960
R^2	0.7765	0.7765	0.7765	0.7783
Within R^2	–	–	–	0.7763

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A30: Indirect Price Effects - Positive Shocks

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Comps in w-1}_i + \beta_2 \times log(p_j) \times \textit{Comps in w+1}_i + \beta_3 \times log(p_j) \times \textit{Comps in all weeks}_i + \gamma_1 \times \textit{Comps in w-1}_i + \gamma_2 \times \textit{Comps in w+1}_i + \gamma_3 \times \textit{Comps in all weeks}_i + \textit{Controls} + \varepsilon_i$, where q_i is the listed price for property i which has been set/updated during the second week following the most recent price data publication date, p_j is the transaction price for a comparable property j sold in the previous month, $\textit{Comps in w-1}_i$ is a dummy that turns on if the listing-transaction pair have at least one other comparable match in the week before the price publication date but none in the week after, $\textit{Comps in w+1}_i$ is a dummy for pairs that have at least one match in the week after but none in the week before, and $\textit{Comps in all weeks}_i$ is a dummy for listing-transaction pairs with at least one match in each week. The sample of sold properties includes only those whose transaction price is above the predicted price obtained using a hedonic regression of the form: $log(p_j) = FE + \varepsilon_j$, where the fixed effects refer to the month-year of the transaction, location measured as the address outcode, property type (detached, semi-detached, terraced house or a flat) and number of rooms. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1), while controls for the number of comparable quotes in the current and each of the two previous weeks are included in columns (3)-(4). Listing month-year fixed effects are included in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price	0.8926***	0.8871***	0.8868***	0.8811***
	(0.0026)	(0.0058)	(0.0058)	(0.0058)
Price × Comps in week -1	0.0144***	0.0145***	0.0143***	0.0134***
	(0.0041)	(0.0041)	(0.0041)	(0.0041)
Price × Comps in week +1	0.0199***	0.0199***	0.0199***	0.0192***
	(0.0037)	(0.0036)	(0.0036)	(0.0036)
Price × Comps in all weeks	0.0304***	0.0305***	0.0314***	0.0299***
	(0.0031)	(0.0031)	(0.0031)	(0.0031)
Comps in week -1	-0.1847***	-0.1853***	-0.1814***	-0.1709***
	(0.0501)	(0.0501)	(0.0500)	(0.0499)
Comps in week +1	-0.2498***	-0.2499***	-0.2475***	-0.2391***
	(0.0448)	(0.0448)	(0.0447)	(0.0447)
Comps in all weeks	-0.3826***	-0.3832***	-0.3850***	-0.3715***
	(0.0386)	(0.0386)	(0.0385)	(0.0385)
Controls				
Price x Time distance	No	Yes	Yes	Yes
Nb. of comps per week	No	No	Yes	Yes
Fixed-Effects				
Month-year	No	No	No	Yes
Observations	1,409,595	1,409,595	1,409,595	1,409,595
R^2	0.7688	0.7688	0.7689	0.7704
Within R^2	–	–	–	0.7631

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A31: Indirect Price Effects - Negative Shocks

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Comps in w-1}_i + \beta_2 \times log(p_j) \times \textit{Comps in w+1}_i + \beta_3 \times log(p_j) \times \textit{Comps in all weeks}_i + \gamma_1 \times \textit{Comps in w-1}_i + \gamma_2 \times \textit{Comps in w+1}_i + \gamma_3 \times \textit{Comps in all weeks}_i + \textit{Controls} + \varepsilon_i$, where q_i is the listed price for property i which has been set/updated during the second week following the most recent price data publication date, p_j is the transaction price for a comparable property j sold in the previous month, *Comps in w-1*$_i$ is a dummy that turns on if the listing-transaction pair have at least one other comparable match in the week before the price publication date but none in the week after, *Comps in w+1*$_i$ is a dummy for pairs that have at least one match in the week after but none in the week before, and *Comps in all weeks*$_i$ is a dummy for listing-transaction pairs with at least one match in each week. The sample of sold properties includes only those whose transaction price is below the predicted price obtained using a hedonic regression of the form: $log(p_j) = FE + \varepsilon_j$, where the fixed effects refer to the month-year of the transaction, location measured as the address outcode, property type (detached, semi-detached, terraced house or a flat) and number of rooms. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1), while controls for the number of comparable quotes in the current and each of the two previous weeks are included in columns (3)-(4). Listing month-year fixed effects are included in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price	0.8923***	0.8944***	0.8943***	0.8971***
	(0.0025)	(0.0054)	(0.0054)	(0.0054)
Price × Comps in week -1	0.0112***	0.0111***	0.0112***	0.0101**
	(0.0040)	(0.0040)	(0.0040)	(0.0039)
Price × Comps in week +1	0.0176***	0.0176***	0.0176***	0.0169***
	(0.0036)	(0.0036)	(0.0036)	(0.0036)
Price × Comps in all weeks	0.0341***	0.0341***	0.0339***	0.0331***
	(0.0030)	(0.0030)	(0.0030)	(0.0030)
Comps in week -1	-0.1373***	-0.1368***	-0.1389***	-0.1254***
	(0.0475)	(0.0475)	(0.0475)	(0.0474)
Comps in week +1	-0.2080***	-0.2079***	-0.2109***	-0.2047***
	(0.0436)	(0.0436)	(0.0436)	(0.0435)
Comps in all weeks	-0.3994***	-0.3994***	-0.4048***	-0.3987***
	(0.0365)	(0.0365)	(0.0365)	(0.0364)
Controls				
Price x Time distance	No	Yes	Yes	Yes
Nb. of comps per week	No	No	Yes	Yes
Fixed-Effects				
Month-year	No	No	No	Yes
Observations	1,522,663	1,522,663	1,522,663	1,522,663
R^2	0.7498	0.7499	0.7499	0.7513
Within R^2	–	–	–	0.7504

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A32: Indirect Price Effects - Large Positive Shocks

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Comps in w-1}_i + \beta_2 \times log(p_j) \times \textit{Comps in w+1}_i + \beta_3 \times log(p_j) \times \textit{Comps in all weeks}_i + \gamma_1 \times \textit{Comps in w-1}_i + \gamma_2 \times \textit{Comps in w+1}_i + \gamma_3 \times \textit{Comps in all weeks}_i + \textit{Controls} + \varepsilon_i$, where q_i is the listed price for property i which has been set/updated during the second week following the most recent price data publication date, p_j is the transaction price for a comparable property j sold in the previous month, *Comps in w-1*$_i$ is a dummy that turns on if the listing-transaction pair have at least one other comparable match in the week before the price publication date but none in the week after, *Comps in w+1*$_i$ is a dummy for pairs that have at least one match in the week after but none in the week before, and *Comps in all weeks*$_i$ is a dummy for listing-transaction pairs with at least one match in each week. The sample of sold properties includes only those for which the log transaction price is by more than 0.20 greater than the predicted price obtained using a hedonic regression of the form: $log(p_j) = FE + \varepsilon_j$, where the fixed effects refer to the month-year of the transaction, location measured as the address outcode, property type (detached, semi-detached, terraced house or a flat) and number of rooms. This cutoff roughly corresponds to the top 20% of the residual price distribution. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1), while controls for the number of comparable quotes in the current and each of the two previous weeks are included in columns (3)-(4). Listing month-year fixed effects are included in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price	0.8987***	0.8854***	0.8851***	0.8788***
	(0.0036)	(0.0085)	(0.0085)	(0.0084)
Price × Comps in week -1	0.0137**	0.0138**	0.0134**	0.0128**
	(0.0057)	(0.0057)	(0.0057)	(0.0057)
Price × Comps in week +1	0.0234***	0.0234***	0.0234***	0.0226***
	(0.0050)	(0.0050)	(0.0050)	(0.0050)
Price × Comps in all weeks	0.0297***	0.0299***	0.0306***	0.0292***
	(0.0043)	(0.0043)	(0.0043)	(0.0043)
Comps in week -1	-0.1794**	-0.1808**	-0.1722**	-0.1652**
	(0.0706)	(0.0706)	(0.0704)	(0.0703)
Comps in week +1	-0.2977***	-0.2977***	-0.2966***	-0.2869***
	(0.0625)	(0.0625)	(0.0624)	(0.0625)
Comps in all weeks	-0.3800***	-0.3817***	-0.3803***	-0.3656***
	(0.0539)	(0.0539)	(0.0538)	(0.0538)
Controls				
Price x Time distance	No	Yes	Yes	Yes
Nb. of comps per week	No	No	Yes	Yes
Fixed-Effects				
Month-year	No	No	No	Yes
Observations	571,002	571,002	571,002	571,002
R^2	0.7804	0.7804	0.7805	0.7820
Within R^2	–	–	–	0.7738

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A33: Indirect Price Effects - Large Negative Shocks

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Comps in w-1}_i + \beta_2 \times log(p_j) \times \textit{Comps in w+1}_i + \beta_3 \times log(p_j) \times \textit{Comps in all weeks}_i + \gamma_1 \times \textit{Comps in w-1}_i + \gamma_2 \times \textit{Comps in w+1}_i + \gamma_3 \times \textit{Comps in all weeks}_i + \textit{Controls} + \varepsilon_i$, where q_i is the listed price for property i which has been set/updated during the second week following the most recent price data publication date, p_j is the transaction price for a comparable property j sold in the previous month, *Comps in w-1*$_i$ is a dummy that turns on if the listing-transaction pair have at least one other comparable match in the week before the price publication date but none in the week after, *Comps in w+1*$_i$ is a dummy for pairs that have at least one match in the week after but none in the week before, and *Comps in all weeks*$_i$ is a dummy for listing-transaction pairs with at least one match in each week. The sample of sold properties includes only those for which the log transaction price is by more than 0.20 lower than the predicted price obtained using a hedonic regression of the form: $log(p_j) = FE + \varepsilon_j$, where the fixed effects refer to the month-year of the transaction, location measured as the address outcode, property type (detached, semi-detached, terraced house or a flat) and number of rooms. This cutoff roughly corresponds to the bottom 20% of the residual price distribution. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1), while controls for the number of comparable quotes in the current and each of the two previous weeks are included in columns (3)-(4). Listing month-year fixed effects are included in column (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)	(4)
Price	0.8945***	0.8943***	0.8942***	0.8970***
	(0.0033)	(0.0074)	(0.0074)	(0.0073)
Price × Comps in week -1	0.0118**	0.0118**	0.0118**	0.0106**
	(0.0052)	(0.0052)	(0.0052)	(0.0052)
Price × Comps in week +1	0.0189***	0.0189***	0.0189***	0.0184***
	(0.0048)	(0.0048)	(0.0048)	(0.0048)
Price × Comps in all weeks	0.0449***	0.0449***	0.0445***	0.0440***
	(0.0040)	(0.0040)	(0.0040)	(0.0040)
Comps in week -1	-0.1431**	-0.1429**	-0.1449**	-0.1302**
	(0.0623)	(0.0623)	(0.0623)	(0.0619)
Comps in week +1	-0.2179***	-0.2181***	-0.2214***	-0.2176***
	(0.0569)	(0.0569)	(0.0569)	(0.0567)
Comps in all weeks	-0.5155***	-0.5160***	-0.5218***	-0.5200***
	(0.0474)	(0.0474)	(0.0473)	(0.0471)
Controls				
Price x Time distance	No	Yes	Yes	Yes
Nb. of comps per week	No	No	Yes	Yes
Fixed-Effects				
Month-year	No	No	No	Yes
Observations	674,737	674,737	674,737	674,737
R^2	0.7622	0.7622	0.7623	0.7640
Within R^2	–	–	–	0.7617

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A34: Indirect Price Effects by Sales Volume

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Comps in w-1}_i + \beta_2 \times log(p_j) \times \textit{Comps in w+1}_i + \beta_3 \times log(p_j) \times \textit{Comps in all weeks}_i + \gamma_1 \times \textit{Comps in w-1}_i + \gamma_2 \times \textit{Comps in w+1}_i + \gamma_3 \times \textit{Comps in all weeks}_i + \textit{Controls} + \varepsilon_i$, where q_i is the listed price for property i which has been set/updated during the second week following the most recent price data publication date, p_j is the transaction price for a comparable property j sold in the previous month, *Comps in w-1*$_i$ is a dummy that turns on if the listing-transaction pair have at least one other comparable match in the week before the price publication date but none in the week after, *Comps in w+1*$_i$ is a dummy for pairs that have at least one match in the week after but none in the week before, and *Comps in all weeks*$_i$ is a dummy for listing-transaction pairs with at least one match in each week. The sample is split based on sales activity: the first two columns show the effect on quotes of transaction prices in months and regions with sales volume below the sample average, while the last two columns show the results for the subsample with above-average sales volume. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price, as well as controls for the number of comparable quotes in the current and each of the two previous weeks are included in all columns. Listing month-year fixed effects are included in columns (2) and (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	Low Volume		High Volume	
	(1)	(2)	(3)	(4)
Price	0.8306***	0.8194***	0.8223***	0.8237***
	(0.0076)	(0.0075)	(0.0053)	(0.0053)
Price × Comps in week -1	0.0071	0.0064	0.0087**	0.0082**
	(0.0055)	(0.0055)	(0.0040)	(0.0040)
Price × Comps in week +1	0.0128**	0.0113**	0.0182***	0.0175***
	(0.0051)	(0.0051)	(0.0036)	(0.0036)
Price × Comps in all weeks	0.0220***	0.0199***	0.0177***	0.0166***
	(0.0042)	(0.0042)	(0.0031)	(0.0031)
Comps in week -1	-0.0959	-0.0873	-0.1108**	-0.1044**
	(0.0664)	(0.0662)	(0.0481)	(0.0480)
Comps in week +1	-0.1567**	-0.1378**	-0.2201***	-0.2131***
	(0.0614)	(0.0614)	(0.0434)	(0.0433)
Comps in all weeks	-0.2779***	-0.2539***	-0.2080***	-0.1977***
	(0.0508)	(0.0506)	(0.0373)	(0.0373)
Controls				
Price x Time distance	Yes	Yes	Yes	Yes
Nb. of comps per week	Yes	Yes	Yes	Yes
Fixed-effects				
Month-year	No	Yes	No	Yes
Observations	897,043	897,043	2,035,215	2,035,215
R^2	0.7174	0.7194	0.6978	0.6989
Within R^2	–	0.7057	–	0.6969

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A35: Indirect Price Effects by Excess Volume at the Local Level

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_0 \times log(p_j) + \beta_1 \times log(p_j) \times \textit{Comps in w-1}_i + \beta_2 \times log(p_j) \times \textit{Comps in w+1}_i + \beta_3 \times log(p_j) \times \textit{Comps in all weeks}_i + \gamma_1 \times \textit{Comps in w-1}_i + \gamma_2 \times \textit{Comps in w+1}_i + \gamma_3 \times \textit{Comps in all weeks}_i + \textit{Controls} + \varepsilon_i$, where q_i is the listed price for property i which has been set/updated during the second week following the most recent price data publication date, p_j is the transaction price for a comparable property j sold in the previous month, *Comps in w-1*$_i$ is a dummy that turns on if the listing-transaction pair have at least one other comparable match in the week before the price publication date but none in the week after, *Comps in w+1*$_i$ is a dummy for pairs that have at least one match in the week after but none in the week before, and *Comps in all weeks*$_i$ is a dummy for listing-transaction pairs with at least one match in each week. The sample is split based on relative sales activity: the first two columns show the effect on quotes of transaction prices in months with below-average sales volume for the local area, while the last two columns show the results for months with above-average sales volume. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price, as well as controls for the number of comparable quotes in the current and each of the two previous weeks are included in all columns. Listing month-year fixed effects are included in columns (2) and (4). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	Low Relative Volume		High Relative Volume	
	(1)	(2)	(3)	(4)
Price	0.8315***	0.8063***	0.8178***	0.8163***
	(0.0063)	(0.0062)	(0.0060)	(0.0060)
Price × Comps in week -1	0.0072	0.0079*	0.0105**	0.0085*
	(0.0046)	(0.0046)	(0.0045)	(0.0044)
Price × Comps in week +1	0.0131***	0.0148***	0.0205***	0.0183***
	(0.0042)	(0.0042)	(0.0040)	(0.0040)
Price × Comps in all weeks	0.0199***	0.0212***	0.0245***	0.0203***
	(0.0035)	(0.0035)	(0.0035)	(0.0035)
Comps in week -1	-0.0940*	-0.1012*	-0.1339**	-0.1130**
	(0.0558)	(0.0553)	(0.0542)	(0.0540)
Comps in week +1	-0.1600***	-0.1781***	-0.2501***	-0.2265***
	(0.0508)	(0.0506)	(0.0492)	(0.0491)
Comps in all weeks	-0.2435***	-0.2603***	-0.2988***	-0.2527***
	(0.0421)	(0.0422)	(0.0424)	(0.0422)
Controls				
Price x Time distance	Yes	Yes	Yes	Yes
Nb. of comps per week	Yes	Yes	Yes	Yes
Fixed-effects				
Month-year	No	Yes	No	Yes
Observations	1,384,793	1,384,793	1,547,465	1,547,465
R^2	0.7035	0.7080	0.7073	0.7104
Within R^2	–	0.6688	–	0.6928

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A36: Chain Effects of Prices on Quotes by Order of Match

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_1 \times log(p_j) + \sum_{k=2}^{10} \beta_k^{pre} \times log(p_j) \times Comp\ Order\ k\ Pre_i + \sum_{k=2}^{10} \gamma_k^{pre} Comp\ Order\ k\ Pre_i + \sum_{k=1}^{10} \beta_k^{post} \times log(p_j) \times Comp\ Order\ k\ Post_i + \sum_{k=1}^{10} \gamma_k^{post} Comp\ Order\ k\ Post_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and *Comp Order k Pre (Post)*$_i$ is a dummy that turns on when quote i is the k-th sequential match to transaction j in the period before (after) the price data publication date. The sample includes listings in the one-month period surrounding the publication date that have a comparable transaction which has at least one treated and one untreated match. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1). Column (3) also includes listing month-year fixed effects. Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)
Price	0.8388***	0.8404***	0.8401***
	(0.0011)	(0.0020)	(0.0020)
Price × 2nd Untreated	0.0010	0.0011	0.0011
	(0.0016)	(0.0016)	(0.0016)
Price × 3rd Untreated	0.0043**	0.0045**	0.0044**
	(0.0020)	(0.0020)	(0.0020)
Price × 4th Untreated	0.0054**	0.0057**	0.0056**
	(0.0023)	(0.0023)	(0.0023)
Price × 5th Untreated	0.0060**	0.0063**	0.0062**
	(0.0027)	(0.0027)	(0.0027)
Price × 6th Untreated	0.0039	0.0042	0.0041
	(0.0033)	(0.0033)	(0.0033)
Price × 7th Untreated	0.0068*	0.0071*	0.0070*
	(0.0039)	(0.0039)	(0.0039)
Price × 8th Untreated	0.0021	0.0024	0.0025
	(0.0044)	(0.0044)	(0.0044)
Price × 9th Untreated	0.0036	0.0040	0.0041
	(0.0055)	(0.0055)	(0.0055)
Price × 10th or more Untreated	0.0000	0.0004	0.0003
	(0.0038)	(0.0038)	(0.0038)
Price × 1st Treated	0.0015	0.0023	0.0025
	(0.0014)	(0.0017)	(0.0017)
Price × 2nd Treated	0.0039**	0.0048**	0.0050***
	(0.0016)	(0.0019)	(0.0019)
Price × 3rd Treated	0.0054***	0.0064***	0.0065***
	(0.0019)	(0.0022)	(0.0022)

Continued on next page

Table A36 – *Continued from previous page*

	(1)	(2)	(3)
Price × 4th Treated	0.0106***	0.0116***	0.0116***
	(0.0022)	(0.0025)	(0.0025)
Price × 5th Treated	0.0082***	0.0093***	0.0092***
	(0.0026)	(0.0029)	(0.0029)
Price × 6th Treated	0.0098***	0.0109***	0.0107***
	(0.0031)	(0.0033)	(0.0033)
Price × 7th Treated	0.0117***	0.0129***	0.0126***
	(0.0037)	(0.0039)	(0.0039)
Price × 8th Treated	0.0177***	0.0188***	0.0186***
	(0.0044)	(0.0046)	(0.0046)
Price × 9th Treated	0.0099*	0.0110**	0.0108**
	(0.0051)	(0.0052)	(0.0052)
Price × 10th or more Treated	0.0101***	0.0113***	0.0112***
	(0.0036)	(0.0039)	(0.0039)
2nd Untreated	-0.0115	-0.0131	-0.0122
	(0.0198)	(0.0198)	(0.0198)
3rd Untreated	-0.0508**	-0.0535**	-0.0524**
	(0.0236)	(0.0237)	(0.0237)
4th Untreated	-0.0635**	-0.0669**	-0.0661**
	(0.0280)	(0.0282)	(0.0281)
5th Untreated	-0.0686**	-0.0726**	-0.0720**
	(0.0328)	(0.0330)	(0.0330)
6th Untreated	-0.0406	-0.0449	-0.0437
	(0.0395)	(0.0397)	(0.0396)
7th Untreated	-0.0749	-0.0797*	-0.0777*
	(0.0469)	(0.0471)	(0.0470)
8th Untreated	-0.0158	-0.0207	-0.0210
	(0.0529)	(0.0532)	(0.0531)
9th Untreated	-0.0269	-0.0319	-0.0328
	(0.0659)	(0.0661)	(0.0661)
10th or more Untreated	0.0194	0.0140	0.0174
	(0.0457)	(0.0462)	(0.0461)
1st Treated	-0.0165	-0.0288	-0.0305
	(0.0174)	(0.0200)	(0.0200)
2nd Treated	-0.0442**	-0.0582**	-0.0590***
	(0.0199)	(0.0229)	(0.0228)
3rd Treated	-0.0599**	-0.0751***	-0.0747***
	(0.0233)	(0.0264)	(0.0263)
4th Treated	-0.1213***	-0.1372***	-0.1360***
	(0.0269)	(0.0300)	(0.0299)

Continued on next page

Table A36 – *Continued from previous page*

	(1)	(2)	(3)
5th Treated	-0.0894***	-0.1059***	-0.1035***
	(0.0318)	(0.0347)	(0.0346)
6th Treated	-0.1101***	-0.1269***	-0.1229***
	(0.0377)	(0.0403)	(0.0403)
7th Treated	-0.1278***	-0.1451***	-0.1399***
	(0.0449)	(0.0472)	(0.0471)
8th Treated	-0.1997***	-0.2172***	-0.2118***
	(0.0529)	(0.0551)	(0.0551)
9th Treated	-0.1079*	-0.1256**	-0.1199*
	(0.0612)	(0.0632)	(0.0631)
10th or more Treated	-0.0996**	-0.1177**	-0.1125**
	(0.0438)	(0.0470)	(0.0469)
Controls			
Price x Time distance	No	Yes	Yes
Fixed-Effects			
Month-year	No	No	Yes
Observations	11,292,009	11,292,009	11,292,009
R^2	0.7081	0.7081	0.7089
Within R^2	–	–	0.7063

Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A37: Chain Effects of Prices on Quotes by Order of Match - Linear Effects

The table presents the results of the following regression: $log(q_i^s) = \alpha^s + \beta_0^s \times log(p_j) + \beta_1^s \times log(p_j) \times$ $\textit{Nb. prior comps}_i^s + \gamma_1 \times \textit{Nb. prior comps}_i^s + \textit{Controls} + \varepsilon_i^s$, where q_i^s is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and $\textit{Nb. prior comps}_i^s$ is the number of previous listings that have been matched to the same transaction. The superscript s is an indicator for whether the regressions use the sample of treated listings, i.e., listings posted in the period after the price data publication date (last three columns), or the set of untreated ones (first three columns). Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1) and (4) and listing month-year fixed effects are included in columns (3) and (6). Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	Untreated			Treated		
	(1)	(2)	(3)	(4)	(5)	(6)
Prior × Nb. prior comps	0.0001	0.0001	0.0001	0.0007***	0.0008***	0.0007***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Price	0.8409***	0.8414***	0.8408***	0.8420***	0.8435***	0.8436***
	(0.0011)	(0.0022)	(0.0022)	(0.0011)	(0.0030)	(0.0030)
Nb. prior comps	0.0005	0.0003	0.0008	-0.0072**	-0.0076**	-0.0067**
	(0.0034)	(0.0035)	(0.0034)	(0.0032)	(0.0033)	(0.0033)
Controls						
Price x Time distance	No	Yes	Yes	No	Yes	Yes
Fixed-Effects						
Month-year	No	No	Yes	No	No	Yes
Observations	5,577,243	5,577,243	5,577,243	5,714,766	5,714,766	5,714,766
R^2	0.7068	0.7068	0.7076	0.7094	0.7094	0.7103
Within R^2	–	–	0.7048	–	–	0.7076

Two-way (Transaction ID & Listing ID) standard-errors in parentheses

*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A38: Chain Effects of Prices on Quotes by Order of Match - Before March 2012

The table presents the results of the following regression: $log(q_i) = \alpha + \beta_1 \times log(p_j) + \sum_{k=2}^{10} \beta_k^{pre} \times log(p_j) \times Comp\ Order\ k\ Pre_i + \sum_{k=2}^{10} \gamma_k^{pre} Comp\ Order\ k\ Pre_i + \sum_{k=1}^{10} \beta_k^{post} \times log(p_j) \times Comp\ Order\ k\ Post_i + \sum_{k=1}^{10} \gamma_k^{post} Comp\ Order\ k\ Post_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and *Comp Order k Pre (Post)*$_i$ is a dummy that turns on when quote i is the k-th sequential match to transaction j in the period before (after) the (placebo) price data publication date. The regressions use only data from the period before March 2012. The sample includes listings in the one-month period surrounding the publication date that have a comparable transaction which has at least one treated and one untreated match. Controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price are included in all columns but (1). Column (3) also includes listing month-year fixed effects. Standard errors double-clustered at the transaction and listing ID levels are reported in parentheses.

