University of Reading

Herding in US Housing Markets

Doctor of Philosophy

Henley Business School

Matthew Pollock

October 2023

Contents

Declaration of Originality

Declaration: I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged

Matthew Pollock

October 2023

Acknowledgments

I thank my supervisors, Yi Wu and Masaki Mori, for their professionalism and patience during my studies.

I thank Reading Real Estate Foundation and Nottingham Trent University for their financial and economic assistance during my studies.

I thank Cornelia Agyenim-Boateng, Kaio Nogueira, Nuno Machado and Jorn van de Wetering for their friendship and moral support during the past years.

I thank Joseph, Olivia and Alma for their patience, understanding and provision of a country summer home.

Most importantly, another thesis would not be sufficient space to thank my parents for all they have given me. I am everything I am because of, and in spite of, them.

Abstract

This thesis investigates herding behaviour across US metropolitan housing markets. Herding, the conscious or unconscious copying of behaviour by individuals on a large scale, is a phenomenon commonly observed across markets and locations.

Firstly, the presence of herding and reverse herding, the latter a rarely discussed phenomenon, is identified across major US urban housing markets. There is broad spatial and temporal variation, and reverse herding is found to be far more prevalent than herding. Market inefficiencies and the cost of obtaining private information, along with innate homeowner overconfidence, are presented as the theoretical context for these results.

Secondly, the commonly discussed connection between herding and price bubbles is estimated using several measures of bubble activity to ensure robustness. There is evidence that extreme price growth may be a significant determinant of herding behaviour, with stronger effects observed in larger gateway cities that draw investment. The results indicate that herding behaviour may be triggered by speculation rather than information asymmetries.

Lastly, the determinants of herding and reverse herding behaviour at the urban market level are estimated using a discrete choice model. Various social and housing factors are significant determinants of rational behaviour, as opposed to the overall economic conditions.

The scale and role of housing demonstrates the significance of these findings. Whilst some results are in line with previous literature, there are sufficient variations to illustrate the importance of market structure and location. It also illustrates that further work is required to disentangle rational and irrational dynamics, and that leading indicators could be constructed for practical use in risk management and portfolio construction, in addition to any value for policy making.

1. Introduction

1.1. General Introduction

Herding broadly describes the tendency of individuals to copy each other, or some observable group trend. When this mimetic action occurs, then highly correlated behaviour is observed across the group, which can often result in excessive risk taking and poor decision making. Herding can have both rational and irrational explanations. Rational herding may result from information asymmetries when individuals have limited access to clear information. In this situation, it makes sense to copy the actions of those around you who you believe to have more or better quality information. Irrational herding is likely to result from behavioural biases, for example where individuals desire to conform or engage in conspicuous consumption. Conversely, under certain conditions then the behaviour of individuals may diverge widely from that of the group as a whole, a situation labelled as reverse herding.

This thesis considers herding where the actions of economic agents in markets are highly correlated and may not follow a pattern that would be expected from any assumption of rational choice. Specifically, the research focusses on identifying and understanding herding and reverse herding behaviour in housing market across the United States (USA). The motivation for this derives both from the significance of housing markets to individuals and the economy, and also from the unique characteristics that inform behaviour in these markets. The consumption function and social value of housing suggests that a behavioural analysis of market dynamics may provide significant insights. When the apparent information asymmetries in property are included, this shows the potential contribution of this analysis. These also demonstrate an expectation for significant evidence of rational and irrational herding in housing markets.

It follows that these irrational behaviours, both herding and reverse herding, could result in temporal mispricing, an extreme case of which would be a price bubble. More formally, this would be where prices deviate from a fundamentally-derived rational value for a significant time period. The consequence is that bubbles burst, with potentially substantial losses in value to investors and creditors. Whilst the desire to understand more about rational and irrational mechanisms in both herding and reverse herding initially motivates this thesis, there is a connected interest in estimating the wider implications of such behaviour on market dynamics.

The thesis is formed around three main essays that approach, in turn, the identification of herding behaviour, the connection of herding with price bubbles, and the determinants of herding behaviour.

1.2. Motivation

For the majority of people, home ownership is the largest financial decision they will ever engage in, considering the size of individual assets and corresponding liabilities, the latter represented by a mortgage with a term measured in decades. Housing also consequently acts as the individuals' dominant store of wealth. Housing therefore easily represents the largest asset in a global perspective, and so the necessity of housing leads to an enormous global market, both in transaction volumes and asset valuations. As a result, housing maintains great importance in the macroeconomy through wealth effect and consumption channels, and through the development and mortgage lending sectors.

This clearly motivates the importance of residential real estate as a research topic for (credit) risk management, debt investment and policymaking (macro-prudential risk), beyond the wider interests of deepening the understanding of residential property market dynamics, especially in the growing debate about housing affordability. Herding can also make markets more volatile, which has very wide ranging implications on market stability and other factors.

1.3. Significance

As a whole range of behaviours can impact market dynamics, several aspects of herding motivate the specific focus on this topic. In addition to adding to the existing research on a spatially and temporally recurrent phenomenon in investment markets, it is evident that herding can have wider repercussions for individuals, markets, the financial system and the macroeconomy. This derives largely from its proposed connection to bubble formation and market volatility.

The housing market has numerous macroeconomic functions as well as being a core underpinning of household economics. Housing can drive wealth portfolio effects in terms of investment and consumption, in addition to its importance via collateralised lending for the functioning of the financial system and providing sentiment for the wider economy. These channels have wider implications for economic performance and welfare loss, particularly via the link to price bubbles.

	GDP (2001)	GDP (2021)	% change
New York	1,197,117,064	1,598,387,648	34
Los Angeles	632,418,122	950,157,776	50
Chicago	513,554,905	630,126,315	23
Dallas	273,531,523	513,979,216	88
Houston	294,040,871	463,233,301	58
Washington	332,261,059	511,253,994	54
Philadelphia	307,311,547	399,782,262	30
Miami	227,290,316	341,292,101	50
Atlanta	254,139,990	399,129,858	57

Boston	284,104,016	444,402,874	56
Phoenix	150,332,518	261,707,170	74
San Francisco	308,080,917	577,347,865	87
Riverside	101,690,806	171,399,926	69
Detroit	220,103,043	241,602,714	10
Seattle	197,751,593	413,816,976	109
Minneapolis	176,318,625	249,963,205	42
San Diego	142,064,204	224,954,460	58
Tampa	100,297,272	158,130,174	58
Denver	134,531,914	214,520,784	59
Baltimore	131,684,133	185,182,016	41
USA	11,874,235,672	17,481,998,373	47

Table 1: Urban Economic Output
Economic output for 20 major cities over last 20 years with aggregate percentage change over that period. Data is in thousands of chained 2012 dollars and provided by the Bureau of Economic Analysis

Housing requires major investment and borrowing decisions for most individuals and households, decisions that are nuanced by the dual investment and consumption functions. As discussed, in addition to the investment function of financial instruments, housing has an additional function for consumption and social value, suggesting that the pattern of rationality will be significantly affected. As the contemporary USA is an urbanised society, this thesis focuses on major urban areas. Table 1 illustrates the economic scale of these centres and the relative variation in growth performance, which also motivates the emphasis on spatial variation and local market characteristics.

In addition, the information asymmetries and other idiosyncratic market characteristics demonstrate that herding in housing markets may differ markedly from exchange traded securities. When the size of the market and systemic importance of housing as a store of wealth and collateral for credit are included, then it shows the motivation for pursuing this thesis.

There is also the non-economic importance of the housing market, such as the control it has over national and local politics and the obvious impact on public health. Therefore, the systemic importance of the housing market, such as its role in the Global Financial Crisis, and the sheer size of the asset class demonstrate that the research can provide applicable information for policy making. In addition, as housing markets are ubiquitous across countries then the potential impact is not constrained to only one market.

Finally, there is validity in research that is targeted at understanding the unique market structure of real estate, and specifically housing, and how that impacts the market dynamics, especially as empirical investigations of property herding are fairly limited.

Specifically, the USA was chosen because it gives an opportunity to study a significant sample of large diversified urban areas. In addition, the decentralised nature of administration allows for some consideration of spatial planning restrictions on market dynamics (the "laboratories of democracy"). Furthermore, with an estimated value of $40 trillion, the American housing market exceeds both the New York Stock exchange and US national debt in size, and is responsible for in excess of $4 trillion in associated annual mortgage originations. This demonstrates the systemic importance of the asset class, largely in its relation to the commercial banking system.

In addition, the geographic scale of the USA and presence of multiple large urban centres allows an appreciation of the impact of local determinants on behaviour relative to national metrics, which has implications for risk and portfolio management and also for some elements of public policy. This is particularly relevant considering the presence of substantial spatial and temporal variation in behaviour. In line with previous research, the thesis finds strong empirical evidence of temporal and spatial variation which motivates the focus on estimating the significant determinants of herding behaviour, whilst giving some information for the role of spatial variation on investment diversification in the context of risk and portfolio management.

As the USA has a developed market that shares similarities with many other large housing markets, the findings from this study may be generalised to other countries, or at least provide a basis for more detailed investigation into behavioural dynamics in other countries. Additionally, herding is a generalised behavioural topic, so whilst it is important in housing it can also have some lessons for other asset classes.

Having seen mixed evidence on herding in institutional investors, and limited research on herding in direct residential real estate markets, it appears that research into this topic is required to fill a research gap, as the role of residential property both as an investment good, store of wealth and consumption good is highly significant in the global economy. For example, Deng et al (2018) find that herding in mutual funds leads to more extreme crashes in stock prices afterwards.

1.4. Contribution

This thesis contributes to the literature on herding in housing and direct real estate, which is relatively under-researched considering its role in very powerful market forces. The specific market characteristics are discussed throughout as an explanation for some of the unique findings, and the potential implications for the housing market are considered.

Behavioural asset pricing or, more widely, incorporating behavioural forces into market dynamics, is increasingly accepted as the direction for research. Real estate requires more targeted research in terms of behaviour of investors and consideration of the unique property aspects such as information asymmetries, asset heterogeneity, location factors, innate overconfidence and positional motivations. Previous research on homeownership is combined with behavioural theory to construct a property-specific context for herding. Reverse herding is equally considered and the findings are reviewed critically. Some practical uses for these findings, especially in the context of bubble formation, are proposed.

Having reviewed the existing literature on herding, both in general and specific to real estate, having established a theoretical basis for rational and irrational herding, and having provided evidence of its existence in various asset classes and markets, there is a clear potential for policy implications, especially due to the sheer size of the owner-occupied housing sector. In addition, the presence of herding behaviour feeds into bubble formation, housing affordability, housing policy and risk management, which motivates its relevance for research.

1.5. Structure and Content

The Introduction is formed by this current chapter. The subsequent three chapters form the empirical content of the thesis, all closely grouped around herding. The first empirical chapter identifies herding and reverse herding, their spatial and temporal variation and their market conditionality. The next empirical chapter estimates the connection between herding and price bubbles, and how this connection operates across time, cities of different sizes and states of market turbulence. The final empirical chapter considers the social, economic and housing characteristics that determine the rationality of urban areas and what drives them to herd or reverse herd. The final chapter forms the Conclusion and summarises the findings, assesses the significance and contribution, and discusses the limitations whilst setting out avenues of future research.

2. Theory

2.1 Rationality, Information and Market Structure

The term "rational" may be problematic if behavioural motivations for market dynamics are accepted. One definition of rationality is conditioned on whether some outcomes conform with the axioms of rational economics (Baddeley 2010) rather than a deeper appreciation of social, psychological and neuroscientific considerations. Whilst any behaviour that does not fall within a traditional asset pricing model may be seen as "irrational", it can be shown that much of this irrational behaviour is motivated by the market structure and information asymmetries.

To illustrate the role of information asymmetries, consider a major securities exchange where all the relevant market information is freely and easily available, and where over 75% of the assets may be held by large institutional managers (Glossner et al. 2021) with well-resourced research teams and decades of institutional knowledge. Even with access to all necessary information, an individual retail investor may not possess the skills and knowledge to act upon this information. From a practical perspective, they may not have the time to absorb this information and act before the dominant institutions have traded and market prices have restored to equilibrium. This is before any consideration that the investor may feel they do not actually have access to all relevant information. The individual may therefore feel, or know, they are at an informational disadvantage.

If the institutions start to sell large holdings of a major firm, the individual will observe these price movements. The actual sales may be unobservable, but the fall in price is easily detectible. The retail investor may assume that, as the large managers driving price movements at the margins are better informed or better analysts, then there is a valid reason for this sell off. Therefore, it follows that there are perfectly rational motivations for mimicking the actions of others and selling their own holdings in the affected firm. This stylised scenario may provide an explanation of rational motivations for herding, which are driven by asymmetric information.

Real estate markets, specifically the housing context of this thesis, are likely to exhibit a greater degree of market inefficiency than centralised exchanges. Whilst the prices of financial instruments such as securities are easy to observe in real time, there is an obvious lag in housing markets. Perhaps more importantly, whilst homogeneous securities can be repeatedly and frequently traded, individual properties are traded very infrequently, therefore making exact price information challenging to observe. As these conditions suggest that housing markets are likely to be relatively information inefficient, then consequently it is likely that these markets will be more prone to rational herding.

Herding by definition is mimetic, and so there must be some observable metric that can be copied. Whilst in a basic example such as flocks of sheep, shoals of fish, or public voting, then it is easy to understand what is observed, it is impossible to observe individual buy and sell decisions in modern asset markets. One signal to observe would be a proxy for aggregated individual behaviour, such as a price index. But there are other, less institutional, mechanisms for information transmission. The shoe shine boy of 1929 may be apocryphal but it serves as an example of how information can cascade and also symbolises the motivator for mimetic behaviour.

The 1929 crash may have been driven by the speculative inflows of retail investors into closed-ended funds but even in a more institutionalised market in the late 1990-early 2000s, fund managers sank vast sums into internet-related ventures with extreme valuations, which would also suggest that they followed some common metric which created a herd mentality. Likewise, many private firms will announce loudly that they have received institutional investment (as a means perhaps of attracting further equity but also some element of self-promotion for the founders) and so it is easy to observe the scale of investment into certain assets. These observations become self-reinforcing as the agent believes that all the others cannot all be wrong, and this validates the decision to mimic the market.

More recently, there have been obvious signs of herding in cryptocurrency asset markets, with resultant volatility in asset prices. For example, Bitcoin, which may well be considered the “safest” and most established cryptocurrency, has seen extreme volatility resulting from media coverage (Bakas et al 2022), much of which has focused on a narrative that non-digital currencies will soon be obsolete and that savers will be wiped out.

There are considerable well-developed arguments for the existence of herding in manager-driven institutionalised securities markets, which will be discussed in depth later on. However, some of the factors that drive herding in securities markets may not be equally important in understanding the mimetic behaviour found in an individually-driven market with a dual investment and consumption function for heterogenous assets, a market that also lacks a central clearing place for the creation and dissemination of transaction pricing information.

One of the proposed rationales for the presence of irrational behaviour is that housing markets, due to their informational asymmetries resulting from asset heterogeneity, possess strong private information. This can then motivate reverse herding (Avery and Chevalier, 1999). Another consideration is that, when market structure determines the cost of acquiring information, it can trigger herding. Housing markets are hugely segmented and dispersed, and so the time and cost of physically travelling to view a property creates a high cost of acquiring information. As investors want to minimise such costs, this leads to a dependence on word-of-mouth information and information cascades. Information cascades occur when a group of individuals make the same decision sequentially, as subsequent decisions are informed by the previous individual's choice, to such an extent that it overrides the followers' own private information (Bikchandani et al 1992). Clearly this phenomenon has similarities with herding, and therefore forms a version of rational herding.

Housing markets are unlikely to fulfil the symmetrical information assumptions of the efficient markets hypothesis (EMH). At the market level, as property values are highly locational, then information about local areas is required to make effective investment decisions. Clearly, this requires acquisition costs and therefore may not be realistic for an individual. At the asset level, as property is highly heterogeneous, then acquiring information is an intensive process. Whilst regulatory requirements such as surveyor reports may ameliorate some of this by assigning a proportion of the costs to the seller, there is still an element of "buyer beware".

As the name suggests, the EMH is a theory that considers the efficiency and rationality of the complete market, so that the individual investors can still be irrational. At the aggregate market level, the competition between buyers and sellers will remove any opportunities to derive abnormal profits and prices will settle at their “fair” or “correct” value.

If long and short positions can be taken on assets easily and cheaply, then the market cannot maintain mispricing and prices will quickly revert to the correct level. Whilst risk arbitrage can be executed on exchanges due to the availability of derivatives, this cannot be practically executed in direct real estate. Whilst some housing futures are available in the USA (see the CME Metro Area Housing Index Futures), the transaction volumes are very limited and they are likely to be of limited interested to the majority of owner-occupiers due to the knowledge barriers. The lack of a shorting mechanism in direct housing markets allows for extended periods of mispricing. In addition, there are numerous technical barriers to the calculation of imputed rents.

Beside the significant information asymmetries in property markets, largely driven by the highly heterogenous nature of physical assets, there are wider asymmetries in property. The common example of asymmetric preferences between winning and losing have been revealed empirically (Anenberg 2011) and the theory gives great insight into house purchase decisions.

Clearly, much of this herding behaviour links directly with excessive price movements, especially where very strong price appreciation is observed, to a degree that suggests the presence of a price bubble. Evidently, price bubbles are often followed by price crashes. Therefore, beyond the potential losses for an individual, these aggregate (and dissipate through transmission mechanisms) into wider market losses which explains at a conceptual level why bubbles are often followed by crashes.

This mimetic behaviour depends upon the existence of an information transmission mechanism. This mechanism is itself dependent on how agents in housing markets access relevant information. As market are highly localised, then a significant proportion of the market context that sellers and purchasers collect will be derived from local media sources. Whilst this will primarily focus on pricing and volumes, buyers and sellers will also seek non-price information such as time-on-market, lending criteria, sentiment and the development pipeline. Clearly there will be some input from national media, but when actively considering purchase and sale decisions, the comparable assets considered will be very specific to the sub-location. In terms of the how this mechanism operates, Romer (1993) assumes that the behaviour of others does not need to be directly observed but that the information from this can be generated via the development of the price, suggesting that the behaviour of others can be condensed into a single, easily accessible metric.

However, this makes an assumption that both buyers and sellers of housing would approach the transaction as a purely investment-based decision, and that the only consideration for price would be valuing the asset as a discounted cash flow of imputed rents. If this was an accurate reflection of reality then it would follow that the market is, in the long run at least, rational. However, a rational market would not allow for the existence of consistently irrational behaviour that would lead to a phenomenon such as bubbles. Indeed, even in the conceptually more efficient equity markets there has been evidence for more than 40 years investors are not rational (Ritter 2003). A behavioural assessment of economics far pre-dates any mathematical modelling of behaviour, and is evident in the work of Adam Smith, whilst more recently the work that Kahnemann and Tversky (1979) started has motivated an appreciation of cognitive psychology and bias as additional explanatory factors in market dynamics, work that was soon expanded into real estate analysis (Shiller 1982).

Unlike many other investment assets, real estate possesses a dual nature, in that it is an investment that exhibits both income producing qualities and capital appreciation and in addition provides a consumption good which is accommodation for the occupiers, whether they own or rent, whether for living, working, selling or producing (Miller and Pandher 2008). Housing in particular is a good required by all people, placing it somewhere in the basic safety and physiological needs described by Maslow.

In addition, when assuming that irrational motivations justify herding as a result of conformity or conspicuous consumption, then these motivators are strong in housing markets. Housing's function as a positional good derives from its physical characteristics and location, themselves clear signals of value. Whether via the investment or consumption function, housing can therefore operate as an obvious and easily understandable signal of wealth. The desire for social position and conformity can easily lead to herding at the aggregate level. Whilst this may induce conformity in terms of purchase decisions, it could also lead to significant deviations from average behaviour. Under certain conditions, an investor seeking some status may wish to show substantial distance from the average investor and therefore, if this contrary behaviour is common enough, trigger reverse herding. When investors deviate from the crowd and instead go their own way, there are several additional behavioural explanations such as localised herding, a flight to quality, or overconfidence (Gebka and Wohar 2013).

There are commonly accepted characteristics of real estate markets that differ markedly from equity and fixed income markets, and therefore whilst finance-based assumptions are valid in these asset classes they may not be in property markets. In addition, the unique aspects of real estate markets such as informational asymmetries, heterogeneity, illiquidity, lumpiness, and carrying and transaction costs have limited the relevance of rational market analysis. The market is also subject to time lags. Additionally housing is a consumption good for the purposes of shelter, and so investment considerations will be secondary for most transactions. Property is a heavily leveraged asset class, and so housing costs will also impact consumption and investment decisions.

Not only may someone buying a house be at the informational disadvantages already mentioned for a retail investor in securities, but they may also have a desire to maintain some social standing. Housing is a clear signal of wealth and taste, and has a high locational value, therefore forming a strong positional good. Furthermore, as all housing is occupied but not all is held for investment purposes then the primary driver of transactions must be the demand for space. Demand for living space will be heavily influenced by household structure and finances but also personal tastes and social considerations, such as the previously mentioned positional goods. It can already been seen that the consumption nature of real estate will introduce significant psychological biases into the transaction decision making process.

Clearly, the empirical evidence is that people do not always make the most rational decision (Lux 1995), even in the cases where they have all available information. For example, people may be emotional, impulsive, influenced by environment, circumstance and others around them, not exercise self-control and apply short-term considerations to long-term decisions, and continue to repeatedly make the same errors. In many cases, these inherent or situational aspects lead to poor choices by those involved, choices which can result in sub-optimal outcomes for themselves and more market volatility for the wider economy. To reiterate, as real estate, and specifically housing, has a clear consumption function but also has social value, it can be seen that the influence of behavioural biases is potentially much greater and that an appreciation of behavioural analysis is particularly useful.

2.2. Intuition Specific to Real Estate

Before reviewing the extant literature and constructing an empirical framework of analysis, some comment should be made on the intuition behind herding mechanisms specific to real estate and housing in particular.

The spatial fixity of real estate, relative to other investable asset classes, introduces the importance of locational characteristics. From a herding perspective, this may demonstrate some function of behavioural or social bias, as some locational attribute may be a driving force in the mimetic motivation. In a non-physical securitised market, location is irrelevant beyond regulatory conditions, and therefore there is no concept of locational bias. In housing markets, this reduction in potential transactors may increase the prevalence of local cultural or demographic biases, and reduce liquidity, therefore impacting market rationality and increasing the potential prevalence of herding.

The fixity also suggests that the ability to quickly flock in and out of markets is restrained as the potential investor pool is limited, and the flight-to-safety impetus may be unrealistic due to the transaction times and costs involved. Similarly, the limited mobility of households reduces the purely investment driven aspect of real estate decision making as homeowners have reduced asset choices as there are fixed in employment.

This also relates to limited substitutability of product (Ren Ren paper), as the locational fix of employment may mean that housing is substitutable within a city and not at the larger regional level.

The spatial context of real estate defines that supply is a function of a physical process, and the nature of real estate development leads to inelastic supply. The concept of equilibrium in economics dictates that demand imbalances should motivate a supply response, but this will be delayed in a housing market, allowing for greater persistence in irrational dynamics.

Other than the physical and spatial nature specific to real estate, as rational herding may be driven significantly by imperfections in market efficiency, then the specific nature of information transmission in residential markets is an important determining force. Buyers and sellers may observe a variety of direct or price signals. However, the exact signals and the relative importance of them may in fact differ greatly between individuals, their side of the transaction and also the location they are in. For example, a local investor can physically view a property and neighbourhood whilst an investor moving city is limited to word-of-mouth or online information, before making the cost and effort of traveling for an in-person viewing. Additionally, efficiency may have increased with online listing platforms whilst opening up local markets to a greater investor pool, further increasing efficiency.