	(1)	(2)	(3)
Price	0.7875***	0.7777***	0.7786***
	(0.0023)	(0.0042)	(0.0042)
Price × 2nd Untreated	0.0057*	0.0052	0.0047
	(0.0034)	(0.0034)	(0.0034)
Price × 3rd Untreated	0.0094**	0.0086**	0.0077**
	(0.0038)	(0.0038)	(0.0038)
Price × 4th Untreated	0.0076*	0.0066	0.0054
	(0.0045)	(0.0045)	(0.0045)
Price × 5th Untreated	0.0102**	0.0090*	0.0076
	(0.0051)	(0.0051)	(0.0051)
Price × 6th Untreated	0.0073	0.0060	0.0045
	(0.0057)	(0.0057)	(0.0057)
Price × 7th Untreated	0.0047	0.0033	0.0018
	(0.0064)	(0.0064)	(0.0064)
Price × 8th Untreated	0.0021	0.0007	-0.0009
	(0.0075)	(0.0075)	(0.0075)
Price × 9th Untreated	-0.0011	-0.0027	-0.0044
	(0.0085)	(0.0086)	(0.0086)
Price × 10th or more Untreated	-0.0175***	-0.0193***	-0.0213***
	(0.0058)	(0.0059)	(0.0059)
Price × 1st Treated	0.0014	-0.0034	-0.0036
	(0.0030)	(0.0034)	(0.0034)
Price × 2nd Treated	0.0028	-0.0024	-0.0030

Continued on next page

Table A38 – *Continued from previous page*

	(1)	(2)	(3)
	(0.0033)	(0.0038)	(0.0038)
Price × 3rd Treated	-0.0002	-0.0058	-0.0068
	(0.0038)	(0.0043)	(0.0043)
Price × 4th Treated	0.0071*	0.0013	0.0001
	(0.0042)	(0.0047)	(0.0047)
Price × 5th Treated	0.0070	0.0011	-0.0001
	(0.0048)	(0.0053)	(0.0053)
Price × 6th Treated	-0.0031	-0.0092	-0.0105*
	(0.0054)	(0.0059)	(0.0059)
Price × 7th Treated	0.0070	0.0008	-0.0005
	(0.0061)	(0.0065)	(0.0065)
Price × 8th Treated	0.0036	-0.0027	-0.0040
	(0.0069)	(0.0073)	(0.0073)
Price × 9th Treated	-0.0051	-0.0114	-0.0127
	(0.0077)	(0.0081)	(0.0081)
Price × 10th Treated	-0.0183***	-0.0249***	-0.0263***
	(0.0056)	(0.0062)	(0.0062)
2nd Untreated	-0.0666*	-0.0606	-0.0543
	(0.0397)	(0.0398)	(0.0397)
3rd Untreated	-0.1110**	-0.1009**	-0.0895**
	(0.0453)	(0.0454)	(0.0454)
4th Untreated	-0.0867	-0.0739	-0.0584
	(0.0533)	(0.0536)	(0.0536)
5th Untreated	-0.1160*	-0.1011*	-0.0824
	(0.0600)	(0.0604)	(0.0602)
6th Untreated	-0.0808	-0.0646	-0.0452
	(0.0677)	(0.0681)	(0.0680)
7th Untreated	-0.0485	-0.0311	-0.0109
	(0.0760)	(0.0764)	(0.0763)
8th Untreated	-0.0128	0.0052	0.0255
	(0.0887)	(0.0892)	(0.0892)
9th Untreated	0.0310	0.0495	0.0725
	(0.1015)	(0.1020)	(0.1016)
10th or more Untreated	0.2482***	0.2696***	0.2952***
	(0.0692)	(0.0702)	(0.0701)
1st Treated	-0.0201	0.0375	0.0404
	(0.0355)	(0.0404)	(0.0403)

Continued on next page

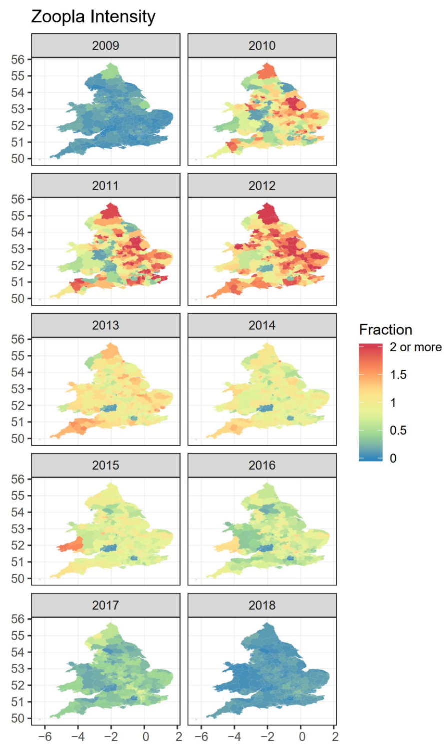
Table A38 – *Continued from previous page*

	(1)	(2)	(3)
2nd Treated	-0.0336	0.0302	0.0385
	(0.0395)	(0.0452)	(0.0451)
3rd Treated	0.0038	0.0718	0.0854*
	(0.0452)	(0.0512)	(0.0512)
4th Treated	-0.0815	-0.0108	0.0050
	(0.0501)	(0.0560)	(0.0559)
5th Treated	-0.0762	-0.0033	0.0131
	(0.0567)	(0.0630)	(0.0630)
6th Treated	0.0479	0.1223*	0.1405**
	(0.0644)	(0.0698)	(0.0698)
7th Treated	-0.0713	0.0041	0.0228
	(0.0722)	(0.0773)	(0.0773)
8th Treated	-0.0251	0.0513	0.0702
	(0.0824)	(0.0873)	(0.0871)
9th Treated	0.0810	0.1579	0.1764*
	(0.0918)	(0.0966)	(0.0967)
10th or more Treated	0.2571***	0.3363***	0.3567***
	(0.0674)	(0.0748)	(0.0746)
Controls			
Price x Time distance	No	Yes	Yes
Fixed-Effects			
Month-year	No	No	Yes
Observations	4,772,899	4,772,899	4,772,899
R^2	0.60688	0.60689	0.60811
Within R^2	–	–	0.60669

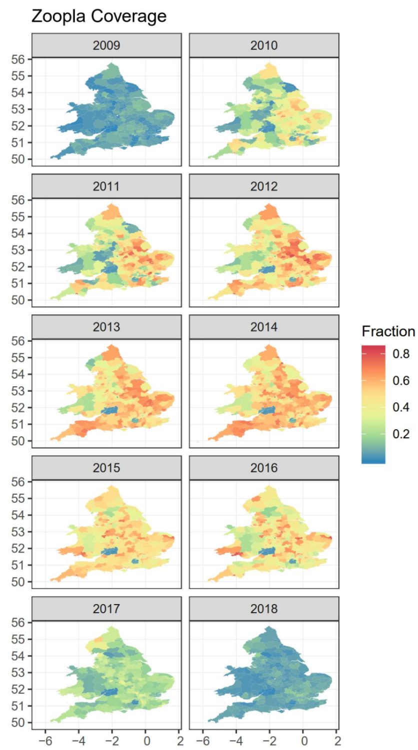
Two-way (Transaction ID & Listing ID) standard-errors in parentheses
*Signif Codes: ***: 0.01, **: 0.05, *: 0.1*

Figure A17: Relative Coverage of Listing Data

The figure plots heat maps of the relative geographic coverage of the Zoopla listing data between 2009-2018 by year across England and Wales. Figure A17a displays the total number of listings as a fraction of transactions shifted by six months (average TOM), while Figure A17b the fraction of transactions that were matched to their respective listings in the Zoopla data.



(a) Listings-to-Transactions Ratio

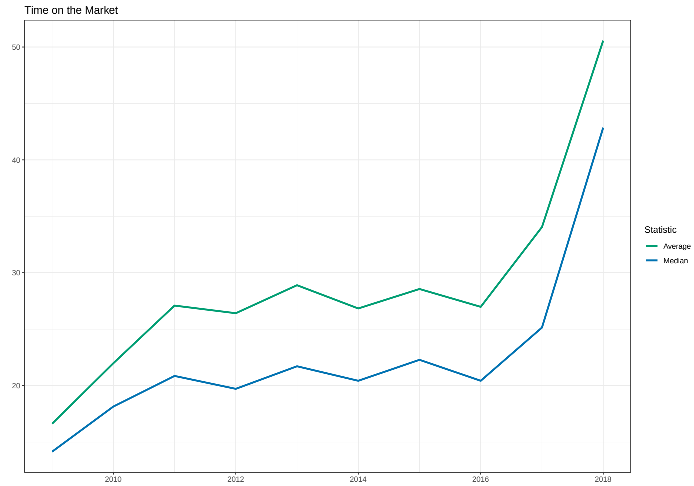


(b) % of Transactions Listed on Zoopla

Figure A18: Time-series of Price Discount and Time on the Market

The figure displays the time-series of price discount and time on the market (TOM) for the set of property listings that were matched to their respective ex-post transactions in the sample from 2009 to 2018. Figure A18a plots the time-series of the percentage difference between the first listed price and the final transaction price, while Figure A18b shows the time-series of time on the market measured as the number of weeks from listing to sale completion. The green lines show the time-series of the average values and the blue lines represent the median values.

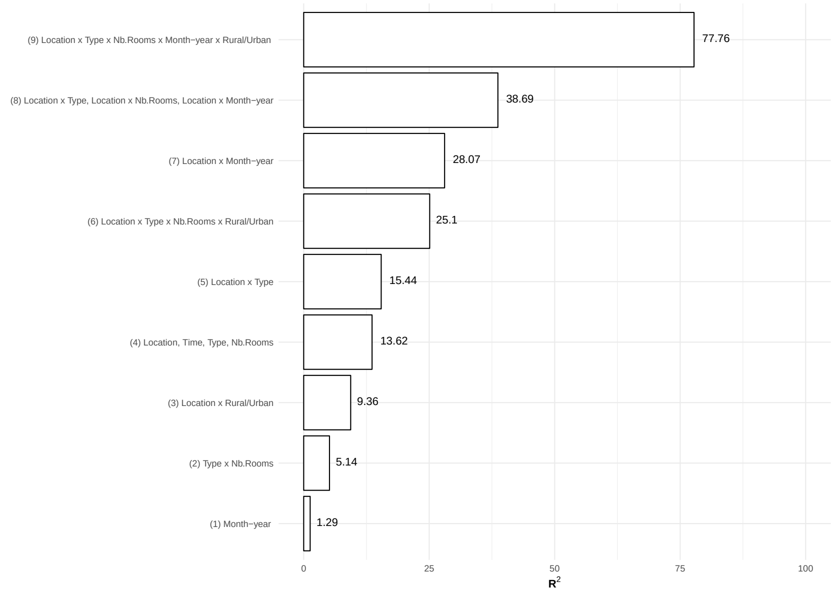
(a) Price Discount



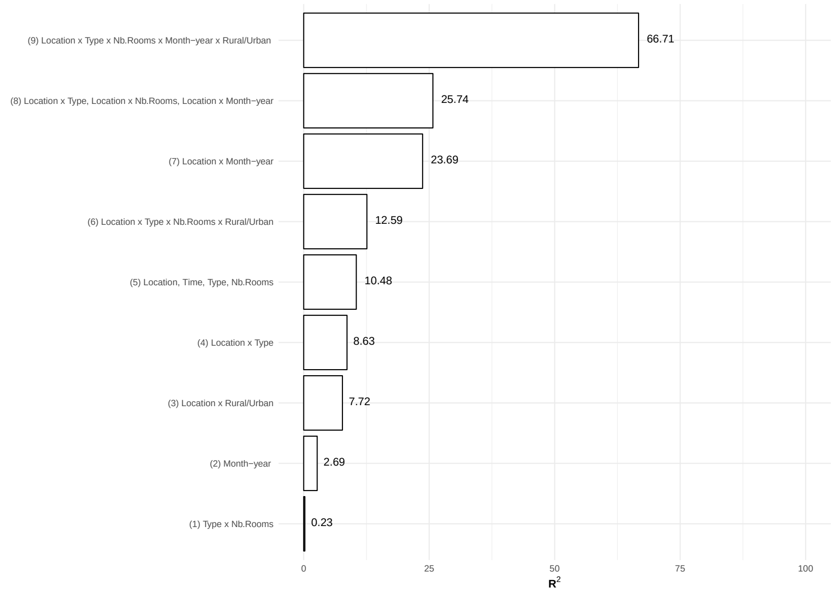
(b) Time on the Market

Figure A19: Fraction of Explained Variation in Price Differences

The figure displays the percentage of variation in differences between listing and transaction prices that is explained by observable characteristics, measured as the R-squared from a regression of price differences on various fixed effects, for the set of property listings that were matched to their respective ex-post transactions in the 2009-2018 sample. Figure A19a shows the explained variation in level price differences and Figure A19b in percentage price differences. Fixed effects included are: month-year of the listing or transaction; property type (detached, semi-detached, terraced house or a flat); number of rooms in the property, where properties with between 6 and 10 rooms are placed in one bucket and properties with more than 10 rooms in another; location, measured as the address outcode, and; a rural/urban area indicator from the 2011 Census classification of Output Areas.



(a) Price Difference in pounds (£)



(b) Price Difference in percent (%)

Figure A20: Variation In Quotes Around Publishing Dates - Before March 2012

The figure plots the results from a regression of listing prices on dummies for the signed number of days between the listing date and the closest (placebo) price data publication date for the sample before March 2012. The regression is specified as follows: $q_i = \alpha + \sum_{\Delta=-15}^{15} \gamma_\Delta \Delta_i + FE + \varepsilon_i$, where the fixed-effects correspond to the characteristics the matching is based on, i.e., location, property type, number of rooms and month-year, and Δ_i is a dummy for the signed difference in days between the date on which a listing is posted and the closest publication date. The baseline coefficient is the one for listings posted exactly on the publication date. The vertical lines represent the 95% confidence bounds for the point estimates.

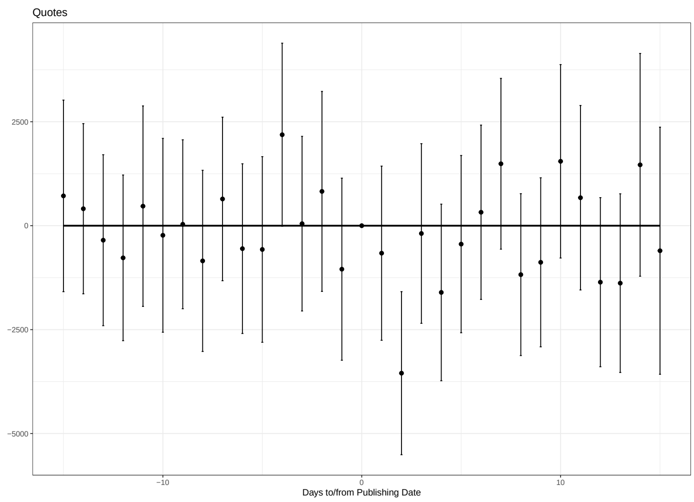
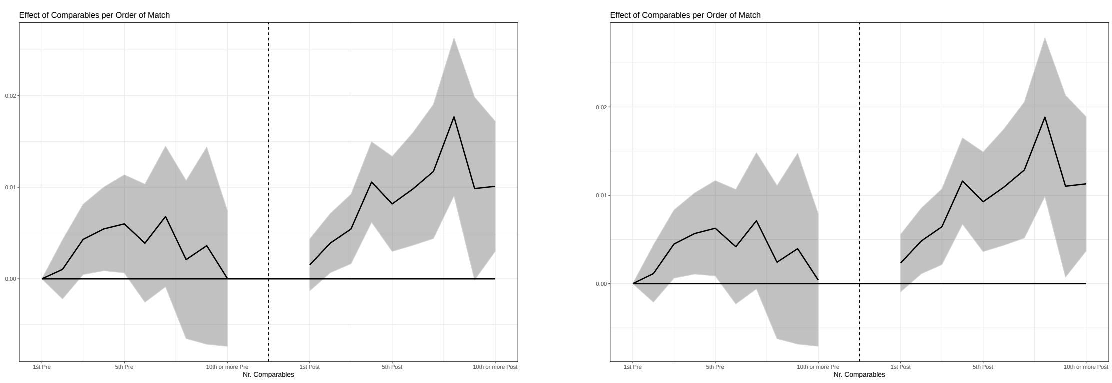


Figure A21: Effect of Transaction Prices by Order of Match

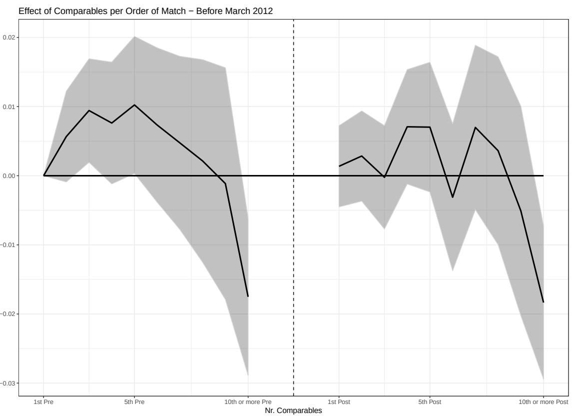
The figure plots the price coefficients from the following regression along with their 95% confidence bounds: $log(q_i) = \alpha + \beta_1 \times log(p_j) + \sum_{k=2}^{10} \beta_k^{pre} \times log(p_j) \times Comp\ Order\ k\ Pre_i + \sum_{k=2}^{10} \gamma_k^{pre} Comp\ Order\ k\ Pre_i + \sum_{k=1}^{10} \beta_k^{post} \times log(p_j) \times Comp\ Order\ k\ Post_i + \sum_{k=1}^{10} \gamma_k^{post} Comp\ Order\ k\ Post_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and *Comp Order k Pre (Post)*$_i$ is a dummy that turns on when quote i is the k-th sequential match to transaction j in the period before (after) the price data publication date. The sample includes listings in the one-month period surrounding the publication date that have a comparable transaction which has at least one treated and one untreated match. Figure A21a is the baseline regression with no controls and Figure A21b includes controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price.



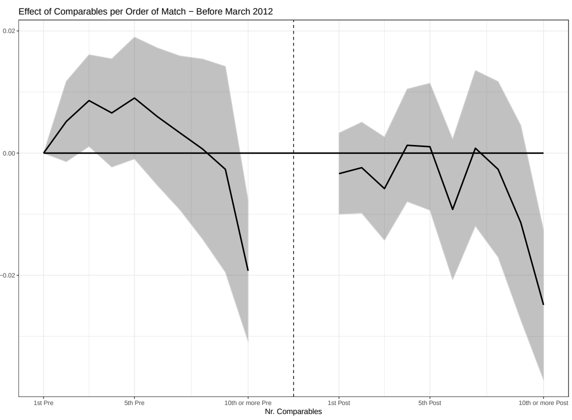
(a) Baseline regression (b) Time distance controls

Figure A22: Effect of Transaction Prices by Order of Match - Before March 2012

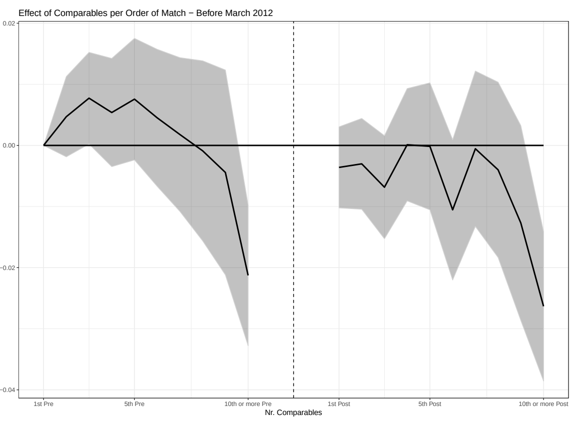
The figure plots the price coefficients from the following regression along with their 95% confidence bounds: $log(q_i) = \alpha + \beta_1 \times log(p_j) + \sum_{k=2}^{10} \beta_k^{pre} \times log(p_j) \times Comp\ Order\ k\ Pre_i + \sum_{k=2}^{10} \gamma_k^{pre} Comp\ Order\ k\ Pre_i + \sum_{k=1}^{10} \beta_k^{post} \times log(p_j) \times Comp\ Order\ k\ Post_i + \sum_{k=1}^{10} \gamma_k^{post} Comp\ Order\ k\ Post_i + Controls + \varepsilon_i$, where q_i is the listed price for property i, p_j is the transaction price for a comparable property j sold in the previous month and *Comp Order k Pre (Post)$_i$* is a dummy that turns on when quote i is the k-th sequential match to transaction j in the period before (after) the (placebo) price data publication date. The sample includes only listings before March 2012 posted in the one-month period surrounding the publication date that have a comparable transaction which has at least one treated and one untreated match. Figure A22a is the baseline regression with no controls, Figure A22b includes controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price and Figure A22c includes month-year fixed effects in addition to time distance controls.



(a) Baseline regression



(b) Time distance controls



(c) Time distance controls and month-year fixed effects

Figure A23: Number of Quote Updates per Listing

The figure plots the histogram of the total number of price changes per listing. The sample includes listings posted after March 2012 that have at least one comparable transaction.

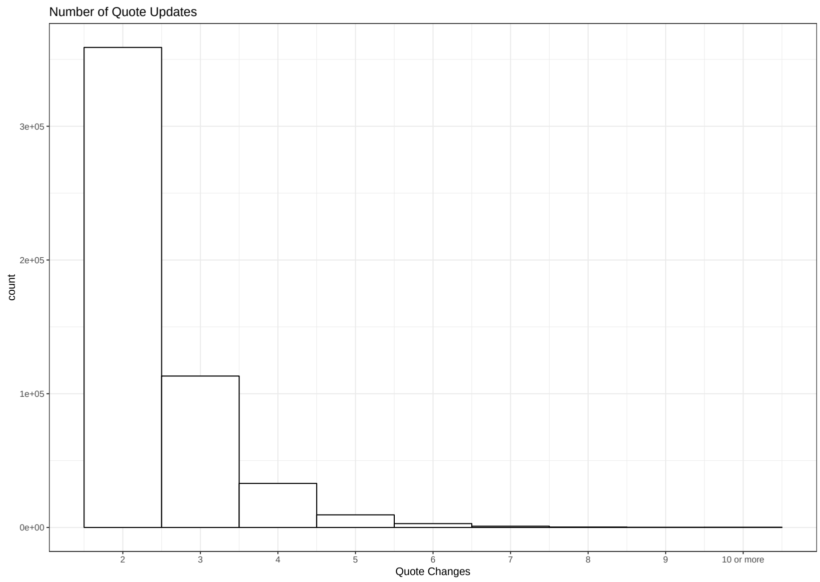


Figure A24: Distance in Days between Initial Listing Date and Subsequent Quote Changes

The figure plots a histogram of the difference in days between quote changes and the initial date of the listing. The sample includes listings posted after March 2012 that have at least one comparable transaction.

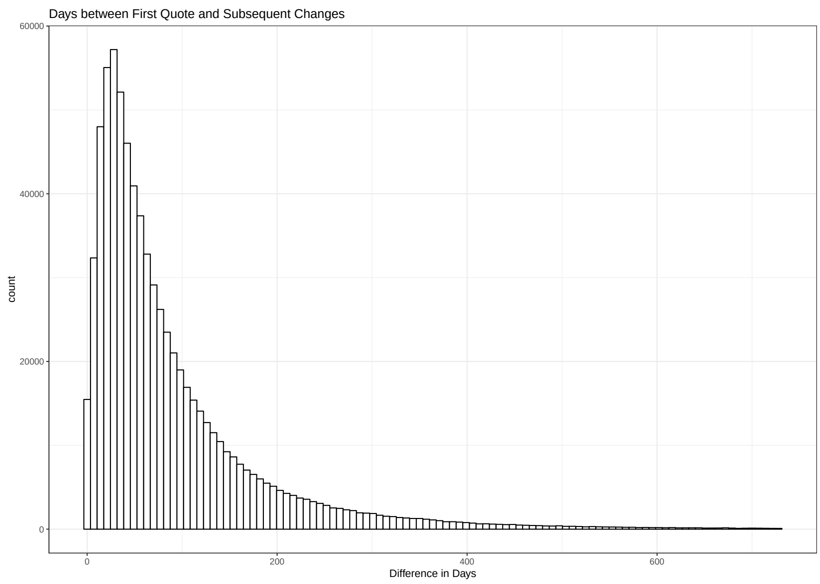
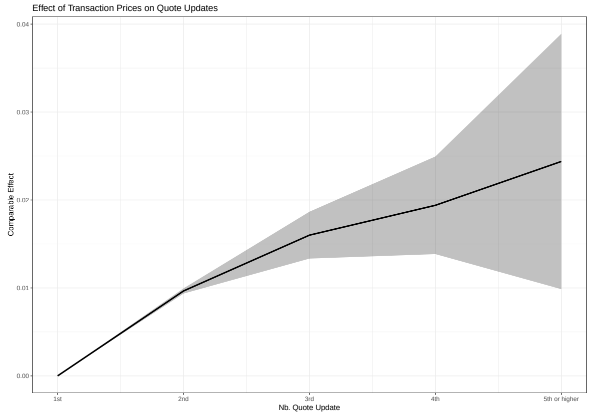
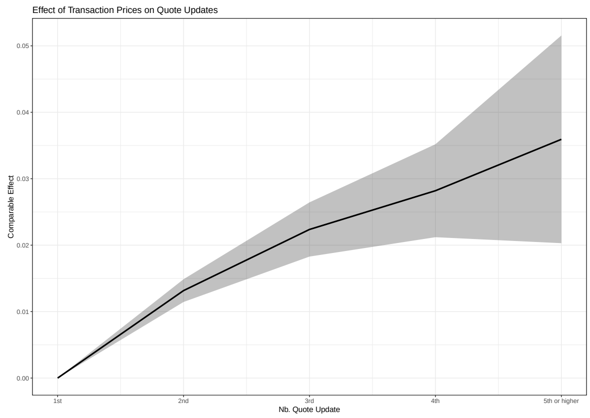


Figure A25: Effect of Transaction Prices on Quote Updates

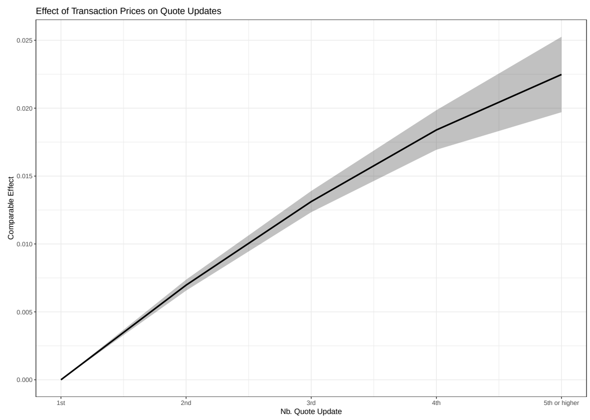
The figure displays the price coefficients from the following regression along with their 95% confidence bounds: $log(q_i^n) = \alpha + \beta_1 \times log(p_j) + \sum_{n=2}^{5} \beta_n \times log(p_j) \times Update\ Number\ n_i + \sum_{n=2}^{5} \gamma_n Update\ Number\ n_i + Controls + \varepsilon_i^n$, where q_i^n is the n-th listed price update for property i, p_j is the transaction price for a comparable property j sold in the month before property i was initially listed and *Update Number* n_i is a dummy that turns on when q_i^n is the n-th consecutive quote update for property i. The sample includes listings in the post March 2012 period that have at least one price change and a comparable transaction that has been published just before the listing has been first posted. Figure A25a is the baseline regression with no controls, Figure A25b includes controls for the time distance between the listing and the comparable transaction measured in days and its interaction with log price and Figure A25c includes month-year and listing ID fixed effects in addition to time distance controls.



(a) Baseline Regression



(b) Time distance controls



(c) Time distance controls, month-year and listing ID fixed effects

B. Appendix to Revealed Expectations and Learning Biases: Evidence from the Mutual Fund Industry

B.1 Tables

Table B1: Summary Statistics

The table reports summary statistics for the data used. Column $\bar{x}$ reports the sample average of each variable, column σ its standard deviation, Min the smallest observation, Q1 the first quartile, Median the $50th$ percentile, Q3 the third quartile, Max the largest observation and N the number of observations. The first panel reports summary statistics regarding average and median past returns experienced by managers. The second panel reports six measures of expected excess returns computed as $\hat{\Sigma}_t \boldsymbol{w_{i,t}}$. Rows (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; rows (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Rows (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, rows (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and rows (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$. The third panel reports summary statistics on managers' careers; experience refers to the number of quarters since the first time a certain stock appeared in the manager's portfolio; max.experience refers to the maximum experience achieved for each manager-stock pair; tenure refers to the number of quarters since the first time the manager appeared in sample.