The long holding periods of property ownership amplify the fundamental uncertainty of future market developments. It may be that long holding periods will reduce any short-term frothy herding, however it will not reduce any social bias that the investor is susceptible to. Certainly, the literature review does show a relationship between herding and uncertainty, or volatility.

Furthermore, the predominance of owner occupiers in housing markets may influence herding behaviour where they over-invest or over-consume housing in their chosen location, or by crowding into "trendy" new locations. Indeed, previous literature on the topic suggests owner occupiers may be more susceptible to herding as their social biases are stronger and their reactions may be less rational, relative to professional investors.

There can be a tendency toward post hoc labels of rationality and irrationality in general, and this can be common when analysing herding. Therefore, more careful consideration of the causal channels is required. Whilst one approach would be further use of causal econometrics on the available datasets, there may also be an argument for understanding the underlying practical mechanism and therefore designing a series of testable hypotheses, potentially in an experimental context.

Both rational and irrational herding are predicated on the assumption that the relevant agents are acting in a conscious and deliberate manner. In the former case, it is assumed that agents copy others due to a lack of clear information, whilst the latter assumes that the mimetic action is motivated by a deliberate social bias. However, some further attention is required on this starting point, as these actions may not necessarily be deliberate.

Another perspective may be that herding can be intentional or spurious. Clearly, both rational and irrational herding in the form presented would fit within the concept of intentional herding. In line with comments on the challenges of proving the causality of herding for closely correlated price movements, it may be difficult to ascertain if the causation is deliberate by the involved agents. This leaves the possibility that the observed behaviour is an almost coincidental movement, and the statistical significance of the responses may be a result of a confounding factor.

These issues can therefore demonstrate the specific characteristics of real estate and housing investment that will impact the particular nature of herding dynamics in housing markets.

3. Literature Review

3.1. General Introduction to Herding and Reverse Herding

Herding is a recurrent phenomenon in investment markets, reaching back to the Dutch tulip crash of the 17th century, and so an appreciation of the existing theoretical and empirical work will illustrate the approach and method of the thesis.

As discussed, herding can be defined as behaviour that is correlated across individuals (Devenow and Welch 1996) specifically where it leads to sub-optimal investment decisions and bubble formation. Agents form their decisions on the basis of other agents' behaviour, rather than any apparent fundamental market conditions. This characteristic results in the critical mass of investors allocating capital to the same or similar assets purely because they observe others allocating their respective capital in such a manner. As they implicitly believe in the wisdom of the crowd, they assume this is an optimal investment decision and do not wish to miss out. A mechanism must be present that coordinates investors' actions, such that a herd can be formed. Such a mechanism is most likely to be either observable asset prices that act as a form of signalling, or the ability to directly observe the actions of other investors.

Herding in its basic, non-asset specific form may be rational or irrational. In the former, the presence of externalities in the form of informational inefficiency and agency issues leads to investors making sub-optimal allocations. Specifically, investors may not have, or not believe they have, the best information and so "follow the crowd" by copying aggregate trading (Bikhchandani et al. 1998). In the latter case, investors are liable to psychological biases that prevent rational analysis of decision making, and essentially follow the animal form of herding, following others into investing without any fundamental analysis (Devenow and Welch 1996).

Lux (1995) took strides toward formalising herding behaviour by recognising that the EMH is not empirically supported, due to evidence that stock markets exhibit far more volatility than the fundamental values and expected returns would warrant. He therefore states that some speculative force must operate in the market dynamic. This was previously emphasised by West (1988) who showed that an appreciation of sociological and psychological factors would provide an understanding of market dynamics beyond the EMH. Lux continued by setting out that extreme market conditions such as booms and busts are not permanent situations, and therefore a changing bias between optimistic and pessimistic outlooks was present. In a market with limited information, investors can form this bias on the only information that is concrete and available, more specifically actual returns (which in a real estate context, would be transacted house prices), and so they have an opportunity to assess the current market sentiment against actual market conditions. This is countered by a desire to extract any excess profit available by trading whilst expectations of capital appreciation are optimistic. Eventually, much like a pyramid scheme, the supply of new market entrants dries up and so trading starts to get thinner and expected and realised returns start to fall, leading to a switch to a pessimistic bias and a collapse in the market. Lux concludes that the excess volatility caused by herding (see also Hott 2012) is due to swings away from the fundamental equilibrium caused by the over-reaction of investors to small deviations from the equilibrium price. This is countered by mean-reversion where the excess returns experienced by investors disappear as the bubble collapses. More formally, there is positive autocorrelation in the short run and negative autocorrelation in the long run. Individually, this behaviour may be caused by irrationality, rational copying of the crowd or even reputational consideration. Lux makes the important point that in aggregate all of these behaviours may be present and not in conflict.

An important factor for real estate, especially when held in non-securitised form such as owner-occupied housing, is that a short-selling constraint applies. Baker and Stein (2004) showed that markets without short-selling can use liquidity as a measure of investor sentiment. If the market has a high proportion of irrational investors, which is believed to be the case in real estate due to behavioural biases, then these investors will not react correctly to the market information available. If investor sentiment is high, then liquidity will be high as a result and therefore returns will be lower than rationally expected as investors overvalue and overpay.

Several possible concepts could explain why, in a market with a large increase in returns, cross-sectional dispersion is greater than would be estimated by a rational asset pricing analysis (Gębka and Wohar 2013). Firstly, localised herding is present when small groups of investors move in and out of assets against the wider market direction in some attempt to take advantage of market movements, which leads to greater dispersion as they are going against the main consensus. However, the illiquidity and transactions costs in real estate make this impractical and costly.

More possible is a flight to safety, where investors liquidate assets to rebalance their portfolios into safe assets, which is often associated with housing. This could be a rational behaviour on an individual basis, conditional on the individual risk tolerance and capacity for loss, as they rebalance portfolios in up and down markets. If assets are coming from non-housing liquidations, the money could be redeployed quite quickly into property. However, fairly limited numbers of people have liquid assets sufficient to buy an additional house. In addition, housing is not a costless or quick asset to liquidate if market conditions justify rebalancing into other asset classes.

3.2. Identification and Prevalence of Herding

If people are presented with incomplete information or lack confidence in their abilities, then they are more likely to herd in extreme market conditions (Christie and Huang 1995) so as to take advantage of the potential excess returns (or avoid excess losses). Also, as they are in abnormal market conditions, they feel less confident in their own abilities and more inclined to follow the crowd. Therefore, greater evidence of herding is expected in boom and bust periods.

Indeed, herding is more prevalent when the return distribution has extreme tails (Ngene et al. 2017), and during periods of high uncertainty in both financial markets and economic policy (EPU). The authors further find that when herding behaviour is present, the dispersion in returns will subsequently reduce by a significant amount, as investors mimic each other and individual observations become highly correlated. This suggests that herding also has a spatial dimension, as if residential markets in different geographies have high levels of co-movement then the potential for geographical diversification is limited.

Further evidence that herding is asymmetrical is demonstrated amongst US homebuilders (Ro et al. 2019), as over a 24 year period these developers displayed much stronger herding tendencies in up markets. Indeed, the authors believe their evidence is stronger than that found for real estate mutual funds (Ro and Gallimore 2014) and they theorise that this is motivated by a scarcity of independent information in buoyant markets. The further fund evidence that stronger herding has a positive impact on house price growth in the next period. Developers may herd both because they learn from others in the market and also because they seek to minimise any reputational loss (as seen with forecasters) that may come from iconoclastic investment choices (DeCoster and Strange 2012), resulting in overbuilding. Forecasting specific to housing starts in Canada, Japan and the United States does not show evidence of herding (Pierdzioch et al. 2012), and in fact shows evidence of reverse herding where forecasters deliberately make forecasts that differ from the consensus, motivated by individual loss functions, but conclude by raising the point that if public policy is formed on forecasts of housing starts, then policymaker should be aware of herding issues and conduct further research on this behaviour.

The evidence that does exist on herding in residential real estate (Hott 2012) found that house prices fluctuate more than justified by the market fundamentals (when the fundamental value is derived from fundamentally derived imputed rents) in that there are significant over- and under-valuations at different periods of time. Hott modelled herding based on investor sentiment being contagious across investors, a behaviour which leads to a general market overreaction. The inclusion of a herding aspect into the empirical model increases the estimated volatility of theoretical prices and therefore more accurately matches the observed volatility in actual house prices, an observation that does not hold constant across all countries nor explain all of the observed bubbles. Hott further posits that the willingness of banks to provide housing credit will largely explain housing demand and therefore housing prices, especially for the pre-GFC market.

Lan (2014) uses least squares and quantile regression methods to examine herding behaviour in Chinese residential property between 1998 and 2013, finding that herding is stronger in an up market than a down market, as shown by Hyun and Milcheva (2018). Lan derives the empirical methodology from academic research in the equity sector, using the cross-sectional absolute standard deviation to measure the spread of returns around the average market return. The initial summary statistics showed that the returns exhibit a non-normal distribution, so that a least squares estimation would be biased. Lan, following from Barnes and Hughes (2002), utilised a quantile regression estimation to determine that herding shows strong evidence in an up market, and is also present in a down market when there is significant turbulence. They also found herding was more prevalent before the GFC, and is not apparent during and after the GFC.

Investor overconfidence can exist if returns have recently been strong and individuals feel this will continue, which can take place at the expense of market signals. If these reactions to market signals are heterogeneous this will result in a greater cross-sectional dispersion in returns.

Although previous studies have suggested that herding is present when market signals are unclear, an inefficient market is not necessarily turbulent, if investors think they have the information they need, or strong private information

Whilst reverse herding has been identified in REITs (Phillipas et al 2013, Zhou and Anderson 2013) and in housing (Ngene et al 2017), the context and explanation has been little discussed. By definition, if herding can exist then so can reverse herding (Hwang and Salmon, 2004) and therefore its existence must be considered equally.

3.3. Herding and Market Outcomes

Although herding provides an interesting academic query in its own right, it can produce a systemic issue for the functioning of asset markets. The major consideration is the role it plays in causing and sustaining price bubbles, and the potential economic losses that come from the following price collapse. There are two issues to handle; the mechanism by which herding and bubbles interact, and quantifying the outcomes of bubbles (and their bursting).

Whether it be sustained price appreciation followed by a price implosion Kindleberger (1978), price movement that is not clearly based on changes in fundamental considerations (Garber 2000) or similarly when asset prices are detached from their fundamentally-derived values for sustained periods (Hott 2009), there are broadly agreed definitions of bubbles such that they represent some price disconnect from the level that is presumably rational. For example, if house prices consistently exceed the sum of imputed rents discounted at an appropriate required return, can it be reasonably claimed that there is a price bubble? However, the main challenge is measuring these deviations, and in fact price bubbles tend to be identified after the fact as a qualitative description.

The mechanism can be simple. Herding is a sustained deviation in pricing behaviour from rational responses to market movements, which could be a causal explanation for sustained deviations in price levels from the assumed rational level. If prices deviate for prolonged periods from fundamental valuations then its suggests an emotional expectation of excess price appreciation, or speculation. Another way would be to state that the investor has been motivated purely by an assumption of price appreciation to purchase the asset. Clearly this would not yield positive profits in a perfectly efficient market, but as it has been shown that the assumption of efficiency in housing markets is perhaps not a valid one then this strategy may have more hope. How does the consumption function impact speculation? Ultimately, people still expect some financial return from property investment even as a home so there is likely to be a desire for price appreciation even amongst non-investors.

Whilst speculation may explain the presence of bubbles, they may also be explained purely by herding (Hott 2009). In addition, if there is some common consensus that the continued price appreciation does in fact result from the fundamentals then trend-chasing may result which in turn will lead to herding. It then follows that, as the assumptions behind the appreciation are incorrect, a destructive bubble will result.

An experimental approach by Lei et al (2001) demonstrates that bubbles occur even when speculation is not allowed, which is an important consideration for the argument between speculation and rationality.

Clearly, if a market is more efficient as information is more symmetrically distributed then the expectation of rational herding is reduced. However, the likelihood of perfect information symmetry in a property market is low and, as shorting is not possible, then herding and bubbles may still persist.

The main empirical issue that this thesis faces (and does not claim to overcome) is that bubbles do not come with a comprehensive statistical test. Indeed, they are much easier to identify after the event (Litimi 2016). Harras and Sornette (2011) investigated the theoretical origins of a bubble and asserted that the adaptive and imitative mechanism that underpins herding does in fact fuel bubbles. As bubble formation is enhanced by high returns, the empirical testing therefore, also in line with Kindelberger, is based on the assumption is that some measure of "high" returns is used as a proxy for bubble formation.

There is evidence that herding leads to price instability (Deng et al. 2018), although this conflicts with evidence that, although herding is internationally widespread in institutional investors, herding by these agents reduces volatility in future prices (Choi and Skiba 2015). The authors find, in line with their expectations, herding is more prevalent in informationally efficient markets. As investors are better able to observe fundamental market information, there is more correlated trading, and these markets also experience faster price adjustments. This would suggest that herding has some form of stabilising effect. The evidence of demand being significantly correlated to the subsequent period's returns is not surprising. The authors conclude that the presence of herding in more informationally efficient markets suggests that herding is a result of a mechanism where fundamental information is fed into market prices. Rather than in an informational cascade, where investors essentially find fundamental information through the wider aggregate trades of others when they cannot deign this information for themselves, the fact that this herding behaviour is more prevalent with efficient information means that the behaviour is not copying but a similar response to the same information. However, this is largely opposed to the main concept of herding being motivated by information inefficiency, and further demonstrates the complexity of understanding behavioural market dynamics.

Following on from informational efficiency, by introducing liquidity (Galariotis et al. 2016), it can be shown that herding behaviour is more prevalent in high-liquidity equities. It may then follow that housing, as a relatively illiquid asset, may see a comparatively lower prevalence of herding and potentially greater likelihood of reverse herding.

There is evidence that investors make trading decisions based on information they gather from the investments of other institutions (Banerjee 1992). The impact of this behaviour on actual prices is harder to gauge, as transaction volumes have little predicative power for commercial property prices in USA from 2001 to 2015 (Wiley 2017), and instead the author found evidence that appreciation was driven by credit and the market share taken up by active investors.

Although not direct property, there has been some investigation into herding specifically in Real Estate Investment Trusts (REITs). Using a quantile regression approach, evidence exists (Zhou and Anderson 2013) that investors in US REITs will exhibit herding behaviour when market conditions are more volatile. The research further find evidence of asymmetry in this behaviour, as herding is both more common and stronger when the market is in bearish phase as opposed to a bull market. REITs however operate as equities rather than direct real estate, so the relevance to a direct market should not be over-argued, although the long time series used (1980-2010) would pick up the underlying property market behaviour of REITs prices and returns, it may not impact the herding behaviour of investors.

Looking specifically at institutional behaviour in the REIT market (Lantushenko and Nelling 2017) through autocorrelation analysis, there is strong evidence of herding by property type as investors copy the investment decisions of other institutions that they observe in a form of positive feedback. As expected, the impact of herding is quicker in public markets than in private. Continuing in the institutional market, using a survivorship-bias free dataset of daily US REIT returns covering 2004-2011 (Philippas et al. 2013), there is evidence that herding behaviour is intensified by disruption in the wider macroeconomy and also by worsening investor sentiment.

Sentiment must be a key factor in herding as shown by the role of optimistic and pessimistic biases in creating market conditions, however again the bulk of property sentiment literature deals with the commercial property sector (Freybote and Seagraves 2017). More generally than real estate, evidence of significant herding in US closed-ended funds between 1992-2016 suggests a motivation from non-fundamentals (Cui et al. 2019), such as previously mentioned noise trading, fund size and possibly feedback trading. The body of evidence also suggests that herding is more prevalent when the macroeconomy or particular market is experiencing uncertainty or in the aftermath of some form of shock.

Thoma (2013) provided evidence through an asset pricing approach that, when market experts from industry and academia declare that house prices will continue to appreciate because "this time it is different", then herding is more likely to occur, contributing to bubble formation and an increased likelihood of a severe correction. Thoma attributed the persistent failure of experts to identify bubble formation to two main factors; a lack of economic history knowledge in economic education, so that researchers fail to recognise warning signs (especially in light of relatively few major housing bubbles in recent decades), and an over-reliance on mathematical and theoretical models that fail to capture the real-life factors that create bubbles. The result is that few experts foresee a bubble, and instead give a list of reasons, such as demographics and financial innovation, as to why price appreciation will continue unabated. Therefore, any rational investor who trusted this information would invest in housing in the expectation of further capital growth, thus further adding to the herding behaviour and bubble formation.

3.4. Causes of Herding

Existing literature suggests that institutional investors may herd due to informational cascades (Bikhchandani et al. 1992, Welch 1992), where they can see the behaviour of other investors and believe that the information gathered from this is valuable, in line with the "wisdom of the crowds". However, although private investors can view aggregate trades (prices and transaction volumes are publicly available), it is impossible to distinguish the exact trades of others, in addition to which it has been shown that equity trading by private investors has a wealth reducing effect and so is not a desirable strategy (Odean 1999, Barber and Odean 2000, Barber et al. 2009).

Otherwise, herding may be motivated by the principle-agent problem (Scharfstein and Stein 1990), as asset managers are assessed on returns relative to other asset managers, so they seek to "run with the crowd" even if they sacrifice potential outperformance in order to minimise any downside risk. Individual investors are not assessed by an external principle so they have no principle-agent problem, and so this cannot motivate the observed herding. However, as seen in seminal research on loss aversion by Kahneman and Tversky (1979) people feel losses much more than they value the same sized gain, which cannot be discounted as a motivating factor in herding behaviour.

However, for private investors, the motivation is more likely to be due to behavioural biases (Barber et al. 2009), as private investors cannot observe the trades of others nor suffer from a principle-agent problem. Scharfstein and Stein (1990) mention that, if herding is rational when information is lacking, then it may still be rational if the agent possesses better knowledge if they are concerned about reputational issues, which could equally apply to professional asset managers and private investors. Barber et al. (2009) state that correlated trading in institutional investors may be driven by some form of limit order program that copies market trades to stay within risk boundaries and asset allocations, or from common responses to tax regulations or from systematic changes in risk attitudes. The former set of factors are unlikely to apply to any private investors.

There is mixed empirical evidence on whether institutional investors herd (Barber et al. 2009). For example, looking specifically at investments into US closed-ended mutual funds (Cui et al. 2019), there is strong evidence of herding. However, as retail investors form the bulk of investors and act as “noise traders”, then these funds are heavily exposed to this herding-style of investing (Flynn 2012, Huang 2015), which makes fund pricing more volatile. Noise trading refers to investing without professional advice or fundamental analysis, and is therefore impulsive and based on irrational exuberance, fear, or greed. Consequently, the traders follow trends and overreact to news, especially to evidence of transaction volumes. This clearly has strong parallels with herding, and could be an integrated factor. As retail investors, they have access to less information (due to the skills and cost of research) and also may not have the necessary experience or time to form investment strategies (this relates to herding being rational when you lack information) and so make less sophisticated decisions than institutional investors.

This is further supported by Flynn (2012) who found that in the UK markets, with a much larger proportion of institutional investment than the US mutual fund market, there was less noise trading and less volatility in pricing. Presumably, due to the lack of information and also the lack of investment discipline found in professional investors, retail investors are more likely to be suspect to behavioural biases with the result that they buy based on recent past performance, hold onto assets rather than realise a loss and purchase assets with unusually large recent transaction volumes (Barber et al. 2009). Private investors also dispose of well-performing assets and repeat past behaviour that coincided with good returns in an effort to achieve replication, therefore underperforming the benchmark (Barber and Odean 2013). The importance of country and time-specific factors in herding behaviour (i.e. whether herding is based on fundamentals or linked to a macroeconomic shock) was further commented on by Galariotis et al (2015).

Other than in investment, when analysing Investment Property Forum consensus forecasts, McAllister et al. (2008) found significant evidence of errors and agreement in the results. They reasoned that this pattern would be due to common factors, and provided evidence of such behaviour in other economic forecasting. The apparent consensus in the forecasts could potentially be due to a behavioural bias, most likely herding. The motivation for this herding behaviour may be explained by the need for forecasts to be acceptable to the intended users, which incentivises self-censorship of the results produced by econometric methods, and therefore ensures published results will fall broadly in line with other forecasters (Gallimore and McAllister 2004). This form of herding behaviour, where analysts replicate forecasts already published by others even if those results are contrary to the analysts own beliefs, is motivated by a desire for approval and to look consistent with other forecasts (Trueman 1994). Even if the forecaster ignores private information and simply replicates what others are doing, this behaviour can be seen as rational where market or industry reputation is paramount to long-term success (Scharfstein and Stein 1990). Lastly, there is a potential for "regulatory herding" where demonstrating the existence of assets in other investors' portfolios may minimise the risk of litigation from private or public agencies (Sias 2004).

The accuracy of commercial real estate forecasters in the UK between 1999 and 2011 exhibited distinct herding behaviour (Papastamos et al. 2015), with a conservative bias whereby capital appreciation and total returns were underestimated in growth markets and overestimated in a downturn, performing worse than a naïve forecast in three-quarters of estimates. Therefore, the estimates were very smoothed and the authors believed this showed evidence of herding.

There is more evidence of herding in credit rating agencies (An et al. 2019), who have a similar motivation to revise ratings when competing agencies have adjusted their respective ratings on a bond. If there is a disagreement, agencies show a strong pattern of convergence in their ratings in the next period, which demonstrates that they incorporate public information as well as their underlying ratings model.

Furthermore, in an experimental setting, people are willing to default on mortgages when they observe similar behaviour amongst other homeowners (Seiler et al. 2014). The motivation may be that when an agent's private information differs from that implied by the behaviour of other agents observed in the market, then it is rational to instead believe in the collective wisdom of the crowd and follow their behaviour, and indeed this choice is also rational when an agent feels they do not have sufficient private information to form a complete choice (Bikhchandani, Hirshleifer et al.

1998), or when market outcomes are determined by the behaviour of the herd (Ro and Gallimore 2014).

The existing literature on herding in residential property investment is less extensive, but evidence from the stock market found significant evidence that systematic behaviour such as buying based on recent performance, not selling assets at a loss and buying stocks that are heavily traded has a psychological driver, as the behaviour could not be explained by changes in risk aversion or taxation (Barber et al. 2009).

There is a long line of research starting with Shiller (1982) and LeRoy and Porter (1981) that demonstrates fluctuations in equity prices beyond that justified by the asset fundamentals. If this excess volatility was due to speculative investing, it would apparently be less likely in property than the equity markets, for two particular reasons; firstly, real estate requires large transaction costs, which require long hold periods for amortisation and excess returns to compensate (Miller and Pandher 2008), and secondly, holding property as an investment or as an occupied home comes with significant carrying costs, such as utilities, mortgage payments, taxes, insurance and maintenance. Both transactions costs and carrying costs in the equity markets are non-existent (such as utilities) or minimal (such as brokerage costs).

Lan made an interesting comment that asymmetry in herding behaviour may stem from the flow of information, as developers and agents issue buy recommendations extensively in up market conditions and so create a flow of positive information. Another contributing factor would be the pro-cyclical policies followed by many governments and central banks in the early part of an economic expansion.

It is important not to overstate the direct relevance of this body of literature, as it applies to very liquid and efficient equity markets, and so cannot be implied directly to real estate. However, the evidence would imply the importance of issues such as liquidity, macroeconomic events, country specific events and differences over time periods.

3.5. Social Standing and Positional Goods

Research into relative, as opposed to absolute, economic standing and consumption has developed as a new field in the past few decades, but the idea that relative consumption impacts utility as well as absolute consumption reaches back as far as the 19th century (Veblen 1899). The outcome of status-concerned consumption is that individuals spend too much on status goods (Hopkins and Kornienko 2004), causing negative externalities which in turn lead to sub-optimal allocations of capital and lower-than-possible utility (Frank 2005, Frank 2008). Higher rates of income inequality lead to higher positional stresses as a means of social status, resulting in higher house prices, higher rates of bankruptcy and higher rates of marital breakdown (Frank 2008).

If all goods were impacted equally by consumer positional concerns, then utility would be reduced due to over-consumption, but the overall allocation of resources to different aspects of consumption would be unchanged (Arrow et al. 2004). However, evidence (Solnick and Hemenway 1998, Solnick and Hemenway 2005, Carlsson et al. 2007) shows that not all things have the same position. Even within Western societies, the relative positions can differ quite significantly (Henrich et al. 2001).

Solnick et al (2007) are motivated by what they perceive as a lack of empirical investigation into this culturally motivated social position and conducted a comparative analysis of US and Chinese populations. The authors found that the former placed an emphasis on physical attractiveness, intelligence and educational attainment, whilst the latter were concerned about income and holiday time, lending further weight to the idea that different goods have different positions.

Using survey-experimental methods, Alpizar et al (2005) find that, whilst both relative income and consumption are important, some goods such as insurance, which is hard to frame as positional, is strongly driven by relative consumption. The sensitivity of "social spending" to the spending of others suggest either strong positional concerns or herding (Brown et al. 2011). Overall, it is clear that the exact mix of relative and absolute income and consumption is not a homogenous pattern across all goods and places, and that there must be cultural and institutional drivers in positional concerns.

Housing can be framed as a positional good as it is a clear physical illustration of social status and consumption preferences (Frank 2005). Using a spatial Durbin model to estimate positional concerns in 20,000 transactions in Ohio, Leguizamon and Ross (2012) find that, whilst relative concerns are important, they are still dominated by absolute well-being. In Columbus, individuals were willing to pay $1103 for an additional 100sqft in their own house but only $400 for a 100sqft decrease in a neighbour's house.

Whilst housing has already been framed as an investment and consumption good, it has clearly a role as a status good. When considering the role of housing as a marker of marital attractiveness, Wei et al (2012) found that increased marital competition in China as measured by a rise in the sex ratio was responsible for between a third and a half of urban house price increases between 2003 and 2009.

Zahirovic-Herbert and Chatterjee (2011) developed a standard hedonic price model in Baton Rouge, Louisiana, which found that between 1984 and 2005, the inclusion of “country” or “country club” within a neighbourhood name lead to price premiums of 4.2% and 5.1% respectively. This label premium is not unique to housing, as Dermisi and McDonald (2010) found a 44% price uplift for commercial property assessed as Class A.

4. Analysis

4.1. Measurement of herding

To empirically assess the presence of herding, some measure of correlation or convergence is required. As herding is the existence of correlated behaviour, or behaviour that is not widely dispersed, then this thesis bases the core of the empirical analysis around the concept of dispersions. As dispersion can be measured, then a judgement on the presence of herding can be made. The starting point is to take a simple asset pricing model, and adjust it to account for large changes in average, or market-level, returns. If the large mass of individual returns moves largely in line with the average, then there is an assumption that investors (buyers and sellers) are acting rationally. If in fact the larger mass reacts differently then there is evidence of irrational behaviour. This can be formalised with reference to traditional asset pricing models.

The estimation approach of this thesis is based upon the work influenced by Christie and Huang (1995). This does not assess herding as a direct observation of the actions of individuals, but as a broader analysis of how the components of the market move in relation to the average. Conceptually, the concept of reaction to the average, or the crowd, fits the mimetic narrative of herding. Specifically, by assessing individual transactions relative to the wider market and presuming that responses will be perfectly rational, it is possible to identify any transactions that do not fit this pattern.

The approach initialised by Christie and Huang in 1995 was further developed in Chang et al (hereafter CCK) in 2000. As the initial cross-sectional standard deviation approach developed by

Christie and Huang was found to be sensitive to outliers due to the use of squared deviations, then CCK modified this to use the cross-sectional absolute deviation (CSAD);

$$\boldsymbol{CSAD_t = \frac{1}{N}\sum_{i=1}^{N}|R_{i,t} - R_{m,t}|} \tag{1}$$

Here, $R_{i,t}$ refers to the return in any period t of the individual asset and $R_{m,t}$ refers to the equivalent for the market. Herding cannot be explicitly tested, rather the response of deviations to changes in returns can be evaluated to estimate the presence of herding behaviour.

CCK derive the intuition from a rational context, where the market returns result from an approach such as the Capital Asset Pricing Model. They demonstrate that the dispersions of the returns, measured by the CSAD, are a linear function of the market returns.

More specifically, Christie and Huang (1995) stated that there was a conflict between rational asset pricing and behavioural analysis. The former starts with the assumption that individual returns are influenced by a set of common factors, core of which would be the market return as the most observable variable. However, the individual returns are different in their sensitivity to the market return. Therefore, large increases in the market return will lead to an increase in dispersion of individual returns. However, the idea of herding counters this by demonstrating that dispersions can, in many situations, be significantly reduced. If this behaviour is present, then dispersions resulting from differing sensitivities will increase at a decreasing rate, and indeed severe herding may lead to an actual decrease in dispersion. Having established that dispersions are an increasing function of market returns, the CSAD model of CCK goes further to establish the relationship as linear. Therefore, under conditions of herding, this linear and increasing relationship no longer holds and instead the function is non-linearly increasing or decreasing.

To demonstrate more formally the linearity, if R_i is the return of any asset *i*, R_m is the return of the market *m*, and $E_t(\bullet)$ is the expectation in time *t*, consider the conditional CAPM;

$$\boldsymbol{E_t(R_i) = \gamma_0 + \beta_i E_t(R_m - \gamma_0)} \tag{2}$$

Where γ_0 is the return of the zero-beta portfolio and β_i is the security-level systematic risk, whilst β_m is the systematic risk of the equally-weighted market portfolio.

For any security *i*, in period *t,* the absolute value of the deviation of expected returns from the portfolio expected return can be expressed as;

$$\boldsymbol{AVD_{i,t} = |\beta_i - \beta_m| E_t(R_m - \gamma_0)} \tag{3}$$

Following this, again in period *t,* the expected cross-sectional absolute deviation (ECSAD) can be defined as;

$$\boldsymbol{ECSAD_t = \frac{1}{N}\sum_{i=1}^{N} AVD_{i,t} = \frac{1}{N}\sum_{i=1}^{N} |\beta_i - \beta_m| E_t(R_m - \gamma_0)} \tag{4}$$

Finally, to show that dispersion and expected market returns have an increasing and linear relationship;

$$\boldsymbol{\frac{\delta ECSAD_t}{\delta E_t(R_m)} = \frac{1}{N}\sum_{i=1}^{N} |\beta_i - \beta_m| > 0} \tag{5}$$

$$\boldsymbol{\frac{\delta^2 ECSAD_t}{\delta E_t(R_m)^2} = 0} \tag{6}$$

CSAD is merely a statistic of distribution, and herding itself cannot be directly observed, rather the nature of the responses to market changes needs to be analysed. In others words, the relationship between CSAD and the market return is used to test for herding. Therefore, following the results outlined above, a regression-based model can test for any non-linear relationship between market returns and return dispersions. The realised CSAD and market returns are used to proxy unobserved ECSAD and expected market returns ex post.

Consulting the regression-estimated response;

$$\boldsymbol{CSAD_t = \propto_t + \gamma_1 |R_{m,t}| + \gamma_2 R^2_{m,t} + \varepsilon_t} \quad \textbf{(7)}$$

And having established the presence of a linear relationship between ECSAD and expected market returns, it can be seen that when there is a large increase in absolute market returns, simulated by the absolute market return, then the impact on CSAD can be estimated. As the rational asset-pricing framework assumes a linear response of dispersion to increases in the market return, then (as per CCK) a non-linear market return term ($R^2_{m,t}$) is included. This allows testing for the presence of herding under the condition that the coefficient for this estimated non-linear coefficient γ_2 is negative and significant. This would give evidence that as market returns increase, and if CSAD decreases and if the response is statistically significant then there is evidence that investors are acting homogeneously and therefore there is evidence for herding behaviour. The individual security returns are now more correlated than would be expected under rational conditions, and so the cross-sectional dispersions have not increased as much as the market return and may possibly even have declined. Assuming that this is statistically significant, the relationship is now non-linear and therefore violates the assumptions of rationality.

Likewise, a significant positive estimated coefficient would give evidence of reverse herding, as it suggests an increase in dispersion when there is a large increase in the market return. Reverse herding is also an irrational response to increases in the market return, as the same non-linear response exists in the opposite direction, suggesting that returns are driven systematically by factors other than the market risk. On the contrary, if the estimated coefficients for γ_2 are not statistically different from zero, there is no evidence to reject the existence of a rational pricing model for generating market returns.

The CSAD method proposed by CCK is widely accepted as a clear and robust approach for estimating any herding behaviour. Firstly, it is well grounded in financial approaches to asset pricing and matches intuitively with common asset pricing models. Secondly, the estimated results from the model are easy to interpret in terms of the qualitative narrative of herding behaviour. In addition, it is widely applicable across asset classes, allowing for discussion in terms of differences between securities markets in relation to market structure. Finally, it has the benefit of being computationally fast and easy to modify, with a range of estimation methods available to account for the sample parameters.

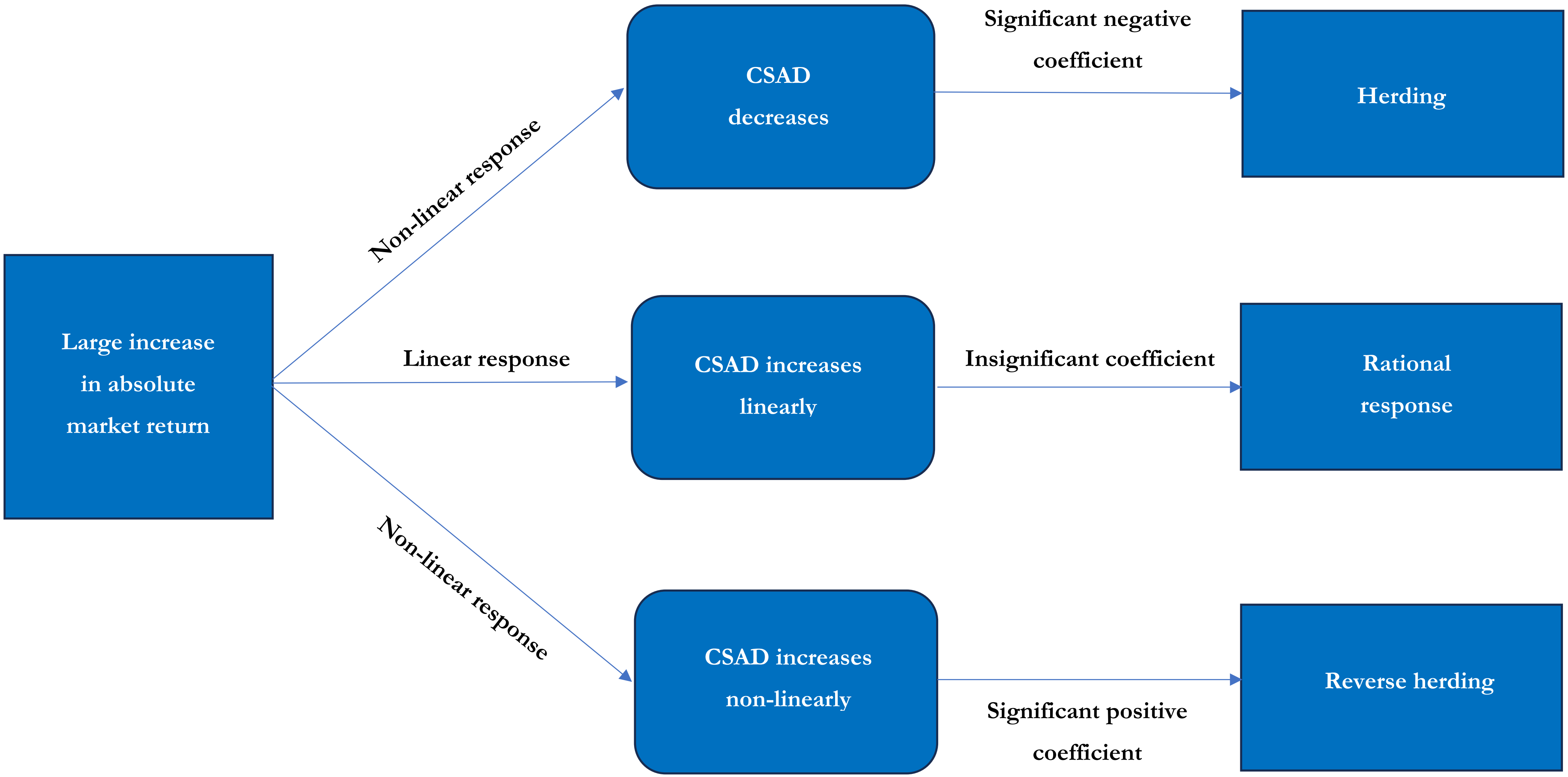
The nature of housing data and the spatial scale used lends itself to this quantitative approach, however there are alternative models of herding available and the considerations for future research will be discussed in the Conclusion.

One of the unique characteristics of property, namely the heterogeneity of the assets, is a limitation of the approach. CCK based their model on equities that are homogenous and frequently priced. Unlike most exchange traded securities, specific assets are not transacted constantly. Instead, a proxy for individual securities is used, which is the local neighbourhood index. In the CCK use of the model, the underlying security prices where the individuals. However, this has been adjusted to consider neighbourhood average prices as the "individual" and the individual security returns will be proxied by the local house market returns.

Of course, a CSAD-based estimate of herding is not the only option available to researchers. Most notably, Lantushenko and Nelling (2017) provided a volume-based measure of herding. The limitations of this come mainly from the lack of appropriate date for housing markets, and in this thesis the CSAD-based model is used as it matches the data and market under investigation. In addition, it is in line with commonly accepted asset pricing models, which aides interpretation as well as positioning the research within mainstream finance approaches.

Furthermore, the concept of herding has clearly been defined with respect to data dispersion and the CSAD therefore aligns with this base concept as well as allowing for a clear operation between the theory and the empirical analysis.

Finally, a methodology derived from the CAPM also allows for a clear test of reverse herding, which is core to the thesis. Firstly, some other herding measures cannot test empirically for reverse herding, and the ability to simultaneously test for herding and reverse herding allows for concise and intuitive discussion.


Large increase
in absolute
market return
Non-linear response
Linear response
Non-linear response
CSAD
decreases
CSAD increases
linearly
CSAD increases
non-linearly
Significant negative
coefficient
Insignificant coefficient
Significant positive
coefficient
Herding
Rational
response
Reverse herding

4.2. Data

The main testing models for herding require regular and fairly frequent price data to assess how dispersions react to strong price appreciation. There are several very high-quality sources of transaction-based price indices for US house prices (Case-Shiller being but the most famous).

As property is a highly localised asset then large regional or state housing markets have little practical interpretation, and so the metropolitan statistical area (MSA) is the most logical geography for analysis. As herding is a measure of intra-market dynamics, a full set of data is required for all the components of the market, which in turn needs a measurement of space at a finer level, broadly the neighbourhood.

	MSA	Major City	Ranking	Relativity
New York	20,140,470	8,467,513	1	42
Los Angeles	13,200,998	3,849,297	2	29
Chicago	9,618,502	2,696,555	3	28
Dallas	7,637,387	1,288,457	9	17
Houston	7,122,240	2,288,250	4	32
Washington	6,385,162	689,545	23	11
Philadelphia	6,245,051	1,576,251	6	25
Miami	6,138,333	442,241	44	7
Atlanta	6,089,815	496,461	38	8
Boston	4,941,632	654,776	24	13
Phoenix	4,845,832	1,624,569	5	34
San Francisco	4,749,008	815,201	17	17
Riverside	4,599,839	314,998	61	7
Detroit	4,383,041	632,464	27	14
Seattle	4,018,762	733,919	18	18
Minneapolis	3,693,261	425,336	46	12
San Diego	3,298,634	1,381,611	8	42
Tampa	3,175,275	387,050	52	12
Denver	2,963,821	711,463	19	24
Baltimore	2,844,510	576,498	30	20
Total	**126,091,573**	**30,052,455**		**24**

Table 2: MSA and Core City Sizes

Population figures for largest MSAs and relevant incorporated core city, data take from the Census Bureau. The Ranking is for the core city and the Relativity is the core city size as a percentage of the MSA

Table 2 shows how administrative boundaries may not be effective for measuring integrated housing markets as the core incorporated city may in many cases account for a very minor part of the urban geography.

The main transactions-based indices are available at the metropolitan level at the smallest spatial scale. Therefore, the pricing data for this thesis is based on Zillow data because it is accurate, comprehensive, well-documented, consistent and available at the required spatial level.

The transition of house listings onto online platforms provides a parallel insight into market efficiency. Herding is suspected to be a rational result of the cost of acquiring private information. The presence of online listings significantly reduces this cost, increasing efficiency, which should result in a reduced prevalence of herding over time.

The initial task is to understand the scale and persistence of herding in housing markets, and one advantage of this data is that it provides 25 years of data which allows some consideration of property cycles and structural market change over an extensive time period to find more robust cross-cycle patterns.

As mentioned, as houses are not transacted frequently and have significant heterogeneity, then the individual return cannot be measured as with corporate securities. Instead, the ZIP code data is used as a proxy for the individual components and allows for a robust estimation of market-level herding, with the MSA-level returns forming the market measure.

House price data is extracted from Zillow, the largest online listing platform in the USA. Their research and data division create a "Zestimate" for different spatial levels across the USA, and the ZIP-code level value) is used in this analysis. The dollar prices are automated values estimated by a neural-network model. The inputs are taken from public records, other listings services and user-generation, which allows for the incorporation of location and property specifics. In each area, a weighted average of all applicable Zestimates is used to create a relevant Zillow Home Value Index (ZHVI). To ensure comparability, the All Homes ZHVI mid-tier series is used which captures homes in the 35th to 65th percentiles. This is the flagship ZHVI and forms the basis for Zillow's forecasts and consumer material.

Naturally, this leads to extreme aggregation of data for a process that is purely individualistic. This poses an issue of epistemology, and certainly there will always be some disconnect between the true underlying individual mechanism when using any aggregated data. However, aattempting to optimise the data available does motivate the choice of urban areas for the spatial definitions rather than wider regions. The data has been used in peer-reviewed articles (Baldauf et al. 2020; Bernstein et al. 2019; Damianov and Escobari 2016; Giglio et al. 2021; Holt and Borsuk 2020; Joshi 2016; Rivas et al. 2019). The choice of data results from its accuracy (using Zestimates for over 100 million properties, the Zillow-declared median error rate is 1.9% for on market homes and 6.9%

for off-market homes), in addition to providing the only comprehensive ZIP-level value source which allows an MSA-level analysis. Values are only presented where there are at least two years of data, and does not capture appreciation resulting from property changes, therefore representing only market appreciation for the typical property. A LOESS-based seasonal decomposition is used and the index chained backwards, then smoothed via a three-month moving average.

Therefore, referring back to equation 1, N is the number of ZIPs in each MSA in month t, $R_{i,t}$ is the return of any ZIP in month t, and $R_{m,t}$ is the equally-weighted average return over all ZIPs in the MSA. This measure is similar to the standard deviation and specifically measures the cross-sectional deviation of returns in any MSA at one time period.

Returns are calculated by differences in the natural logs;

$$\mathbf{R_t = 100\ x\ (log(P_t) - log(P_{t-1})),} \qquad \textbf{(8)}$$

where P_t denotes the ZIP level price index.

5. Identification of Herding and Reverse Herding

5.1. Abstract

This study is the first to examine herding and reverse herding in US metropolitan housing markets based on Zillow ZIP-level house price indices. Reverse herding is found to be more prevalent than herding, which differs markedly from equity markets. Also, herding and reverse herding show strong dependency on market conditions. These results suggest that the overconfidence of homeowners and the presence of private information in local housing markets may be driving these behaviours. Wide spatial and temporal variation in herding and reverse herding suggests the importance of local characteristics as determinants of the rationality of market responses.

Key Words: Herding; Reverse Herding; Housing; Overconfidence

5.2. Introduction

This study is the first to examine herding and reverse herding in the United States (US) housing markets at the Metropolitan Statistical Area (MSA)-level, which allows identification of local variation in herding and reverse herding. Housing markets exhibit unique characteristics, such as local variation and information inefficiency, which are distinct from equity markets. These characteristics suggest the importance of examining the possible evidence of reverse herding as well as herding at the local level. The environments under which herding and reverse herding are observed are examined, focusing on market conditions (up and down markets), volatility environments (high and low volatility), a major crisis period, and individual overconfidence. These sub-analyses shed some light on the determinants of herding and reverse herding behaviours.

5.2.1. Herding

Herding has been defined as the existence of correlated behaviour across individuals, especially where it leads to sub-optimal investment decisions and bubble formation (Devenow and Welch 1996). This can result from investors abandoning a rational asset pricing approach and copying others (Banerjee 1992).

Rational herding is a response from investors with limited information who "follow the herd" as they believe the crowd has superior knowledge or information and they rationally copy others (Bikhchandani et al. 1992; Bikhchandani et al. 1998; Welch 1992). Irrational herding exists when behavioural biases overcome the rational decision-making processes of investors (Barber et al.

2009), for example where a social or personal requirement to keep up with some defined cultural group causes them to copy others e.g. the much-discussed “keeping up with the Joneses”.

When individual investors follow a collective metric, returns will cluster around the market average, meaning the dispersion of returns will be smaller than expected under a rational asset-pricing model (Chang et al. 2000). Herding can then lead to bubble formation, resulting in price collapse and systemic issues in the wider financial and economic systems (Lux 1995).

The evidence for herding in previous studies is largely dependent on exogenous factors (Goodfellow et al. 2009). For example, there is evidence that herding exists around major data releases (Galariotis et al. 2015) and that this behaviour can spill over into other countries, the latter finding aligning with evidence of significant co-movement in herding across European markets (Economou et al. 2011). Chang and Lin (2015) found herding to be dependent on local culture and market sophistication, whilst Lam and Qiao (2015) showed a decline in herding over a 30-year period. Herding has been also reported in Real Estate Investment Trust (REIT) markets (Lantushenko and Nelling 2017; Philippas et al. 2013; Zhou and Anderson 2013). Thus, herding is consistently identified in various asset classes and geographical markets. However, as Griffin et al. (2003) conclude, herding is neither universal nor similar across assets and markets and is heavily influenced by country and time-specific factors (Galariotis et al. 2015).

5.2.2. Real Estate Context

Estimated to exceed $40 trillion of value in the US and representing the largest lifetime financial decision for most individuals, there is valid motivation to examine herding behaviour that can lead to bubble formation. Housing also differs from securitised markets because real estate markets are local and possess significant information asymmetries, which will impact the nature and motivation of mimetic actions.

While research for herding in housing is limited, Hott (2012) looks at housing and finds movements beyond that justified by the fundamentals. Ngene et al. (2017) look at regional US housing markets and find extensive evidence for herding across various market conditions and geographies. Lan (2014) finds herding in the Chinese national housing market.

Ngene et al. (2017) established some evidence of variation between regions. However, the examination at the MSA-level is more appropriate considering the body of research on MSA-level dynamics and its role as an integrated real estate market and economic unit. Prior studies suggest the importance of local variation in housing (Gray 2018; Hortas-Rico and Gómez-Antonio 2020; Lerbs and Oberst 2014; Palomares-Linares and van Ham 2020; Tsai 2015; Zhang and Fan 2019) and the smaller spatial scale allows identification of local variation in herding. This is the first study to investigate herding at the MSA-level.