	$\bar{x}$	σ	Min	Q1	Median	Q3	Max	N
				Experienced Returns				
average	0.024	0.100	-0.557	-0.010	0.026	0.063	0.607	13,912,677
median	0.014	0.111	-0.871	-0.026	0.021	0.062	1.198	13,912,677
				Expected Excess Returns				
(1)	0.012	0.015	-0.282	0.004	0.007	0.014	1.336	5,416,032
(2)	0.011	0.014	-0.208	0.003	0.006	0.012	0.806	5,416,032
(3)	0.011	0.015	-0.161	0.003	0.006	0.013	0.764	5,416,032
(4)	0.012	0.015	-0.278	0.004	0.007	0.014	0.766	12,707,119
(5)	0.011	0.015	-0.292	0.003	0.006	0.012	1.086	12,707,119
(6)	0.011	0.015	-0.319	0.003	0.006	0.013	1.034	12,707,119
				Managers Careers				
experience	13.158	12.853	1	4	9	17	139	13,912,677
max. experience	13.884	11.981	1	6	11	17	139	1,223,610
tenure	26.896	21.943	1	10	21	39	139	75,179

Table B2: The Effect of Average Experienced Returns

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \beta \bar{r}_{i,j,t} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t}$ is the standardised equal-weighted average experienced return, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.103***	0.103***	0.105***	0.149***	0.148***	0.151***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
N	1,270,823	1,270,823	1,270,823	2,856,830	2,856,830	2,856,830
R^2	0.781	0.765	0.773	0.709	0.692	0.695
Within-R^2	0.006	0.006	0.006	0.009	0.009	0.009
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B3: The Effect of Experienced Returns - Five Buckets

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{5} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1,2,3,4,5\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager-stock pair must have at least 5 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.276***	0.287***	0.272***	0.275***	0.273***	0.281***
	(0.008)	(0.013)	(0.008)	(0.006)	(0.007)	(0.006)
β_2	0.134***	0.132***	0.136***	0.134***	0.132***	0.136***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.004)
β_3	0.041***	0.043***	0.040***	0.042***	0.042***	0.046***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)
β_4	0.073***	0.073***	0.077***	0.075***	0.072***	0.078***
	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)
β_5	0.238***	0.237***	0.241***	0.238***	0.237***	0.237***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
N	796,021	796,021	796,021	1,958,072	1,958,072	1,958,072
R^2	0.798	0.786	0.792	0.720	0.705	0.708
Within-R^2	0.042	0.043	0.043	0.043	0.042	0.042
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B4: The Effect of Experienced Returns - Ten Buckets

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{10} \beta_q \bar{r}_{i,j,t\in\Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t\in\Delta T^q_{i,j,t}}$, $q \in \{1, 2, ..., 10\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager-stock pair must have at least 10 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.276***	0.293***	0.258***	0.271***	0.290***	0.268***
	(0.012)	(0.039)	(0.010)	(0.010)	(0.017)	(0.007)
β_2	0.147***	0.163***	0.141***	0.149***	0.157***	0.148***
	(0.008)	(0.023)	(0.009)	(0.006)	(0.009)	(0.005)
β_3	0.100***	0.102***	0.098***	0.100***	0.102***	0.096***
	(0.006)	(0.011)	(0.006)	(0.004)	(0.005)	(0.004)
β_4	0.060***	0.058***	0.067***	0.059***	0.066***	0.061***
	(0.006)	(0.008)	(0.006)	(0.004)	(0.004)	(0.003)
β_5	0.028***	0.023***	0.021***	0.029***	0.030***	0.025***
	(0.005)	(0.008)	(0.005)	(0.003)	(0.003)	(0.003)
β_6	0.022***	0.024***	0.019***	0.021***	0.027***	0.024***
	(0.004)	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)
β_7	0.020***	0.027***	0.026***	0.024***	0.020***	0.023***
	(0.004)	(0.005)	(0.004)	(0.002)	(0.002)	(0.003)
β_8	0.043***	0.045***	0.040***	0.045***	0.046***	0.046***
	(0.004)	(0.006)	(0.005)	(0.002)	(0.003)	(0.003)
β_9	0.080***	0.088***	0.087***	0.086***	0.088***	0.087***
	(0.006)	(0.007)	(0.004)	(0.004)	(0.003)	(0.003)
β_{10}	0.206***	0.204***	0.206***	0.208***	0.216***	0.215***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
N	442,353	442,353	442,353	1,073,779	1,073,779	1,073,779
R^2	0.824	0.812	0.820	0.750	0.736	0.738
Within-R^2	0.039	0.041	0.039	0.039	0.042	0.039
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B5: The Effect of Experienced Returns by Number of Managers

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_{q,n} \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1,2,3,4,5\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Each column reports the results for the sub-sample of managers working in a team of $n_{i,t} \in \{1,2,3,4 \text{ or more}\}$, members at time t. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. The first four columns report results for measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; the last four columns report results for measure (4), using sample covariance matrices $\hat{\Sigma}^1_t$ and including zero weights on stocks that belong to the manager's investment universe.

	Expected Returns							
	(1)				(4)			
Nr. Managers	1	2	3	≥ 4	1	2	3	≥ 4
β_1	0.276***	0.114***	0.006**	0.015***	0.275***	0.114***	0.006***	0.005***
	(0.008)	(0.014)	(0.003)	(0.003)	(0.006)	(0.008)	(0.002)	(0.002)
β_2	0.133***	0.053***	0.004**	0.008***	0.134***	0.047***	0.004***	0.001
	(0.005)	(0.005)	(0.002)	(0.002)	(0.003)	(0.004)	(0.001)	(0.002)
β_3	0.040***	0.011***	0.004**	0.006***	0.041***	0.010***	0.006***	0.003**
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)
β_4	0.072***	0.014***	0.000	0.001	0.074***	0.015***	0.003***	0.001
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
β_5	0.237***	0.017***	0.002**	0.001	0.237***	0.019***	0.004***	0.001
	(0.004)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
N	796,021	580,367	1,000,968	790,078	1,958,072	1,455,284	2,773,180	2,181,406
R^2	0.798	0.912	0.991	0.989	0.720	0.866	0.984	0.978
Within-R^2	0.042	0.002	0.000	0.001	0.043	0.003	0.001	0.000
$w_{i,j,t} = 0$	No	No	No	No	Yes	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B6: Managers Who Have Switched Funds - Five Buckets

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{5} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1,2,3,4,5\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager must have in his current investment universe a stock that he has previously held in a different fund. A manager-stock pair must have at least 5 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.224***	0.272***	0.209***	0.240***	0.214***	0.250***
	(0.018)	(0.031)	(0.023)	(0.014)	(0.014)	(0.022)
β_2	0.133***	0.116***	0.112***	0.125***	0.117***	0.124***
	(0.015)	(0.016)	(0.015)	(0.009)	(0.009)	(0.011)
β_3	0.048***	0.046***	0.031***	0.063***	0.049***	0.066***
	(0.011)	(0.014)	(0.011)	(0.008)	(0.007)	(0.009)
β_4	0.066***	0.071***	0.051***	0.078***	0.065***	0.073***
	(0.010)	(0.010)	(0.011)	(0.007)	(0.006)	(0.007)
β_5	0.199***	0.199***	0.202***	0.216***	0.219***	0.211***
	(0.011)	(0.013)	(0.013)	(0.007)	(0.007)	(0.009)
N	110,037	110,037	110,037	225,676	225,676	225,676
R^2	0.892	0.885	0.889	0.843	0.834	0.842
Within-R^2	0.034	0.038	0.034	0.040	0.040	0.040
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B7: Managers Who Have Switched Funds - Ten Buckets

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{10} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1, 2, ..., 10\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager must have in his current investment universe a stock that he has previously held in a different fund. A manager-stock pair must have at least 10 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.246***	0.253***	0.252***	0.281***	0.212***	0.198***
	(0.026)	(0.026)	(0.027)	(0.055)	(0.017)	(0.021)
β_2	0.129***	0.148***	0.127***	0.154***	0.124***	0.116***
	(0.018)	(0.018)	(0.020)	(0.023)	(0.011)	(0.013)
β_3	0.071***	0.102***	0.106***	0.101***	0.097***	0.071***
	(0.015)	(0.014)	(0.018)	(0.011)	(0.010)	(0.010)
β_4	0.054***	0.066***	0.074***	0.060***	0.055***	0.042***
	(0.015)	(0.012)	(0.017)	(0.009)	(0.008)	(0.009)
β_5	0.027**	0.035***	0.030***	0.040***	0.029***	0.022***
	(0.013)	(0.011)	(0.012)	(0.008)	(0.007)	(0.008)
β_6	0.026**	0.025**	0.015	0.029***	0.012*	0.023***
	(0.011)	(0.011)	(0.013)	(0.007)	(0.007)	(0.006)
β_7	0.011	0.019*	0.013	0.027***	0.020***	0.027***
	(0.011)	(0.010)	(0.011)	(0.006)	(0.007)	(0.007)
β_8	0.044***	0.031***	0.033**	0.056***	0.040***	0.038***
	(0.012)	(0.011)	(0.013)	(0.007)	(0.006)	(0.007)
β_9	0.090***	0.073***	0.085***	0.084***	0.086***	0.077***
	(0.012)	(0.013)	(0.012)	(0.008)	(0.007)	(0.007)
β_{10}	0.183***	0.169***	0.180***	0.195***	0.193***	0.200***
	(0.014)	(0.014)	(0.016)	(0.010)	(0.009)	(0.009)
N	78,920	78,920	78,920	160,237	160,237	160,237
R^2	0.914	0.915	0.914	0.869	0.865	0.867
Within-R^2	0.038	0.037	0.039	0.044	0.040	0.039
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B8: Learning Parameters

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \beta\left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{j,t+1-k}\right) + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $r_{j,t+1-k}$ is the realised return of stock j from time $t-k$ to $t+1-k$, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Weights are represented by the following functional form : $\omega_{i,j,t,k} = \frac{(T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}$. Clustered standard errors are in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.146***	0.139***	0.144***	0.205***	0.205***	0.207***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
λ_1	-1.901***	-1.838***	-1.873***	-1.663***	-1.700***	-1.683***
	(0.068)	(0.064)	(0.064)	(0.034)	(0.038)	(0.035)
λ_2	-1.659***	-1.487***	-1.563***	-1.574***	-1.610***	-1.590***
	(0.108)	(0.116)	(0.108)	(0.053)	(0.061)	(0.053)
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B9: Risk Aversion - Pooled Regressions

The table reports the parameter estimates obtained from the following pooled regression: $r_{j,t+1} - r_f = \alpha + \gamma(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \boldsymbol{w}^*_{i,t})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. α is the pooled estimated bias across managers and time, γ is the pooled estimated risk aversion across managers and time. Standard errors are clustered at the manager-time and stock-time level and reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
α	0.011***	0.011***	0.011***	0.010***	0.010***	0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
γ	0.915***	0.999***	0.958***	1.204***	1.283***	1.255***
	(0.079)	(0.082)	(0.080)	(0.077)	(0.079)	(0.078)
N	5,383,850	5,383,850	5,383,850	12,545,295	12,545,295	12,545,295
R^2	0.004	0.004	0.004	0.006	0.006	0.006
$w_{ijt} = 0$	No	No	No	Yes	Yes	Yes
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B10: Risk Aversion and Bias - Summary Statistics

The table reports the summary statistics of the parameter estimates $\hat{\alpha}_i$ and $\hat{\gamma}_i$ obtained by running one regression per manager with the following specification: $r_{j,t+1} - r_f = \alpha_i + \gamma_i (\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \boldsymbol{w}^*_{i,t})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. The reported results are obtained under measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights.

	$\hat{\alpha}_i$	$\hat{\gamma}_i$
mean	0.007	1.236
standard deviation	0.068	5.850
median	0.010	1.117
min	-0.676	-44.666
max	0.736	48.631
skewness	-0.626	1.075
kurtosis	27.395	13.200

B.2 Figures

Figure B1: Explained R^2

The figure reports the fraction of variation in expected excess returns explained by various fixed effects. For (1), (2) and (3) we report the R^2 of the following regression $\mu_{i,j,t} - r_f = H_k + \epsilon_{i,j,t}$. (1) reports results for manager fixed effects, i.e., $H_k = H_i$; (2) for stock fixed effects $H_k = H_j$; (3) for time fixed effects $H_k = H_t$. (4) reports the R^2 for separate manager, stock and time fixed effects, i.e., $\mu_{i,j,t} - r_f = H_i + H_j + H_t + \epsilon_{i,j,t}$. (5) reports the results for manager-time and stock fixed effects, i.e., $\mu_{i,j,t} - r_f = H_{i,t} + H_j + \epsilon_{i,j,t}$. (6) reports the results for manager and stock-time fixed effects, i.e., $\mu_{i,j,t} - r_f = H_i + H_{j,t} + \epsilon_{i,j,t}$. (7) reports the results for manager-time and stock-time fixed effects, i.e., $\mu_{i,j,t} - r_f = H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$.

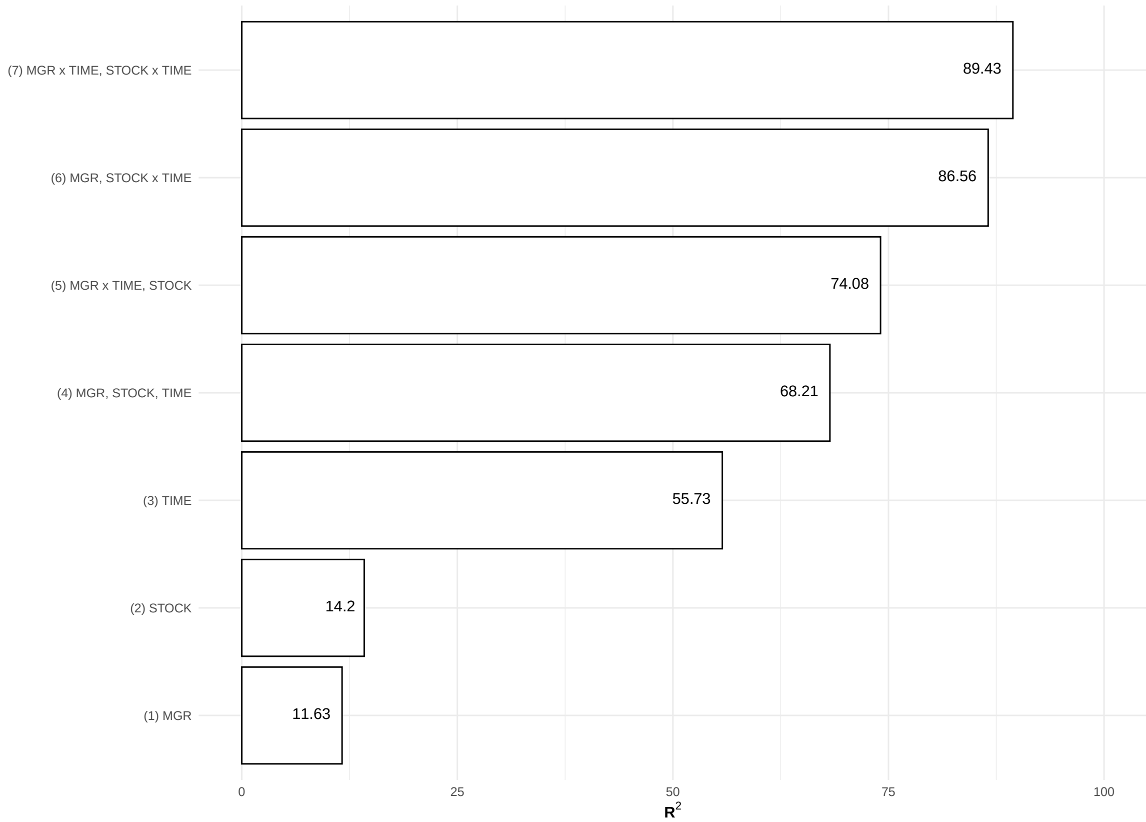


Figure B2: Managers' Careers

The upper panel shows the distribution of starting date for the managers' careers, as the first date we can track the manager in sample. The bottom panel shows the distribution of tenure across managers and dates as the difference between the current date and the starting date in quarters.

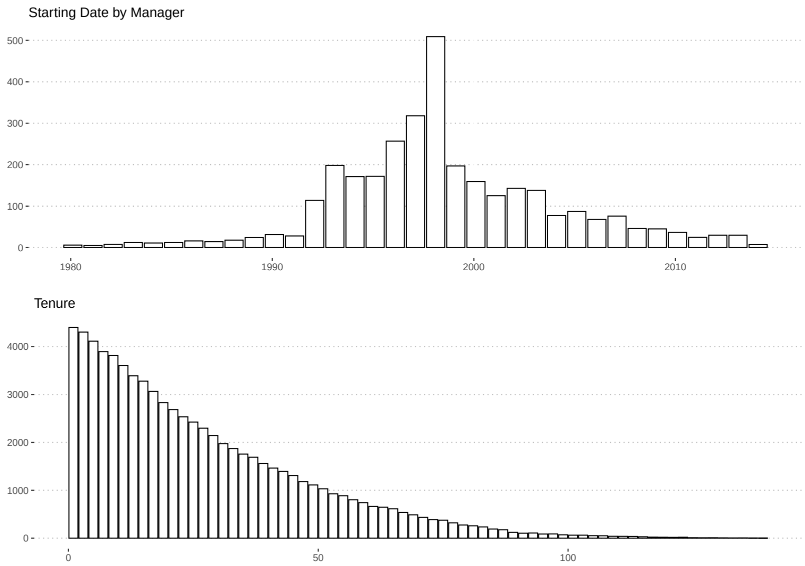


Figure B3: Stock-Manager Experience

The upper panel depicts the starting date of each manager-stock pair $t_{i,j,0}$, as the first date in which we observe a certain manager i holding a certain stock j. The middle panel shows the distribution of stock-manager experience, i.e., for any date t, manager i and stock j experience$_{i,j,t} = t - t_{i,j,0}$. The bottom panel reports the distribution of the maximal experience achieved for each manager-stock pair, i.e., max. experience$_{i,j} = \max_t\{\text{experience}_{i,j,t}\}$.

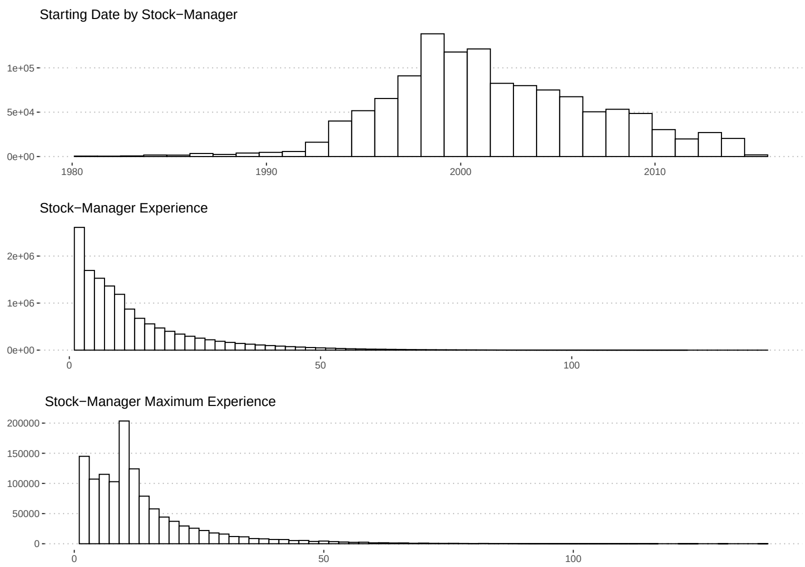


Figure B4: Weights on Past Experience

The figure reports the parameter estimates for β_q obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$ is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. The upper panel reports the results for $Q = 5$, while the bottom panel for $Q = 10$. To be included in the upper panel, a manager-stock pair must have at least 5 quarters of experience, while 10 quarters are needed for the bottom panel. Measures (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; measures (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Measures (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, measures (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and measures (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

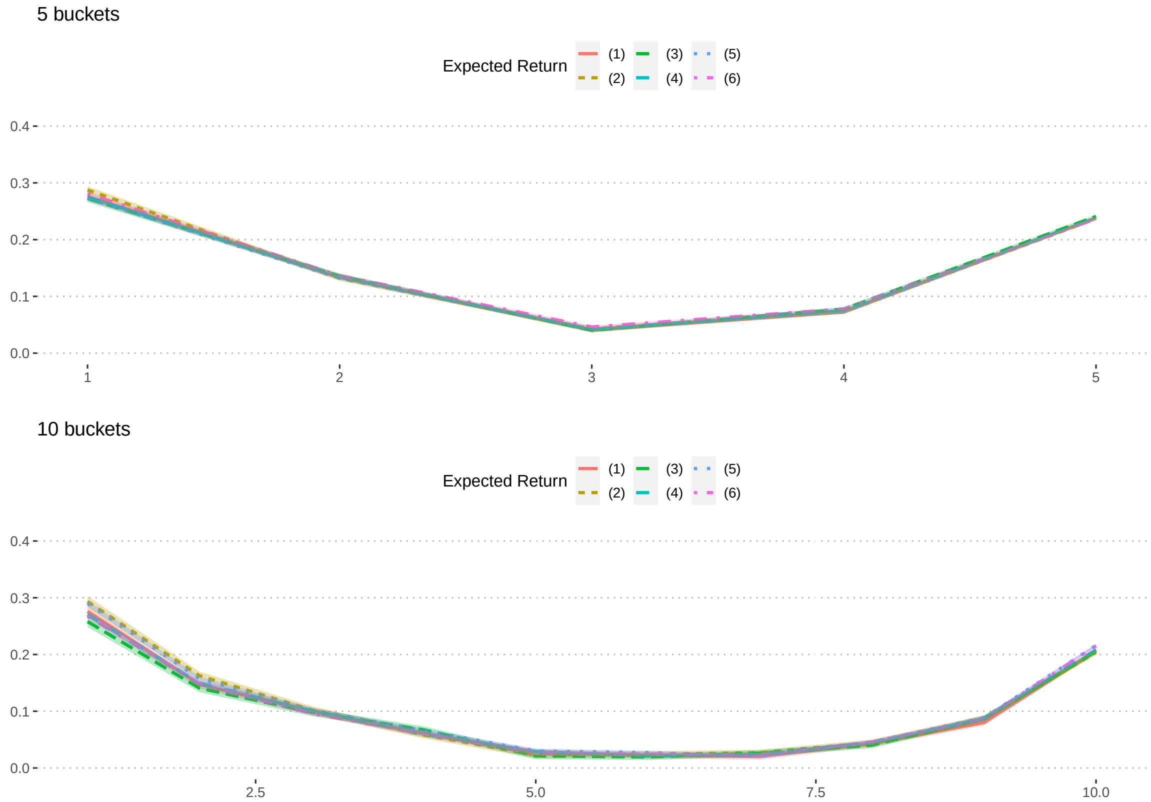


Figure B5: Weights on Past Experience by Number of Managers

The figure reports the parameter estimates for $\beta_{q,n}$ obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_{q,n} \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. The horizontal axis refers to q, while each line to $n_{i,t} \in \{1, 2, 3, 4 \text{ or more}\}$. The top row reports the results for $Q = 5$, the bottom for $Q = 10$. The left column plots coefficients for measure (1), namely expected excess returns are computed without $w_{i,j,t} = 0$ and using the sample covariance matrix $\hat{\Sigma}^1_t$; the right column for measure (4), namely expected excess returns are computed with $w_{i,j,t} = 0$ and using the sample covariance matrix $\hat{\Sigma}^1_t$.

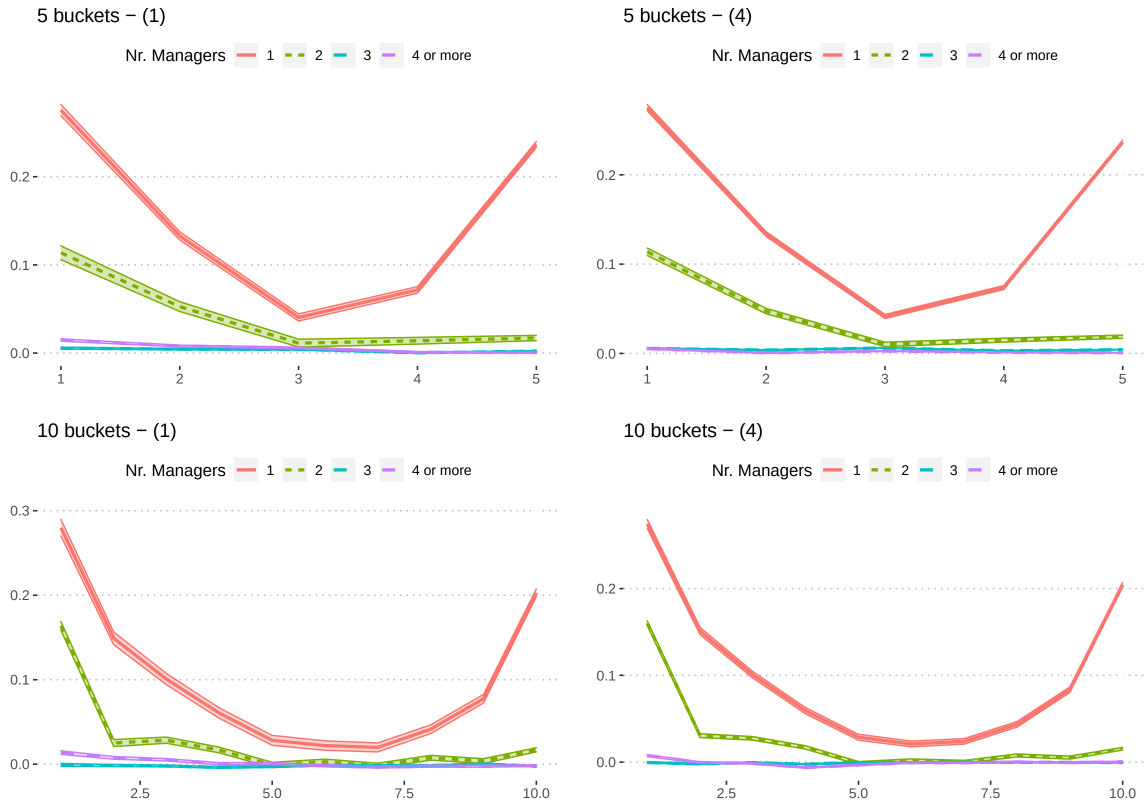


Figure B6: Weights on Past Experience - Managers Who Have Switched Funds

The figure reports the parameter estimates for β_q obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$ is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. The upper panel reports the results for $Q = 5$, while the bottom panel for $Q = 10$. To be included, a manager must have in his current investment universe a stock that he has previously held in a different fund. In the upper panel, manager-stock pairs have at least 5 quarters of experience, while 10 quarters are needed for the bottom panel. Measures (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; measures (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Measures (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, measures (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and measures (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_{i,t}$.