5.2.3. Reverse Herding

Under some market conditions, rather than assigning more weight to the market consensus, investors follow their own opinion and actively deviate from the market average. As individual returns will not cluster around the market return but will disperse more widely, greater cross-sectional dispersion of returns will be observed, leading to reverse herding (Bekiros et al. 2017). Hwang and Salmon (2004) state that reverse herding must exist by definition if herding exists, and so it should be equally considered.

This reverse herding behaviour has been identified in equity markets (Chang et al. 2000; Galariotis et al. 2015; Hwang and Salmon 2004) and in REIT markets (Philippas et al. 2013; Zhou and Anderson 2013). When studying herding in housing on a sub-national level, Ngene et al. (2017) found fairly extensive evidence of herding and reverse herding under various market states.

Christie and Huang (1995) and Gleason et al. (2004) found more evidence of reverse herding in developed markets which, combined with Chang et al. (2000) identifying reverse herding in similarly developed US, Japanese, and Hong Kong markets, suggests that herding is more common in developing markets and reverse herding in developed markets. Klein (2013) proposes that behaviour is linked to market sophistication and that markets may progress in the long-term from herding to reverse herding as they mature, a development also seen by Lam and Qiao (2015).

5.2.4. Asymmetric Results and Volatility

In addition to evidence that prices respond asymmetrically to market conditions (Bekaert and Wu 2000; Conrad et al. 1991; Hong et al. 2007), herding also displays asymmetry (Hyun and Milcheva 2018; Lan 2014; Ro and Gallimore 2014; Ro et al. 2018).

Herding may be more present in extreme market conditions (Christie and Huang 1995) as people are somewhat overwhelmed by noisy information and struggle to process price signals. As a result, people follow the lead of others believing they are better informed, often referred to as "the wisdom of the crowd" (Bikhchandani et al. 1992; Welch 1992).

5.2.5. Housing Market Structure

As market maturity may be accompanied by relatively more reverse herding (Klein 2013), this phenomenon is expected to be present in the US housing market which is considered as a developed market. In addition, the existence of more prevalent reverse herding specifically in a housing context can come from the nature of the market itself, which is characterised by low transparency and a lack of easily accessible and frequent pricing, culminating in strong private information.

The motivation for reverse herding may be reputational (Effinger and Polborn 2001; Levy 2004), due to strong private information (Avery and Chevalier 1999), or resulting from bullish sentiment (Sibande et al. 2021). In line with Avery and Chevalier (1999), Hwang et al. (2020) make the argument that the importance assigned to information in trading decisions is dependent on whether the information is public or private, as profit can be derived from private information in inefficient markets. An investor could be rational to deviate from the public information represented by the market average when they possess strong private information. Assigning more weight to private information and trading on it would lead to greater dispersions, resulting in reverse herding.

5.2.6. Overconfidence

Ekholm and Pasternack (2008) present evidence that individuals may be less likely to herd as they are supremely confident in their abilities. Daniel et al. (1997) show that in an overconfident context, individuals overreact to private information and underreact to public information. Bao and Li (2020) find a conspicuous overconfident effect during booms and inefficient periods, and simulations suggest that this leads to excessive trading. Chuang et al. (2014) and Griffin et al. (2007) find that inefficient markets are prone to overconfident, excessive trading, which can lead to reverse herding. In addition, Hwang et al. (2020) state that homeowners are generally overconfident in the United Kingdom (UK), a market similar in maturity to the US housing market. By November 2012, all the top 20 MSAs had returned to consistent house price appreciation, which would trigger the overconfident response for the post-Global Financial Crisis (GFC) period.

5.2.7. Contributions

This paper contributes to the existing literature in several ways. Firstly, this study adds to the fairly limited research on herding in direct real estate and housing in particular. Also, investigating a largely owner-occupier market assesses herding in unsophisticated investors who are also consumers of the investment good. Secondly, herding is analysed at a new spatial level by employing a unique database of local house price indices to mimic the behaviour of individuals and test irrational responses at the city level. Thirdly, reverse herding, a little researched phenomenon in investment markets generally, is considered as an equally important outcome, and the context of its prevalence in real estate is discussed. Finally, a unique measure of individual overconfidence is proposed by combining a national-level economic sentiment measure with a national-level housing market sentiment measure, which allows an examination of the potential role of overconfidence as one of the driving factors of reverse herding.

A significant part of equity ownership is through institutions who are sophisticated and less prone to irrational psychological biases (although there is evidence of herding in funds (Cui et al. 2019; Zhou and Anderson 2013)). However, in a market such as housing that is predominantly held by individual owner-occupiers, more irrational responses would be apparent (Flynn, 2012). Within the irrational responses identified, as housing clearly demonstrates the characteristics of an inefficient market, and as the US is a developed economy, previous studies would suggest that reverse herding will be relatively more prevalent than herding. Following the existing literature, a potential change in responses after the GFC is expected, along with more prevalent herding in turbulent markets.

5.2.8. Summary of the Results

In line with previous findings, evidence is found that markets often react irrationally to large increases in price. Specifically, this study identifies significantly more reverse herding than herding, which may be due to the innate overconfidence of homeowners, the presence of strong private information in local housing markets, or the general level of market maturity. Herding is found to be more prevalent in down markets, in volatile markets, before the GFC, and when individual confidence in the housing market is low. Conversely, reverse herding is more common in up markets, in stable market conditions, after the GFC, and when individual confidence in the housing market is high.

5.3. Data and Methodology

5.3.1. Market

Unlike the central clearing place of a stock market, housing is local and therefore, rather than testing for national-level herding, a smaller spatial scale is employed, namely the metropolitan statistical area (MSA). Herding has been tested on the regional-level in the US, however, herding has not been tested on the MSA-level before. Due to the interconnected socio-economic nature of MSAs, much empirical analysis of housing dynamics is done on an MSA-level. Table 3 presents the list of the 20 largest MSAs by population with basic market descriptions. Strong population growth in Southern and Western cities has not translated into the greatest price appreciation, which has been in California and other technology-centred economies such as Denver and Seattle. Other variation may derive from states such as California possessing stronger regulatory and geographical impediments to development.

5.3.2. Herding and Reverse Herding Identification

	Population			Real per capita income ($)			House price ($)			Ownership (%)	ZIP
MSA	**1996**	**2020**	%	**1996**	**2020**	%	**1996**	**2020**	%	**2020**	**2021**
New York	17,681,708	19,124,359	8	53,574	82,322	54	308,039	509,356	65	51	939
Los Angeles	11,771,038	13,109,903	11	42,852	69,805	63	304,629	735,356	141	48	362
Chicago	8,782,253	9,406,638	7	47,952	67,671	41	243,076	257,714	6	68	387
Dallas	4,622,564	7,694,138	66	43,728	61,554	41	185,483	274,597	48	64	267
Houston	4,314,589	7,154,478	66	42,664	59,893	40	181,549	232,626	28	62	237
Washington	4,549,151	6,324,629	39	54,711	76,771	40	295,432	471,701	60	66	322
Miami	4,652,414	6,173,008	33	44,818	64,190	43	183,825	321,994	75	61	182
Philadelphia	5,602,154	6,107,906	9	46,179	69,705	51	198,865	274,637	38	71	388
Atlanta	3,765,817	6,087,762	62	44,470	58,773	32	200,935	264,610	32	67	206
Phoenix	2,855,711	5,059,909	77	38,982	51,851	33	187,034	326,891	75	68	154
Boston	4,265,564	4,878,211	14	52,503	85,724	63	275,812	535,789	94	60	284
San Francisco	3,923,208	4,696,902	20	57,510	111,050	93	419,320	1,178,986	181	52	168
Riverside	2,990,316	4,678,371	56	32,395	45,365	40	205,586	422,649	106	66	149
Detroit	4,433,102	4,304,136	++	45,302	58,356	29	171,830	198,541	16	72	209
Seattle	2,856,795	4,018,598	41	48,579	80,420	66	273,083	583,855	114	59	159
Minneapolis	2,846,496	3,657,477	28	48,083	67,214	40	197,545	319,088	61	72	231
San Diego	2,651,549	3,332,427	26	42,215	66,266	57	289,310	678,553	135	56	99
Tampa	2,256,460	3,243,963	44	39,732	52,291	32	147,493	253,548	72	69	130
Denver	1,959,552	2,991,231	53	47,885	69,822	46	241,171	484,473	101	67	127
St Louis	2,640,161	2,805,473	6	43,481	60,844	40	161,032	196,929	22	69	220
USA	**269,390,000**	**331,500,000**	**23**	**42,588**	**61,674**	**45**	**177,761**	**268,690**	**51**	**66**	

Table 3: MSA Descriptive Statistics

Descriptive statistics for the twenty largest urban areas in the USA by population, using the metropolitan statistical area defined by the Census Bureau. The population and per capita income data are provided by the Bureau of Economic Analysis and house prices by Zillow Research, and the homeownership rate by the Census Bureau. Per capita income and house price figures are all in 2020 dollars using the Consumer Price Index for all Urban Consumer from the Bureau of Labor Statistics. ZIPs are the number of ZIP codes with price data in the MSA in January 2021.

Returns and dispersions are measured on a monthly frequency and on a month-to-month basis. Data is available from January 1996 to January 2021, and losing one observation to calculate differences leaves 300 observations for each MSA. For each MSA, the MSA itself is defined as the market and the ZIPs that aggregate to form the MSA are defined as the individuals, and so a monthly time series can be constructed for each of the 20 MSAs. The 20 largest MSAs cover urban areas with populations greater than 3 million inhabitants and account for approximately 45% of the total urban population. In addition, these MSAs have at least 71 ZIPs within their boundaries in January 1996, and data coverage increases significantly over time, which provides enough observations to make robust estimates of herding behaviour. This data is extracted from Zillow as it is the sole provider of ZIP-level house price estimates.

To ensure the robustness of the estimated results, the parameters are estimated using quantile regression (QR), which better accounts for observations in the extreme tails of the distribution than the standard ordinary least squares (OLS) approach. This is more appropriate for non-normal distributions and investigating non-linear relationships, as the theory suggests herding is more commonly observed in extreme tails of the distributions. Whilst OLS coefficients are estimated by minimising the squared deviations from the conditional sample mean, QR coefficients are estimated by minimising the weighted sum of absolute errors, where weights are defined by the quantiles.

$$\boldsymbol{Q_t(\tau|CSAD_t) = \theta_\tau + \gamma_{1,\tau}|R_{m,t}| + \gamma_{2,\tau}R^2_{m,t} + \varepsilon_{t,\tau}} \qquad \textbf{(9)}$$

A range of percentiles are used to perform the quantile estimation; 0.025, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95 and 0.975. As irrational and non-normal behaviour, herding is assumed to take place in the tails and so estimating responses across the full range of quantiles identifies the exact presence of irrational behaviour. Models for all the empirical analysis are estimated with robust standard errors to account for any potential heteroskedasticity.

5.3.3. Testing for Asymmetric Responses to Market Conditions

Having established that responses to market conditions are often asymmetric, estimating the role of market conditions can be most effectively modelled using a dummy variable approach to test for herding under up and down markets;

$$\begin{aligned} Q_t(\tau|CSAD_t) = \theta_\tau + \gamma_{1,\tau} D^{down}\left|R_{m,t}\right| + \gamma_{2,\tau} D^{up}\left|R_{m,t}\right| + \gamma_{3,\tau} D^{down} R_{m,t}^2 \\ + \gamma_{4,\tau} D^{up} R_{m,t}^2 + \varepsilon_{t,\tau}, \end{aligned} \tag{10}$$

where D^{down} is 1 where $R_{m,t}<0$ and D^{up} is 1 where $R_{m,t}>0$.

Estimated via a quantile regression, the significance and sign of the respective quadratic coefficients (γ_3 and γ_4) will give evidence for the existence of herding or reverse herding under either market condition. As market states are MSA-specific, there is some variation in sample size in each estimation. These range from 46 down months (15% of months) for Houston to 86 months (29%) for San Diego. The average for all MSAs is around 67 (22%) months.

This same model is employed for the further sub-analysis, with the definition of the dummies being changed appropriately. When investigating the role of volatility as a proxy for market signals of risk, high volatility is defined as any state with a standard deviation greater than the mean of the standard deviation of the previous 12 months and low volatility as a standard deviation less than the mean standard deviation. Again, these are MSA-specific, although the samples are more balanced than for market-state modelling and the data ranges from 135 high volatility months for Los Angeles (47% of the months) to Detroit where 175 months are counted as high volatility (61% of months). The model estimation in Equation (7) is adjusted such that, rather than up and down dummies, low and high volatility dummies are used. Due to the measurement approach, the total sample size is reduced by 12 from 300 to 288, which was not deemed to make a material difference to estimations or comparisons with other market conditions.

For the GFC-based analysis, the data set is split into before and after the Federal Reserve definition of the recession which provides nearly equal samples (142 months pre-GFC and 139 months post-GFC).

5.3.4. Overconfidence Measure

A unique measure of individual overconfidence is proposed by combining a national-level economic sentiment measure with a national-level housing market sentiment measure. Baker and Wurgler (2007) define sentiment as “a belief about future cash flows and investment risks that is not justified by the facts at hand”. Economic sentiment is measured by the Daily News Sentiment Index produced by the Federal Reserve Bank of San Francisco (FRBSF) (Shapiro et al. 2022). The news aggregator service Factiva collects articles of at least 200 words from 24 major US newspapers where the main topic was US economics. These sources cover all major regions and include several national papers. Publicly available lexicons are combined with a news-specific lexicon created by the FRBSF and trained on a historical archive of 16 major US newspapers to create a newspaper-specific sentiment-scoring model.

This model correlates highly with human-derived sentiment scores and outperforms some current machine learning techniques. The index is produced daily and is converted to monthly averages for this analysis.

House buying sentiment is measured by the Survey Research Center at the University of Michigan, which surveys a minimum of 500 households monthly to ask around 50 core questions. The core questions cover personal finances, and business and buying conditions. To measure house buying sentiment specifically, the percentage of respondents who believe that now is a good time to buy a house is collected.

To allow comparison between the measures, each observation is transformed into a percentile ranking based on the whole period. The ratio of house buying sentiment to economic sentiment then serves as a proxy for how overconfident prospective house purchasers are relative to the wider economy. Specifically, if the ratio is greater than one, the market is classed as overconfident and if the ratio is less than one then it is classed as unconfident.

5.4. Empirical Results

5.4.1 Descriptive Statistics

The CSAD for each MSA is calculated on a monthly basis with the MSA as the market and the ZIPs as the individual observations. As expected, Table 4 shows that house price growth is high in urban areas such as San Francisco and Seattle which have been outsized beneficiaries of growth in technology-based industries. The relationship with dispersion measured by CSAD is less clear, as Los Angeles, despite being the fastest-growing city, has seen relatively low dispersion of responses whereas Chicago has seen low growth but much higher dispersion than Los Angeles. This may suggest that responses are not purely driven by pricing but also other market conditions and motivates further analyses.

MSA	Metric	Mean	Median	Min	Max	SD	Skewness	Kurtosis	Obs
NYC	Return	0.33	0.30	-0.81	1.30	0.54	-0.12	2.13	300
	CSAD	0.39	0.38	0.20	0.57	0.05	0.28	3.48	300
LAX	Return	0.50	0.65	-2.47	2.50	0.86	-0.79	4.29	300
	CSAD	0.29	0.27	0.13	0.93	0.12	2.40	10.35	300
CHC	Return	0.20	0.32	-1.30	1.39	0.52	-0.96	3.31	300
	CSAD	0.36	0.32	0.18	0.89	0.13	1.29	4.56	300
DFW	Return	0.29	0.26	-0.78	1.31	0.39	-0.09	3.12	300
	CSAD	0.22	0.20	0.10	0.62	0.08	1.54	6.32	300
HOU	Return	0.26	0.26	-0.67	0.98	0.32	-0.28	3.43	300
	CSAD	0.32	0.31	0.20	0.53	0.05	0.91	4.15	300
WDC	Return	0.32	0.25	-1.47	1.87	0.63	-0.06	3.45	300
	CSAD	0.37	0.35	0.14	0.68	0.11	0.77	3.10	300
MIA	Return	0.35	0.52	-2.76	2.35	0.98	-1.05	4.30	300
	CSAD	0.31	0.29	0.10	0.73	0.13	1.09	3.95	300
PHD	Return	0.26	0.21	-0.83	1.30	0.47	0.13	2.63	300
	CSAD	0.33	0.34	0.16	0.49	0.05	-0.18	3.35	300
ATL	Return	0.28	0.41	-1.51	1.27	0.57	-1.44	4.70	300
	CSAD	0.35	0.30	0.15	0.86	0.14	1.11	3.65	300
PHN	Return	0.37	0.48	-2.71	3.53	1.05	-0.38	4.55	300
	CSAD	0.33	0.28	0.15	0.86	0.14	1.49	5.14	300
BOS	Return	0.39	0.46	-0.70	1.44	0.52	-0.29	2.33	300
	CSAD	0.34	0.33	0.17	0.65	0.08	0.73	3.50	300
SFR	Return	0.49	0.62	-1.61	1.96	0.75	-0.46	2.62	300
	CSAD	0.48	0.45	0.20	1.01	0.15	0.73	3.01	300
RIV	Return	0.42	0.50	-3.24	2.46	1.04	-1.27	5.48	300
	CSAD	0.38	0.35	0.16	0.87	0.14	1.37	4.70	300
DTR	Return	0.24	0.39	-1.65	1.66	0.65	-1.03	3.78	300
	CSAD	0.39	0.35	0.17	1.08	0.15	1.14	4.53	300
STL	Return	0.42	0.58	-1.78	1.64	0.70	-1.01	3.43	300
	CSAD	0.28	0.26	0.11	0.61	0.10	1.05	3.94	300
MNN	Return	0.33	0.46	-1.05	1.21	0.53	-0.98	3.22	300
	CSAD	0.33	0.30	0.18	0.73	0.11	0.99	3.43	300
SDG	Return	0.45	0.60	-2.19	2.18	0.86	-0.69	3.22	300
	CSAD	0.29	0.25	0.12	0.93	0.13	2.78	11.56	300
TMP	Return	0.36	0.58	-2.07	2.31	0.88	-0.93	3.63	300
	CSAD	0.31	0.29	0.13	0.64	0.08	0.94	4.41	300
DNV	Return	0.40	0.42	-0.64	1.22	0.45	-0.25	2.33	300
	CSAD	0.26	0.24	0.11	0.59	0.08	1.12	4.64	300
SLS	Return	0.22	0.27	-0.70	0.82	0.32	-0.79	2.94	300
	CSAD	0.39	0.38	0.25	0.60	0.06	0.51	3.22	300

Table 4: Descriptive and Distributional Statistics
For each MSA, descriptive and distributional statistics for both price returns (from equation 4) and the cross-sectional absolute deviation (CSAD) (from equation 3) are calculated from Zillow data for the period January 1996 to January 2021 on a month-to-month basis and on a monthly frequency (authors' own calculations)

5.4.2. Initial Herding Analysis

Table 5 is a summary of quantiles with significant evidence of herding or reverse herding for each MSA, collated by the count of quantiles where the γ_2 response coefficient on the non-linear term in Equation (9) is statistically significant. The initial analysis estimates responses for the entire period of available price data.

The estimated coefficient γ_2 is significantly negative, suggesting herding, at the 10% level in at least one quantile in seven out of twenty MSAs, and indeed two markets show evidence in only one quantile. Conversely, γ_2 is significantly positive, suggesting reverse herding, in 15 markets and is more persistent across quantiles within the MSAs. For example, there is significantly positive evidence in 10 or more out of 13 quantiles in Chicago, Atlanta, Riverside, Detroit, and Minneapolis. Overall, there are 20 quantiles of herding and 81 quantiles of reverse herding. When not accounting for market conditions, there is more than four times as much evidence that cross-sectional dispersion increases non-linearly in response to increases in market returns as there is evidence of decreases in cross-sectional dispersion, and so there is substantially more evidence of reverse herding than of herding. This differs markedly from Ngene et al. (2017) who found approximately three times as much herding as reverse herding in the US regional house markets using 50 states as individuals and 9 census regions as markets. This may result from the existence of stronger private information in locally defined markets (i.e. MSAs used in this study instead of regions used in Ngene et al. (2017)), leading to greater reverse herding. As 101 out of 260 quantiles overall show some non-linear response, there is evidence that around three-fifths of responses can be explained by a rational asset-pricing model. Again, Ngene et al. (2017) found that more than half of market responses in total were irrational. Broadly similar rates of rationality are observed (roughly speaking, somewhere towards half in both cases). Chiang and Zheng (2010) found around half of the equity responses to be rational but all are herding, and indeed this may support the idea that the specific market characteristics of real estate tend toward reverse herding.

Concerning the fairly low persistence of herding across quantiles, Ngene et al. (2017) found weak evidence of persistent herding in the US regional housing markets across long periods, and behavioural motivators may only be evident under certain market conditions. Therefore, further analysis to examine the effects of market conditions is required as suggested by previous studies.

	τ=0.025	τ=0.05	τ=0.1	τ=0.2	τ=0.3	τ=0.4	τ=0.5	τ=0.6	τ=0.7	τ=0.8	τ=0.9	τ=0.95	τ=0.975	Herding	Reverse Herding
NYC	-0.075	-0.079	-0.033	-0.060	-0.077	-0.055	-0.057	-0.089**	-0.081**	-0.108**	-0.149**	-0.027	0.051	4	
	(0.101)	(0.074)	(0.066)	(0.054)	(0.073)	(0.042)	(0.040)	(0.036)	(0.037)	(0.051)	(0.058)	(0.102)	(0.085)		
LAX	-0.013	-0.007	0.000	0.007	0.014	0.011	0.013	0.041	0.052	0.100**	0.148***	0.108	0.109		2
	(0.013)	(0.014)	(0.009)	(0.011)	(0.014)	(0.017)	(0.028)	(0.048)	(0.044)	(0.048)	(0.051)	(0.143)	(0.126)		
CHC	0.293***	0.379***	0.399***	0.426***	0.497***	0.599***	0.584***	0.505***	0.430***	0.320***	0.360*	0.297	1.349**		12
	(0.075)	(0.067)	(0.075)	(0.078)	(0.097)	(0.081)	(0.080)	(0.096)	(0.085)	(0.115)	(0.201)	(0.301)	(0.542)		
DFW	0.090	0.117***	0.150**	0.197***	0.193***	0.189***	0.156***	0.127***	0.108	0.088	0.074	-0.152	-0.331		7
	(0.061)	(0.039)	(0.059)	(0.049)	(0.032)	(0.039)	(0.034)	(0.047)	(0.089)	(0.075)	(0.122)	(0.169)	(0.243)		
HOU	-0.033	-0.019	-0.023	0.035	0.033	0.058	0.028	0.006	0.008	-0.026	0.072	-0.113	-0.113		
	(0.071)	(0.074)	(0.065)	(0.084)	(0.063)	(0.046)	(0.070)	(0.089)	(0.077)	(0.130)	(0.175)	(0.169)	(0.141)		
WDC	-0.006	-0.028	-0.045	-0.014	-0.026	-0.060**	-	-	-	-0.169**	-0.101	-0.015	-0.048	5	
	(0.039)	(0.045)	(0.032)	(0.027)	(0.028)	(0.027)	(0.024)	(0.032)	(0.031)	(0.070)	(0.132)	(0.109)	(0.079)		
MIA	-0.021	-0.015	-0.009	-0.019	0.001	-0.010	0.003	0.0081	0.009	0.046	0.054*	-0.018	-0.090**	1	1
	(0.017)	(0.015)	(0.016)	(0.018)	(0.027)	(0.035)	(0.029)	(0.022)	(0.031)	(0.032)	(0.028)	(0.042)	(0.038)		
PHD	-	-0.113**	-0.081	-0.041	-0.007	-0.016	-0.002	-0.034	-0.025	-0.046*	-0.095	-0.113	-0.001	3	
	(0.073)	(0.055)	(0.051)	(0.039)	(0.034)	(0.039)	(0.036)	(0.022)	(0.027)	(0.027)	(0.064)	(0.072)	(0.066)		
ATL	0.026	0.171	0.295***	0.295***	0.256***	0.232***	0.207***	0.209***	0.216***	0.164	0.421**	0.351*	0.373***		10
	(0.139)	(0.160)	(0.086)	(0.042)	(0.037)	(0.035)	(0.045)	(0.044)	(0.058)	(0.215)	(0.177)	(0.194)	(0.136)		
PHN	0.010	0.010	0.008	0.007	0.001	-0.003	-0.004	0.006	0.002	-0.009	-0.015	-0.081	-0.057		
	(0.008)	(0.008)	(0.006)	(0.009)	(0.015)	(0.022)	(0.025)	(0.026)	(0.024)	(0.033)	(0.046)	(0.056)	(0.053)		
BOS	0.083*	0.108***	0.101*	0.046	0.070*	0.049	0.037	0.010	-0.015	-0.010	-0.019	-0.052	-0.106		4
	(0.048)	(0.033)	(0.054)	(0.036)	(0.038)	(0.041)	(0.057)	(0.041)	(0.042)	(0.060)	(0.058)	(0.054)	(0.104)		
SFR	-0.003	0.0172	0.050	0.035	0.065	0.068*	0.066	0.051	0.200**	0.142*	0.120	0.014	-0.001		3
	(0.076)	(0.053)	(0.044)	(0.022)	(0.041)	(0.039)	(0.060)	(0.087)	(0.081)	(0.081)	(0.088)	(0.151)	(0.086)		
RIV	0.0249	0.055***	0.050***	0.058***	0.070***	0.069***	0.072***	0.090***	0.082***	0.080***	0.053**	0.044**	0.033		11
	(0.015)	(0.015)	(0.017)	(0.014)	(0.012)	(0.015)	(0.019)	(0.017)	(0.015)	(0.012)	(0.023)	(0.023)	(0.029)		
DTR	0.178***	0.088***	0.137***	0.153***	0.161***	0.132***	0.103*	0.139*	0.087	0.268**	0.248	0.377**	0.475**		11

	(0.058)	(0.031)	(0.036)	(0.039)	(0.044)	(0.047)	(0.052)	(0.079)	(0.107)	(0.135)	(0.177)	(0.152)	(0.200)		
STL	-	-0.057**	-0.080**	-0.116**	-0.074	-0.065	-0.028	0.014	0.064	0.063	0.229**	0.199	0.083	4	1
	(0.018)	(0.025)	(0.035)	(0.050)	(0.049)	(0.051)	(0.068)	(0.062)	(0.073)	(0.081)	(0.099)	(0.121)	(0.122)		
MNN	0.187***	0.188***	0.189***	0.213***	0.221***	0.251***	0.257***	0.280***	0.505***	0.488***	0.264*	-0.041	-0.124		11
	(0.028)	(0.023)	(0.032)	(0.074)	(0.064)	(0.074)	(0.087)	(0.090)	(0.136)	(0.132)	(0.143)	(0.133)	(0.251)		
SDG	-0.016	-0.014	-0.006	-0.008	0.010	0.017	0.021	0.037	0.069	0.225***	0.244***	0.161	0.114		2
	(0.019)	(0.013)	(0.016)	(0.018)	(0.012)	(0.011)	(0.018)	(0.027)	(0.063)	(0.079)	(0.062)	(0.113)	(0.183)		
TMP	-0.011	-0.006	-0.006	-0.015	-0.004	0.019	0.008	0.006	0.003	0.005	0.032	0.106**	0.157***		2
	(0.025)	(0.024)	(0.019)	(0.021)	(0.024)	(0.026)	(0.022)	(0.019)	(0.031)	(0.038)	(0.065)	(0.054)	(0.035)		
DNV	0.056	0.009	0.065**	0.050*	0.039	-0.005	-0.004	-0.012	-0.041	-0.076	-0.057	-0.261	-0.231*	1	2
	(0.067)	(0.038)	(0.029)	(0.030)	(0.043)	(0.028)	(0.042)	(0.052)	(0.063)	(0.085)	(0.122)	(0.181)	(0.135)		
SLS	-0.483**	-0.478*	-0.318	0.175	0.173*	0.132	0.108	0.205	0.212	0.357**	0.312	0.117	-0.170	2	2
	(0.209)	(0.249)	(0.331)	(0.129)	(0.096)	(0.113)	(0.138)	(0.170)	(0.203)	(0.156)	(0.198)	(0.194)	(0.171)		

Table 5: Base Results
Estimated via equation 4 for a range of quantiles across the distribution, with standard errors provided in parenthesis. * denotes 10%, ** denotes 5%, *** denotes 1% significance. A significantly negative (positive) coefficient provides evidence of (reverse) herding, cumulative counts for each MSA are also provided

5.4.3. Up and Down Markets

Panel A of Table 6 shows that, in at least one quantile, there is evidence for herding in 10 down markets and seven up markets and evidence of reverse herding in six down markets and in 15 up markets. In terms of intensity, there is more evidence of persistence of herding in down markets and reverse herding in up markets. The latter point is in line with Duffee (2001) who found that stock returns are more dispersed in a rising stock market than when the market falls. When markets appreciate, investors diverge from the market return as they may be experiencing overconfidence and feel they can outperform the market. It may be that in benign market conditions, investors assign more weight to any private information they possess.

Herding in a down market may be rationally motivated when uninformed investors observe a declining market and, as they are unsure of the exact scale of the market disruption, copy the actions they can observe. In an environment of poor market conditions, investors may feel that any private information is not worth trading on and is overwhelmed by the negative signals shown in public information.

	Panel A - Up and Down Markets				Panel B - Volatility			
	Down		Up		Low volatility		High volatility	
	Herding	Reverse	Herding	Reverse	Herding	Reverse	Herding	Reverse
New York			1		9			6
Los Angeles		1		2		7		1
Chicago	1			11		11		8
Dallas			1	9		7		6
Houston				2		4	1	
Washington	9		1	1	1		1	1
Miami	12		7			5	10	
Philadelphia	1		3		6			
Atlanta	9			9		10		7
Phoenix	8			9		8	2	2
Boston		3		7		5		3
San Francisco				6		5		5
Riverside	2	4		7		5		11
Detroit		3		5		4		6
Seattle	6		6	2		3	3	
Minneapolis		4		11		10		5
San Diego		10				1		3
Tampa	4					1		
Denver	1			4		7	4	
St Louis			3	4		1	2	
Total	**53**	**25**	**22**	**89**	**16**	**94**	**23**	**64**

Table 6: Market Condition Results

Panel A estimates responses to market returns in up and down markets, and based on high and low volatility for Panel B. A cumulative count is presented due to space limitations. The coefficients and standard errors can be produced on request

Lan (2014) finds herding in up markets, not down, as do Hyun and Milcheva (2018) under a different empirical framework. However, Lan looks at China, one large national market, which may not possess the same structure as a metropolitan housing market, and previous studies have suggested that herding is more common in developing markets such as China. In addition, Lan did find evidence of herding in a down market if it was also turbulent.

5.4.4. Volatility

Volatility was defined as the standard deviation of the previous 12 months of returns. In line with the base and asymmetric analyses, Panel B of Table 6 first shows that herding is again less common than reverse herding. More interestingly, there is a 50% increase (23 against 16 quantiles) in the prevalence of herding when volatility is high, in line with expectations drawn from Christie and Huang (1995). However, this cannot be overstated and herding is still only present in around 10% of quantiles. It is possible that in structures where market signals are clearer, it is easier for traders to herd around the index as the index is published and current, whereas house market prices are much more lagged and not always for the exact asset as housing is a highly heterogeneous investment asset. Although herding can result from information asymmetry, the investor needs a minimum level of market information to actually copy. Herding is still observed in low volatility states and indeed Hwang and Salmon (2004) find herding in tranquil market conditions, and Zhou and Anderson (2013) suggest that wider market conditions are also a determinant of whether herding exists in turbulent markets.

Reverse herding is 50% more common in low volatility than high volatility markets (94 against 64 quantiles) as expected. Low volatility may create some complacency as conditions are unchanging and investors become overconfident, which has been theorised as a motivator for reverse herding. Fairly significant reverse herding is found also in turbulent periods, at a rate almost three times as prevalent as herding, which is generally consistent with findings of Philippas et al. (2013) with REITs and Chang et al. (2000) in wider equity markets.

In the cryptocurrency market, Coskun et al. (2020) found the existence of herding under low volatility conditions and reverse herding under high volatility states. This could explain why the presence of herding is not overwhelming in more volatile conditions (16 up to 23 counts) and why reverse herding is much more prevalent. The market structure of housing is more akin to cryptocurrency with high levels of individual ownership and information asymmetries. Hwang and Salmon (2004) also said that herding can take place under non-extreme conditions of normality, and so as with previous studies (Griffin et al., 2003), no behaviour is completely unobserved under any conditions.

5.4.5. Global Financial Crisis

There is evidence of long-term changes in herding behaviour (Klein, 2013) which may be related to the level of market sophistication. However, previous studies find evidence of change in behaviour after the GFC (Zhou and Anderson, 2013), an event that may have served as a catalyst as the fiscal and monetary action was accompanied by regulatory change.

Using the Federal Reserve definition of the recession lasting from December 2007 until June 2009, estimated results for cross-sectional responses demonstrate the existence of irrational behaviour both before and after the GFC. The use of the economic defined recession has been chosen to allow for clear comparison of concurrent responses to national conditions, rather than MSA-specific housing cycles. The "during" period is too short to draw any significant economic conclusions from, but the pre- and post-GFC periods are almost identical in size (142 and 139 months respectively), which allows for easy comparison of behaviour.

As anticipated, both herding and reverse herding behaviours changed significantly after the GFC. Panel A from Table 7 shows a marked decline in herding from 47 to 13 quantiles, suggesting the GFC did cause some structural changes to house market behaviour. The occurrence of reverse herding almost doubled after the GFC as seen by the increase in quantile evidence from 59 to 107, such that almost eight times as much reverse herding as herding was recorded post-GFC. As consistent house price appreciation was seen from late 2012, which should be highly correlated with overconfidence, the presence of substantial reverse herding is expected.

There is persistent evidence of herding in Miami, Philadelphia, Riverside, and Washington both before and after the GFC, whilst persistent reverse herding is observed both before and after the

GFC in Atlanta, Chicago, Dallas, Houston, Minneapolis, Phoenix, Riverside[1] , and San Francisco. On the contrary, some cities saw marked changes in behaviour after the crisis, such as Detroit

[1] Riverside, although showing much more prevalence of reverse herding, still exhibits some evidence of herding in both periods.

	Panel A - GFC				Panel B - Overconfidence			
	Pre-GFC (142)		**Post-GFC (139)**		**Unconfident (139)**		**Overconfident (161)**	
	Herding	**Reverse**	**Herding**	**Reverse**	**Herding**	**Reverse**	**Herding**	**Reverse**
New York	1			6	4		3	
Los Angeles	1			3		1		4
Chicago		7		4		4		7
Dallas		6	1	8	1	4		8
Houston		5		2		1	2	
Washington	8		5		7			
Miami	7		1	1	6			6
Philadelphia	6		2	1	6		2	
Atlanta		2		10				11
Phoenix		5	1	7		9		4
Boston		4	2			6	6	
San Francisco		8		11	1			8
Riverside	1	2	1	6		5		11
Detroit		9				4		5
Seattle	9			11	4	1		1
Minneapolis		11		10		7		8
San Diego	3			4	3	1		8
Tampa				5	1			8
Denver	9			7	3	3		
St Louis	2			11			1	1
Total	**47**	**59**	**13**	**107**	**36**	**46**	**14**	**90**

Table 7: Market Condition Results (continued)

Panel C aggregates results split between pre- and post-GFC periods, and Panel D estimates behaviour according to overconfidence or unconfidence. A cumulative count is presented due to space limitations. The coefficients and standard errors can be produced on request.

which went from strong evidence of reverse herding to showing no irrational market responses after. This may result from the context that the pre-GFC housing bubble was national whilst the recovery has been more geographically varied, and indeed it seems that Detroit never recovered in housing or economic terms as opposed to other major MSAs. This suggests variation between MSAs that motivates further research on the impact of local characteristics.

5.4.6. Overconfidence

Previous studies suggest that both market structure and inherent behavioural characteristics lead to generally overconfident conditions in housing markets. The results in Panels A and B of Table 6 and Panel A of Table 7 also suggest the potential effect of overconfidence, especially on reverse herding behaviour. Following prior studies (Blasco et al., 2012; Liao et al., 2011; Philippas et al., 2013; Simões Vieira and Valente Pereira, 2015) that suggest sentiment may determine herding behaviour, an innovative combination of sentiment measures is used as a proxy for overconfidence.

The data is split into overconfident and unconfident groups based on the ratio of house buying to economic sentiment, assuming that overconfidence (unconfidence) is present when house buying sentiment is more positive (negative) than economic sentiment. This allows for a clear comparison of how dependent behavioural responses are on relative confidence in the housing markets.

Panel B of Table 7 shows evidence that herding is more than twice as prevalent in unconfident markets, in line with equity market results (Bekiros et al., 2017; Choi and Yoon, 2020). In a real estate context, Ngene et al. (2017) also showed that regional housing markets herd in times of uncertainty in both financial markets and economic policy. Consistent with expectations, more than twice as much reverse herding is observed when markets are overconfident. As with previous sub-analyses, there is still herding and reverse herding in both market states and indeed prior findings (Choi and Yoon, 2020; Ngene et al., 2017; Simões Vieira and Valente Pereira, 2015) are mixed.

Note that Panel B shows broad disparities between MSAs, as they respond heterogeneously to overconfidence in a manner suggestive of local variation in behaviour. Previous literature (Carlino and DeFina, 1998; Carlino and DeFina, 1999; Giannakis and Bruggeman, 2017; Gupta and Kabundi, 2010; Hwang and Quigley, 2006) shows that, due to differing economic structures, sub-national markets react heterogeneously to exogenous shocks, motivating further research into local measures of overconfidence.

5.5. Conclusion

This paper examines herding and reverse herding at the MSA-level and found extensive evidence of potentially irrational responses to large increases in absolute market returns. Analysing the phenomena on an appropriate spatial level, and across a variety of market conditions, has contributed to herding research.

As expected from the review of existing literature, herding exists primarily in downturns whereas reverse herding exists under more bullish market conditions. Likewise, in line with rational herding under information asymmetries or inefficiencies, herding is high in volatile markets and reverse herding is somewhat more common in stable conditions. In terms of temporal change, the GFC may have caused some permanent change in behaviour as herding became sparse whilst the occurrences of reverse herding doubled. Wide spatial and temporal variation in herding and reverse herding behaviour warrants further investigation to isolate the MSA-specific characteristics that determine the rationality of market responses.

Existing theory demonstrates that inefficient markets and innate homeowner overconfidence may contribute to reverse herding, which is supported by evidence from using a sentiment-derived proxy for overconfidence, motivating future research into the potential link between confidence and irrational behaviour, especially by establishing a good measure of confidence at the MSA-level, if possible.

In this relatively local geographical context, individuals may be better informed than stylised facts on real estate information asymmetries suggest and indeed it can be assumed they possess significant knowledge on local housing markets. It would follow then that, due to strong private information, markets are more overconfident than expected, therefore motivating reverse herding. Additionally, housing markets still exhibit a low level of institutional involvement relative to securitised investment classes, and so homeowners are not at an informational disadvantage.

Lastly, consumption is always the primary driver for housing and therefore investment must take a secondary role, especially for the owner-occupiers who still constitute around two-thirds of the asset holders. Therefore, further research should consider herding not only relative to investment considerations but also in the context of consumption-driven behaviour, especially homeownership. Also worthy of consideration is that regional variation may be due to spatial constraints common in real estate markets.

These findings have policy implications as, because cities display these irrational behaviours under certain market conditions, these results may have use as leading indicators, especially considering the link between herding and bubble formation. These behaviours, therefore, operate as important warning signs for lenders, investors, and policymakers.

6. Estimating the Connection between Herding and Price Bubbles

6.1. Abstract

This paper examines the connection between herding and house price bubbles in US metropolitan markets. Contrary to widely accepted theory on herding as a prime motivator for price bubbles, limited evidence is found to support this proposition. The estimated results suggest that price bubbles, proxied by excess returns, are a significant motivator for herding. There are substantial time delays in the mechanism which are particularly relevant for discussions of market efficiency and persistence in behaviour. Sub-analyses suggest that the Global Financial Crisis may have caused structural changes in the mechanism between herding and price bubbles, and that size effects may result from greater investor focus in larger cities.

Key Words: Herding; Reverse Herding; Housing; Bubbles

6.2. Introduction

This study examines price bubbles as an outcome of herding in urban housing markets across the United States (US). Herding is proposed as a theoretical explanation for price bubbles, and this potential relationship is estimated via a vector autoregression (VAR) framework. Excess returns are used to proxy the presence of a speculative price bubble at the metropolitan level and the potential impacts of structural breaks, along with size effects and idiosyncratic volatility, on the relationship between herding and price bubbles are also considered. A panel of the fifty largest Metropolitan Statistical Areas (MSAs) is used to incorporate both national and local level controls on house price changes. The idiosyncratic impacts of price volatility on market dynamics are also included and the study assesses the relative impacts of local and national determinants on the extent of herding and price bubbles.

Devenow and Welch (1996) broadly define herding as significant correlated behaviour amongst individuals. In an asset market, investors may not value their investment choices using a rational asset pricing model and instead purely copy the actions of other investors that can be observed (Banerjee 1992). This irrational behaviour can in many cases lead to sup-optimal investment decisions by individuals (Devenow and Welch 1996). These irrational decisions by investors can then feed irrational market outcomes such as price bubbles, with a resulting price crash and loss of value (Lux, 1995).

The psychological explanation behind this mimetic behaviour can be broadly categorised into rational and irrational motivations. In the former, if an investor or consumer believes that others possess more (or more accurate) information then it is optimal to copy others, most commonly by following a market metric such as a price index (Bikhchandani et al 1992, Bikhchandani et al 1998, Welch 1992). This is also the case when the market is experiencing high price volatility, creating uncertainty and pushing investors toward following the herd (Christie and Huang 1995). Irrational herding is likely to result from biases (Barber et al 2009) geared toward a social or personal desire to conform. When the role of real estate as an indicator of wealth is considered then the potential significance of irrational motivations is clear (Hopkins and Kornienko 2004, Brown et al 2011, Frank 2005).

This process also suggests the assumption of continued price appreciation, which in itself will fuel price bubble formation until the flow of new capital is exhausted. This demonstrates a transmission mechanism between behavioural phenomenon and systemically important market outcomes (Lux 1995).

Prior research has demonstrated that, as with many market dynamics, herding is commonly observed but that the prevalence is inconsistent across space and time. Specifically, Griffin et al. (2003) find that herding is neither universal not similar across assets and markets and Galariotis et al. (2015) find that it is heavily influenced by country and time-specific factors. For example, herding is heavily determined by exogeneous factors (Goodfellow et al 2009). This demonstrates a requirement for better understanding of the asset- and market-specific characteristics that drive herding and robust estimation of the association between herding and bubbles. This is especially relevant in the context of the unique dynamics of housing markets as housing is a substantial and universal asset class. There is evidence of more extreme house price movements than what would be expected from fundamental dynamics (Hott 2012), as well as extensive evidence of herding across various market conditions and across US regions (Ngene et al 2017).

The role of real estate as a core part of an individual's investment portfolio is complicated further by housing primary function as a consumption good (Miller and Pandher 2008), which then involves non-financial decision making, social concerns and psychological biases.

Whilst herding can be clearly defined, price bubbles are a fuzzier concept and are often identified post-event in a more qualitative manner. Kindleberger (1978) says "an upward price movement over an extended range that then implodes", Garber (2000) says a bubble is the part of the price movement that cannot be explained the fundamentals and Hott (2009) says "an asset price bubble exist when the price of an asset deviates from its fundamental value over a longer period".

Overall, a common definition is that a bubble is some deviation from an assumed rational level. Considering that herding is a deviation from rational asset pricing, then the potential causality from herding to price bubble formation is evident. The existence of bubbles suggests some form of speculation, whereby an investor purchases an asset only in the assumption of price appreciation. Whilst this strategy would not be profitable in an efficient market, there is limited evidence that housing is a perfectly efficient market. In addition, this speculative approach is complicated in housing investment due to its primary purpose as a consumption good.

Even without a speculative motive, Hott (2009) models a theoretical proof that price bubbles can be explained by herding behaviour and Thoma (2013) shows that greater herding (caused by trend-chasing when sustained price appreciation is declared to result from fundamentals) leads to an increased likelihood of a destructive bubble.

It could be assumed that a more efficient market would be less susceptible to herding and the resultant bubble formation, as rational herding would have no rationale. However, housing suffers from informational asymmetries, restricting efficiency, and is still an illiquid market. Furthermore, the inability to short direct real estate allows for persistent mispricing.

Whilst there is no statistical test for bubbles, and they are much easier to describe and identify after the event, Harras and Sornette (2011) used a theoretical model to determine the origin of a bubble and assert that herding is an adaptive and imitative mechanism that fuels bubbles. Increasing bubbles are enhanced by high returns, and so high returns, estimated by squared returns, form an effective proxy for a bubble, which aligns with the definition provided by Kindleberger (1978). However, there is some limitation in arguing that high returns are themselves irrational so reverting to the concept of speculative behaviour may be captured by assessing the impact of excess returns. The empirical challenge is to define "excess" returns, which themselves also may require some definition of a rational or fundamental return. In addition, brief periods of high or excess returns may be noise, and so it is important to isolate the persistence of these dynamics to understand the behavioural forces.

US household wealth is approximately $140 trillion, of which housing itself is valued at approximately $40 trillion, a figure which has more than doubled over the past decade. This illustrates the systemic importance to the general economy and the financial sector, as clearly seen in the Global Financial Crisis. In addition, whilst also being the single largest borrowing and investment decision most individuals make in their lifetime, the requirement of all individuals to access housing as shelter provides a parallel and primary consumption demand for housing, with crises in the housing sector directly impacting individual welfare.

The systematic importance of housing markets, and the potential link that shows from theoretical and empirical analysis, shows that further investigation of herding and bubbles is justified in housing markets. Whilst there is extensive discussion of these issues, much of the bubble identification is made after the event which limits its effectiveness as a leading indicator and use for any practical decision making, either for investors or policy makers.

As mentioned, it is challenging to empirically estimate bubbles and additionally so to construct a link with herding. However, the potential impacts on the economy and individuals justify continued empirical investigation as, if these behaviours can be forecast, then more effective policy making, and more robust portfolio and risk management, can be undertaken.

As there are theoretical justifications for price bubbles resulting from herding, and bubbles can lead to significant issues, this also provides motivation in finding if there is a relationship between herding and bubbles in housing markets.

6.2.1. Contribution

Firstly, a commonly discussed link between herding and bubble formation is tested to add to the literature on these systemically important market dynamics. Secondly, this study adds to the fairly limited research on herding in housing markets, especially as real estate provides a dual investment and consumption function that complicates the behavioural motivations. Lastly, the study separately considers the rational motivations for herding and the speculative elements of price bubbles, as well as the relative role of investment- and consumption-driven markets.