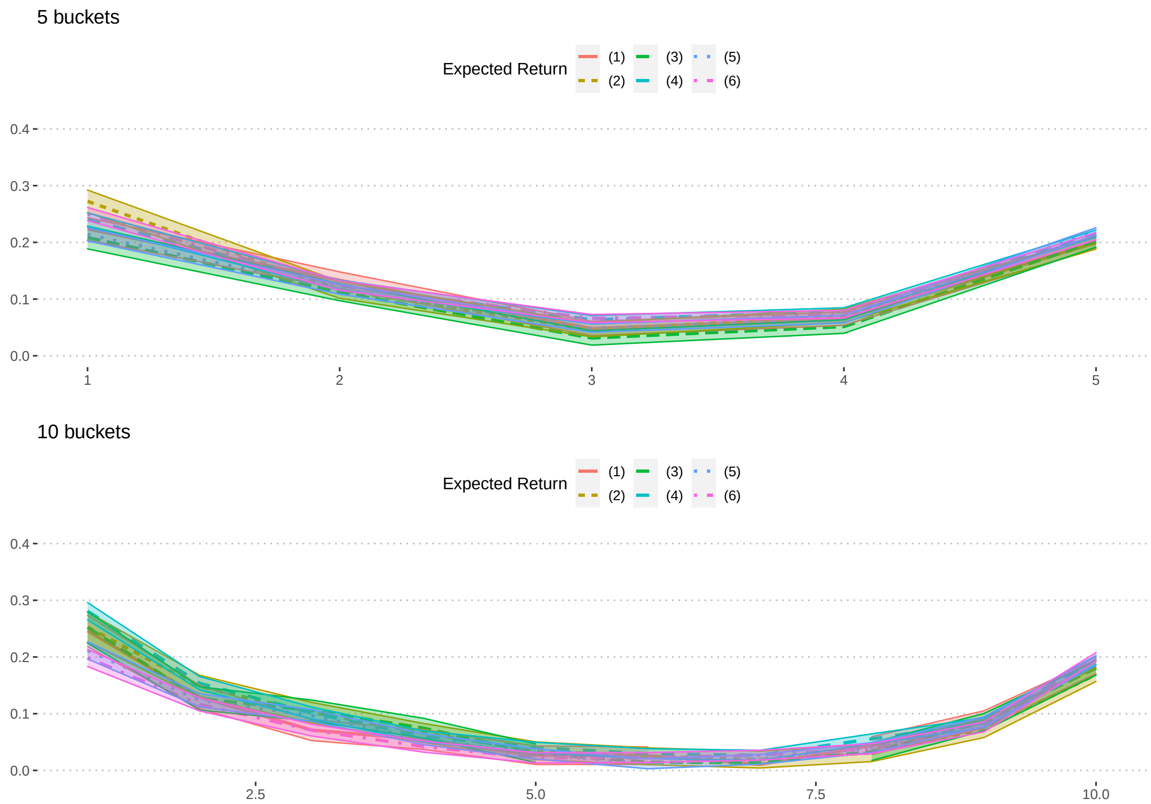


Figure B7: Weighting Functions - Various Examples

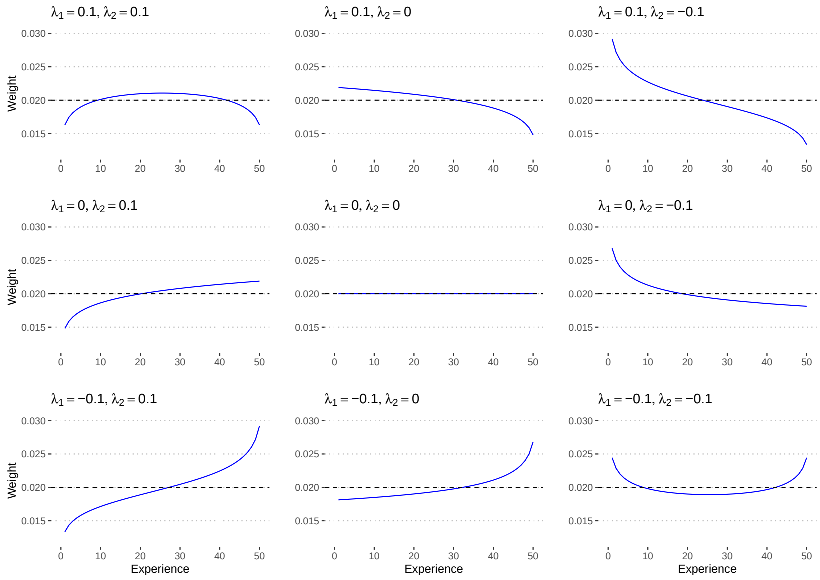


Figure B8: Empirical Weighting Function

The figure plots the weights implied by the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \beta \left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{j,t+1-k} \right) + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t according to measure (1), $r_{j,t+1-k}$ is the realised return of stock j from time $t-k$ to $t+1-k$, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Weights are represented by the following functional form : $\omega_{i,j,t,k} = \frac{(T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}$. The upper panel reports weights for a manager with stock-specific experience of 9 quarters and the lower for 13 quarters.

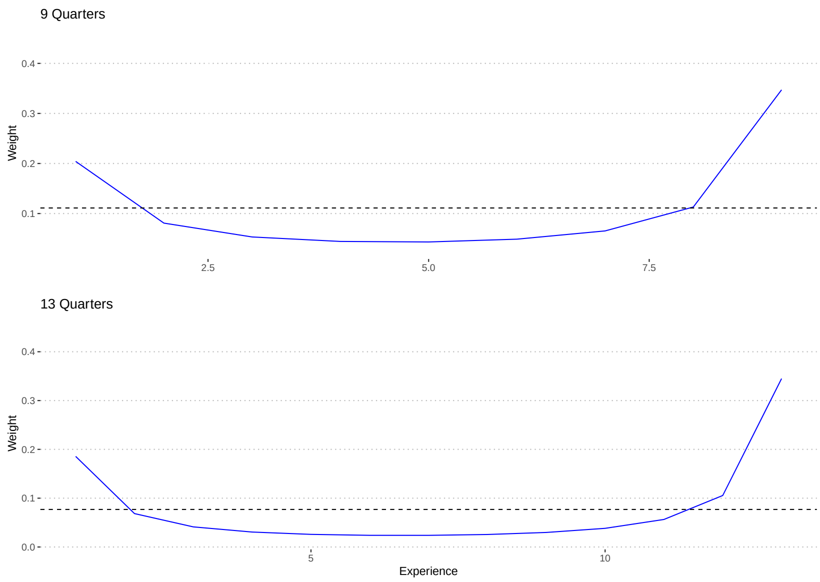


Figure B9: Bias and Risk Aversion

The figure shows the empirical distribution of the parameter estimates $\hat{\alpha}_{i,t}$ and $\hat{\gamma}_{i,t}$ obtained by running one regression per manager with the following specification: $r_{j,t+1} - r_f = \alpha_i + \gamma_i(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \boldsymbol{w}^*_{i,t})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. The dashed lines represent the median bias and risk aversion, respectively. The histograms are trimmed for outliers.

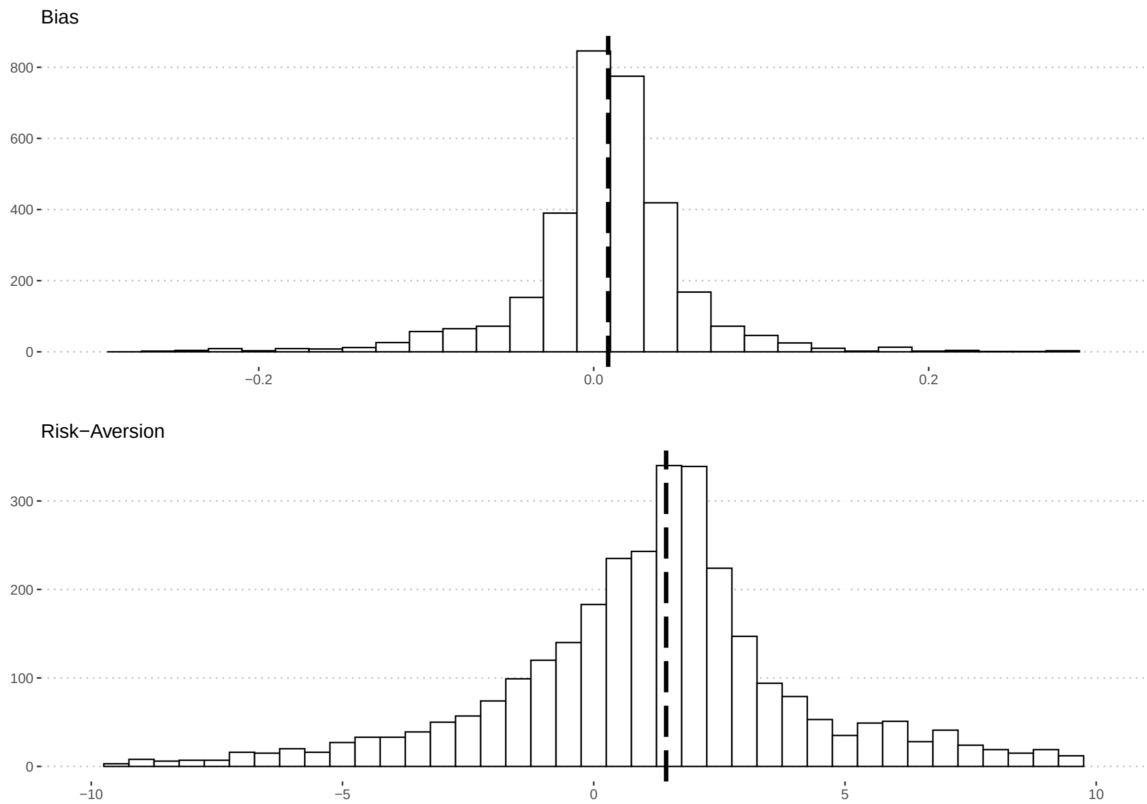


Figure B10: Bias and Risk Aversion by Tenure

The figure plots the parameter estimates $\hat{\alpha}_\tau$ and $\hat{\gamma}_\tau$ obtained by running one regression per tenure τ with the following specification: $r_{j,t+1} - r_f = \alpha_\tau + \gamma_\tau(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \boldsymbol{w}^*_{i,t})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. Bias is the estimated parameter $\hat{\alpha}_\tau$, while Risk Aversion is the estimated parameter $\hat{\gamma}_\tau$. Tenure is measured in quarters since the first observation where we can identify the manager. The shaded grey area covers two standard deviations around the point estimate. The left panel reports results for measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; the right panel reports results for measure (4), using sample covariance matrices $\hat{\Sigma}^1_t$ and including zero weights on stocks that belong to the manager's investment universe.

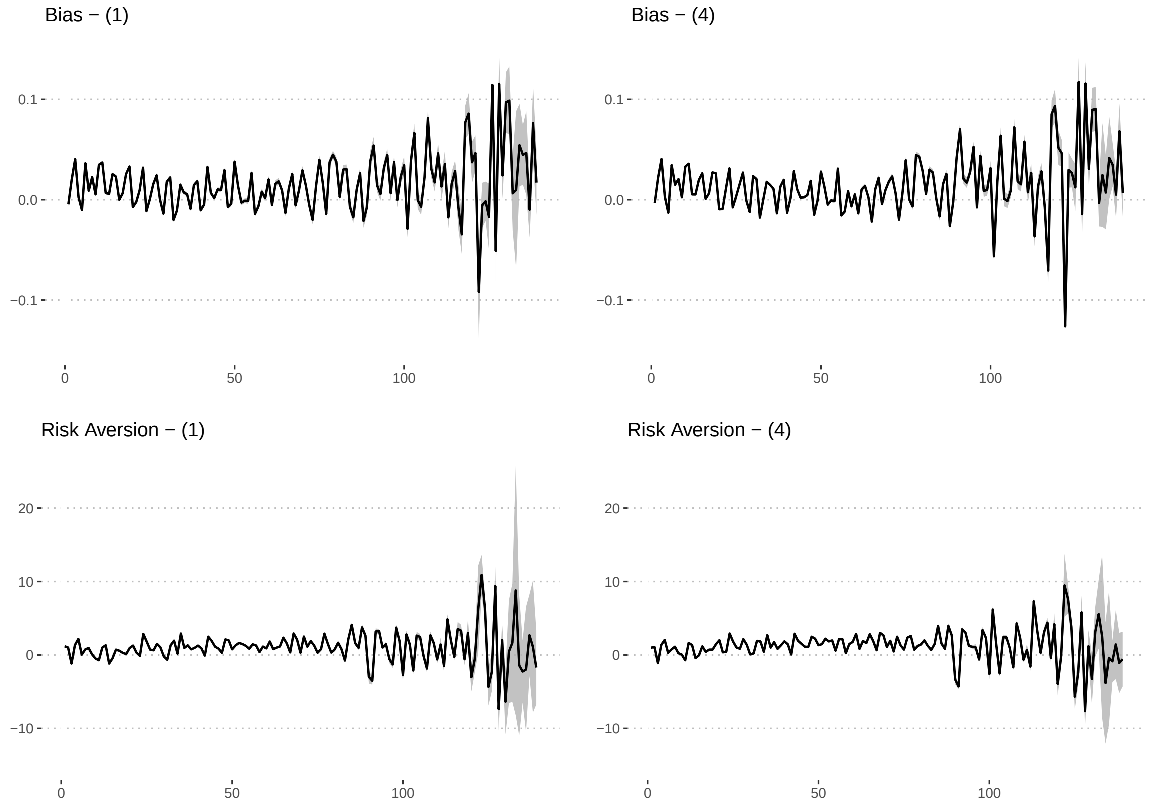
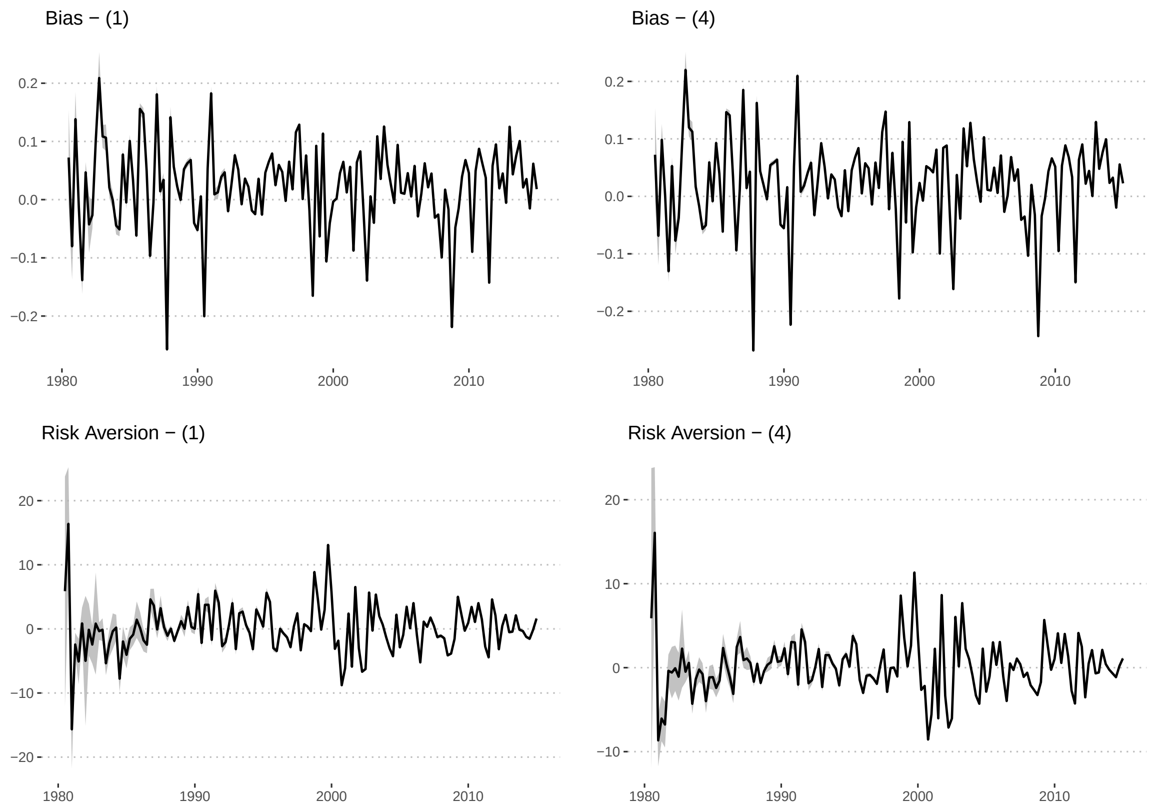


Figure B11: Bias and Risk Aversion by Date

The figure plots the parameter estimates $\hat{\alpha}_t$ and $\hat{\gamma}_t$ obtained by running one regression per date with the following specification: $r_{j,t+1} - r_f = \alpha_t + \gamma_t(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \boldsymbol{w}^*_{i,t})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. Bias is the estimated paramater $\hat{\alpha}_t$, while Risk Aversion is the estimated parameter $\hat{\gamma}_t$. The shaded grey area covers two standard deviations around the point estimate. The left panel reports results for measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; the right panel reports results for measure (4), using sample covariance matrices $\hat{\Sigma}^1_t$ and including zero weights on stocks that belong to the manager's investment universe.



B.3 Optimal Portfolio Choice

In what follows we provide four examples of optimal portfolio choice and describe how we can (or cannot) achieve identification of beliefs. We first look at an investor facing borrowing constraints, second an investor facing short sale constraints, third we look at an investor worried about model misspecification and, finally, an investor who is tracking a benchmark. We show that we can identify beliefs in the first three cases, while the last one requires us to make additional assumptions.

B.3.1 Borrowing Constraint

We follow the approach of Cvitanic and Karatzas (1992), Xu and Shreve (1992) and Tepla (2000). There exists a standard filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,\infty)}, \mathbb{P})$ where all the *usual* regularity conditions are satisfied. We assume that the investor maximises his expected utility over terminal wealth $\mathbb{E}_0[U(W_T)]$. Returns follow a geometric Brownian motion and the investor faces a borrowing constraint. He solves the following problem:

$$\sup_{\{\boldsymbol{w_s}\}_{s\in[0,T]}} \mathbb{E}_0\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right] \quad \text{s.t.} \tag{B.1}$$

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{B.2}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t} dt + \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} \tag{B.3}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w_t}'\left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\boldsymbol{1}\right) \tag{B.4}$$

$$\boldsymbol{w_t}'\boldsymbol{1} \leq k \tag{B.5}$$

where B_t is the price of a risk-free bond, $\boldsymbol{S_t}$ is a vector of stock prices, $\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \left[\frac{dS_{1,t}}{S_{1,t}}, \dots, \frac{dS_{j,t}}{S_{j,t}}, \dots, \frac{dS_{N,t}}{S_{,N,t}}\right]'$, r_f is the instantaneous risk-free rate, $\boldsymbol{\mu_t}$ is the vector of stock return drifts, $\boldsymbol{w_t}$ is the vector of stock portfolio weights, $\Sigma_t^{\frac{1}{2}}$ is the matrix of instantaneous loadings on the Brownian motion processes $\boldsymbol{Z_t}$, $\boldsymbol{1}$ is a vector of ones and k is a real number. Cvitanic and Karatzas (1992) show that the problem in (B.1)-(B.5) is equivalent to an unconstrained problem with modified drifts, i.e., where (B.2) and (B.3) are replaced by:

$$\frac{dB_t}{B_t} = (r_f + \delta(\boldsymbol{v_t}))dt \tag{B.6}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = (\boldsymbol{\mu_t} + \boldsymbol{v_t} + \delta(\boldsymbol{v_t})\boldsymbol{1})dt + \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} \tag{B.7}$$

where the support function $\delta(\boldsymbol{x}) = \sup_{\boldsymbol{w}'\mathbf{1}\leq k}(-\boldsymbol{w}'\boldsymbol{x})$, $\boldsymbol{v_t}$ is such that $\delta(\boldsymbol{v_t}) < \infty$. Cvitanic and Karatzas (1992) show that the optimal $\boldsymbol{v_t^*}$ and portfolio weights $\boldsymbol{w_t^*}$ can be obtained by solving the 'dual' Hamilton-Jacobi-Bellman equation[1]. In particular, the optimal portfolio weights are:

$$\boldsymbol{w_t^*} = \frac{1}{\gamma}\Sigma_t^{-1}(\boldsymbol{\mu_t} - r_f\mathbf{1} - \boldsymbol{v_t^*}) \tag{B.8}$$

where $\boldsymbol{v_t^*} = \arg\min_{\{\boldsymbol{v} \text{ s.t. } \delta(\boldsymbol{v})<\infty\}}\left[||\boldsymbol{\theta_t} + \Sigma_t^{-\frac{1}{2}}\boldsymbol{v_t}||^2 + 2\gamma\delta(\boldsymbol{v_t})\right]$ and $\boldsymbol{\theta_t} = \Sigma_t^{-\frac{1}{2}}(\boldsymbol{\mu_t} - r_f\mathbf{1})$. Tepla (2000) shows that $\boldsymbol{v_t^*} = \bar{v}^*\mathbf{1}$ with $\bar{v}^* = \frac{\gamma(1-\gamma)-\mathbf{1}'\Sigma_t^{-1}(\boldsymbol{\mu_t}-r_f\mathbf{1})}{\mathbf{1}'\Sigma_t^{-1}\mathbf{1}}$ when the borrowing constraint binds, and zero otherwise. Notice that the above result implies that the solution to the constrained optimisation problem is equivalent to that of an unconstrained problem with a risk-free rate shifted by the scalar $\bar{v}^*$. Identification of beliefs is easily achieved in (B.8) by saturating the model with manager-time fixed effects in order to absorb any variation in manager-specific borrowing constraints. Specifically, for each manager *i* solving the above problem, the subjective beliefs can be expressed as:

$$\boldsymbol{\mu_{i,t}} - r_f\mathbf{1} = \gamma_i\Sigma_t\boldsymbol{w_{i,t}^*} + \boldsymbol{H_{i,t}} \tag{B.9}$$

where the manager-time fixed effect is equal to $\boldsymbol{H_{i,t}} = \bar{v}_{i,t}^*\mathbf{1}$.

B.3.2 Short Sale Constraints

The manager solves the following problem[2]:

$$\sup_{\{\boldsymbol{w_s}\}_{s\in[0,T]}} \mathbb{E}_0\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right] \quad \text{s.t.} \tag{B.10}$$

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{B.11}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t}dt + \Sigma_t^{\frac{1}{2}}\boldsymbol{dZ_t} \tag{B.12}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w_t'}\left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1}\right) \tag{B.13}$$

$$-w_{j,t} \leq 0 \quad \forall j = 1, 2, ..., N \tag{B.14}$$

The problem (B.10)-(B.14) can be solved by using Cvitanic and Karatzas (1992) and Xu and Shreve (1992)'s dual approach, similarly to the previous section. The support function now becomes $\delta(\boldsymbol{x}) = \sup_{\{-w_{j,t}\leq 0 \quad \forall j=1,2,...,N\}}(-\boldsymbol{w}'\boldsymbol{x})$. As before, we can

[1] See Sections 12 and 15 of Cvitanic and Karatzas (1992). In particular, see equations (15.1), (15.2) and (15.10).

[2] This problem is similar to the discrete problem analyzed by Koijen and Yogo (2019) as $\gamma \to 1$.

find $\boldsymbol{v}_t^*$ by solving:

$$\min \left[||\boldsymbol{\theta_t} + \Sigma_t^{-\frac{1}{2}} \boldsymbol{v_t}||^2 + 2\gamma\delta(\boldsymbol{v_t}) \right] \text{ s.t.} \tag{B.15}$$

$$-\boldsymbol{v_t} \leq \mathbf{0} \tag{B.16}$$

Denote the vector of Lagrange multipliers on the the constraint in equation (B.16) by $\boldsymbol{\lambda_t} = [\lambda_{1,t}, ..., \lambda_{N,t}]'$. Taking first-order conditions of the above minimisation problem yields:

$$\Sigma_t^{-1}(\boldsymbol{\mu_t} - r_f \mathbf{1} + \boldsymbol{v_t^*}) + \boldsymbol{\lambda_t} = \mathbf{0} \tag{B.17}$$

Consider the following partitions: $\boldsymbol{v_t^*} = [\mathbf{0}' \quad \boldsymbol{v_t^{(2)*'}}]'$, $\boldsymbol{\lambda_t} = [\boldsymbol{\lambda_t^{(1)'}} \quad \mathbf{0}']'$, where we have divided between assets for which the short sale constraint does not bind and those for which it does. We can also partition the vector of expected excess returns and the covariance matrix: $\boldsymbol{\mu_t} - r_f\mathbf{1} = [(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1})' \quad (\boldsymbol{\mu_t^{(2)}} - r_f\mathbf{1})']'$,

$$\Sigma_t = \begin{bmatrix} \Sigma_t^{(1,1)} & \Sigma_t^{(1,2)} \\ \Sigma_t^{(2,1)} & \Sigma_t^{(2,2)} \end{bmatrix},$$

Standard results imply that the inverse of the covariance matrix can be partitioned as:

$$\Sigma_t^{-1} = \begin{bmatrix} \Omega_t^{(1)} & -\Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}\Omega_t^{(2)} \\ -\Sigma_t^{(2,2)-1}\Sigma_t^{(2,1)}\Omega_t^{(1)} & \Omega_t^{(2)} \end{bmatrix}$$

where

$$\Omega_t^{(1)} = \left(\Sigma_t^{(1,1)} - \Sigma_t^{(1,2)}\Sigma_t^{(2,2)-1}\Sigma_t^{(2,1)}\right)^{-1}$$

$$\Omega_t^{(2)} = \left(\Sigma_t^{(2,2)} - \Sigma_t^{(2,1)}\Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}\right)^{-1}$$

Using the above, rewrite equation (B.17) as:

$$\mathbf{0} = \begin{bmatrix} \Omega_t^{(1)}(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}) - \Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}\Omega_t^{(2)}\left(\boldsymbol{\mu_t^{(2)}} - r_f\mathbf{1} + \boldsymbol{v_t^{(2)*}}\right) + \boldsymbol{\lambda_t^{(1)}} \\ -\Sigma_t^{(2,2)-1}\Sigma_t^{(2,1)}\Omega_t^{(1)}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) + \Omega_t^{(2)}\left(\boldsymbol{\mu_t^{(2)}} - r_f\mathbf{1} + \boldsymbol{v_t^{(2)*}}\right) \end{bmatrix} \tag{B.18}$$

Multiplying the second row of (B.18) by $\Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}$ and adding it to the first row allows us to solve for the Lagrange multipliers:

$$\boldsymbol{\lambda_t^{(1)}} = -\Sigma_t^{(1,1)-1}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) \tag{B.19}$$

Insert the multipliers into the first-order condition in equation (B.17) to obtain:

$$\boldsymbol{v_t^*} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{v_t^{(2)*}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \Sigma_t^{(1,1)-1}\Sigma_t^{(2,1)}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) - \left(\boldsymbol{\mu_t^{(2)}} - r_f\mathbf{1}\right) \end{bmatrix} \tag{B.20}$$