6.2.2. Summary of Results

This study finds that, contrary to expectations, herding is not a significant determinant of bubble formation but that excessive price appreciation is often a significant factor in herding, and this result is more prevalent in larger cities. However, the role of asymmetric information and excessive volatility is inconsistent with the theory and the relationship breaks down after the Global Financial Crisis.

6.3. Data and Methodology

6.3.1. Price Bubbles

Following the discussion of existing literature, a bubble can be broadly defined as some form of excess return. Incorporating this into the model requires a realistic proxy drawn from market data. Additionally, it is relevant that the use of the squared return metric in the adjusted asset pricing model aligns with previously mentioned definitions of bubbles as some form of excessive return. As strong upward returns may be indicative of abnormal positive returns, so squared returns can also proxy any price bubble (Litimi 2016). From a behavioural perspective, large (or non-linear) returns may be a relevant motivator as they give evidence that supports a belief in speculative returns therefore, in line with the original CCK estimation approach to herding identification, a squared return metric is employed.

Most definitions of asset bubbles relate to a positive deviation in price from a rational level, which can be interpreted as an excess return driven by speculation. Clearly, the rational level is built on some fundamental market measures which would generate a "fair" return. The measurement of "excess" returns is somewhat complicated by the requirement for a definition of "fair" returns. There are various data and computational barriers that may make it unlikely that the vast majority of house purchasers calculate a time series of imputed rents and discount them at an appropriate rate. Instead, it is likely that the investment decision part of the dual investment-consumption decision is based on the expected return from acquiring the asset. In line with most investment strategies, this is usually based on forming expected returns from historic observations. Rather than deriving a fundamental return from estimates of price level, housing is liquid and homogeneous enough to allow purchasers to consider recent housing price movements and construct a simple expected return value.

Expected returns will adjust over time and so they can reasonably be based on fairly recent price data. The historic return will come from price movements they observe on recent transactions, much like comparable evidence used for a valuations or appraisals, and so the expected return is set to equal the average of the last year of price movements.

When price appreciation exceeds the expected level, it means there is evidence of some level of excess return. This relatively high rate of return could result from strong price appreciation driven by fundamentals, but this argument is often invalid and has often been disproven post-event. Instead, as herding can result from speculative expectations of returns (Hott 2009) then follows that the presence of speculative returns may reasonably suggest that herding may follow. If returns exceed the expected level then it may show evidence of accelerating returns, which may also be indicative of a bubble. This would also clearly be dependent upon the degree of persistence.

From the perspective of market efficiency, it may be difficult to achieve information symmetry in a particular transaction, although presumably the role of broker will ameliorate some of this. However, the relative accessibility of cheap, consistent and regular general price data means that investors can adjust their expectations frequently.

To operationalise this, the moving average of twelve months of returns is used as a measure of the expected return. As mentioned, a typical investor is more likely to derive expectations from recent historic data than from any more developed analysis. Therefore, if the expected return is likely to be a function of historic data, then any return above the historic rate, as measured by the moving average, could be indicative of speculative returns and so this represents a good proxy for the speculative element of herding.

Generally, it is assumed that reduced dispersions of returns, measured by lower values of CSAD, will lead to increased probability of a price bubble, which is proxied with both squared and excess returns. However, it is worth noting that even without a speculative motivation, it is possible that herding will lead to a price bubble. For robustness, the coefficients are estimated with both non-linear squared results and excess results.

6.3.2. Control Data

The literature review suggests that herding should be a motivating factor in the formation of price bubbles, the latter of which can be proxied by non-linear or excess returns. Clearly, a substantial part of house price movements will be a reaction to the broader fundamental conditions. In order to control for this core response and to isolate the impact of herding on bubbles, the model controls for the fundamental determinants of house price movements. Having removed the impact of the fundamentals then an irrational movement remains which is explained as a price bubble. One of the potential causes for herding has been discussed as a speculative momentum in investor behaviour, which itself would be any expected price appreciation not justified by movements in the fundamental characteristics of the market.

Previous studies suggest that several economic conditions may be determinants of house price movements; real domestic gross product, bank credit growth, long-term bond yields, real effective exchange rates (Vogiazas and Alexiou 2017). Land supply constraints and mortgage rates are also included to control for the property market characteristics. Volatility is also included to account for its role in driving rational herding and is included as a trigger (Yao et al 2014).

Bank credit issued by all commercial banks, collected by the Federal Reserve, is measured by the seasonally adjusted monthly percentage change. The real broad effective exchange rate, from the Bank for International Settlements, is also measured by the monthly percentage change. Data for yields on ten year government bonds comes from the Organization for Economic Co-Operation and Development and the monthly change is used. Likewise, the monthly change in 30-year fixed-rate mortgage rates are collected by Freddie Mac. All these variables are gathered via FRED. Real gross domestic product is collected at the MSA-level from the BEA and transformed into annual log differences. Land supply elasticities are collected from Saiz (2008), and these are time invariant.

Other than economic growth, land supply constraints and volatility, these controls are national-level variables. The financial aspects of house price dynamics are set at the national level as financial capital is not spatially fixed, and the main local property market factors of demand and supply have been incorporated. A significant part of price dynamics results from the inelastic supply response which it is believed will capture the core of the local market character and its role in price movements.

The general a priori assumption derived from the theory is that the presence of herding may lead to bubble formation. This would be identified by a lower value of CSAD being a significant determinant of higher non-linear or excess returns. Broader expectations are that less elastic supply functions would also contribute to more significant herding and bubble formation, as the supply response is too constrained to allow for a quick price adjustment. If there is a speculative element to herding, then stronger economic growth and easier access to credit will, respectively, stimulate this expectation and provide the necessary financing conditions, also leading to lower CSAD and higher non-linear or excess returns.

6.4. Empirical Results

6.4.1. Descriptive Statistics

	Mean	Median	Min.	Max.	Skew.	Kurt.	SD
Returns	0.255	0.301	-4.051	3.859	-0.699	6.814	0.701
Returns^2	0.556	0.191	0.000	16.411	5.728	51.086	1.133
Excess	0.011	0.006	-2.714	2.725	0.070	7.546	0.399
CSAD	0.402	0.371	0.002	1.775	1.427	7.255	0.165

Table 8: Descriptive Statistics for Herding and Bubble Measures
Descriptive statistics for the price returns (from equation 7) and their squared values, excess returns as described in "Method" and the cross sectional-sectional absolute deviation (CSAD) (from equation 8). All variables have 11,450 observations and form a perfectly balanced panel

Returns are measured from the difference in natural logged values of the monthly ZHVI, and data is used from January 2002 until January 2021. Descriptive statistics for returns, extreme returns, excess returns, volatility and dispersions are presented in Table 8 and show little evidence of extreme outliers. The correlations in Table 9 are relatively low and do not suggest any concern of multicollinearity, so the base conditions for VAR modelling should be satisfied.

	Returns	**Returns^2**	**Excess Ret.**	**CSAD**
Returns	1.000			
Returns^2	0.012	1.000		
Excess Ret.	0.428	0.048	1.000	
CSAD	-0.325	0.252	-0.007	1.000

Table 9: Correlations for Herding and Bubble Measures
Pairwise correlations of price returns (from equation 7) and their squared values, excess returns as described in "Method" and the cross sectional-sectional absolute deviation (CSAD) (from equation 8)

6.4.2. Empirical Results

There was no evidence of a cointegrating relationship between the variables and no evidence for unit roots in the panels. For robustness, Im-Pesaran-Shin, Levin-Lin-Chu and Harris-Tzavalis unit root tests were used and all found strong evidence of stationarity in panels for CSAD, squared, and excess returns. A panel VAR is estimated in the form;

$$\boldsymbol{CSAD_{t,j}} = \propto_{1,j} + \sum_{i=1} \boldsymbol{\beta_{1i,j}\, CSAD_{t-1,j}} + \sum_{i=1} \boldsymbol{\gamma_{1i,j}\, bubble_{t-1,j}} + \boldsymbol{\delta_{1,j} X_{t,j}} + \boldsymbol{\varepsilon_{1t,j}}$$

$$\boldsymbol{bubble_{t,j}} = \propto_{2,j} + \sum_{i=1} \boldsymbol{\beta_{2i,j}\, CSAD_{t-1,j}} + \sum_{i=1} \boldsymbol{\gamma_{2i,j}\, bubble_{t-1,j}} + \boldsymbol{\delta_{2,j} X_{t,j}} + \boldsymbol{\varepsilon_{2t,j}} \qquad (11)$$

where the bubble measure can be estimated via squared returns or excess returns to ensure some robustness of results and X represents a vector of exogenous controls, seen in Table 10. All the control variables follow a monthly frequency, other than supply elasticities which are time invariant.

	Mean	Median	Min	Max	Skew	Kurt	SD
Real GDP	1.956	2.200	-23.300	13.300	-1.040	8.853	3.050
Bank credit growth	0.480	0.434	-1.056	5.136	2.499	17.694	0.651
Long-term bond yields	-0.018	-0.010	-1.110	0.650	-0.515	6.574	0.207
Real effective exchange	-0.046	-0.098	-3.565	5.637	0.511	4.504	1.246
Mortgage rates	-0.019	-0.035	-0.802	0.632	0.548	6.714	0.165
Supply elasticities	1.661	1.500	0.600	4.000	0.716	2.604	0.850

Table 10: Descriptive Statistics for VAR Control Variables
Descriptive statistics for the control variables used to estimate equation 10, as described in "Control Data". All growth values are monthly, supply elasticities are time invariant values

For all estimated models, three lags are presented. Most previous VAR-based analysis has used annual frequencies, and this would dictate 12, 24 or 36 lags which would be impractical. However, an analysis of these lags (not presented) demonstrates that the estimated coefficients for longer lags are not statistically significant, suggesting that the mechanism is relatively quick. There could be more work on the dynamic structure to understand some of the lag interactions.

6.4.3. Base Model

Table 11 shows that, for both measures of bubble formation, there is no evidence that herding, as measured by CSAD, is a significant determinant of bubble formation. Rather, both bubble measures are a significant determinant of herding. The unit root circles for the companion matrix eigenvalues suggest the model is stable.

Specifically, when using squared returns in Panel A to proxy for extreme price appreciation, the impact of herding is statistically significant and suggests in the short run that greater non-linear returns result in higher levels of CSAD. This indicates that returns are more dispersed that would be expected from the initial asset pricing model, creating a phenomenon of reverse herding. There is some evidence that higher CSAD leads to increased squared returns via the coefficient of the first lag, however this displays a low level of statistical significance.

	Panel A - Squared Returns		**Panel B - Excess Returns**	
	Bubble	**Herding**	**Bubble**	**Herding**
Bubble (-1)	**0.883*** **	**0.008*** **	**0.911*** **	**-0.011*** **
	(0.009)	(0.002)	(0.009)	(0.003)
Bubble (-2)	**0.205*** **	**0.004** **	**0.114*** **	**-0.008* **
	(0.012)	(0.002)	(0.011)	(0.004)
Bubble (-3)	**-0.459*** **	**-0.007*** **	**-0.515*** **	**0.017*** **
	(0.012)	(0.002)	(0.011)	(0.004)
CSAD (-1)	**0.106* **	**0.762*** **	0.005	**0.768*** **
	(0.056)	(0.009)	(0.025)	(0.009)
CSAD (-2)	-0.110	**0.036*** **	-0.003	**0.038*** **
	(0.070)	(0.012)	(0.032)	(0.012)
CSAD (-3)	0.049	**-0.028** **	0.033	**-0.028** **
	(0.070)	(0.012)	(0.032)	(0.012)
Real GDP	-0.000	**-0.001*** **	**-0.006*** **	**-0.001*** **
	(0.003)	(0.000)	(0.001)	(0.000)
Bank credit growth	-0.005	**-0.004*** **	-0.002	**-0.004*** **
	(0.006)	(0.001)	(0.003)	(0.001)
Long-term bond yields	**-0.136*** **	**-0.016*** **	**0.069*** **	**-0.012** **
	(0.037)	(0.006)	(0.017)	(0.006)
Real effective ex. rates	-0.005	**0.007*** **	**0.008*** **	**0.002*** **
	(0.004)	(0.001)	(0.002)	(0.007)
Mortgage rates	**0.093* **	0.011	**-0.091*** **	0.007
	(0.048)	(0.008)	(0.022)	(0.008)
Supply elasticities	-0.006	0.000	0.001	0.000
	(0.005)	(0.001)	(0.002)	(0.001)
Volatility	**0.263*** **	0.007	0.002	**0.017*** **
	(0.030)	(0.005)	(0.013)	(0.005)
Constant	**0.061*** **	**0.040*** **	0.003	**0.039*** **
	(0.017)	(0.003)	(0.007)	(0.003)
Adjusted R^2	0.858	0.787	0.764	0.786
Observations	11,025		11,025	

Table 11: Results from VAR(3) Estimation
Both panels are estimated using equation 10. Figures in parentheses are standard errors. * denotes 10%, ** denotes 5%, *** denotes 1% significance

Strong non-linear price appreciation may exhibit as a boom condition and create an extreme adjustment in expected returns. This in itself may motivate more excessive risk taking via a behavioural reaction. If herding is a reaction to weaker market conditions or high price volatility, then it could follow that, when markets are strong then reverse herding will be present. If there is some excessive level of expected return resulting from speculation, then investors will attach value to their own abilities and deviate from the market average, leading to greater than expected price dispersion.

Panel B shows that, when price bubbles are proxied by excess returns, the estimated coefficients still run counter to the a priori hypothesis, but contrary to non-linear proxies they show that greater excess returns result in lower rates of CSAD, indicative of herding, in the initial few lags. Intuitively this suggests that the formation of a price bubble leads to increased herding. If the potential speculative element of a bubble is considered to be a motivator, excess returns may indicate greater than expected appreciation, encouraging a general belief in the advantage of following the market trend due to the assumption of more price appreciation.

When the linkages are estimated for the market as a whole, then there is no comprehensive evidence that CSAD, whether increasing or decreasing, is a significant determinant of bubble formation. Table 11 presents relatively short autoregressive coefficients, however a check of longer lags (not presented here) shows that statistical significance is not present at any longer lags. Overall, contrary to the hypothesis, herding may be the result, rather than the cause, of price bubbles.

The estimated coefficients for volatility suggest that more volatile markets are in fact less likely to herd, which runs counter to expectations and suggest the need for further sub-analysis to better understand the dynamics. Elasticity is neither economically nor statistically significant for either measure of bubble formation.

6.4.4. Granger Causality Analysis

Whilst some economic and statistical values have been attached to the determinants, the causality of the influences is also worthy of estimation. Specifically, the literature review suggested that price bubbles, regardless of definition, may be triggered by herding behaviour. Therefore, some measure of causality may help to assess if this theoretical mechanism is shown in the data.

	Obvs.	**F-statistic**	**P-value**
		Excess Returns	
Bubbles do not Granger cause herding	11,350	6.786	0.001
Herding does not Granger cause bubbles		9.667	0.000
		Squared Returns	
Bubbles do not Granger cause herding	11,350	28.574	0.000
Herding does not Granger cause bubbles		5.720	0.003

Table 12: Pairwise Granger Causality Tests
Pairwise Granger causality tests for bubbles (as measured by excess returns and squared returns, as described in "Method") and cross-sectional absolute deviation (CSAD) (from equation 8)

In all cases shown in Table 12, the null hypothesis is rejected, and there is evidence of Granger causality in all four directions. Regardless of the specific measure of price bubble, the evidence supports the contention that herding behaviour can drive excessive price appreciation. However, the evidence that price bubbles also Granger cause herding (or at least, lower price dispersion), may suggest some self-reinforcing loop.

Of course, other estimates of causation would be possible with more intensive econometric analysis, but the results do have implications for any forecasting models that seek to estimate market conditions.

6.4.5. Global Financial Crisis

There is limited evidence of consistent herding behaviour across space and time (Galariotis et al 2015) and herding behaviour may have changed after the GFC (Zhou and Anderson 2013), both as it is dependent on market sophistication (Chang and Lin 2015, Klein 2013) and as herding may have generally declined over time (Lam and Qiao 2015). Considering that the GFC led to significant changes in market structure via changes in leverage regulations and lower yields then this may also have changed the mechanism between herding and price bubbles.

The Federal Reserve definition of the recession from December 2007 until June 2009 is used as an assumed structural break caused by the GFC and the dataset is split into pre-GFC and post-GFC sub-samples.

	Mean	Median	Min.	Max.	Skew.	Kurt.	SD
			Pre-GFC (3550 observations)				
Returns	0.501	0.389	-2.080	3.859	0.663	5.356	0.618
Squar. Ret.	0.633	0.180	0.000	14.893	4.620	37.238	1.155
Excess	-0.064	-0.031	-2.314	2.053	-0.689	8.184	0.350
CSAD	0.366	0.341	0.005	1.339	1.578	8.521	0.144
			Post-GFC (6900 observations)				
Returns	0.291	0.323	-2.823	2.792	-0.162	4.521	0.554
Squar. Ret.	0.932	0.183	0.000	7.971	4.339	31.467	0.636
Excess	0.084	0.058	-2.714	2.725	0.392	7.741	0.396
CSAD	0.402	0.377	0.002	1.216	0.923	4.687	0.151

Table 13: Descriptive Statistics for Herding and Bubble Measures by Period
Descriptive statistics for the price returns (from equation 7) and their squared values, excess returns as described in "Method" and the cross sectional-sectional absolute deviation (CSAD) (from equation 8). The samples are defined as pre-GFC before December 2007 and post-GFC after June 2009

The hypothesis here is that the GFC reduced the prevalence of herding, which was demonstrated in the first paper (see pages 77-80), and any relationship between herding and bubble formation may also have changed. Considering the descriptive statistics in Table 13, although returns declined after the GFC, excess returns actually increased, suggesting there was more evidence of bubble formation in the later period. However, the level of dispersion has increased which suggests there may have been less herding, although CSAD-specific skewness and kurtosis did reduce. This does, at a basic descriptive level, support the contention that markets became more rational after the GFC with less herding and more normal data distributions.

	Panel A - Pre-GFC		**Panel B - Post-GFC**	
	Bubble	**Herding**	**Bubble**	**Herding**
Bubble (-1)	**0.940*** **	0.001	**0.893*** **	-0.005
	(0.017)	(0.007)	(0.011)	(0.004)
Bubble (-2)	**0.157*** **	-0.005	**0.065*** **	-0.002
	(0.022)	(0.009)	(0.015)	(0.005)
Bubble (-3)	**-0.517*** **	0.003	**-0.488*** **	**-0.010* **
	(0.022)	(0.009)	(0.015)	(0.005)
CSAD (-1)	0.040	**0.839*** **	0.029	**0.730*** **
	(0.041)	(0.018)	(0.033)	(0.012)
CSAD (-2)	-0.069	**0.057** **	-0.015	0.010
	(0.054)	(0.023)	(0.042)	(0.015)
CSAD (-3)	0.010	**-0.084*** **	0.027	0.002
	(0.054)	(0.023)	(0.042)	(0.015)
Real GDP	**0.002** **	**-0.001** **	**-0.007*** **	**-0.001** **
	(0.001)	(0.000)	(0.001)	(0.000)
Bank credit growth	-0.003	-0.001	**0.010** **	**-0.008*** **
	(0.006)	(0.002)	(0.005)	(0.002)
Long-term bond yields	0.008	-0.002	**0.074*** **	**-0.029*** **
	(0.028)	(0.012)	(0.028)	(0.010)
Real effective ex. rates	**0.005** **	0.001	**0.012*** **	0.001
	(0.003)	(0.001)	(0.002)	(0.001)
Mortgage rates	-0.026	-0.006	**-0.103*** **	**0.015* **
	(0.035)	(0.015)	(0.038)	(0.014)
Supply elasticities	-0.003	0.002	0.004	-0.001
	(0.003)	(0.001)	(0.003)	(0.001)
Volatility	**-0.149*** **	**0.020** **	**0.054*** **	-0.002
	(0.022)	(0.009)	(0.018)	(0.001)
Constant	0.009	**0.036*** **	0.012	**0.048*** **
	(0.011)	(0.005)	(0.011)	(0.004)

Adjusted R^2	0.819	0.796	0.727	0.743
Observations	3283		6566	

Table 14: Results from Structural Break Model using Excess Returns
Both panels are estimated using equation 10. Figures in parentheses are standard errors. * denotes 10%, ** denotes 5%, *** denotes 1% significance. The samples are defined as Pre-GFC before December 2007 and post-GFC after June 2009

As with the base model, Table 14 shows no evidence of herding being a significant determinant of price bubbles in either time period. Pre-GFC, Panel A shows evidence that herding is more persistent but this seems to disappear after the GFC, as seen in Panel B. However, the overall pattern of links between bubbles causing herding has broken down. The fact that excess returns are no longer a significant determinant of herding behaviour is understandable post-GFC as herding is potentially no longer prevalent. This possibility is supported by the lack of persistence in the autoregressive structure of herding post-GFC, in accordance with the assumption that herding is not prevalent.

However, the lack of a relationship pre-GFC has a less ready explanation, suggesting there is some additional factor that may explain what was driving herding. It may have been purely an issue of price appreciation as excess returns were actually lower, and this is a dynamic that may need more analysis. There is still an economically and statistically significant autoregressive part to bubble formation in both time periods, however the economic definition of the recession is being used for the sub-sample measurement rather than the peak and trough of the housing market.

The impacts of volatility are more confused as, pre-GFC, higher volatility is suggestive of more dispersed returns which is counter to the idea of rational herding and requires more sub-analysis.

The robustness of these results is shown in Table 15, where squared returns show broadly similar results to excess return proxies.

	Panel A - Pre-GFC		Panel B - Post-GFC	
	Bubble	**Herding**	**Bubble**	**Herding**
Bubble (-1)	**0.919***	-0.001	**0.789***	**0.007***
	(0.017)	(0.003)	(0.012)	(0.003)
Bubble (-2)	**0.328***	0.002	**0.148***	-0.002
	(0.022)	(0.004)	(0.015)	(0.004)
Bubble (-3)	**-0.538***	0.000	**-0.280***	-0.004
	(0.022)	(0.004)	(0.015)	(0.004)
CSAD (-1)	0.126	**0.840***	-0.002	**0.726***
	(0.105)	(0.018)	(0.051)	(0.012)
CSAD (-2)	**-0.285****	**0.057****	0.016	0.010
	(0.138)	(0.023)	(0.063)	(0.015)
CSAD (-3)	**0.236****	**-0.085***	-0.048	0.003
	(0.139)	(0.023)	(0.063)	(0.015)
Real GDP	**0.009***	**-0.001***	**-0.006***	-0.001
	(0.003)	(0.000)	(0.001)	(0.000)
Bank credit growth	0.000	-0.001	**-0.015****	**-0.008***
	(0.014)	(0.002)	(0.007)	(0.002)
Long-term bond yields	-0.061	0.003	0.017	**-0.030***
	(0.072)	(0.012)	(0.042)	(0.010)
Real effective ex. rates	0.007	0.001	**-0.007****	0.000
	(0.007)	(0.001)	(0.004)	(0.001)
Mortgage rates	0.008	-0.007	-0.017	**0.026***
	(0.087)	(0.015)	(0.057)	(0.014)
Supply elasticities	**-0.027***	0.001	0.001	-0.001
	(0.009)	(0.001)	(0.005)	(0.001)
Volatility	**0.130****	**0.017***	**0.302***	0.000
	(0.054)	(0.009)	(0.028)	(0.007)
Constant	**0.060****	**0.035***	**0.072***	**0.047***
	(0.029)	(0.005)	(0.017)	(0.004)
Adjusted R^2	0.895	0.796	0.753	0.744
Observations	3283		6566	

Table 15: Results from Structural Break Model using Squared Returns
Both panels are estimated using equation 10. Figures in parentheses are standard errors. * denotes 10%, ** denotes 5%, *** denotes 1% significance. The samples are defined as Pre-GFC before December 2007 and post-GFC after June 2009

6.4.6. Size Effects

There is ample evidence of different price dynamics in cities of different sizes which may conceivably also link with different herding and bubble effects. Larger cities may act as gateways for international capital, diversifying the investor base, increasing liquidity, reducing information asymmetries and risk, and therefore having more rational dynamics. This would exhibit both in less evidence of herding and excess returns and also in a weaker relationship between herding and price bubbles. More liquid and investment-focussed markets will have greater market efficiency due to more available information, and therefore the rational argument for herding will also be much reduced.