We can now substitute $\boldsymbol{v}_t^*$ into equation (B.8) and solve for the optimal weights:

$$\boldsymbol{w}_t^* = \begin{bmatrix} \boldsymbol{w}_t^{(1)*} \\ \mathbf{0} \end{bmatrix} = \frac{1}{\gamma} \begin{bmatrix} \Omega_t^{(1)} \left(\boldsymbol{\mu}_t^{(1)} - r_f \mathbf{1}\right) - \Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}\Omega_t^{(2)} \left(\Sigma_t^{(1,1)-1}\Sigma_t^{(2,1)}(\boldsymbol{\mu}_t^{(1)} - r_f\mathbf{1})\right) \\ -\Sigma_t^{(2,2)-1}\Sigma_t^{(2,1)}\Omega_t^{(1)} \left(\boldsymbol{\mu}_t^{(1)} - r_f \mathbf{1}\right) + \Omega_t^{(2)} \left(\Sigma_t^{(1,1)-1}\Sigma_t^{(2,1)}(\boldsymbol{\mu}_t^{(1)} - r_f\mathbf{1})\right) \end{bmatrix} \tag{B.21}$$

Multiplying the second row by $\Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}$ and adding the two rows together gives the optimal weights on the unconstrained assets:

$$\boldsymbol{w}_t^{(1)*} = \frac{1}{\gamma}\Sigma_t^{(1,1)-1} \left(\boldsymbol{\mu}_t^{(1)} - r_f \mathbf{1}\right) \tag{B.22}$$

Intuitively, the optimisation program of a short sale constrained investor results in an unconstrained portfolio allocation over the set of assets for which the constraint does not bind. For each manager i, identification of beliefs can be achieved by inverting equation (B.22):

$$\boldsymbol{\mu}_{i,t}^{(1)} - r_f \mathbf{1} = \gamma_i \Sigma_t^{(1,1)} \boldsymbol{w}_{i,t}^{(1)*} \tag{B.23}$$

B.3.3 Model Misspecification

In this section we follow the approach of Maenhout (2004) and analyse the behaviour of an investor worried about model misspecification. The investor solves the following problem:

$$J_0 = \sup_{\{\boldsymbol{w_s}, C_s\}} \mathbb{E}_0 \left[\int_0^\infty f(C_s, J_s) ds\right] \quad \text{s.t.} \tag{B.24}$$

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{B.25}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t} dt + \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} \tag{B.26}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w}_t' \left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1}\right) - \frac{C_t}{W_t} dt \tag{B.27}$$

where we are vague about the functional form of the value function. Standard dynamic optimisation arguments yield the following HJB equation:

$$0 = \sup_{\{\boldsymbol{w_t}, C_t\}} \{f(C_t, J_t)dt + \mathbb{E}_t [dJ_t]\} \tag{B.28}$$

Equation (B.28) assumes that the investor is certain about the value of $\mathbb{E}_t [dJ_t]$ and chooses his portfolio accordingly. An investor worried about model misspecification will choose the optimal allocation given the worst-case scenario. Following Anderson et al. (2003), Maenhout (2004) shows that the wealth of the investor under the endogenously chosen model for $u(W_t)$ will evolve according to:

$$dW_t = W_t\left(r_f + \boldsymbol{w_t'}(\boldsymbol{\mu_t} - r_f\mathbf{1}) - \frac{C_t}{W_t}\right)dt + W_t\boldsymbol{w_t'}\Sigma_t^{\frac{1}{2}}\boldsymbol{dZ_t} + W_t^2\boldsymbol{w_t'}\Sigma_t\boldsymbol{w_t}u(W_t)dt \tag{B.29}$$

where $u(W_t)$ is a drift term chosen by the investor to minimise the following expression:

$$u^*(W_t) = \inf_{u_t}\left\{\mathbb{E}_t[dJ_t|u_t] + \frac{1}{2\Psi}u_t^2W_t^2\boldsymbol{w_t'}\Sigma_t\boldsymbol{w_t}dt\right\} \tag{B.30}$$

where $\mathbb{E}_t[dJ_t|u_t]$ is computed under the law of motion in equation (B.29). Among all the models for $u(W_t)$ the investor chooses the least favourable one in terms of its effect on $\mathbb{E}_t[dJ_t|u_t]$, subject to the entropy constraint $\frac{1}{2\Psi}u_t^2W_t^2\boldsymbol{w_t'}\Sigma_t\boldsymbol{w_t}dt$. The HJB equation thus becomes:

$$\begin{aligned} 0 = \sup_{\{\boldsymbol{w_t},C_t\}}\inf_{u_t} \quad & f(C_t,J_t) + \frac{\partial J_t}{\partial t} + J_{W_t}W_t\left(r_f + \boldsymbol{w_t'}(\boldsymbol{\mu_t} - r_f\mathbf{1}) - \frac{C_t}{W_t}\right) + \\ & J_{W_t}W_t^2\boldsymbol{w_t'}\Sigma_t\boldsymbol{w_t}u_t + \frac{1}{2\Psi}u_t^2W_t^2\boldsymbol{w_t'}\Sigma_t\boldsymbol{w_t} + \frac{1}{2}J_{W_tW_t}W_t^2\boldsymbol{w_t'}\Sigma_t\boldsymbol{w_t} \end{aligned} \tag{B.31}$$

The agent will choose $u(W_t)^* = -J_{W_t}\Psi$. The optimal portfolio, therefore, will be:

$$\boldsymbol{w_t^*} = -\frac{J_{W_t}}{[J_{W_tW_t} - J_{W_t}^2\Psi]W_t}\Sigma_t^{-1}(\boldsymbol{\mu_t} - r_f\mathbf{1}) \tag{B.32}$$

An investor concerned about model misspecification will behave like an otherwise identical investor with relative risk aversion of $\gamma_{i,t} = -\frac{[J_{W_{i,t}W_{i,t}} - J_{W_{i,t}}^2\Psi_i]W_{i,t}}{J_{W_{i,t}}}$. In this case, identification follows in a way similar to the standard model presented in the main text.

B.3.4 Benchmarking

In the spirit of van Binsbergen et al. (2008), consider an investor who has his objective function defined over his terminal wealth W_T relative to a benchmark portfolio M_T. He will solve the following problem:

$$J_0 = \sup_{\{\boldsymbol{w_s}\}}\mathbb{E}_0\left[f\left(\frac{W_T}{M_T^\beta}\right)\right] \quad \text{s.t.} \tag{B.33}$$

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{B.34}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t}dt + \Sigma_t^{\frac{1}{2}}\boldsymbol{dZ_t} \tag{B.35}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w_t'}\left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1}\right) \tag{B.36}$$

Assume that the benchmark has weights $\boldsymbol{\theta_t}$ in the N risky assets and therefore evolves according to:

$$\frac{dM_t}{M_t} = \frac{dB_t}{B_t} + \boldsymbol{\theta}_t' \left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1} \right) \tag{B.37}$$

The problem can be recast in terms of the state variable $X_t = \frac{W_t}{M_t^\beta}$ with the following law of motion:

$$\begin{aligned} \frac{dX_t}{X_t} =& ((1-\beta)r_f + (\boldsymbol{w_t} - \beta\boldsymbol{\theta_t})'(\boldsymbol{\mu_t} - r_f\mathbf{1}))dt - \frac{1}{2}\beta(\beta-1)\boldsymbol{\theta}_t'\Sigma_t\boldsymbol{\theta_t}dt + \\ & (\boldsymbol{w_t} - \beta\boldsymbol{\theta_t})'\Sigma_t^{\frac{1}{2}}\boldsymbol{dZ_t} - (\boldsymbol{w_t} - \beta\boldsymbol{\theta_t})'\Sigma_t\beta\boldsymbol{\theta_t}dt \end{aligned} \tag{B.38}$$

If we set up the HJB equation and take first-order conditions, we obtain the optimal weights:

$$\boldsymbol{w}_t^* = -\frac{J_{X_t}}{J_{X_tX_t}X_t}\Sigma_t^{-1}(\boldsymbol{\mu_t} - r_f\mathbf{1}) + \beta\boldsymbol{\theta_t}\left(1 + \frac{J_{X_t}}{J_{X_tX_t}X_t}\right) \tag{B.39}$$

In this case, it is not obvious that we can identify beliefs. However, if there is no variation in the objective function in the cross-section of managers adopting the same benchmark portfolio $\boldsymbol{\theta_t}$, stock-time fixed effects would suffice to recover expectations. Although the above model requires an additional assumption to achieve identification, this is consistent with the common practice of evaluating managers using summary statistics such as CAPM alphas (Berk and van Binsbergen, 2016; Barber et al., 2016). For instance, set $f\left(\frac{W_T}{M_T^\beta}\right) = \frac{1}{1-\gamma}\left(\frac{W_T/W_0}{(M_T/M_0)^\beta}\right)^{1-\gamma} = \frac{1}{1-\gamma}\left(\frac{R_{W,T}}{R_{M,T}^\beta}\right)^{1-\gamma}$. That would be equivalent to solving:

$$\sup_{\{\boldsymbol{w_s}\}} \quad \mathbb{E}_0[r_{W,T}] - \beta\mathbb{E}_0[r_{M,T}] - \frac{(\gamma-1)}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T}) \tag{B.40}$$

where $r_{W,T} = \log W_T/W_0$ and $r_{M,T} = \log M_T/M_0$ are log-returns. The manager is maximising $\alpha = \mathbb{E}_0[r_{W,T}] - \beta\mathbb{E}_0[r_{M,T}]$ subject to the tracking error penalisation $\frac{(\gamma-1)}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T})$[3]. In this case $-\frac{J_{X_t}}{J_{X_tX_t}X_t} = 1/\gamma$ and we could recover beliefs using:

$$\boldsymbol{\mu_{i,t}} - r_f\mathbf{1} = \gamma\Sigma_t\boldsymbol{w}_{i,t}^* + \boldsymbol{H_t} \tag{B.41}$$

Notice that each element of the vector $\boldsymbol{H_t}$ varies at the stock-time level, i.e.: $\boldsymbol{H_t} = (1-\gamma)\beta\Sigma_t\boldsymbol{\theta_t}$.

[3]As it is well known, the agent penalises tracking error for any value of $\gamma > 0$, even for $0 < \gamma \leq 1$. To see this, notice that we can substitute $\mathbb{E}_0[r_{W,T} - \beta r_{M,T}] = \log\mathbb{E}_0\left[\frac{R_{W,T}}{R_{M,T}^\beta}\right] - \frac{1}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T})$ and obtain the following objective:

$$\sup_{\{\boldsymbol{w_s}\}} \quad \log\mathbb{E}_0\left[\frac{R_{W,T}}{R_{M,T}^\beta}\right] - \frac{\gamma}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T})$$

B.4 Data Construction

In this section we provide details on the construction of the data that are used in the paper. We start with the universe of mutual funds in the CRSP database. We remove funds whose manager name clearly does not refer to a person[4]. After having obtained a list of names of managers, we look for cases in which the same manager is spelled differently, e.g. "John Smith", "J. Smith", "J Smith" or just "Smith". To be sure that the pairing is done correctly we proceed in the following way: first, we compute a matrix of distances between names using cosine, Jaccard and Jaro-Winkler methods. We then keep pairs that have a distance below a distance-specific threshold (0.10, 0.17, 0.10 for the cosine, Jaccard and Jaro-Wrinkler methods, respectively) that is set to make sure that we avoid false negatives. We then proceed to manually check over 15,000 pairs to guarantee proper matching with the help of online resources and common sense. After having obtained a list of managers with the dates in which they manage a specific fund, we follow Evans (2010) and Benos et al. (2010) to screen for equity mutual funds. First, if available, we keep funds with the following Lipper class: EIEI, G, LCCE, LCGE, LCVE, MCCE, MCGE, MCVE, MLCE, MLGE, MLVE, SCCE, SCGE, SCVE. We then keep the funds with missing Lipper class and the following Strategic Insight Objective Code: AGG, GMC, GRI, GRO, ING, SCG. If neither of the previous are available, we use the following Wiesenberger Fund Type Codes: G, G-I, AGG, GCI, GRI, GRO, LTG, MCG, and SCG. We then keep all the funds with policy equal to CS. Finally, we remove funds with less than 80% of holdings in common equity, similarly to Kacperczyk et al. (2006). To check for possible mistakes we keep funds with CRSP objective code starting with E and M and remove those starting with EF. This provides us with a manager-by-manager history of the funds managed that we subsequently match with the S12 type1 file from the Thomson-Reuters Institutional Holdings database, using Russ Wermer's MFLinks tables. We then proceed by joining with the S12 type2 and type3 files to obtain a history of holdings.

We continue by adding stock return and balance sheet data using CRSP and Compustat, respectively. From the CRSP Compustat Merged Database we select LinkTypes LU and LC and LinkPrim P and C for stocks with share codes of 10 and 11. After we have merged the two datasets, we compute dividends using CRSP returns and returns not including distributions, similarly to Koijen and Yogo (2019). From Compustat we compute the following quantities: *me* as market equity, *beme* as the book to market equity ratio, *dp* as the ratio between dividends and market prices, *profitability* as the ratio between operating profits and book equity and *investment* as the growth rate of assets similarly to Fama and French (2015).

We then proceed with the construction of the scaled demands $\hat{\Sigma}_t w_{i,t}$. We start

[4]We use various automatic screens like "advisors", "ltd", "limited", etc..., paired with manual inspection.

from CRSP daily return data and compute covariance matrices using the previous year. We compute three daily covariance matrices: $\hat{\Sigma}_t^{d,1}$ which is the sample covariance matrix, and two Bayesian shrinkage estimates. The first one follows Touloumis (2015) and shrinks the daily sample covariance towards a target diagonal matrix with the sample variances on the diagonal, i.e., the resulting estimator is $\hat{\Sigma}_t^{d,2} = \lambda\hat{\Sigma}_t^{d,1} + (1-\lambda)\Sigma_t^{target}$, with $\Sigma_t^{target} = I_N * \hat{\Sigma}_t^{d,1}$, where $*$ denotes the Hadamard product and I_N is an $N \times N$ identity matrix with N being the number of stocks. The third covariance estimator follows Ledoit and Wolf (2004) and shrinks the daily covariance matrix towards a diagonal matrix with the average variance on the diagonal, i.e., $\hat{\Sigma}_t^{d,3} = \lambda\hat{\Sigma}_t^{d,1} + (1-\lambda)\tilde{\Sigma}_t^{target}$, where $\tilde{\Sigma}_t^{target} = \frac{tr(\hat{\Sigma}_t^{d,1})}{N} I_N$, where $tr(\hat{\Sigma}_t^{d,1})$ is the trace of the daily sample covariance matrix. The shrinkage intensity λ is chosen similarly to Touloumis (2015) to minimise the risk function $\mathbb{E}[||\hat{\Sigma}_t^{d,k} - \Sigma_t^d||_F^2]$ where $||S||_F^2 = \frac{tr(S'S)}{dim(S)}$ denotes the Frobenius norm of matrix S, which results in $\lambda = \frac{Y_{2,T} + Y_{1,T}^2}{TY_{2,T} + \frac{N-T+1}{N} Y_{1,T}^2}$, where $Y_{1,T} = \frac{1}{T}\sum_{s=1}^T X_s' X_s - \frac{1}{P_2^T}\sum_{s\neq h} X_h' X_s$, $Y_{2,T} = \frac{1}{P_2^T}\sum_{s\neq h}(X_h' X_s)^2 - 2\frac{1}{P_3^T}\sum_{s\neq h\neq k} X_s' X_h X_s' X_k + \frac{1}{P_4^T}\sum_{s\neq h\neq k\neq w} X_s X_h' X_k X_w'$ with X_j being the vector of stock returns for which we have T observations and $P_a^b = \frac{b!}{(b-a)!}$. Finally, we can scale the matrices $\hat{\Sigma}_t^{d,k}$ by the average number of trading days in a quarter, which in our sample is equal to $\frac{num.obs}{num.quarters} = 63.07$ to obtain our quarterly estimators $\hat{\Sigma}_t^k = \frac{num.obs}{num.quarters} \times \hat{\Sigma}_t^{d,k}$. We can then proceed to compute scaled demands as $\hat{\Sigma}_t^k \boldsymbol{w_{i,t}}$. We compute two vectors of scaled demands for each estimator: one that does not include stocks that currently have zero weights, but belong to the investment opportunity set of the manager, and one that does, i.e., in the first case all the $w_{i,j,t}$ in $\boldsymbol{w_{i,t}}$ are different from zero, while in the second $\boldsymbol{w_{i,t}}$ has some zero elements. The investment opportunity set is constructed similarly to Koijen and Yogo (2019) and includes all stocks that are currently held or have ever been held by the manager in the past 11 quarters.

B.5 Parametric Estimation

As described in Section 2.5, we estimate the model in equation (2.22) via non-linear least squares (NLS). In particular we obtain the coefficients $\hat{\boldsymbol{\theta}} = (\hat{\beta}, \hat{\lambda}_1, \hat{\lambda}_2)'$ by minimising the sum of squared errors:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_i \sum_j \sum_t \left(\mu_{i,j,t} - r_f - \beta \left(\sum_{k=1}^{T_{i,j,t}} \frac{(T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}} r_{j,t+1-k} \right) - H_{i,t} - H_{j,t} \right)^2 \tag{B.42}$$

We perform the minimisation with $(\hat{\lambda}_1, \hat{\lambda}_2) \in [-5, 5] \times [-5, 5]$ via Simulated Annealing and limited-memory BFGS[5]. Fixed effects are partialled out by demeaning $\mu_{i,j,t} - r_f$ and $\left(\sum_{k=1}^{T_{i,j,t}} \frac{(T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}} r_{j,t+1-k} \right)$. To compute standard errors, we can rewrite (B.42) as:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2} \sum_{p=1}^{P} (y_p - \varphi(\boldsymbol{x_p}; \boldsymbol{\theta}))^2 \tag{B.43}$$

where the index p is a short-hand for all the P combinations of i, j, t. We next follow the approach of Davidson and MacKinnon (2001) and recover standard errors using Gauss-Newton Regressions. Consider the 3×1 gradient vector $\Psi(\boldsymbol{x_p}; \boldsymbol{\theta}) = \frac{\partial \varphi(\boldsymbol{x_p}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ and the following regression:

$$y_p - \varphi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}}) = \Psi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}})' \boldsymbol{b} + u_p \tag{B.44}$$

where we regress the residuals $y_p - \varphi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}})$ on the estimated gradient $\Psi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}})$[6]. Denote the $P \times 3$ matrix of gradient observations as $\hat{\Psi} = [\Psi(\boldsymbol{x_1}; \hat{\boldsymbol{\theta}}), ..., \Psi(\boldsymbol{x_P}; \hat{\boldsymbol{\theta}})]'$, then we can estimate the covariance matrix of the coefficients $\boldsymbol{b}$ using the standard clustered "sandwich" estimator:

$$S(\hat{\boldsymbol{b}}) = (\hat{\Psi}'\hat{\Psi})^{-1} \hat{\Psi}' \hat{\Omega} \hat{\Psi} (\hat{\Psi}'\hat{\Psi})^{-1} \tag{B.45}$$

Davidson and MacKinnon (2001) show that the covariance matrix of $\boldsymbol{b}$ in (B.45) is a consistent estimator for the covariance of $\boldsymbol{\theta}$.

[5] Notice that, conditional on λ_1 and λ_2, β can be estimated via OLS and, therefore, is left unconstrained.

[6] For expositional reasons we exclude the estimated fixed effects from $\boldsymbol{\theta}$. Given that they enter linearly in $\varphi(\boldsymbol{x_p}; \boldsymbol{\theta})$, their gradients are identical to the matrix containing the full set of dummies and, therefore, can be taken care of by including dummies on the right hand side of (B.44) or by demeaning.

B.6 Additional Tables and Figures

Table B11: The Effect of Average Experienced Returns

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.148***	0.140***	0.146***	0.188***	0.179***	0.189***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
profitability	-0.002	-0.0010	-0.002	-0.003	-0.002	-0.004
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
investment	0.040***	0.032***	0.035***	0.051***	0.039***	0.042***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
BE/ME	0.012	0.020***	0.012*	0.016*	0.019**	0.017**
	(0.008)	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)
ME	0.011	0.009	0.012	0.009	0.0009	0.008
	(0.015)	(0.012)	(0.013)	(0.018)	(0.016)	(0.017)
D/P	-0.019***	-0.017***	-0.018***	-0.005	-0.006	-0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
N	1,153,333	1,153,333	1,153,333	2,596,853	2,596,853	2,596,853
R^2	0.591	0.583	0.588	0.546	0.538	0.536
Within-R^2	0.016	0.014	0.015	0.021	0.019	0.021
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B12: The Effect of Experienced Returns - Five Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.297***	0.283***	0.286***	0.274***	0.264***	0.277***
	(0.009)	(0.010)	(0.008)	(0.006)	(0.007)	(0.006)
β_2	0.137***	0.129***	0.138***	0.125***	0.115***	0.121***
	(0.009)	(0.008)	(0.008)	(0.005)	(0.005)	(0.005)
β_3	0.061***	0.054***	0.057***	0.055***	0.048***	0.054***
	(0.008)	(0.008)	(0.007)	(0.005)	(0.004)	(0.005)
β_4	0.084***	0.083***	0.088***	0.085***	0.078***	0.084***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
β_5	0.266***	0.259***	0.262***	0.267***	0.258***	0.261***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
profitability	-0.005*	0.0009	-0.005	-0.010**	-0.006	-0.008*
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
investment	0.006	0.003	0.002	0.019***	0.011*	0.013**
	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
BE/ME	0.066***	0.072***	0.062***	0.053***	0.056***	0.054***
	(0.014)	(0.015)	(0.016)	(0.010)	(0.010)	(0.010)
ME	-0.009	-0.007	-0.005	-0.012	-0.017	-0.013
	(0.015)	(0.012)	(0.013)	(0.020)	(0.019)	(0.019)
D/P	-0.008	-0.003	-0.006	0.005	0.003	0.004
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)
N	724,999	724,999	724,999	1,783,648	1,783,648	1,783,648
R^2	0.594	0.587	0.591	0.556	0.547	0.545
Within-R^2	0.066	0.064	0.065	0.070	0.067	0.069
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B13: The Effect of Experienced Returns - Ten Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.259***	0.224***	0.237***	0.235***	0.229***	0.242***
	(0.011)	(0.008)	(0.008)	(0.007)	(0.006)	(0.007)
β_2	0.123***	0.116***	0.131***	0.124***	0.115***	0.120***
	(0.007)	(0.006)	(0.007)	(0.005)	(0.005)	(0.005)
β_3	0.098***	0.094***	0.102***	0.088***	0.083***	0.084***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.004)	(0.004)
β_4	0.078***	0.064***	0.073***	0.069***	0.063***	0.065***
	(0.006)	(0.005)	(0.006)	(0.004)	(0.004)	(0.005)
β_5	0.059***	0.048***	0.049***	0.047***	0.041***	0.043***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
β_6	0.061***	0.053***	0.055***	0.053***	0.053***	0.052***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
β_7	0.067***	0.066***	0.066***	0.066***	0.057***	0.063***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)	(0.004)
β_8	0.074***	0.063***	0.067***	0.071***	0.070***	0.074***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
β_9	0.107***	0.109***	0.113***	0.120***	0.114***	0.112***
	(0.005)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
β_{10}	0.243***	0.239***	0.239***	0.243***	0.247***	0.246***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
profitability	-0.005	-0.003	-0.007	-0.013**	-0.009	-0.011*
	(0.004)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)
investment	-0.015*	-0.011	-0.015*	-0.005	-0.011*	-0.007
	(0.008)	(0.007)	(0.008)	(0.007)	(0.006)	(0.007)
BE/ME	0.076***	0.078***	0.065***	0.069***	0.071***	0.066***
	(0.018)	(0.019)	(0.024)	(0.014)	(0.013)	(0.013)
ME	-0.019	-0.014	-0.011	-0.022	-0.028	-0.021
	(0.016)	(0.015)	(0.015)	(0.023)	(0.021)	(0.022)
D/P	-0.001	-0.003	-0.006	0.008	0.006	0.010
	(0.010)	(0.010)	(0.009)	(0.008)	(0.008)	(0.008)
N	403,968	403,968	403,968	980,175	980,175	980,175
R^2	0.598	0.588	0.596	0.567	0.557	0.555
Within-R^2	0.065	0.061	0.063	0.070	0.070	0.071
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock	Mgr×Time Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B14: The Effect of Experienced Returns - Three Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.283***	0.294***	0.288***	0.284***	0.288***	0.288***
	(0.006)	(0.007)	(0.006)	(0.005)	(0.005)	(0.004)
β_2	0.077***	0.082***	0.078***	0.078***	0.079***	0.080***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
β_3	0.229***	0.231***	0.232***	0.231***	0.231***	0.233***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)
N	1,031,564	1,031,564	1,031,564	2,483,275	2,483,275	2,483,275
R^2	0.777	0.762	0.769	0.704	0.688	0.690
Within-R^2	0.039	0.041	0.040	0.040	0.040	0.040
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B15: The Effect of Experienced Returns - Three Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.280***	0.273***	0.277***	0.273***	0.267***	0.277***
	(0.008)	(0.008)	(0.007)	(0.005)	(0.005)	(0.005)
β_2	0.073***	0.069***	0.071***	0.066***	0.060***	0.066***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.004)	(0.005)
β_3	0.236***	0.230***	0.233***	0.238***	0.229***	0.235***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
profitability	-0.001	-0.000	-0.002*	-0.003	-0.002	-0.004
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
investment	0.019***	0.013*	0.014**	0.032***	0.021***	0.024***
	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)
BE/ME	0.048***	0.056***	0.048***	0.043***	0.044***	0.044***
	(0.010)	(0.011)	(0.011)	(0.009)	(0.009)	(0.008)
ME	-0.002	-0.003	0.000	-0.006	-0.014	-0.007
	(0.014)	(0.012)	(0.012)	(0.019)	(0.017)	(0.017)
D/P	-0.009	-0.007	-0.008	0.004	0.002	0.003
	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
N	937,382	937,382	937,382	2,258,925	2,258,925	2,258,925
R^2	0.582	0.573	0.578	0.545	0.536	0.535
Within-R^2	0.056	0.055	0.056	0.058	0.056	0.059
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B16: The Effect of Experienced Returns - Three Buckets and $k = 4$ Quarters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_2	0.016***	0.015***	0.022***	0.008***	0.008***	0.014***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
β_3	0.161***	0.159***	0.160***	0.166***	0.168***	0.165***
	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
N	618,451	618,451	618,451	1,499,594	1,499,594	1,499,594
R^2	0.812	0.799	0.807	0.744	0.729	0.733
Within-R^2	0.021	0.021	0.021	0.021	0.021	0.020
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B17: The Effect of Experienced Returns - Three Buckets and $k = 4$ Quarters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.208***	0.189***	0.198***	0.206***	0.194***	0.205***
	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
β_2	0.093***	0.083***	0.089***	0.077***	0.070***	0.076***
	(0.008)	(0.008)	(0.007)	(0.005)	(0.005)	(0.005)
β_3	0.215***	0.208***	0.209***	0.229***	0.225***	0.223***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
profitability	-0.005	-0.003	-0.007	-0.013**	-0.010	-0.011*
	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)	(0.006)
investment	-0.002	-0.003	-0.006	0.004	-0.001	0.001
	(0.008)	(0.007)	(0.008)	(0.007)	(0.006)	(0.007)
BE/ME	0.056***	0.059***	0.048***	0.052***	0.051***	0.049***
	(0.013)	(0.015)	(0.018)	(0.012)	(0.011)	(0.011)
ME	-0.006	-0.002	0.001	-0.009	-0.012	-0.007
	(0.016)	(0.014)	(0.015)	(0.022)	(0.020)	(0.021)
D/P	-0.008	-0.009	-0.012	0.004	0.002	0.004
	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
N	564,287	564,287	564,287	1,367,732	1,367,732	1,367,732
R^2	0.598	0.590	0.597	0.570	0.560	0.558
Within-R^2	0.042	0.039	0.040	0.046	0.044	0.045
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B18: The Effect of Experienced Returns - Three Buckets and $k = 8$ Quarters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_2	0.020***	0.017***	0.026***	0.014***	0.020***	0.019***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.003)	(0.002)
β_3	0.137***	0.131***	0.135***	0.136***	0.144***	0.141***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
N	343,058	343,058	343,058	753,526	753,526	753,526
R^2	0.870	0.864	0.866	0.834	0.821	0.824
Within-R^2	0.021	0.020	0.021	0.020	0.022	0.021
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B19: The Effect of Experienced Returns - Three Buckets and $k = 8$ Quarters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.168***	0.149***	0.157***	0.165***	0.152***	0.160***
	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.008)
β_2	0.067***	0.058***	0.065***	0.066***	0.063***	0.065***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
β_3	0.179***	0.173***	0.178***	0.183***	0.187***	0.186***
	(0.006)	(0.007)	(0.005)	(0.005)	(0.004)	(0.004)
profitability	-0.003	-0.003	-0.007	-0.016*	-0.014	-0.014
	(0.004)	(0.005)	(0.005)	(0.008)	(0.009)	(0.009)
investment	-0.029***	-0.027***	-0.029***	-0.029***	-0.035***	-0.032***
	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
BE/ME	0.093***	0.100***	0.092***	0.077***	0.074***	0.069***
	(0.027)	(0.027)	(0.026)	(0.019)	(0.019)	(0.018)
ME	-0.015	-0.007	-0.008	-0.023	-0.027	-0.021
	(0.019)	(0.017)	(0.018)	(0.027)	(0.025)	(0.026)
D/P	-0.001	0.004	-0.007	0.014	0.009	0.015
	(0.010)	(0.011)	(0.010)	(0.011)	(0.011)	(0.011)
N	314,557	314,557	314,557	691,634	691,634	691,634
R^2	0.671	0.661	0.666	0.655	0.644	0.644
Within-R^2	0.034	0.031	0.033	0.036	0.036	0.036
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B20: The Effect of Experienced Returns by Number of Managers