	Mean	Median	Min.	Max.	Skew.	Kurt.	SD
				Large (3664 observations)			
Returns	0.279	0.356	-4.051	3.719	-0.816	5.885	0.818
Squar. Ret.	0.747	0.290	0.000	16.411	4.975	37.990	1.378
Excess	0.007	0.000	-2.314	2.593	0.249	6.355	0.450
CSAD	0.406	0.370	0.126	1.311	1.619	6.880	0.163
				Medium (4122 observations)			
Returns	0.247	0.292	-3.933	3.859	-0.681	7.147	0.716
Squar. Ret.	0.573	0.182	0.000	15.471	5.414	44.216	1.222
Excess	0.014	0.008	-2.086	2.219	0.033	6.597	0.396
CSAD	0.412	0.381	0.109	1.775	1.666	8.904	0.165
				Small (3664 observations)			
Returns	0.239	0.269	-3.217	2.190	-0.301	5.042	0.537
Squar. Ret.	0.345	0.140	0.000	10.347	4.741	43.651	0.598
Excess	0.012	0.012	-2.714	2.725	-0.246	10.997	0.347
CSAD	0.386	0.364	0.002	1.339	1.024	5.698	0.166

Table 16: Descriptive Statistics for Herding and Bubble Measures by Size
Descriptive statistics for the price returns (from equation 7) and their squared values, excess returns as described in "Method" and the cross sectional-sectional absolute deviation (CSAD) (from equation 8). The samples are defined as largest 16 MSAs for Large, middle 18 MSAs for Medium and smallest 16 MSAs for Small

To estimate if there is a size effect, the dataset is split into three panels of the largest sixteen, middle eighteen and smallest sixteen MSAs. It is assumed that large cities, acting as gateways and having more liquid dynamics, are less risky and more efficient and so should exhibit less herding and fewer bubbles. The descriptive statistics in Table 16 show relatively limited between-sample variation, but there is slight evidence that larger cities have higher returns but lower excess returns, suggesting some element of less bubble formation.

Considering the estimated coefficients in Table 17, the role of excess returns as a significant determinant of lower CSAD values, itself potentially indicative of herding, holds for the large cities but is not present for smaller MSAs. Again, volatility is clearly important for larger cities, but this is not replicated for medium and small MSAs. As with the previous estimates, there is very little evidence that supply elasticity is a significant factor in market dynamics. The autoregressive structures for the excess returns are consistent across size, however for medium and small cities, the autoregressive structure for CSAD is not significant after the first lag suggesting that there is limited irrational persistence in the behaviour.

	Panel A - Large		Panel B - Medium		Panel C - Small	
	Bubble	**Herding**	**Bubble**	**Herding**	**Bubble**	**Herding**
Bubble (-1)	**0.951*****	**-0.010***	**0.903*****	-0.006	**0.865*****	**-0.017*****
	(0.015)	(0.005)	(0.015)	(0.006)	(0.015)	(0.006)
Bubble (-2)	**0.086*****	**-0.020*****	**0.129*****	-0.011	**0.124*****	0.005
	(0.020)	(0.007)	(0.020)	(0.007)	(0.019)	(0.008)
Bubble (-3)	**-0.513*****	**0.029*****	**-0.494*****	0.012	**-0.534*****	0.010
	(0.020)	(0.007)	(0.020)	(0.007)	(0.019)	(0.008)
CSAD (-1)	-0.016	**0.662*****	0.035	**0.777*****	-0.009	**0.828*****
	(0.047)	(0.017)	(0.044)	(0.016)	(0.041)	(0.017)
CSAD (-2)	-0.043	**0.095*****	0.024	0.015	0.005	0.006
	(0.057)	(0.020)	(0.056)	(0.020)	(0.053)	(0.022)
CSAD (-3)	**0.134****	**0.042****	0.032	-0.021	-0.042	**-0.088*****
	(0.057)	(0.020)	(0.056)	(0.020)	(0.053)	(0.022)
Real GDP	**-0.008*****	**-0.001****	**-0.007*****	**-0.001*****	**-0.003*****	**-0.001****
	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
Bank credit	0.007	**-0.004****	-0.002	**-0.005****	-0.006	-0.003
	(0.005)	(0.002)	(0.005)	(0.002)	(0.005)	(0.002)
LT bond yields	**0.120*****	**-0.026****	**0.054***	**-0.022***	0.0026	0.008
	(0.030)	(0.011)	(0.028)	(0.010)	(0.030)	(0.012)
Real eff. ex. rates	**0.012*****	**0.002***	**0.008*****	**0.002***	0.003	**0.002***
	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)
Mortgage rates	**-0.127*****	0.017	**-0.072****	0.013	-0.062	-0.008
	(0.038)	(0.014)	(0.036)	(0.013)	(0.038)	(0.015)
Supply elasticities	0.001	0.001	0.001	0.001	0.002	**-0.003***
	(0.005)	(0.002)	(0.003)	(0.001)	(0.005)	(0.002)
Volatility	**-0.047****	**0.044*****	-0.015	-0.001	0.034	-0.004
	(0.022)	(0.008)	(0.022)	(0.008)	(0.024)	(0.010)
Constant	-0.013	**0.031*****	0.018	**0.046*****	0.012	**0.052*****
	(0.014)	(0.005)	(0.014)	(0.005)	(0.015)	(0.052)
Adjusted R^2	0.814	0.823	0.763	0.764	0.687	0.772
Observations	3600		3825		3600	

Table 17: Results from Size Ranked Model using Excess Returns

All panels are estimated using equation 10. Figures in parentheses are standard errors. * denotes 10%, ** denotes 5%, *** denotes 1% significance. The panels are defined as largest 16 MSAs for Large, middle 18 MSAs for Medium and smallest 16 MSAs for Small

Overall, it seems that the results do not support the hypothesis that that larger cities will act more rationally. It is possible that there is more impact from the financial channels in larger cities as they are more linked to international capital flows and also more investment heavy. As the markets are less occupier driven and more investment driven (in line with the “gateway” idea) then they may be more susceptible to herding. Indeed, variables such as long term bond yields, effective exchange rates and (to a lesser extent) mortgage rates were more significant for larger MSAs. This may suggest an issue of financialisation, and is certainly worth further investigation.

	Panel A - Large		Panel B - Medium		Panel C - Small	
	Bubble	**Herding**	**Bubble**	**Herding**	**Bubble**	**Herding**
Bubble (-1)	**0.970*** **	**0.007*** **	**0.948*** **	**0.009*** **	**0.536*** **	0.001
	(0.016)	(0.002)	(0.016)	(0.002)	(0.017)	(0.004)
Bubble (-2)	0.025	**0.008*** **	**0.239*** **	**0.006** **	**0.435*** **	-0.004
	(0.022)	(0.003)	(0.020)	(0.003)	(0.018)	(0.005)
Bubble (-3)	**-0.388*** **	**-0.011*** **	**-0.550*** **	**-0.009*** **	**-0.287*** **	0.002
	(0.022)	(0.003)	(0.020)	(0.003)	(0.018)	(0.005)
CSAD (-1)	0.107	**0.652*** **	0.086	**0.766*** **	**0.204*** **	**0.831*** **
	(0.118)	(0.017)	(0.101)	(0.016)	(0.067)	(0.017)
CSAD (-2)	0.085	**0.093*** **	-0.119	0.017	**-0.240*** **	0.006
	(0.140)	(0.020)	(0.128)	(0.020)	(0.087)	(0.022)
CSAD (-3)	-0.126	**0.043** **	-0.024	-0.019	0.101	**-0.090*** **
	(0.140)	(0.020)	(0.128)	(0.020)	(0.087)	(0.022)
Real GDP	0.001	**-0.001** **	0.000	**-0.001*** **	0.000	**-0.001** **
	(0.003)	(0.000)	(0.003)	(0.000)	(0.002)	(0.000)
Bank credit	**-0.032** **	**-0.004** **	0.010	**-0.005*** **	0.003	-0.003
	(0.013)	(0.002)	(0.011)	(0.002)	(0.008)	(0.002)
LT bond yields	**0.392*** **	**-0.033*** **	-0.035	**-0.023** **	0.023	0.007
	(0.074)	(0.011)	(0.064)	(0.010)	(0.048)	(0.012)
Real eff. ex. rates	-0.008	0.001	-0.002	**0.002** **	0.000	0.002
	(0.007)	(0.001)	(0.006)	(0.001)	(0.005)	(0.001)
Mortgage rates	**0.235** **	**0.024* **	0.056	0.016	-0.016	-0.007
	(0.095)	(0.013)	(0.082)	(0.013)	(0.062)	(0.015)
Supply elasticities	-0..04	0.001	-0.009	0.001	**-0.013* **	**-0.003* **
	(0.013)	(0.002)	(0.008)	(0.001)	(0.008)	(0.002)
Volatility	**0.377*** **	**0.028*** **	**0.257*** **	-0.010	**0.229*** **	-0.003
	(0.056)	(0.008)	(0.056)	(0.009)	(0.042)	(0.010)
Constant	**0.121*** **	**0.033*** **	0.044	**0.046*** **	**0.060** **	**0.052*** **
	(0.034)	(0.005)	(0.034)	(0.005)	(0.024)	(0.006)
Adjusted R^2	0.879	0.824	0.872	0.767	0.719	0.772
Observations	3600		3825		3600	

Table 18: Results from Size Ranked Model using Squared Returns
All panels are estimated using equation 10. Figures in parentheses are standard errors. * denotes 10%, ** denotes 5%, *** denotes 1% significance. The panels are defined as largest 16 MSAs for Large, middle 18 MSAs for Medium and smallest 16 MSAs for Small

When using squared returns for robustness, Table 18 shows more mixed results. Panel A shows some variation in the sign of the estimated autoregressive coefficients, which may be a signal for further work on the calibration of bubble measures. The estimated coefficients in Panel B show more statistical significance for higher levels of CSAD, indicative of reverse herding, to be a trigger for bubble formation. This is in line with expectations about overconfidence and excessive appreciation.

6.4.7. Rational Herding

Rational herding is motivated by the presence of information asymmetries, as investors believe that others may possess more, or better, information and so it makes sense to follow some market average. This justification will apply to extreme conditions of high volatility (Christie and Huang, 1995), as the available information is noisy and may be unclear. Therefore, it is expected that the links between herding and bubble formation will be dependent also on the general level of volatility. In high volatility markets, where there may be some rational herding due to noisy information, more herding and potentially more bubble formation may be observed.

	Mean	Median	Min.	Max.	Skew.	Kurt.	SD
				High (3664 observations)			
Returns	0.302	0.452	-4.051	3.859	-0.758	4.456	1.020
Squar. Ret.	1.131	0.496	0.000	16.411	3.632	21.182	1.774
Excess	0.011	0.012	-2.714	2.725	0.030	4.435	0.597
CSAD	0.407	0.347	0.105	1.775	1.637	6.638	0.204
				Medium (4122 observations)			
Returns	0.236	0.288	-1.790	2.079	-0.371	3.914	0.529
Squar. Ret.	0.335	0.157	0.000	4.323	2.971	14.716	0.488
Excess	0.011	0.003	-1.081	1.315	0.242	3.876	0.294
CSAD	0.422	0.385	0.159	1.339	1.264	5.317	0.154
				Low (3664 observations)			
Returns	0.229	0.256	-1.167	1.882	-0.018	3.316	0.421
Squar. Ret.	0.230	0.117	0.000	3.542	3.343	17.607	0.330
Excess	0.012	0.007	-0.894	0.841	0.083	3.716	0.215
CSAD	0.374	0.375	0.002	1.035	-0.019	4.423	0.124

Table 19: Descriptive Statistics for Herding and Bubble Measures by Volatility
Descriptive statistics for the price returns (from equation 7) and their squared values, excess returns as described in "Method" and the cross sectional-sectional absolute deviation (CSAD) (from equation 8). The samples are defined as most volatile MSAs for High, middle 18 MSAs for Medium and least volatile 16 MSAs for Low

The descriptive statistics in Table 19 show little variation between sub-samples in terms of average values for excess returns and CSAD, and the estimated determinants for previous sub-samples show little evidence of volatility as a motivator for rational herding. Generally, the coefficients in Table 20 are more significant at higher rates of volatility as expected, but the between-sample variation is not overwhelming.

	High		Medium		Low	
	Bubble	**Herding**	**Bubble**	**Herding**	**Bubble**	**Herding**
Bubble (-1)	**0.879*** **	**-0.018*** **	**0.955*** **	0.007	**0.998*** **	-0.007
	(0.016)	(0.004)	(0.015)	(0.007)	(0.016)	(0.009)
Bubble (-2)	**0.130*** **	-0.003	**0.091*** **	**-0.035*** **	0.021	0.000
	(0.020)	(0.006)	(0.019)	(0.010)	(0.021)	(0.012)
Bubble (-3)	**-0.496*** **	**0.016*** **	**-0.567*** **	**0.035*** **	**-0.504*** **	-0.004
	(0.020)	(0.006)	(0.019)	(0.010)	(0.021)	(0.012)
CSAD (-1)	-0.061	**0.743*** **	0.019	**0.803*** **	**0.086*** **	**0.747*** **
	(0.061)	(0.017)	(0.031)	(0.016)	(0.028)	(0.016)
CSAD (-2)	0.075	**0.063*** **	-0.064	-0.001	-0.050	0.044
	(0.076)	(0.021)	(0.040)	(0.020)	(0.035)	(0.021)
CSAD (-3)	0.045	0.012	0.059	-0.025	-0.016	-0.092
	(0.076)	(0.021)	(0.040)	(0.020)	(0.035)	(0.021)
Real GDP	**-0.009*** **	**-0.001*** **	**-0.004*** **	**-0.001** **	**-0.004*** **	-0.001
	(0.002)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)
Bank credit	0.005	**-0.007*** **	-0.001	**-0.004** **	0.002	-0.001
	(0.008)	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)
LT bond yields	**0.154*** **	**-0.038*** **	**-0.036***	**-0.025** **	0.022	**0.021** **
	(0.046)	(0.013)	(0.021)	(0.011)	(0.016)	(0.010)
Real eff. ex. rates	**0.018*** **	0.001	0.001	**0.003** **	**0.006*** **	**0.002** **
	(0.004)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Mortgage rates	**-0.207*** **	0.021	-0.038	**0.029** **	**-0.039***	**-0.029** **
	(0.059)	(0.016)	(0.027)	(0.014)	(0.021)	(0.012)
Supply elasticities	**-0.029***	0.002	0.004	**-0.004*** **	0.002	0.001
	(0.017)	(0.005)	(0.003)	(0.001)	(0.002)	(0.001)
Volatility	-0.033	**0.017** **	0.024	**0.054*** **	0.046	0.022
	(0.026)	(0.007)	(0.028)	(0.014)	(0.029)	(0.017)
Constant	**0.046***	**0.034*** **	-0.004	**0.054*** **	0.002	**0.033*** **
	(0.024)	(0.007)	(0.011)	(0.005)	(0.008)	(0.005)
Adjusted R^2	0.768	0.808	0.758	0.772	0.764	0.755
Observations	3375		4050		3600	

Table 20: Results from Volatility Model using Excess Returns
All panels are estimated using equation 10. Figures in parentheses are standard errors. * denotes 10%, ** denotes 5%, *** denotes 1% significance The samples are defined as most volatile MSAs for High, middle 18 MSAs for Medium and least volatile 16 MSAs for Low

However, volatility itself is still not showing as significant, which perhaps demonstrates limited rational herding in real estate. In this situation, investors are not worried about information asymmetries but their behaviour is instead driven by speculation. If herding is driven by excess returns, this supports speculation as the motivation and suggests behaviour, rather than the efficiency of the market structure, is driving the dynamics. A potential additional factor would be the volatility resulting simultaneously from an additional fundamental or behavioural factor.

	High		Medium		Low	
	Bubble	**Herding**	**Bubble**	**Herding**	**Bubble**	**Herding**
Bubble (-1)	**0.868*** **	**0.008*** **	**1.001*** **	-0.004	**1.085*** **	**0.031*** **
	(0.017)	(0.002)	(0.015)	(0.006)	(0.016)	(0.010)
Bubble (-2)	**0.215*** **	**0.005** **	**0.094*** **	**0.021** **	0.004	-0.019
	(0.021)	(0.002)	(0.021)	(0.009)	(0.023)	(0.014)
Bubble (-3)	**-0.457*** **	**-0.005** **	**-0.497*** **	**-0.048*** **	**-0.466*** **	**-0.026* **
	(0.021)	(0.002)	(0.021)	(0.009)	(0.023)	(0.014)
CSAD (-1)	**0.325** **	**0.726*** **	**-0.098** **	**0.805*** **	-0.007	**0.739*** **
	(0.157)	(0.017)	(0.038)	(0.016)	(0.027)	(0.017)
CSAD (-2)	-0.131	**0.059*** **	0.019	-0.008	**-0.106*** **	**0.048** **
	(0.194)	(0.021)	(0.050)	(0.021)	(0.034)	(0.021)
CSAD (-3)	0.083	0.013	-0.011	-0.004	**0.100*** **	**-0.082*** **
	(0.194)	(0.021)	(0.050)	(0.000)	(0.034)	(0.021)
Real GDP	0.002	**-0.001** **	-0.001	**-0.001** **	**-0.002** **	-0.001
	(0.004)	(0.000)	(0.001)	(0.002)	(0.001)	(0.000)
Bank credit	-0.025	**-0.007*** **	0.001	**-0.004** **	0.002	-0.001
	(0.020)	(0.002)	(0.004)	(0.011)	(0.003)	(0.002)
LT bond yields	**-0.485*** **	**-0.041*** **	0.018	**-0.025** **	-0.012	**0.018* **
	(0.117)	(0.013)	(0.025)	(0.001)	(0.016)	(0.009)
Real eff. ex. rates	-0.018	0.000	0.002	**0.002** **	0.000	**0.002** **
	(0.011)	(0.001)	(0.002)	(0.014)	(0.001)	(0.001)
Mortgage rates	**0.449*** **	**0.027** **	**-0.081** **	**0.028** **	-0.008	**-0.024** **
	(0.150)	(0.016)	(0.032)	(0.001)	(0.020)	(0.012)
Supply elasticities	0.049	0.0002	-0.004	**-0.004*** **	0.001	0.001
	(0.043)	(0.005)	(0.003)	(0.001)	(0.002)	(0.001)
Volatility	**0.270*** **	0.008	**-0.206*** **	**0.053*** **	**0.164*** **	0.014
	(0.067)	(0.007)	(0.033)	(0.014)	(0.028)	(0.017)
Constant	0.030	**0.036*** **	**0.051*** **	**0.055*** **	0.013	**0.034*** **
	(0.062)	(0.007)	(0.013)	(0.005)	(0.008)	(0.005)
Adjusted R^2	0.833	0.811	0.874	0.773	0.908	0.755
Observations	3375		4050		3600	

Table 21: Results from Volatility Model using Squared Returns
All panels are estimated using equation 10. Figures in parentheses are standard errors. * denotes 10%, ** denotes 5%, *** denotes 1% significance The samples are defined as most volatile MSAs for High, middle 18 MSAs for Medium and least volatile 16 MSAs for Low

When using squared returns for returns, Table 21 again shows that, much like the size analysis, there is some inconsistency in the sign of the estimated coefficient, although the pattern of statistical significance is broadly similar.

6.5. Conclusion

The significant conclusion to draw is that there was very limited evidence to support the a priori assumption that herding would lead to some element of bubble formation. However, there was evidence that excess returns may be a motivator for herding, which may result from some speculative function of real estate.

Excess returns, as a proxy for a price bubble, may lead to herding as they signal strong and potentially accelerating appreciation. In a core asset class such as real estate, easily observable price appreciation attracts additional investment by indicating to less informed or more susceptible investors that they should follow the market. Clearly, this will create herding. If returns are beyond the expected level derived from the market fundamentals, then this excess may be speculation.

Generally, the base model findings were replicated across various different sub-samples which suggests that, although herding and bubble links are not persistent they is some evidence that they exist across space and time (as with Galariotis).

The results from the sub-analyses were mixed. There was evidence of an investment-driven mechanism in gateway cities, which motivates further research on spatial variation and specifically, as some the local market characteristics may be driving differences in dynamic outcomes, then a deeper dive into spatial variation may yield beneficial results. This motivates a further paper on the determinants of herding (and reverse herding) behaviour.

The GFC was expected to signify some change in behaviour which indeed it seems to have, but the specific motivating factors for this change are not clear and again require further analysis. However the volatility samples had no clear results, suggesting that any herding behaviour is not being driven by rational motivations. This is an important finding and gives impetus to the speculative motive.

One significant empirical conclusion is that the estimated connection between herding behaviour and bubble formation in some sub-analyses was sensitive to the measure of bubble. As discussed, bubbles are challenging to define and therefore this supports the need for further work on the measurement of bubbles and the calibration of these models.

There was extremely limited evidence that supply elasticities are economically or statistically significant determinants of herding or excess returns. The original hypothesis was that more elastic supply responses would prevent extensive bubble formation as prices would return to equilibrium more rapidly. With relatively frequent data for real estate, the more structural issues may have unclear impacts on short-term dynamics. Restrictions on the supply side may be measured more robustly by current vacancy rather than long-term stock, and as herding and excess returns are very quick responses then the supply mechanisms as measured by land elasticities may not capture this dynamic. From a speculative behavioural perspective, investors may ignore the long-term development process and consider only short-term gains.

7. Measuring the Determinants of Herding and Reverse Herding

7.1. Abstract

This paper examines the determinants of rational and irrational market dynamics in US housing markets. A discrete choice model is employed to categorise responses into rational, herding and reverse herding outcomes. Using both linear and non-linear estimations, a range of housing characteristics, especially several costs metrics, are found to be the significant determinants of market responses. These suggest that much of the behavioural dynamics observed in housing markets derives from the unique asset characteristics rather than general economic conditions. Furthermore, evidence is presented that private information, conditioned on property market inefficiencies, requires further incorporation into behavioural analysis.

Key Words: Herding; Reverse Herding; Housing; Discrete Choice Model

7.2. Introduction

This study examines the economic and real estate characteristics that determine local variation in observed herding and reverse herding behaviour in US housing markets. As housing markets possess unique characteristics such as local variation and information asymmetries, they exhibit irrational behaviour that differs markedly from equity markets, hence motivating the equal importance attached to the less researched concept of reverse herding. The study separates which characteristics broadly determine irrational behaviour, as well as the specific herding and reverse herding behaviour.

7.2.1. Herding and Reverse Herding

Herding is correlated behaviour across individuals, such that they appear to, consciously or unconsciously, mimic the actions of others (Devenow and Welch 1996). This action may be rational where an individual lacks confidence or information, or believes that others possess better information, and so copy them (Bikhchandani et al. 1992; Bikhchandani et al. 1998; Welch 1992). Alternatively, it could be an irrational behaviour, motivated by a psychological bias (Barber et al. 2009). In an investment context, this would be apparent from price movements being strongly correlated across different assets resulting in lower than expected dispersions of returns (Chang et al. 2000).

Conversely, if greater than expected price dispersion exists, then reverse herding appears. This may result from excessive trading leading to overly dispersed returns as investors operate from a position of overconfidence. There is a clear relationship between these behavioural phenomena and bubble formation (Devenow and Welch 1996), the latter of which may result in an asset-price bust and significant loss of capital value (Lux 1995).