	Expected Returns							
	(1)				(4)			
Nr. Managers	1	2	3	≥ 4	1	2	3	≥ 4
β_1	0.280***	0.164***	-0.001	0.014***	0.275***	0.160***	-0.001	0.008***
	(0.012)	(0.010)	(0.003)	(0.004)	(0.010)	(0.008)	(0.002)	(0.003)
β_2	0.149***	0.025***	-0.002	0.008**	0.151***	0.031***	-0.002	-0.000
	(0.008)	(0.007)	(0.002)	(0.004)	(0.006)	(0.004)	(0.002)	(0.002)
β_3	0.101***	0.028***	-0.002	0.005	0.101***	0.027***	-0.001	-0.001
	(0.006)	(0.004)	(0.002)	(0.003)	(0.004)	(0.003)	(0.001)	(0.002)
β_4	0.060***	0.017***	-0.004**	0.001	0.059***	0.017***	-0.002**	-0.006***
	(0.006)	(0.003)	(0.002)	(0.003)	(0.004)	(0.002)	(0.001)	(0.002)
β_5	0.028***	-0.001	-0.003**	0.001	0.029***	-0.001	-0.000	-0.003**
	(0.005)	(0.002)	(0.001)	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)
β_6	0.022***	0.003	-0.001	-0.002	0.021***	0.002	-0.001	-0.001
	(0.004)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)
β_7	0.020***	-0.002	-0.002*	-0.003*	0.024***	0.000	0.000	-0.001
	(0.004)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
β_8	0.041***	0.008***	-0.002	-0.002	0.044***	0.008***	-0.000	0.000
	(0.004)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
β_9	0.077***	0.004*	0.000	-0.002*	0.083***	0.005***	-0.000	-0.001
	(0.006)	(0.002)	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)
β_{10}	0.203***	0.017***	-0.002***	-0.002	0.204***	0.016***	-0.001	0.001
	(0.005)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
N	442,353	579,965	558,722	428,591	1,073,779	1,454,292	1,524,108	1,158,163
R^2	0.824	0.912	0.993	0.991	0.750	0.867	0.988	0.982
Within-R^2	0.039	0.010	0.000	0.001	0.039	0.012	0.000	0.001
$w_{i,j,t} = 0$	No	No	No	No	Yes	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B21: Learning Parameters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.203***	0.190***	0.198***	0.246***	0.241***	0.251***
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
λ_1	-2.225***	-2.223***	-2.157***	-1.800***	-1.929***	-1.854***
	(0.128)	(0.119)	(0.114)	(0.066)	(0.068)	(0.065)
λ_2	-2.362***	-2.313***	-2.265***	-1.881***	-2.012***	-1.951***
	(0.137)	(0.126)	(0.121)	(0.073)	(0.074)	(0.072)
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr$\times$Time Stock	Mgr$\times$Time Stock	Mgr$\times$Time Stock	Mgr$\times$Time Stock	Mgr$\times$Time Stock	Mgr$\times$Time Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table B22: Risk Aversion and Bias Including Zero Weights - Summary Statistics

	$\hat{\alpha}_i$	$\hat{\gamma}_i$
mean	0.006	1.501
standard deviation	0.056	5.266
median	0.009	1.441
min	-0.431	-43.532
max	0.398	42.658
skewness	-1.111	0.954
kurtosis	13.375	13.727

Figure B12: Weighting Functions - Various Examples

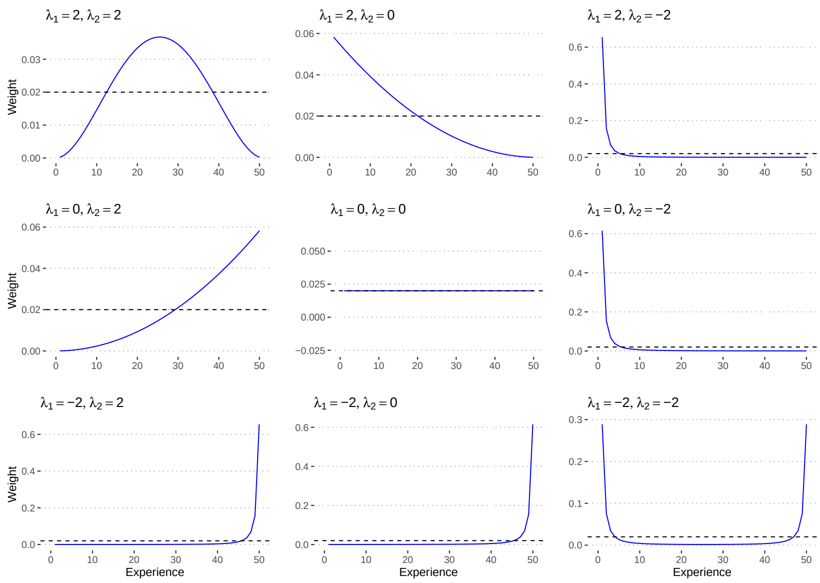
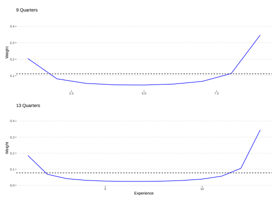
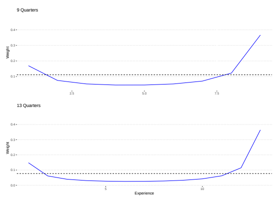


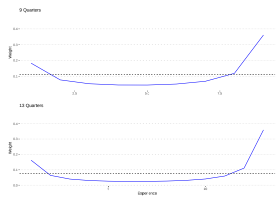
Figure B13: Estimated Weighting Functions - Manager-Time, Stock-Time Fixed Effects



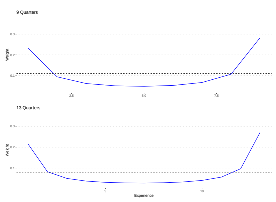
(a) $\hat{\Sigma}_t^1 \boldsymbol{w}_{i,t}^*, w_{i,j,t} > 0$



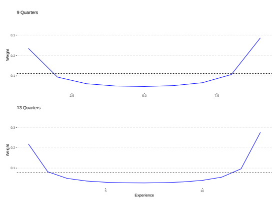
(b) $\hat{\Sigma}_t^2 \boldsymbol{w}_{i,t}^*, w_{i,j,t} > 0$



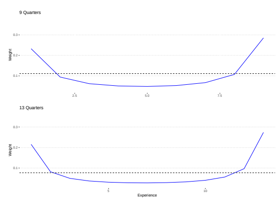
(c) $\hat{\Sigma}_t^3 \boldsymbol{w}_{i,t}^*, w_{i,j,t} > 0$



(d) $\hat{\Sigma}_t^1 \boldsymbol{w}_{i,t}^*, w_{i,j,t} \geq 0$

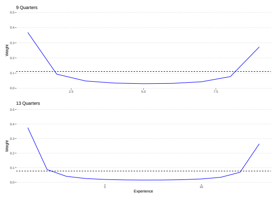


(e) $\hat{\Sigma}_t^2 \boldsymbol{w}_{i,t}^*, w_{i,j,t} \geq 0$

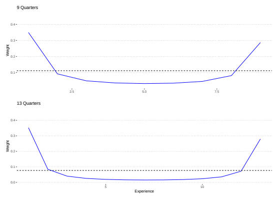


(f) $\hat{\Sigma}_t^3 \boldsymbol{w}_{i,t}^*, w_{i,j,t} \geq 0$

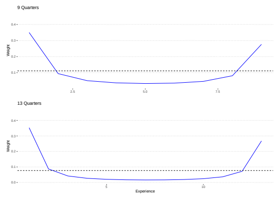
Figure B14: Estimated Weighting Functions - Manager-Time, Stock Fixed Effects



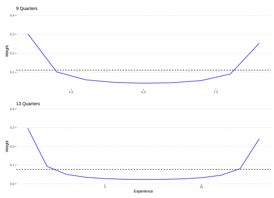
(a) $\hat{\Sigma}_t^1 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} > 0$



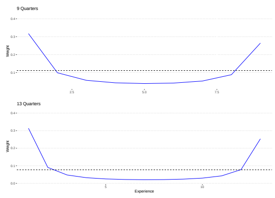
(b) $\hat{\Sigma}_t^2 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} > 0$



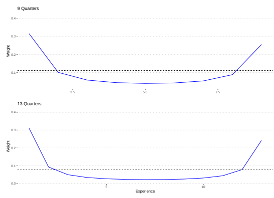
(c) $\hat{\Sigma}_t^3 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} > 0$



(d) $\hat{\Sigma}_t^1 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} \geq 0$

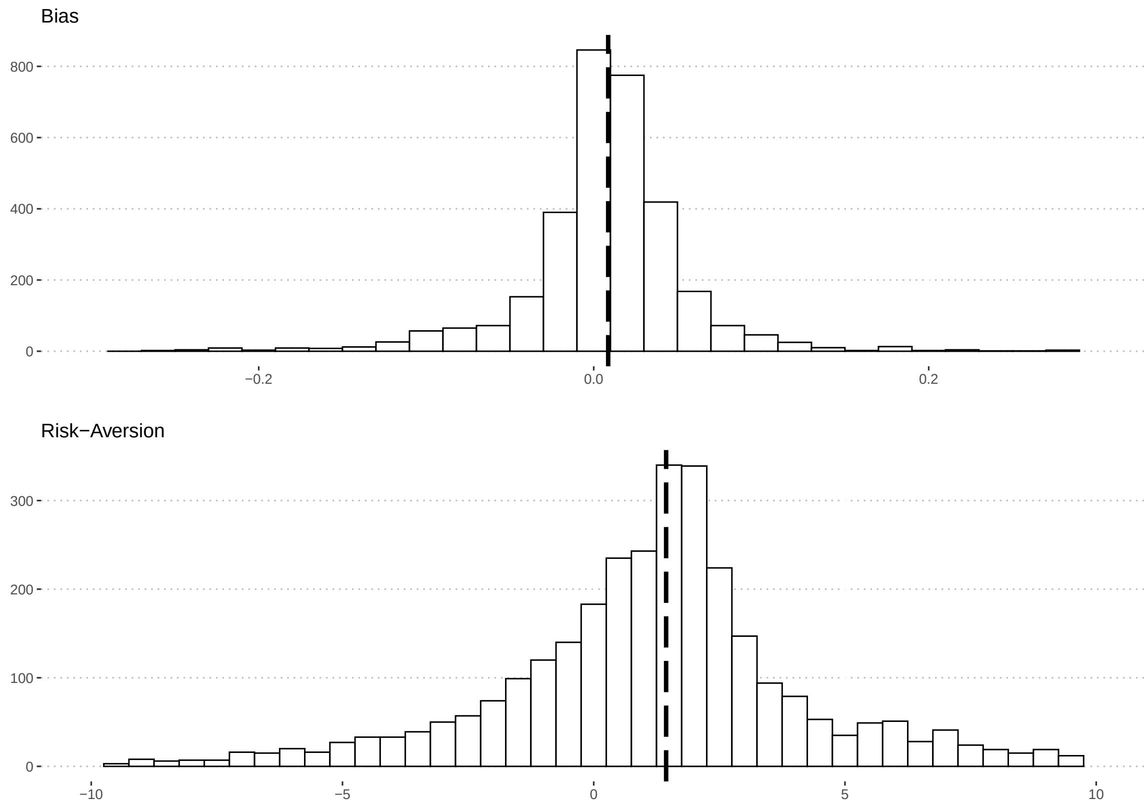


(e) $\hat{\Sigma}_t^2 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} \geq 0$



(f) $\hat{\Sigma}_t^3 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} \geq 0$

Figure B15: Bias and Risk Aversion Including Zero Weights



C. Appendix to Living on the Edge: the Salience of Property Taxes in the UK Housing Market

C.1 Variable Definition

Variable Name	Description
Price	Transaction price for the property as recorded by HM Land Registry
Council Tax	Amount of council tax payable per year
Band	Council tax band. One of: A, B, C, D, E, F, G, H
Year	Calendar year of the transaction
Month	Calendar month of the transaction
Size	Total floor area measured in squared meters
No. Rooms	Number of habitable rooms in the property as defined in the EPC
Property Type	One of: detached, semi-detached or terraced house and flat
Newly-built	Equals 1 if the property is newly-built
Leasehold	Equals 1 if the property is under a leasehold agreement
Energy Cost	Sum of the annual heating, hot water and lighting costs for the property
	One of very low, low, medium, high and very high expenditures
	Baseline = very low
CO_2 Emissions	CO_2 emissions in tonnes/year
	One of very low, low, medium, high and very high
	Baseline = very low
No. Lighting Outlets	Number of fixed lighting outlets in the property, standardised
Energy Rating	A-G energy rating fixed effects with A being the most efficient
Glazed Type	Indicates the type of glazing
	Various categories of single, double or triple glazing according to
	the British Fenestration Rating Council or manufacturer declaration
No. Storeys > 3	Equals 1 if the building has more than 3 storeys
Glazed Area	Estimate of total glazed area of the property
	One of: Normal, Less than Normal, More than Normal
	Baseline = Normal
Fireplaces	Equals 1 if the property has open fireplaces
No. Extensions	Number of extensions added to the property
	One of: 0, 1, 2, 3, 4
Floor Height	Average storey height in metres
	One of: less than 2.3, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3 or more
Built in	Age band when the building was constructed
	One of: before 1949, 1950-1982, 1983-2002, after 2003

Continued on next page

Continued from previous page

Grid ID	An indicator for the grid square in which the property is located
Pair ID	An indicator for the pair of matched properties

C.2 Tables

Table C1: Evidence of Selection

The table shows the estimates of a simple regression of house prices on council tax amounts, namely: $p_{ibdt} = \beta\tau_{bdt} + \delta_{bt} + \zeta' x_{ibdt} + \varepsilon_{ibdt}$ where p_{ibdt} is the price of house i in band b, Borough d at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{bt} are band-year fixed effects; and x_{ibdt} are controls. All columns include band-year and month fixed effects. All other variables are defined in Section C.1. Standard errors double-clustered at the Borough and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	-231.2***	-263.3***	-228.7***	-229.2***
	(71.8)	(86.4)	(78.0)	(78.3)
Size		2,233.7***	2,271.7***	2,270.8***
		(724.4)	(731.2)	(730.9)
Newly-built				14,054.3**
				(5,619.8)
Leasehold				-8,681.7
				(10,801.3)
Fixed-effects				
Band × Year	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Obs.	889,925	889,925	889,925	889,925
R^2	0.530	0.573	0.578	0.578
Within R^2	0.022	0.064	0.058	0.058

Two-way (Borough & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C2: Evidence of Selection - Additional Controls

The table shows the estimates of a simple regression of house prices on council tax amounts, namely: $p_{ibdt} = \beta\tau_{bdt} + \delta_{bt} + \zeta' x_{ibdt} + \varepsilon_{ibdt}$ where p_{ibdt} is the price of house i in band b, Borough d at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{bt} are band-year fixed effects; and x_{ibdt} are controls. All columns include band-year and month fixed effects and control for the property size. All other variables are defined in Section C.1. Standard errors double-clustered at the Borough and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	-255.9***	-225.0**	-259.6***	-220.1**
	(85.6)	(80.0)	(88.2)	(78.2)
Size	2,747.2***	2,266.9***	2,534.4***	2,310.4***
	(911.3)	(734.5)	(780.3)	(784.4)
Energy Cost Low	-26,896.1*			-15,049.6**
	(13,515.0)			(7,107.5)
Energy Cost Medium	-47,312.9**			-24,385.3**
	(22,380.7)			(11,482.6)
Energy Cost High	-69,359.2**			-32,869.7**
	(30,818.3)			(15,291.8)
Energy Cost Very High	-94,269.8*			-39,075.4
	(45,563.6)			(22,987.4)
CO_2 Emisions Low	-17,677.3**			-14,199.3***
	(7,006.9)			(4,857.3)
CO_2 Emissions Medium	-26,558.8**			-23,257.6**
	(11,971.2)			(8,475.0)
CO_2 Emissions High	-36,052.5*			-31,559.2**
	(18,323.2)			(12,521.6)
CO_2 Emissions Very High	-32,523.2			-26,343.9
	(28,385.1)			(17,461.3)
No. Lighting Outlets	20,870.4***			19,659.8***
	(5,833.9)			(5,317.1)
No. Storeys > 3		-3,140.4		632.9
		(5,841.7)		(6,385.2)
Glazed Area Less than Normal		6,923.6		851.3
		(11,930.2)		(10,981.8)
Glazed Area More than Normal		16,669.1***		13,729.2***
		(3,337.1)		(3,490.2)
Fireplaces		42,454.0***		33,624.0***
		(9,985.6)		(9,114.6)
Newly-built			23,567.3***	29,368.6***
			(5,295.0)	(4,958.9)
Leasehold			24,601.4*	-13,104.3
			(13,060.7)	(11,681.6)
Built in 1950-1982			-43,868.7***	-29,435.6***
			(8,169.9)	(5,870.5)
Built in 1983-2002			-22,756.2**	-30,012.4***
			(9,533.2)	(8,919.4)

Continued on next page

Table C2 – *Continued from previous page*

	(1)	(2)	(3)	(4)
Built after 2003			-21,575.1	-31,925.1**
			(13,706.9)	(15,196.7)
Fixed-effects				
Band × Year	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
Energy Rating	Yes	No	No	Yes
Glazed Type	Yes	No	No	Yes
No. Rooms	No	Yes	No	Yes
Property Type	No	Yes	No	Yes
No. Extensions	No	Yes	No	Yes
Floor Height	No	Yes	No	Yes
Obs.	889,925	889,925	889,925	889,925
R^2	0.566	0.580	0.564	0.583
Within R^2	0.095	0.059	0.092	0.063

Two-way (Borough & Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C3: Evidence of Selection - Median Price per Borough, Band, Year

The table shows the estimates of the following regression: $\varepsilon^{med}_{bdt} = \beta\tau_{bdt} + \delta_{bt} + \eta_{bdt}$, where ε^{med}_{bdt} is the median residual price of all houses in band b, Borough d at time t obtained from a hedonic regression of prices on house characteristics; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; and δ_{bt} are band-year fixed effects. The explanatory variables used to computed the hedonic residuals are reported in the panel First-stage controls. All variables are defined in Section C.1. Standard errors double-clustered at the Borough and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	-183.6***	-334.0***	-324.3***	-325.1***
	(56.4)	(84.7)	(83.5)	(83.1)
Fixed-effects				
Band × Year	Yes	Yes	Yes	Yes
First-stage controls				
Month	Yes	Yes	Yes	Yes
Size	No	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Newly-built	No	No	No	Yes
Leasehold	No	No	No	Yes
Obs.	5,014	5,014	5,014	5,014
R^2	0.804	0.501	0.503	0.500
Within R^2	0.055	0.122	0.117	0.118

Two-way (Borough & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C4: Evidence of Selection - Average Price per Borough, Band, Year

The table shows the estimates of the following regression: $\bar{\varepsilon}_{bdt} = \beta\tau_{bdt} + \delta_{bt} + \eta_{bdt}$, where $\bar{\varepsilon}_{bdt}$ is the average residual price of all houses in band b, Borough d at time t obtained from a hedonic regression of prices on house characteristics; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; and δ_{bt} are band-year fixed effects. The explanatory variables used to computed the hedonic residuals are reported in the panel First-stage controls. All variables are defined in Section C.1. Standard errors double-clustered at the Borough and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	-195.6***	-368.4***	-358.6***	-358.9***
	(64.9)	(93.9)	(92.8)	(92.5)
Fixed-effects				
Band $\times$ Year	Yes	Yes	Yes	Yes
First-stage controls				
Month	Yes	Yes	Yes	Yes
Size	No	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Newly-built	No	No	No	Yes
Leasehold	No	No	No	Yes
Obs.	5,014	5,014	5,014	5,014
R^2	0.797	0.512	0.513	0.511
Within R^2	0.053	0.123	0.118	0.118

Two-way (Borough & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C5: Grid Regressions

The table shows the estimates of a regression of house prices on council tax amounts, namely: $p_{ibdgt} = \beta\tau_{bdt} + \delta_{bgt} + \zeta' x_{ibdgt} + \varepsilon_{ibdgt}$, where p_{ibdgt} is the price of house i, in band b, Borough d, grid square g at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{bgt} are band-grid ID-year fixed effects; and x_{ibdgt} are controls. All columns include band-grid ID-year and month fixed effects. The squares are constructed from a 50 × 50 grid of London. All other variables are defined in Section C.1. Standard errors double-clustered at the grid-ID and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	50.3	12.6	13.4	14.3
	(50.9)	(48.0)	(45.3)	(44.7)
Size		4,626.9***	4,547.6***	4,537.0***
		(1,380.6)	(1,368.4)	(1,366.9)
Newly-built				33,398.5***
				(9,937.9)
Leasehold				-75,924.3**
				(27,874.0)
Fixed-effects				
Band × Grid ID × Year	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Obs.	71,734	71,734	71,734	71,734
R^2	0.696	0.771	0.773	0.773
Within R^2	0.000	0.103	0.010	0.101

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C6: Grid Regressions - Additional Controls

The table shows the estimates of a regression of house prices on council tax amounts, namely: $p_{ibdgt} = \beta\tau_{bdt} + \delta_{bgt} + \zeta' x_{ibdgt} + \varepsilon_{ibdgt}$, where p_{ibdgt} is the price of house i, in band b, Borough d, grid square g at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{bgt} are band-grid ID-year fixed effects; and x_{ibdgt} are controls. All columns include band-grid ID-year and month fixed effects and control for the property size. The squares are constructed from a 50 $\times$ 50 grid of London. All other variables are defined in Section C.1. Standard errors double-clustered at the grid-ID and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	7.98	15.0	9.19	17.5
	(42.3)	(45.4)	(40.1)	(43.4)
Size	5,855.9***	4,522.7***	5,318.7***	4,787.6***
	(1,585.1)	(1,366.3)	(1,353.6)	(1,548.8)
Energy Cost Low	-66,317.8**			-35,546.7**
	(23,418.5)			(14,809.1)
Energy Cost Medium	-108,665.0**			-57,508.8**
	(38,438.8)			(23,978.0)
Energy Cost High	-147,580.7**			-77,036.8**
	(52,657.2)			(33,410.5)
Energy Cost Very High	-195,178.8**			-106,670.2*
	(79,418.3)			(52,374.7)
CO_2 Emissions Low	-32,054.7***			-28,497.1***
	(10,727.2)			(9,893.5)
CO_2 Emissions Medium	-48,961.2**			-47,738.6***
	(17,139.3)			(16,123.8)
CO_2 Emissions High	-75,329.8**			-71,914.7***
	(27,138.1)			(24,295.9)
CO_2 Emissions Very High	-69,844.7			-66,123.0*
	(41,209.7)			(33,075.1)
No. Lighting Outlets	21,965.9**			19,370.8**
	(8,176.9)			(7,921.1)
No. Storeys > 3		-20,775.9***		-22,481.8***
		(6,704.8)		(7,545.2)
Glazed Area Less than Normal		-25,624.5		-18,393.5
		(18,311.0)		(17,084.0)
Glazed Area More than Normal		13,298.3		12,980.2
		(8,438.6)		(8,365.9)
Fireplaces		34,202.3***		32,533.3***
		(6,533.8)		(6,832.4)
Newly-built			23,232.4**	23,142.1*
			(9,272.9)	(11,929.6)
Leasehold			-64,925.4***	-81,662.8***
			(17,738.1)	(28,274.4)
Built in 1950-1982			-33,647.4***	-31,823.5***
			(8,513.9)	(9,833.6)
Built in 1983-2002			43,030.8**	5,862.8

Continued on next page

Table C6 – *Continued from previous page*

	(1)	(2)	(3)	(4)
			(19,392.1)	(10,375.6)
Built after 2003			37,169.2**	-1,989.2
			(16,898.5)	(15,825.0)
Fixed-effects				
Band × Grid ID × Year	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
Energy Rating	Yes	No	No	Yes
Glazed Type	Yes	No	No	Yes
No. Rooms	No	Yes	No	Yes
Property Type	No	Yes	No	Yes
No. Extensions	No	Yes	No	Yes
Floor Height	No	Yes	No	Yes
Obs.	71,734	71,734	71,734	71,734
R^2	0.762	0.774	0.759	0.777
Within R^2	0.216	0.010	0.209	0.110

Two-way (Grid ID & Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C7: Grid Regressions for Different Grids

The table shows the estimates of a regression of house prices on council tax amounts, namely: $p_{ibdgt} = \beta\tau_{bdt} + \delta_{bgt} + \zeta' x_{ibdgt} + \varepsilon_{ibdgt}$, where p_{ibdgt} is the price of house i, in band b, Borough d, grid square g at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{bgt} are band-grid ID-year fixed effects; and x_{ibdgt} are controls. The grids divide London into 50 × 50, 100 × 100 and 150 × 150 squares in columns (1), (2) and (3), respectively. All columns include band-grid ID-year, month, number of rooms, property type, newly-built and leasehold fixed effects, as well as a control for the property size. All variables are defined in Section C.1. Standard errors double-clustered at the grid-ID and year level are reported in parentheses.

	(1)	(2)	(3)
Council Tax	14.3	-16.2	28.2
	(44.7)	(58.4)	(32.4)
Size	4,537.0***	6,988.4***	7,737.4***
	(1,366.9)	(1,794.8)	(2,319.6)
Newly-built	33,398.5***	22,929.9	-28,536.5
	(9,937.9)	(20,993.7)	(24,742.0)
Leasehold	-75,924.3**	-82,738.9*	-151,551.3**
	(27,874.0)	(44,068.4)	(69,763.6)
Fixed-effects			
Band × Grid ID × Year	Yes	Yes	Yes
Month	Yes	Yes	Yes
No. Rooms	Yes	Yes	Yes
Property Type	Yes	Yes	Yes
Obs.	71,734	21,446	6,954
R^2	0.773	0.792	0.827
Within R^2	0.101	0.139	0.154
Grid	50 × 50	100 × 100	150 × 150

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C8: Grid Regressions for Different Grids - Additional Controls

The table shows the estimates of a regression of house prices on council tax amounts, namely: $p_{ibdgt} = \beta\tau_{bdt} + \delta_{bgt} + \zeta' x_{ibdgt} + \varepsilon_{ibdgt}$, where p_{ibdgt} is the price of house i, in band b, Borough d, grid square g at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{bgt} are band-grid ID-year fixed effects; and x_{ibdgt} are controls. The grids divide London into 50 $\times$ 50, 100 $\times$ 100 and 150 $\times$ 150 squares in columns (1), (2) and (3), respectively. All columns include band-grid ID-year fixed effects. All control variables are identical across columns and are as defined in Section C.1. Standard errors double-clustered at the grid-ID and year level are reported in parentheses.

	(1)	(2)	(3)
Council Tax	17.5	-11.8	75.4**
	(43.4)	(62.0)	(33.1)
Size	4,787.6***	7,579.1***	7,516.7***
	(1,548.8)	(1,968.4)	(2,014.1)
Newly-built	23,142.1*	23,012.5	-17,873.0
	(11,929.6)	(23,337.2)	(11,886.9)
Leasehold	-81,662.8***	-100,958.8*	-180,856.2*
	(28,274.4)	(48,348.3)	(92,255.8)
Built in 1950-1982	-31,823.5***	-36,770.1**	-46,677.3*
	(9,833.6)	(13,202.8)	(22,562.5)
Built in 1983-2002	5,862.8	26,842.2	-20,068.4
	(10,375.6)	(22,698.3)	(26,632.8)
Built after 2003	-1,989.2	-30,612.1	-68,073.3
	(15,825.0)	(29,315.1)	(66,304.7)
No. Storeys > 3	-22,481.8***	-19,920.1**	-10,496.4
	(7,545.2)	(9,293.2)	(10,564.8)
Glazed Area Less than Normal	-18,393.5	-37,901.5	41,021.3
	(17,084.0)	(25,339.4)	(69,940.7)
Glazed Area More than Normal	12,980.2	4,587.5	-102,420.2
	(8,365.9)	(19,443.3)	(63,902.4)
Fireplaces	32,533.3***	41,107.7***	49,004.1***
	(6,832.4)	(12,251.9)	(16,591.1)
Energy Cost Low	-35,546.7**	-55,601.8**	-62,161.9***
	(14,809.1)	(20,340.8)	(16,726.8)
Energy Cost Medium	-57,508.8**	-93,685.8***	-85,741.3***
	(23,978.0)	(31,367.3)	(23,759.2)
Energy Cost High	-77,036.8**	-141,100.4***	-161,362.9***
	(33,410.5)	(46,540.7)	(41,048.0)
Energy Cost Very High	-106,670.2*	-170,909.0**	-189,343.0**
	(52,374.7)	(66,032.5)	(72,613.7)
CO_2 Emissions Low	-28,497.1***	-46,607.7***	-43,592.6***
	(9,893.5)	(14,538.3)	(13,312.9)
CO_2 Emissions Medium	-47,738.6***	-73,467.2***	-96,311.1**
	(16,123.8)	(23,582.7)	(43,159.3)
CO_2 Emissions High	-71,914.7***	-104,329.4***	-141,727.0**

Continued on next page

Table C8 – *Continued from previous page*

	(1)	(2)	(3)
	(24,295.9)	(32,544.5)	(58,712.5)
CO_2 Emissions Very High	-66,123.0*	-133,060.2***	-150,608.5*
	(33,075.1)	(44,218.2)	(77,511.2)
No. Lighting Outlets	19,370.8**	22,306.9	53,060.6*
	(7,921.1)	(16,072.2)	(30,324.9)
Fixed-effects			
Band × Grid ID × Year	Yes	Yes	Yes
Month	Yes	Yes	Yes
No. Rooms	Yes	Yes	Yes
Property Type	Yes	Yes	Yes
No. Extensions	Yes	Yes	Yes
Floor Height	Yes	Yes	Yes
Energy Rating	Yes	Yes	Yes
Glazed Type	Yes	Yes	Yes
Obs.	71,734	21,446	6,954
R^2	0.777	0.798	0.846
Within R^2	0.110	0.150	0.165
Grid	50 × 50	100 × 100	150 × 150

Two-way (Grid ID & Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C9: Grid Regressions - Without Stamp Duty Notches

The table shows the estimates of a regression of house prices on council tax amounts, namely: $p_{ibdgt} = \beta\tau_{bdt} + \delta_{bgt} + \zeta' x_{ibdgt} + \varepsilon_{ibdgt}$, where p_{ibdgt} is the price of house i, in band b, Borough d, grid square g at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{bgt} are band-grid ID-year fixed effects; and x_{ibdgt} are controls. All columns include band-grid ID-year, month, number of rooms, property type, newly-built and leasehold fixed effects, as well as a control for property size. The squares are constructed from a 50 $\times$ 50 grid of London. Column (1) excludes properties sold at a price between £240, 000 and £270, 000; column (2) properties sold for between £490, 000 and £520, 000; and column (3) excludes both properties sold in the £240, 000 - £270, 000 and £490, 000 - £520, 000 price range. All variables are defined in Section C.1. Standard errors double-clustered at the grid-ID and year level are reported in parentheses.

	(1)	(2)	(3)
Council Tax	16.1	16.7	18.8
	(46.9)	(45.3)	(47.7)
Size	4,715.8***	4,586.7***	4,765.9***
	(1,446.6)	(1,403.6)	(1,487.4)
Newly-built	37,062.1***	30,619.3***	33,964.1***
	(9,998.4)	(9,694.0)	(9,549.4)
Leasehold	-80,083.0**	-75,897.2**	-80,141.1**
	(30,142.0)	(28,682.7)	(31,002.5)
Fixed-effects			
Band $\times$ Grid ID $\times$ Year	Yes	Yes	Yes
Month	Yes	Yes	Yes
No. Rooms	Yes	Yes	Yes
Property Type	Yes	Yes	Yes
Obs.	65,328	70,012	63,606
R^2	0.775	0.776	0.779
Within R^2	0.105	0.102	0.106
$p \notin$	[240k-270k]	[490k-520k]	[240k-270k] & [490k-520k]

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C10: Grid Regressions - Median Price per Borough, Band, Grid, Year

The table shows the estimates of the following regression: $\varepsilon^{med}_{bdgt} = \beta\tau_{bdt} + \delta_{bgt} + \eta_{bdgt}$, where ε^{med}_{bdgt} is the median residual price of all houses in band b, Borough d, grid square g at time t obtained from a hedonic regression of prices on house characteristics; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; and δ_{bgt} are band-grid ID-year fixed effects. The squares are constructed from a 50×50 grid of London. The explanatory variables used to computed the hedonic residuals are reported in the panel First-stage controls. All variables are defined in Section C.1. Standard errors double-clustered at the grid ID and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	92.1*	15.4	19.9	19.1
	(50.8)	(35.4)	(36.3)	(36.6)
Fixed-effects				
Band × Grid ID × Year	Yes	Yes	Yes	Yes
Obs.	19,377	19,377	19,377	19,377
R^2	0.866	0.833	0.825	0.823
Within R^2	0.006	0.000	0.000	0.000
First-stage controls				
Month	Yes	Yes	Yes	Yes
Size	No	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Newly-built	No	No	No	Yes
Leasehold	No	No	No	Yes

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C11: Grid Regressions - Average Price per Borough, Band, Grid, Year

The table shows the estimates of the following regression: $\bar{\varepsilon}_{bdgt} = \beta\tau_{bdt} + \delta_{bgt} + \eta_{bdgt}$, where $\bar{\varepsilon}_{bdgt}$ is the average residual price of all houses in band b, Borough d, grid square g at time t obtained from a hedonic regression of prices on house characteristics; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; and δ_{bgt} are band-grid ID-year fixed effects. The squares are constructed from a 50×50 grid of London. The explanatory variables used to computed the hedonic residuals are reported in the panel First-stage controls. All variables are defined in Section C.1. Standard errors double-clustered at the grid ID and year level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	104.5**	23.6	28.2	26.6
	(48.0)	(32.7)	(33.5)	(33.9)
Fixed-effects				
Band × Grid ID × Year	Yes	Yes	Yes	Yes
Obs.	19,377	19,377	19,377	19,377
R^2	0.875	0.835	0.827	0.825
Within R^2	0.007	0.001	0.001	0.001
First-stage controls				
Month	Yes	Yes	Yes	Yes
Size	No	Yes	Yes	Yes
No. Rooms	No	Yes	Yes	Yes
Property Type	No	No	Yes	Yes
Newly-built	No	No	No	Yes
Leasehold	No	No	No	Yes

Two-way (Grid ID & Year) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C12: Matching Regressions - Euclidean Distance

The table shows the estimates of the following regression: $p_{ibdt} = \beta\tau_{bdt} + \delta_{ij} + \zeta' x_{ibdt} + \varepsilon_{ibdt}$, where p_{ibdt} is the price of house i in band b, Borough d at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{ij} are pair fixed effects; and x_{ibdt} are controls. Housing pairs from opposite sides of a given border are constrained to be no more than 500 metres away, sold in the same year, in the same council tax band and to both be either old or newly-built and freehold or leasehold properties. The closest match for each property is chosen as the one minimising the Euclidean distance $d(i,j) = \sqrt{\sum_{k=1}^{K}(x_{ik} - x_{jk})^2}$. The vectors x_i and x_j in columns (1) and (2) include size and number of rooms, while columns (3) and (4) add the energy cost. All variables are defined in Section C.1. Standard errors clustered at the transaction ID level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	53.8**	12.9	50.7**	9.00
	(23.4)	(18.3)	(23.8)	(18.8)
Size		3,770.6***		3,750.2***
		(763.8)		(734.2)
Fixed-effects				
Pair ID	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
No. Rooms	No	Yes	No	Yes
Property Type	No	Yes	No	Yes
Obs.	115,224	115,224	114,646	114,646
Unique Transaction IDs	71,578	71,578	71,656	71,656
R^2	0.799	0.836	0.796	0.834
Within R^2	0.001	0.042	0.001	0.042
Distance	Euclidean 1	Euclidean 1	Euclidean 2	Euclidean 2

One-way (Transaction ID) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C13: Matching Regressions - Linear Distance

The table shows the estimates of the following regression: $p_{ibdt} = \beta\tau_{bdt} + \delta_{ij} + \zeta' x_{ibdt} + \varepsilon_{ibdt}$, where p_{ibdt} is the price of house i in band b, Borough d at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{ij} are pair fixed effects; and x_{ibdt} are controls. Housing pairs from opposite sides of a given border are constrained to be no more than 500 metres away, sold in the same year, in the same council tax band and to both be either old or newly-built and freehold or leasehold properties. The closest match for each property is chosen as the one minimising the following distance: $d(i, j) = |\hat{p}_{it} - \hat{p}_{jt}|$, where $\hat{p}_{it}$ and $\hat{p}_{jt}$ are model-predicted prices for two matched property transactions i and j based on a linear model: $p_{it} = \alpha + \beta' x_{it} + \varepsilon_{it}$. The vectors x_{it} and x_{jt} in columns (1) and (2) include size and number of rooms, while columns (3) and (4) add the energy cost. All variables are defined in Section C.1. Standard errors clustered at the transaction ID level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	56.8**	15.3	55.7**	14.6
	(23.4)	(18.1)	(23.7)	(18.7)
Size		3,879.2***		3,809.8***
		(778.8)		(762.1)
Fixed-effects				
Pair ID	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
No. Rooms	No	Yes	No	Yes
Property Type	No	Yes	No	Yes
Obs.	114,904	114,904	113,854	113,854
Unique Transaction IDs	71,588	71,588	71,649	71,649
R^2	0.799	0.837	0.798	0.835
Within R^2	0.001	0.045	0.001	0.043
Distance	Linear 1	Linear 1	Linear 2	Linear 2

One-way (Transaction ID) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C14: Matching Regressions - Linear Distance Less than 30% of Predicted Prices

The table shows the estimates of the following regression: $p_{ibdt} = \beta\tau_{bdt} + \delta_{ij} + \zeta' x_{ibdt} + \varepsilon_{ibdt}$, where p_{ibdt} is the price of house i in band b, Borough d at time t; τ_{bdt} is the council tax amount for a house in band b, Borough d at time t; δ_{ij} are pair fixed effects; and x_{ibdt} are controls. Housing pairs from opposite sides of a given border are constrained to be no more than 500 metres away, sold in the same year, in the same council tax band and to both be either old or newly-built and freehold or leasehold properties. Each house i is matched to all possible candidates j that satisfy the following constraint: $d(i,j) = |\hat{p}_{it} - \hat{p}_{jt}| < 0.3 \times \max\{\hat{p}_{it}, \hat{p}_{jt}\}$, where $\hat{p}_{it}$ and $\hat{p}_{jt}$ are model-predicted prices for two matched property transactions i and j based on a linear model: $p_{it} = \alpha + \beta' x_{it} + \varepsilon_{it}$. The vectors x_{it} and x_{jt} in columns (1) and (2) include size and number of rooms, while columns (3) and (4) add the energy cost. All variables are defined in Section C.1. Standard errors clustered at the transaction ID level are reported in parentheses.

	(1)	(2)	(3)	(4)
Council Tax	-8.19	-5.24	-7.65	-8.14
	(10.1)	(9.68)	(11.0)	(10.3)
Size		3,980.1***		3,982.4***
		(295.8)		(349.3)
Fixed-effects				
Pair ID	Yes	Yes	Yes	Yes
Month	Yes	Yes	Yes	Yes
No. Rooms	No	Yes	No	Yes
Property Type	No	Yes	No	Yes
Obs.	175,639	175,639	167,704	167,704
Unique Transaction IDs	59,722	59,722	58,917	58,917
R^2	0.871	0.875	0.855	0.859
Within R^2	0.000	0.017	0.000	0.018
Distance	Linear 1	Linear 1	Linear 2	Linear 2

One-way (Transaction ID) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table C15: Model-averaged Posterior Distributions for the Council Tax Incidence

The table displays 1%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, 99% quantiles, the modal and mean values of the average posterior distribution for the council tax incidence obtained by using the estimates from Tables C5-C9 and C12-C14. The last column reports the attenuation factor γ computed as the ratio of the posterior and prior median. Each row refers to a different choice of prior.

Prior	1%	5%	10%	25%	50%	75%	90%	95%	99%	mode	mean	γ
$\mathcal{N}(-150, 50^2)$	-143.50	-110.66	-93.21	-61.85	-22.87	-1.98	18.67	31.04	51.90	-12.08	-31.85	0.15
$\mathcal{N}(-100, 50^2)$	-116.75	-85.88	-69.23	-39.43	-12.79	7.60	29.51	41.86	62.81	-9.86	-16.81	0.13
$\mathcal{N}(-50, 50^2)$	-90.71	-61.45	-45.54	-20.99	-2.17	20.33	42.09	54.20	75.25	-6.78	-1.76	0.04
$\mathcal{N}(-150, 75^2)$	-126.67	-87.78	-67.60	-32.92	-7.49	16.49	41.31	54.86	78.03	-8.24	-10.46	0.05
$\mathcal{N}(-50, 25^2)$	-82.09	-64.43	-54.54	-36.79	-18.64	-4.40	9.15	17.64	32.93	-13.86	-20.87	0.37

C.3 Figures

Figure C1: A Typical Border

The figure shows an example of a border between two Boroughs in London. Houses on the left side of the West Eaton Place road belong to the Borough of Kensington and Chelsea and have an annual council tax bill of £2, 279, while houses on the right side belong to the Borough of Westminster and have an annual council tax bill of only £1, 421.

Figure C2: Time Series of Council Taxes

The figure reports the time series of council tax amounts payable across Boroughs. Each panel refers to a different band, while the lines in each panel represent different Boroughs.

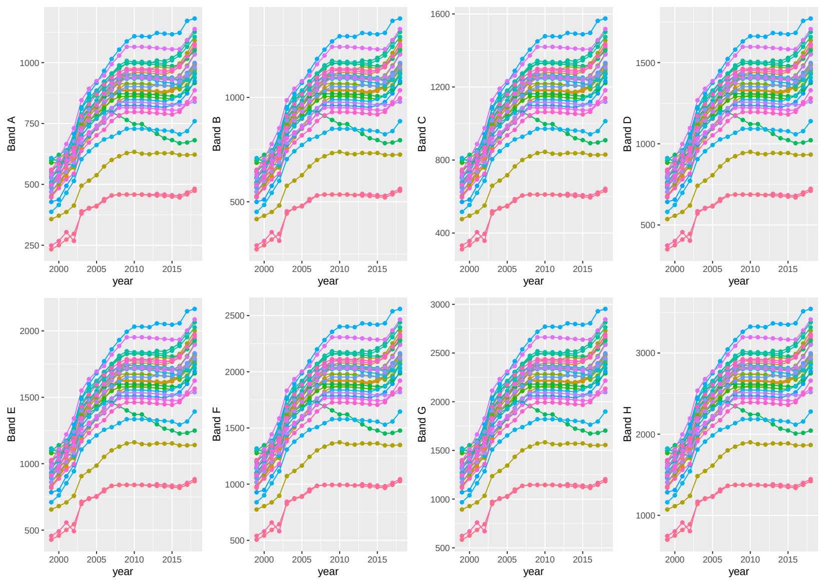


Figure C3: Histogram of Property Prices in London

The figure presents a histogram of the distribution of house transaction prices in London. The distribution is truncated at £1,500,000.

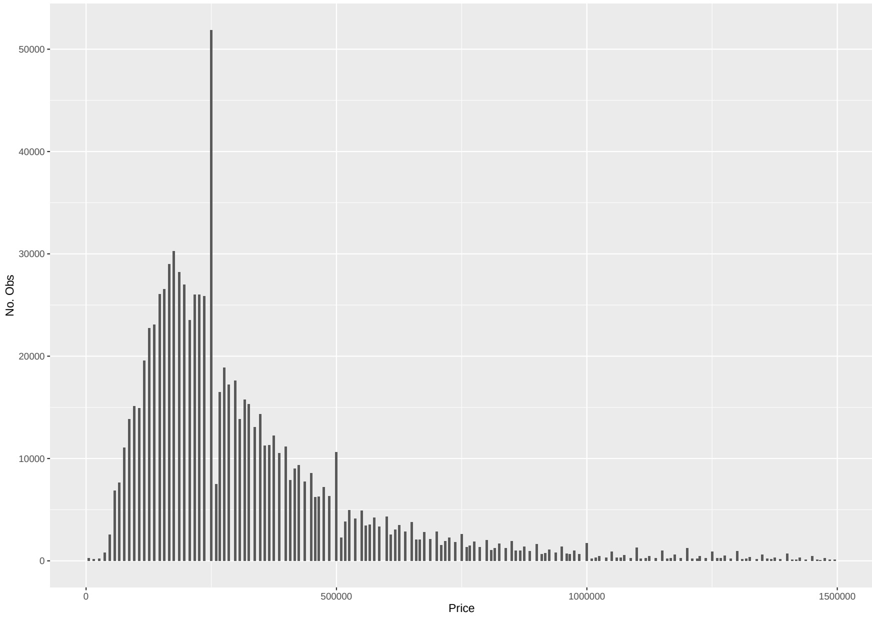


Figure C4: Bunching at Stamp Duty notches

The figure presents a histogram of the distribution of house transaction prices in London around stamp duty notches. Panel (A) refers to the notch at £250,000, while panel (B) at £500,000.

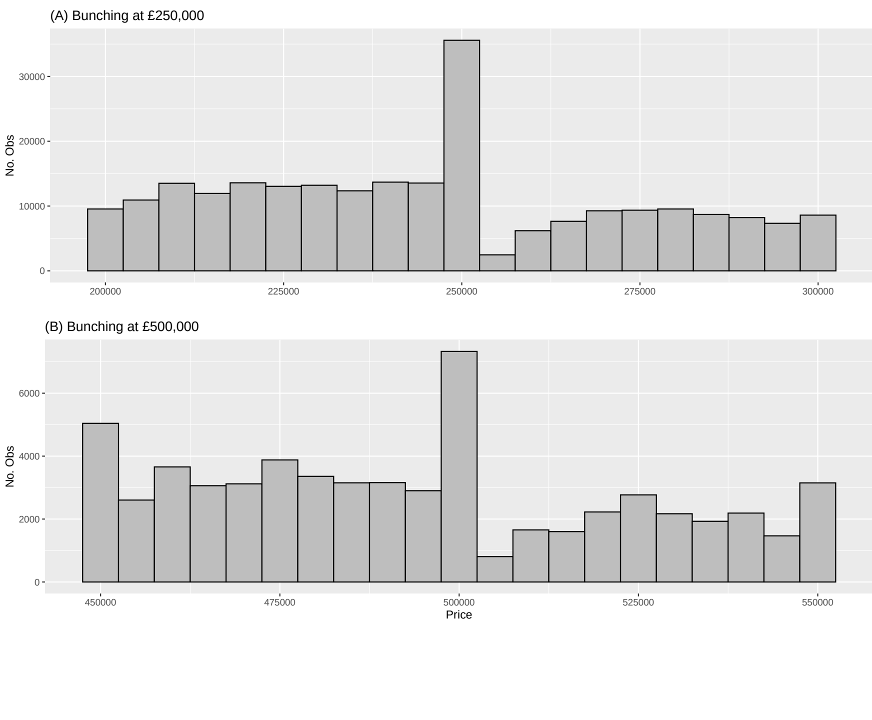


Figure C5: Histogram of Prices by Band

The figure presents a histogram of the distribution of house transaction prices in London per band. Each panel refers to properties belonging to different bands. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

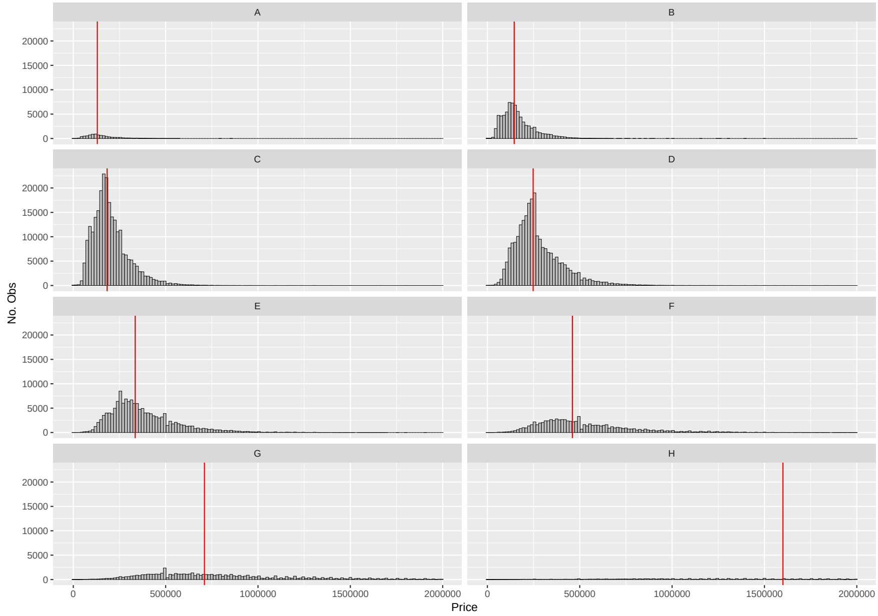


Figure C6: Histogram of Prices by Property Type

The figure presents a histogram of the distribution of house transaction prices in London by property type. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

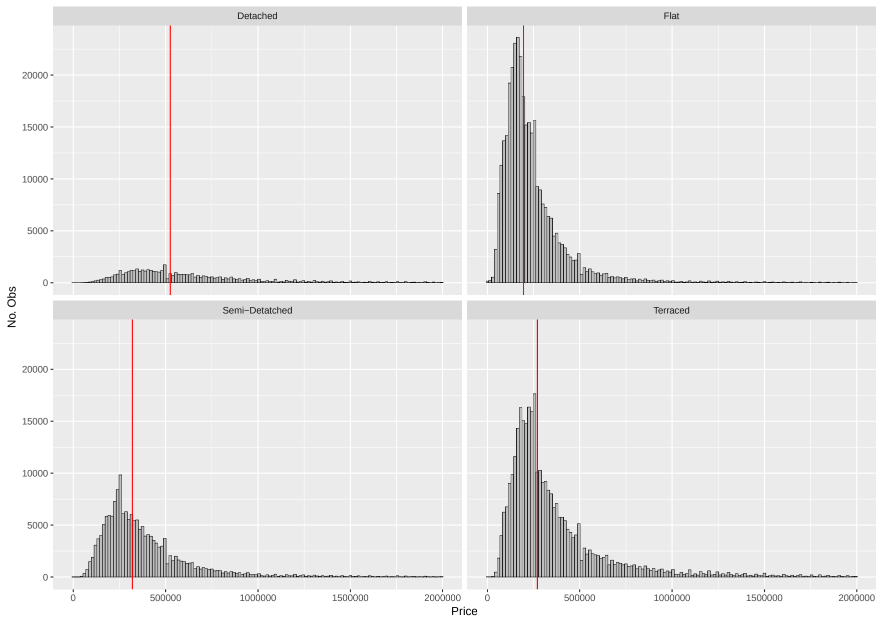


Figure C7: Histogram of Prices by Number of Rooms

The figure presents a histogram of the distribution of house transaction prices in London by number of rooms. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

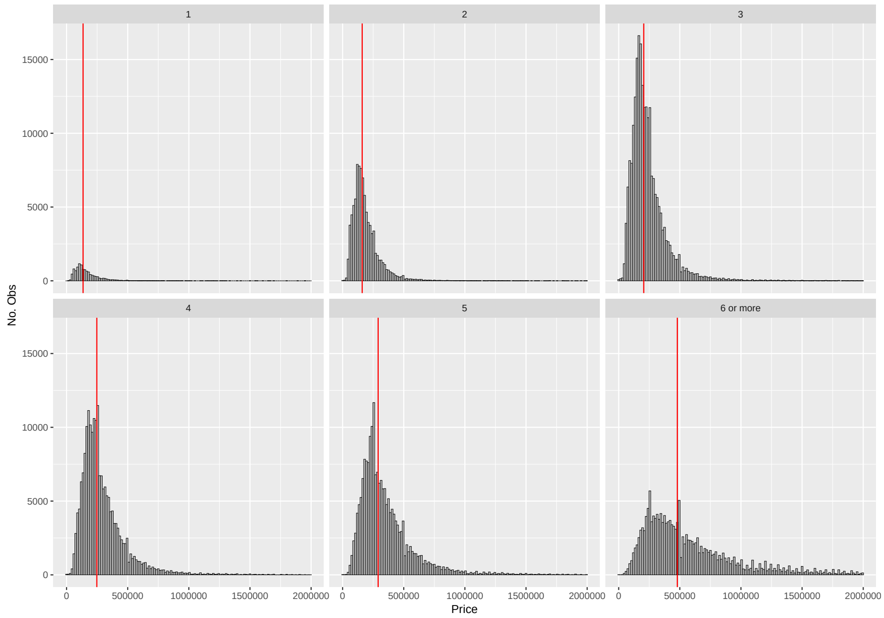


Figure C8: Histogram of Prices by Age

The figure presents a histogram of the distribution of house transaction prices in London by age. The top panel reports the histogram of prices for newly-built properties, while the bottom for established residential buildings. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

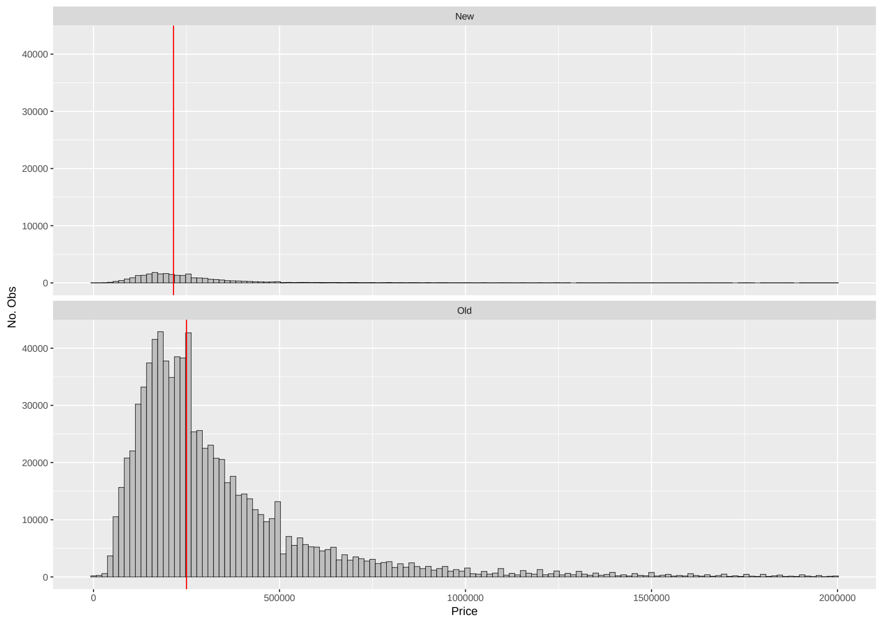


Figure C9: Histogram of Prices by Year of Construction

The figure presents a histogram of the distribution of house transaction prices in London by year of construction. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

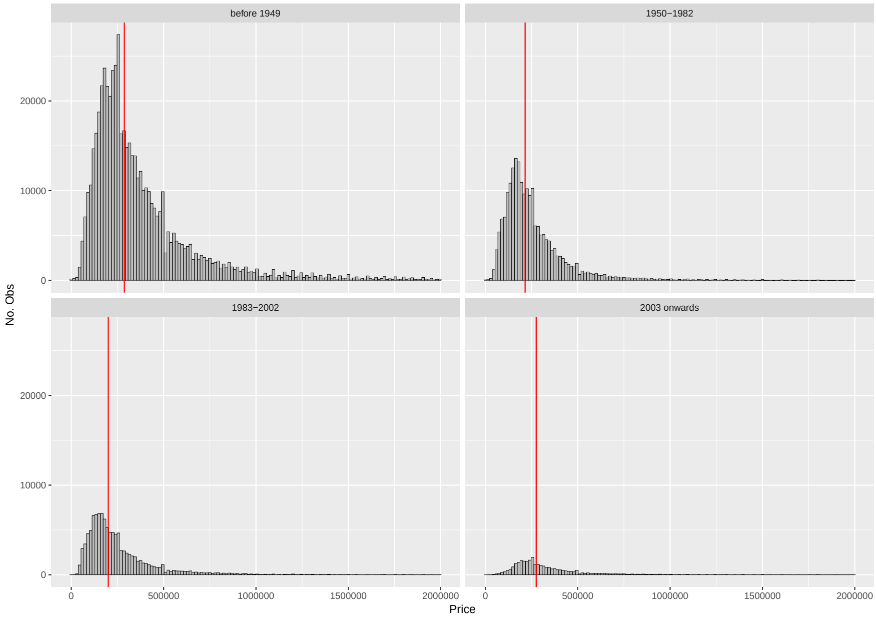


Figure C10: Histogram of Prices by Duration

The figure presents a histogram of the distribution of house transaction prices in London by tenure duration. The top panel reports the histogram of prices for freehold properties, while the bottom for leasehold properties. The distribution is truncated at £2,000,000. The red vertical lines represent the median values computed using the full sample.

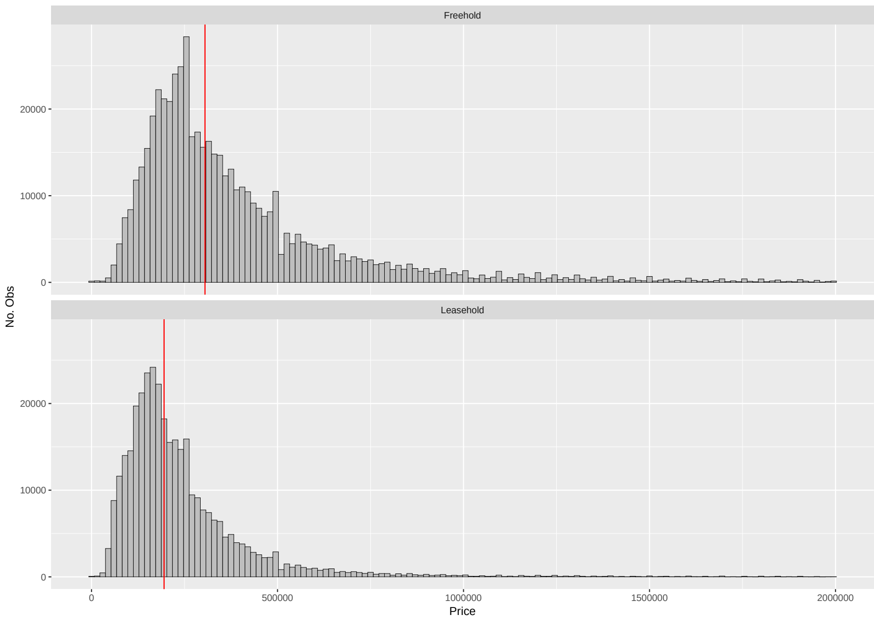
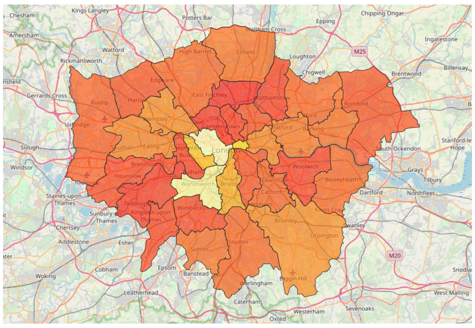
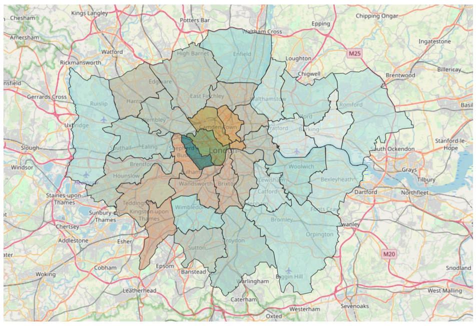


Figure C11: Council Taxes and House Prices

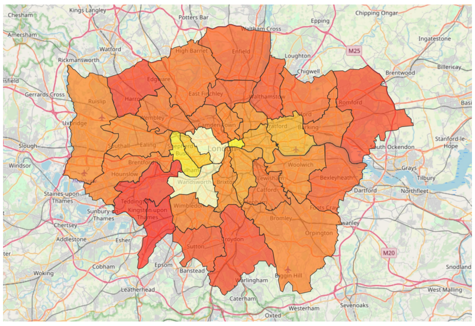
The maps show the distribution of council tax payable for properties in band D for each London Borough, along with the respective distribution of house prices in 2000 and 2018.



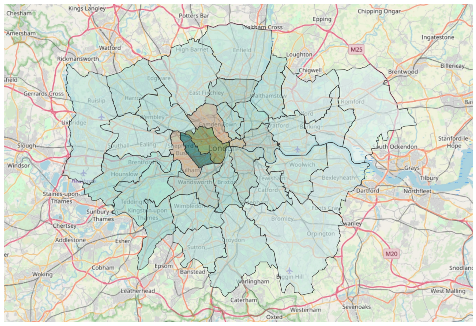
(a) Council Taxes in 2000



(b) House Prices in 2000



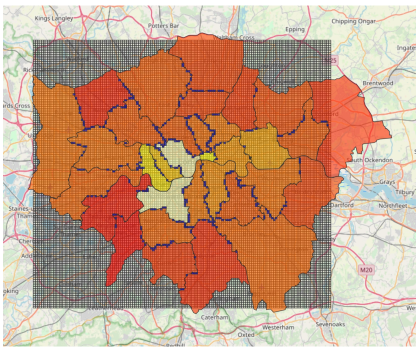
(c) Council Taxes in 2018



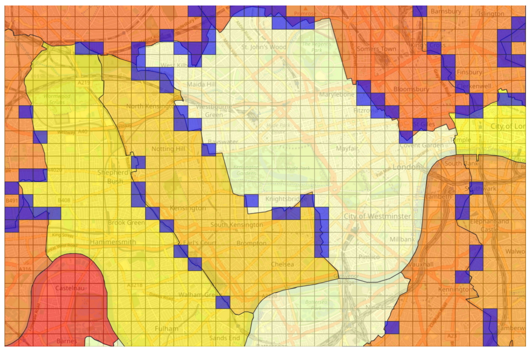
(d) House Prices in 2018

Figure C12: Grids

The maps depict our first identification strategy of dividing London in a grid of equally sized squares. Panel C12a shows a grid of 150 $\times$ 150 squares superposed on the map of the city; Panel C12b shows an enlargement of the central Boroughs. The blue squares denote areas which contain at least two similar properties located on opposite sides of a border.



(a) Grid



(b) Enlargement of the Centre

Figure C13: Distribution of Distances for the Grid Regressions

The figure depicts histograms for the distribution of distances between houses on opposite sides of a border that are used in our grid regressions. We report the distributions for three different grids, namely grids where we have divided London in 50×50 squares, 100×100 and, finally, 150×150. For each histogram we report the approximate size of the square sides in meters.

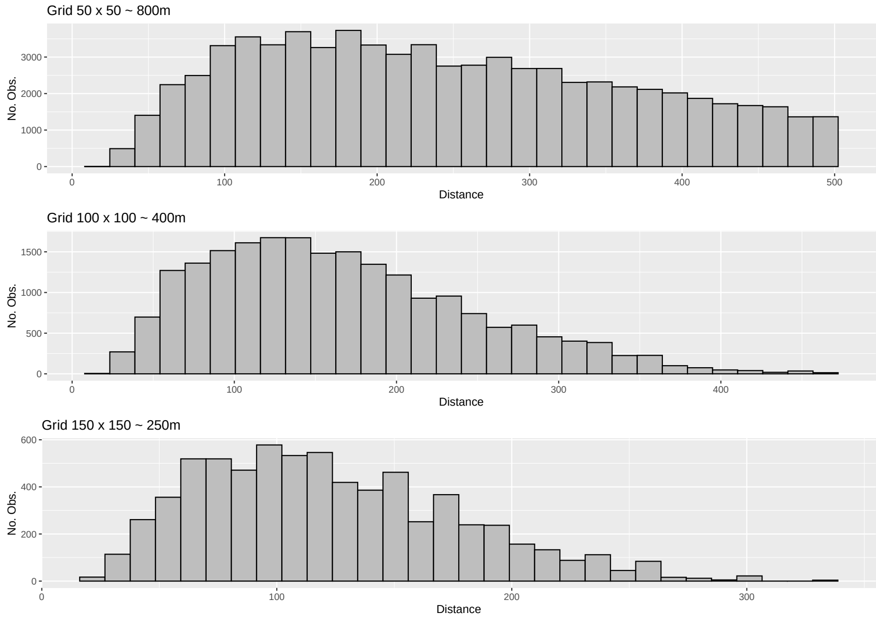


Figure C14: Model-implied Incidence

The figure plots the relationship between tax incidence on house prices and discount rates, where the discount rate is defined as $r + k$ as in Section 3.4. The upper panel shows the incidence of the stamp duty, while the bottom panel the incidence of the council tax.

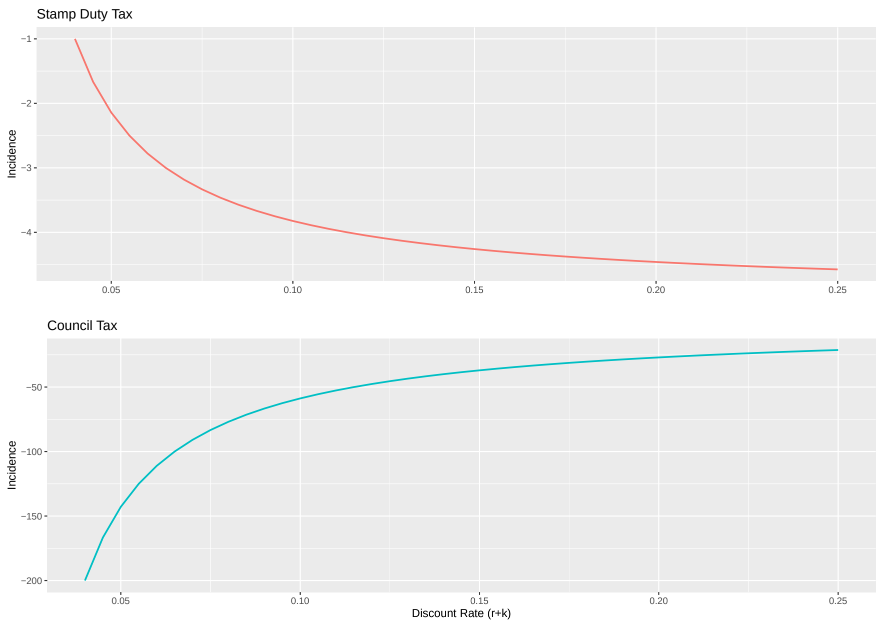
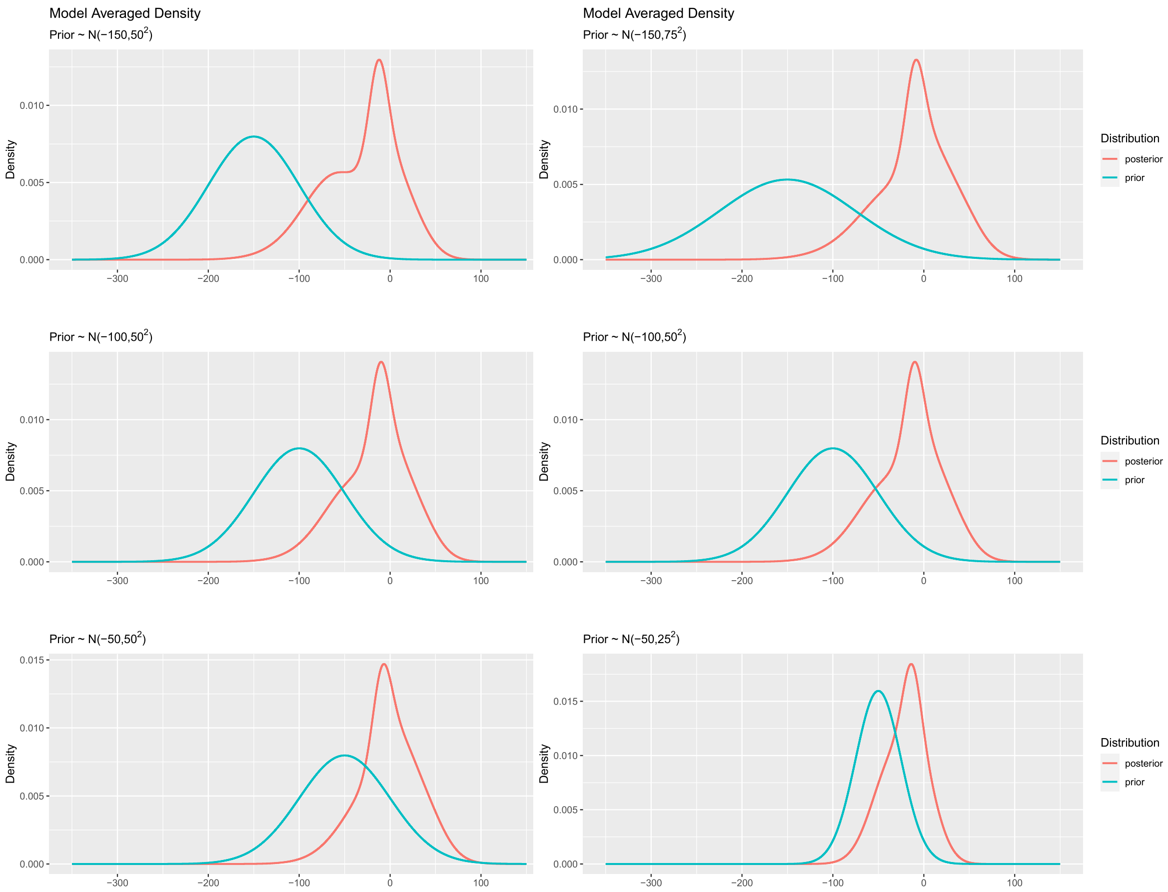


Figure C15: Model-averaged Estimate of the Posterior Council Tax Incidence

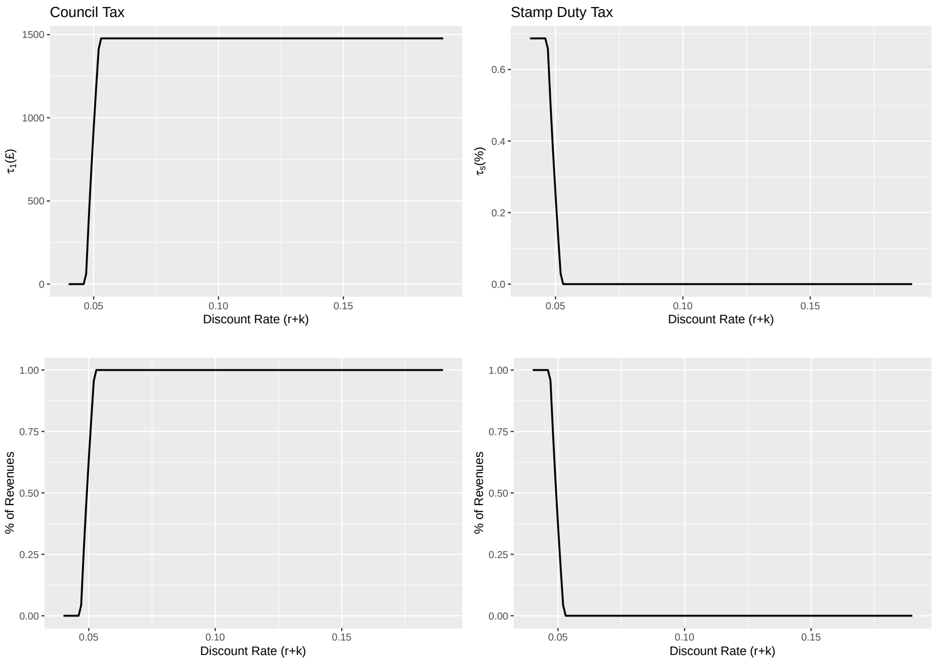
The figure plots the density of the council tax incidence obtained by taking the model-average of the posteriors as described in Sections 3.4.1 and C.4. The priors are normally distributed $\mathcal{N}(b_0, \sigma_0^2)$ in all figures. In panel (a) the priors have constant standard deviation $\sigma_0 = 50$ and varying means of $b_0 = -150, -100, -50$, respectively. In panel (b) the standard deviation of the priors is proportional to the mean, i.e., $\sigma_0 = |b_0|/2$.



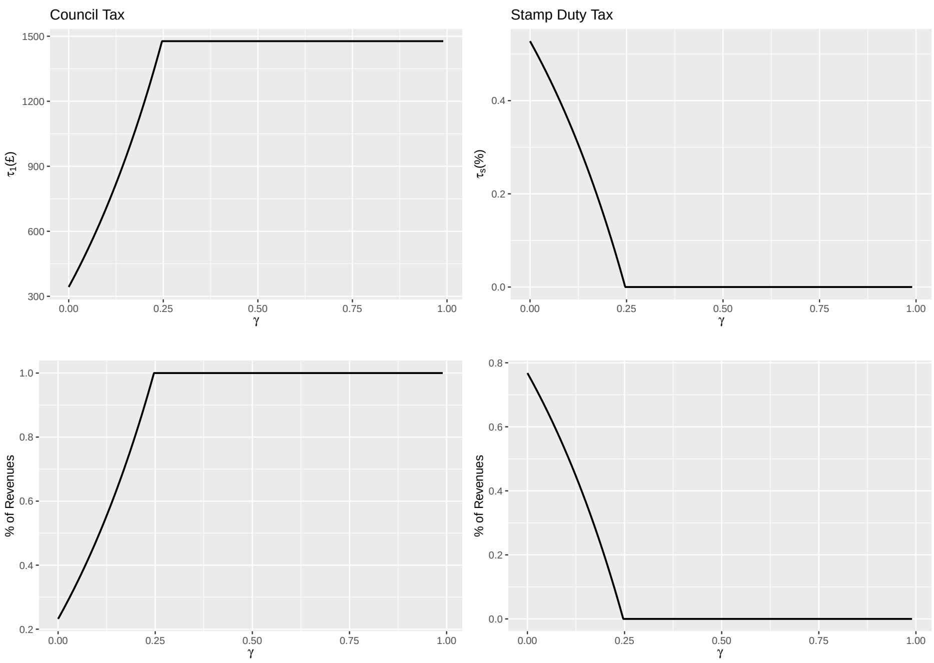
(a) Constant Variance (b) Proportional Variance

Figure C16: Optimal Tax Policy

The figure plots the optimal mix of stamp duty and council tax the Government should choose to maximise the utility of buyers and maintain revenue-neutrality. Panel (a) displays the variables as a function of the discount rate $r + k$, while panel (b) as a function of the attenuation parameter γ. The top plots of each panel show the optimal amount of council tax in £ and stamp duty tax as percentage of house price, respectively. The bottom plots provide the relative percentages of revenue raised through council and stamp duty tax, respectively. In the upper panel we calibrate the parameters as follows: $\alpha = 0.8$, $g = \tilde{g} = 3.5\%$, $\eta_S = 0.5$, $\beta = 0.99$, $\gamma = 0.15$; in the bottom panel: $\alpha = 0.8$, $g = \tilde{g} = 3.5\%$, $\eta_S = 0.5$, $\beta = 0.99$, $r + k = 5\%$.



(a) Optimal Taxes as a Function of $r + k$



(b) Optimal Taxes as a Function of γ

C.4 Computation of the Model-averaged Posterior Incidence of Council Tax

In Section 3.3 we estimate models of the type:

$$y = X^m \beta^m + \varepsilon^m \tag{C.1}$$

where $\varepsilon^m | m \sim \mathcal{N}(0, \Omega^m)$, with Ω^m being the population covariance matrix of the errors under model m. We partition the parameters as $\beta^m = (\beta_0, \beta^m_{-0})$, where $\beta^m_{-0} = (\beta^m_1, \beta^m_2, ...)$ and β_0 is the parameter of interest. We then make the (strong) simplifying assumption that Ω^m is known and assume that the prior distribution of the parameters is: $\beta^m | m \sim \mathcal{N}(b^m, \Sigma^m)$. We also assume that the marginal prior distribution of the parameter of interest is common across models, i.e., $p(\beta_0|m) = p(\beta_0) = \mathcal{N}(b_0, \sigma_0^2)$. It follows that the posterior distribution is: $\beta^m | y, m \sim \mathcal{N}(((\Sigma^m)^{-1} + X^{m\prime}(\Omega^m)^{-1} X^m)^{-1}(X^{m\prime}(\Omega^m)^{-1} y + (\Sigma^m)^{-1} b^m), ((\Sigma^m)^{-1} + X^{m\prime}(\Omega^m)^{-1} X^m)^{-1})$. We then proceed by making the following approximations:

$$((\Sigma^m)^{-1} + X^{m\prime}(\Omega^m)^{-1} X^m)^{-1}_{[1,1]} \approx (\sigma_0^{-2} + \widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]})^{-1} \tag{C.2}$$

$$\begin{aligned} (((\Sigma^m)^{-1} + X^{m\prime}(\Omega^m)^{-1} X^m)^{-1}(X^{m\prime}(\Omega^m)^{-1} y + (\Sigma^m)^{-1} b^m))_{[1]} \approx \\ (\sigma_0^{-2} + \widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]})^{-1}(\widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]} \hat{\beta}^m_0 + \sigma_0^{-2} b_0) \end{aligned} \tag{C.3}$$

where $A_{[i,j]}$ and $a_{[i]}$ indicate the ij-th element of matrix A and the i-th element of vector a, respectively. This leads, therefore, to the following approximate posterior distribution for the parameter of interest:

$$\begin{aligned} & p(\beta_0 | y, m) = \\ & \mathcal{N}\left((\sigma_0^{-2} + \widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]})^{-1}(\widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]} \hat{\beta}^m_0 + \sigma_0^{-2} b_0), (\sigma_0^{-2} + \widehat{Var}(\hat{\beta}^m)^{-1}_{[1,1]})^{-1}\right) \end{aligned} \tag{C.4}$$

After having obtained the posterior distribution for β_0 for each model we average using a flat prior across models to obtain the final density $p(\beta_0|y) = \frac{1}{M}\sum_{m=1}^{M} p(\beta_0|y, m)$.

Returning to the choice of prior distribution for the parameter of interest, we are guided by the model-implied incidence from Section 3.4. We calibrate the following parameters: $g = 0.035$, $\tilde{g} = 0.035$, $r = 0.04$ and $\alpha = 0.8$[1]. Given these values we pick three different means for the prior distribution to match the range of incidence of the stamp duty tax obtained in Best and Kleven (2018), namely, $b_0 = -150, -100, -50$, which roughly correspond to stamp duty incidences of: $\frac{dp}{d\tau_S} = -2, -3, -4$. We

[1]The parameters r and $\tilde{g}$ are consistent with the in-sample average mortgage rate and growth rate of council taxes in the UK, respectively; α is consistent with a downpayment of 20% which is common in the UK. We use a conservative expected growth rate of house prices of 3.5% compared to the in-sample average of 7.3%.

choose the standard deviations of the prior to be equal to $\sigma_0 = 50$ or $\sigma_0 = \frac{|b_0|}{2}$ to obtain five prior distributions.