7.2.2. Herding in a Real Estate Context

In comparison to securitised markets, herding as a real estate phenomena has not been extensively researched. However, Hott (2012) found movements in housing beyond that justified by the fundamentals, Ngene et al. (2017) established extensive evidence for herding in various market conditions and geographies across regional US housing markets, and Lan (2014) discovered herding in the Chinese national housing market.

Market structure is a strong predicator of market dynamics, and therefore patterns of herding and reverse herding in real estate differ from those in equity markets, most importantly the relatively greater prevalence of reverse herding. Klein (2013) demonstrates that this latter behaviour is often found in more mature markets, such as the USA.

It is well established that housing has low levels of price transparency (relative to securities) and lacks frequent pricing which, coupled with heterogeneous assets, results in strong private information. In addition, as short selling is impossible (Galariotis et al. 2016), then this mispricing may continue unlike in more efficient security markets. This results overall in low levels of information efficiency, and indeed Avery and Chevalier (1999) suggest that reverse herding may be motivated by strong private information. Furthermore, Hwang et al. (2020) argue that, as in an inefficient market profit can be derived from private information, then the relative importance of information is determined by whether it is private or public. In other words, if an investor possesses strong private information then it would be rational to deviate from public information as represented in the market average. The resulting deviations from the average would then result in greater dispersions of behaviour and reverse herding.

As individuals have a tendency to be extremely confident in their own abilities, they may be less likely to herd as they do not feel they need to default to the opinions of others (Ekholm and Pasternack, 2008) and when they experience overconfidence, individuals overreact to private information and underreact to public information (Daniel et al. 1997). Periods of strong price appreciation and market inefficiency tend to motivate conspicuous overconfidence (Bao and Li 2020) simulations demonstrate that this leads to excessive trading. Chuang et al. (2014) and Griffin et al. (2007) both found overconfidence in inefficient markets, again leading to excessive trading, which can clearly lead to reverse herding. In the UK, a market of similar maturity and structure to the USA, homeowners are generally overconfident, other than in down markets or recessions (Hwang et al. 2020).

In addition to the comparative market structure, supply of housing mediates demand shocks, determining pricing (Saiz 2008) and potentially impacting the pattern of price responses. In addition, unlike most durable goods, house supply is locally determined as construction costs are largely consistent across space and Gyourko and Saiz (2006) find that divergences in housing supply elasticities, and therefore house prices, must result from land markets. Elastic markets could lead to more development as land can be made available easily, and so investors may have opportunities to invest which creates excessive trading and reverse herding.

7.2.3. Local Variations in Real Estate

Strong private information and the resulting overconfidence may stem from the localised nature of housing and many of the stylised facts of real estate, such as information asymmetries, fixed physical nature and the lack of a central clearing market, can be clearly linked to local conditions. Therefore, as the exact local characteristics differ then it may be assumed that the behavioural outcomes will also differ, and indeed Ngene et al. (2017) established some evidence of variation in herding between regions.

Lerbs and Oberst (2014) demonstrate that differences in local characteristics are the prime determinants of variable rates of regional homeownership. Policy can also be defined by local variation, as Hortas-Rico and Gómez-Antonio (2020) find that competition for economic activity between municipalities is a major determinant of land use conversion and therefore land supply for residential development, which has a clear role in pricing. For example, even in locations with strong land supply such as Finland, local regulations have a significant impact on elasticity (Oikarinen et al. 2015). Most interestingly, the results are in line with those previously found in the USA despite significant differences in geography, culture and political structure.

Having already seen a long line of research that demonstrated how economically diversified urban areas are both more stable and tend towards higher economic growth rates (Attaran 1986; Kort 1981; Parr 1965; Siegel et al. 1994; Siegel et al. 1995; Wagner and Deller 1998), Coulson et al. (2020) demonstrate that economic diversification mitigates the scale and duration of negative exogenous shocks on house prices, and so economic diversification may have some mitigation on undesirable market dynamics such as herding and bubble formation.

7.2.4. Contribution

Herding in real estate is generally under-researched, as is the existence of reverse herding more broadly, and this study contributes to both aspects. It also adds generally to the field of behavioural real estate by establishing determinants of rational and irrational outcomes in housing markets. There is no apparent previous research that has modelled econometrically the determinants of herding behaviour, and this study adds to understanding of local variation, as well as how much behaviour is determined by real estate characteristics relative to general economic conditions.

7.2.5. Summary of Results

This study finds that housing factors are strong determinants of behavioural outcomes. Higher ongoing housing costs and price-income ratios are strongly associated with reverse herding, and that markets with more elastic land supply experience significantly reduced levels of herding. In addition, markets with higher rates of college education are also associated with a much lower prevalence of herding, in line with expectations. As herding has been identified as a factor in bubble formation then quantitative estimates for determining this behaviour may aid in more robust risk management.

7.3. Data and Methodology

7.3.1. Market

Due to the importance of micro location in real estate economics, there is limited relevance in national level research and so this analysis is performed at the MSA-level. In addition, previous research shows significant local variation in behaviour, and a specific interest is the role of local characteristics on pricing behaviour. In addition, as urban-level data is available both for house prices and for local characteristics, this is exploited to understand more micro-level variation. Therefore, the empirical analysis is based on the metropolitan statistical area (MSA) as the spatial level rather than political administrative boundaries. MSAs are defined by the Census Bureau as relatively dense urban cores along with economically integrated surrounding urban areas, and are therefore commonly used in regional analysis.

MSA	Population	Per capita income ($)	House price ($)	Economic output ($)
New York	20,140,470	79,844	509,356	1,861,147,410
Los Angeles	13,200,998	66,684	735,212	1,088,676,191
Chicago	9,618,502	63,500	257,714	709,160,008
Dallas-Fort Worth	7,637,387	58,725	274,597	523,861,973
Houston	7,122,240	58,890	232,626	512,222,304
Washington DC	6,385,162	74,385	471,701	559,061,958
Philadelphia	6,245,051	66,596	274,637	454,692,188
Miami	6,138,333	60,966	321,994	377,531,846
Atlanta	6,089,815	54,557	264,610	422,189,461
Boston	4,941,632	81,498	535,789	484,620,546
Phoenix	4,845,832	48,065	326,891	272,113,776
San Francisco	4,749,008	104,921	1,178,986	591,945,456
Riverside	4,599,839	42,242	422,649	199,640,303
Detroit	4,392,041	78,073	198,541	269,850,041
Seattle	4,018,762	78,073	583,855	424,750,310
Minneapolis	3,690,261	64,255	318,088	274,191,982
San Diego	3,298,634	63,729	678,553	253,117,792
Tampa	3,175,275	48,908	253,548	169,151,030
Denver	2,963,821	67,236	484,473	227,395,640
Baltimore	2,844,510	63,988	318,875	215,487,169
St Louis	2,820,253	56,923	196,929	173,456,187
Orlando	2,673,376	45,156	276,989	147,218,306
Charlotte	2,660,329	53,916	265,022	178,413,790
San Antonio	2,558,143	48,684	222,886	129,394,189
Portland	2,512,859	59,921	455,931	174,937,438
Sacramento	2,397,382	58,843	467,195	153,331,952
Pittsburgh	2,370,930	60,227	176,818	162,171,323
Austin	2,283,371	61,977	384,071	159,361,291
Las Vegas	2,265,461	48,806	311,449	128,496,477
Cincinnati	2,256,884	56,033	207,988	153,931,926
Kansas City	2,192,035	55,009	227,020	138,499,915
Columbus	2,138,926	52,477	234,768	134,342,237
Indianapolis	2,111,040	56,360	204,682	144,806,700
Cleveland	2,088,251	55,451	175,771	135,756,216
San Jose	2,000,468	114,080	1,304,286	334,600,993
Nashville	1,989,519	60,680	310,104	138,558,743
Virginia Beach	1,799,674	52,011	263,026	103,080,568
Providence	1,676,579	56,138	351,350	91,015,993
Jacksonville	1,605,848	51,421	251,581	87,140,015
Milwaukee	1,574,731	58,457	224,380	107,136,039
Oklahoma City	1,425,695	48,860	171,033	79,475,474
Raleigh	1,413,982	57,851	306,151	94,806,039
Memphis	1,337,779	47,985	173,574	78,866,609
Richmond	1,314,434	58,628	267,507	91,359,840
Louisville	1,285,439	52,134	197,132	73,833,749
New Orleans	1,271,845	54,363	222,649	83,557,049

Salt Lake City	1,257,936	54,450	434,631	102,801,166
Hartford	1,213,531	65,132	260,022	105,147,321
Buffalo	1,166,902	52,331	192,891	73,754,236
Birmingham	1,115,289	53,374	186,817	63,573,271

Table 22: MSA Descriptive Statistics
Descriptive statistics for the 50 largest urban areas in the USA ranked by population, using the metropolitan statistical area defined by the Census Bureau. The population data is provided by the Census Bureau as of 2020, per capita income and economic output (in 000s) from the Bureau of Economic Analysis in 2019 current dollars and house prices from Zillow as of December 2020

To capture a sufficient part of the US market, the analysis covers the largest 50 MSAs by population as at 2020, for which some descriptive statistics are presented in Table 22. These MSAs provide a cumulative population of 182 million (55% of the total national population or 67% of urban population) and aggregate economic output of almost $14 trillion, which covers all cities over one million in population and over $60 billion in economic output. These values are highly concentrated, as the top ten urban centres alone account for more than half of the economic output and 89 million residents.

Whilst marked disparities are observed in per capita income (San Jose at $114,080 being almost three times that of Riverside at $42,242), house prices show even greater extremes of distribution (San Jose at $1,304,286 being more than seven times Oklahoma City at $171,033), and so marked variation between locations is apparent in income and price, with obvious repercussions for affordability.

7.3.3. House Price and Behavioural Data

MSA	Metric	Mean	Median	Min	Max	SD	Skewness	Kurtosis	Obs
NYC	Return	0.33	0.31	-0.87	1.37	0.55	-0.13	2.17	300
	CSAD	0.43	0.42	0.23	0.60	0.06	0.21	3.19	300
LAX	Return	0.47	0.62	-2.66	2.54	0.89	-0.78	4.24	300
	CSAD	0.30	0.26	0.13	1.08	0.14	2.71	11.65	300
CHC	Return	0.15	0.32	-2.58	1.85	0.61	-1.13	4.60	300
	CSAD	0.42	0.35	0.19	2.09	0.21	3.68	23.34	300
DFW	Return	0.29	0.28	-0.68	1.35	0.38	-0.19	3.04	300
	CSAD	0.35	0.34	0.21	0.67	0.07	0.74	3.67	300
HOU	Return	0.26	0.28	-0.69	1.02	0.32	-0.29	3.44	300
	CSAD	0.37	0.35	0.25	0.69	0.07	1.17	4.77	300
WDC	Return	0.31	0.26	-1.55	1.89	0.66	-0.13	3.53	300
	CSAD	0.40	0.37	0.18	0.79	0.12	1.00	3.87	300
MIA	Return	0.26	0.21	-0.85	1.31	0.48	0.15	2.63	300
	CSAD	0.36	0.36	0.16	0.53	0.06	-0.06	3.53	300
PHD	Return	0.35	0.52	-2.88	2.36	1.00	-1.09	4.35	300
	CSAD	0.32	0.30	0.10	0.82	0.14	1.32	4.72	300
ATL	Return	0.28	0.43	-1.79	1.39	0.62	-1.50	5.00	300
	CSAD	0.38	0.32	0.18	1.03	0.18	1.39	4.48	300
PHN	Return	0.40	0.48	-0.71	1.43	0.53	-0.32	2.33	300
	CSAD	0.38	0.35	0.17	0.71	0.11	0.66	2.88	300
BOS	Return	0.37	0.49	-3.80	3.72	1.16	-0.69	5.35	300
	CSAD	0.39	0.33	0.17	1.17	0.18	1.79	6.39	300
SFR	Return	0.47	0.66	-2.37	2.21	0.85	-0.63	3.13	300
	CSAD	0.55	0.50	0.23	1.22	0.17	1.04	3.92	300
RIV	Return	0.38	0.50	-4.05	2.73	1.16	-1.46	6.24	300
	CSAD	0.46	0.42	0.20	1.25	0.19	1.96	7.31	300
DTR	Return	0.23	0.41	-2.54	1.77	0.76	-1.37	4.90	300
	CSAD	0.54	0.50	0.25	1.31	0.20	1.43	5.39	300
STL	Return	0.42	0.57	-1.87	1.65	0.72	-1.00	3.43	300
	CSAD	0.29	0.27	0.12	0.63	0.10	1.06	4.21	300
MNN	Return	0.33	0.47	-1.16	1.25	0.55	-1.04	3.40	300
	CSAD	0.36	0.32	0.19	0.78	0.12	1.08	3.50	300
SDG	Return	0.43	0.58	-2.59	2.94	0.91	-0.58	3.41	300
	CSAD	0.33	0.29	0.16	0.92	0.14	2.37	8.93	300
TMP	Return	0.35	0.57	-2.11	2.31	0.90	-0.93	3.60	300
	CSAD	0.32	0.30	0.14	0.65	0.08	1.03	4.40	300
DNV	Return	0.41	0.43	-0.66	1.26	0.45	-0.23	2.34	300
	CSAD	0.28	0.26	0.13	0.67	0.09	1.16	4.46	300
SLS	Return	0.26	0.22	-1.23	1.88	0.59	0.22	3.11	300
	CSAD	0.36	0.35	0.18	0.78	0.09	0.83	4.12	300
BAL	Return	0.22	0.26	-0.72	0.90	0.35	-0.58	2.75	300
	CSAD	0.51	0.51	0.31	0.86	0.09	0.97	5.23	300
ORL	Return	0.31	0.50	-3.16	2.82	0.99	-0.87	4.61	300

	CSAD	0.34	0.31	0.12	0.83	0.12	1.55	5.72	300
CHR	Return	0.28	0.39	-0.88	1.49	0.39	-0.92	4.31	300
	CSAD	0.42	0.40	0.25	1.00	0.11	1.90	8.53	300
SAN	Return	0.26	0.29	-0.71	1.36	0.35	-0.13	3.03	300
	CSAD	0.31	0.30	0.17	0.64	0.07	0.97	4.54	300
POR	Return	0.38	0.43	-1.27	1.96	0.59	-0.70	3.71	300
	CSAD	0.33	0.31	0.16	0.70	0.09	0.83	3.76	300
SAC	Return	0.37	0.53	-2.68	2.64	0.99	-0.62	3.33	300
	CSAD	0.67	0.60	0.32	1.77	0.23	1.77	7.27	300
PIT	Return	0.23	0.23	-0.68	1.05	0.27	-0.23	3.93	300
	CSAD	0.63	0.63	0.28	0.90	0.11	-0.15	3.22	300
AUS	Return	0.30	0.29	-0.65	2.04	0.39	0.78	5.72	300
	CSAD	0.42	0.36	0.19	1.09	0.19	1.66	5.41	300
LVS	Return	0.26	0.44	-3.93	3.86	1.27	-0.90	5.02	300
	CSAD	0.35	0.30	0.11	1.27	0.18	2.13	9.02	300
CIN	Return	0.21	0.28	-1.17	1.32	0.37	-0.63	4.58	300
	CSAD	0.41	0.40	0.24	1.03	0.11	1.07	6.08	300
KAN	Return	0.26	0.34	-0.86	1.23	0.36	-0.78	3.68	300
	CSAD	0.43	0.41	0.24	0.73	0.09	0.74	3.39	300
COL	Return	0.22	0.29	-0.65	1.17	0.33	-0.38	2.92	300
	CSAD	0.47	0.45	0.29	0.76	0.09	0.54	2.80	300
IND	Return	0.18	0.16	-1.20	1.62	0.40	0.05	3.75	291
	CSAD	0.47	0.38	0.24	2.30	0.27	3.16	15.89	282
CLV	Return	0.13	0.23	-1.25	1.30	0.43	-0.75	3.91	300
	CSAD	0.44	0.40	2.00	0.95	0.15	1.29	4.45	300
SJS	Return	0.50	0.61	-3.22	2.66	1.05	-0.59	3.29	300
	CSAD	0.31	0.28	0.12	0.92	0.12	1.68	7.43	300
NAS	Return	0.34	0.39	-0.64	1.07	0.33	-0.64	3.05	300
	CSAD	0.38	0.35	0.20	0.85	0.11	1.06	4.19	300
VIR	Return	0.29	0.25	-0.57	1.88	0.49	0.92	4.13	300
	CSAD	0.38	0.35	0.20	0.96	0.12	1.79	7.65	300
PRO	Return	0.34	0.34	-1.51	2.08	0.67	0.18	2.83	300
	CSAD	0.38	0.36	0.19	0.73	0.11	0.88	3.49	300
JAC	Return	0.30	0.43	-1.44	1.29	0.56	-1.06	3.75	300
	CSAD	0.47	0.45	0.24	0.94	0.13	0.68	3.28	300
MIL	Return	0.25	0.37	-1.40	1.51	0.50	-1.12	4.60	300
	CSAD	0.37	0.33	0.16	1.00	0.16	1.29	4.72	300
OKC	Return	0.29	0.32	-0.39	0.88	0.24	-0.52	3.16	300
	CSAD	0.39	0.38	0.23	0.73	0.09	0.85	3.86	300
RAL	Return	0.22	0.26	-1.07	1.34	0.33	-0.68	4.31	300
	CSAD	0.14	0.11	0.00	0.76	0.13	1.79	7.15	300
MEM	Return	0.16	0.24	-0.88	1.05	0.39	-0.72	3.22	300
	CSAD	0.47	0.44	0.17	1.03	0.13	1.28	5.62	300
RIC	Return	0.30	0.28	-0.77	1.24	0.39	-0.16	3.02	300
	CSAD	0.33	0.30	0.15	0.72	0.11	0.88	3.32	300
LOU	Return	0.25	0.30	-0.71	0.96	0.30	-0.69	3.64	300
	CSAD	0.38	0.36	0.18	1.07	0.13	2.38	11.83	300

NOR	Return	0.26	0.25	-0.55	1.05	0.33	-0.03	2.67	300
	CSAD	0.57	0.53	0.28	1.34	0.18	1.13	4.86	300
SLC	Return	0.30	0.26	-1.25	1.69	0.60	-0.18	3.01	300
	CSAD	0.26	0.24	0.10	0.75	0.09	1.47	7.13	300
HAR	Return	0.22	0.24	-1.23	1.19	0.50	-0.24	2.51	300
	CSAD	0.43	0.39	0.17	1.16	0.17	1.01	3.71	300
BUF	Return	0.26	0.23	-0.78	1.50	0.34	0.47	4.12	300
	CSAD	0.50	0.41	0.22	1.35	0.23	1.67	5.23	300
BIR	Return	0.15	0.24	-1.06	1.15	0.39	-0.73	3.50	300
	CSAD	0.58	0.51	0.32	0.68	0.22	2.24	9.63	300

Table 23: Descriptive and Distributional Statistics for House Price Returns and CSAD
For each MSA, descriptive and distributional statistics for both price returns (from equation 1) and the cross-sectional absolute deviation (CSAD) (from equation 2) are calculated from Zillow data on a month-to-month basis and on a monthly frequency (authors' own calculations)

Table 23 shows descriptive and distributional statistics for price returns and dispersions, which again reveal large variation between MSAs in price and CSAD. Cleveland has grown by only 0.13% monthly, whereas San Jose has grown by 0.5%, resulting in an enormous cumulative difference in price appreciation.

There may be some link between size and returns, as the latter generally decreases as cities decline in size[2], although there is no apparent link between size and CSAD.

[2] This link may be dependent on the smaller number of ZIPs used as MSA size decreases, meaning that results for smaller MSAs may not be as robust. However, all the MSAs have at 100 ZIPs, so in context this would be like using the FTSE 100 in terms of number of individuals and so be unlikely to have an issue of insufficient observations.

Variable	Mean	Median	Min	Max	SD	Obs
All	0.22	0.12	-1.17	3.51	0.41	6600
Rational	0.07	0.03	-0.72	2.00	0.21	3972
Irrational	0.45	0.40	-1.17	3.51	0.51	2628
Negative	-0.15	-0.07	-1.17	-0.00	0.19	1700
Positive	0.35	0.22	0.00	3.51	0.38	4900
Herding	-0.41	-0.36	-1.17	-0.04	0.28	248
Reverse Herding	0.54	0.46	0.03	3.51	0.44	2380

Table 24: Descriptive Statistics for Estimated Responses (by Behavioural Groups)
Descriptive statistics for the estimated responses (from equation 13) grouped by category. Rational (irrational) is all non-significant (significant) responses, whereas Herding and Reverse Herding are further classified by the sign of the estimated coefficient

From Equation 7, a response term is estimated on a monthly basis for each MSA via a rolling regression with robust standard errors and a window of 5 years. These responses then form the dependent variable in the model, and descriptive statistics are presented in Table 24. Of the 6600 estimated responses, 3972 (60%) are statistically insignificant and therefore no evidence of irrational behaviour exists. This fits generally with evidence from Ngene et al. (2017) that around 50-60% of house market responses are rational. 248 (just under 4%) of responses are significantly negative, suggesting herding. Therefore, 2380 (36%) of responses are significantly positive and demonstrate evidence for reverse herding. Clearly the prevalence of herding is much less reverse herding, which does tie with some evidence that developed housing markets show persistent levels of overconfidence, and that the localised structure of housing markets leads to strong private information and excessive trading, which is shown in reverse herding. Table 25 shows low levels of correlation between the terms and so multicollinearity is not suspected as an issue.

Variable	Return	CSAD	Response
Return	1.00		
CSAD	-0.33	1.00	
Response	0.03	0.07	1.00

Table 25: Correlations within the Asset Pricing Model
Pairwise correlations of price returns (from equation 11), the cross-sectional absolute deviation (CSAD) (from equation 12) and the estimated response (from equation 13)

7.4.4. Economics, Housing and Social Data

In addition to being the basis for the estimated response that forms the dependent variable, MSA-level house price growth is included as a regressor (deflated by CPI for All Urban Consumers: All Items) to control for the impact of house prices on the responses. Similarly, economic output is obtained from the Bureau of Economic Analysis (BEA)[3] to capture the role of economic conditions as a driver of house price growth and housing market dynamics in addition to capturing the ability of owner-occupiers to invest and consume housing. Population data is collected from the BEA and transformed into growth rates to reflect core demand for housing.

In addition to these economic measures, property-specific characteristics are collected, mainly around supply and affordability. The main motivation here is that many regulations on housing and development are highly localised, especially in the USA, providing an intuitive explanation for the regional variation in responses previously observed. In addition, some geographic factors (i.e. the physical geographic restraints included in the supply elasticity metric) may also impact the behaviour of MSA-level housing markets.

[3] These variables are modelled as changes via logarithmic differences in line with equation (1). GDP by MSA since 2001 is provided in chained 2012 dollars.

Supply in the built environment has two factors to consider, the supply of land and the supply of physical structures upon it. The former has constraints dependent on physical geography, but also on local political and institutional attitudes towards development. Rather than measuring raw availability of land for housing construction, the intuitively significant factor is the response of land supply. This is captured by the land elasticity measures estimated by Saiz (2008). The land elasticity index is available on the MSA level and so removes any artificial constraints based on administrative boundaries. Secondly, construction activity can be measured by issued permits from the Census Bureau's Building Permits Survey.

In addition, the affordability of housing is incorporated via both housing costs and the price-to-income ratio. For the former, the MSA-median housing cost (a combination of mortgage, bill and tax costs) as a percentage of household income is collected, which reflects some element of affordability and may also marginally account for some of the role of consumption in housing behaviour. As the majority of household purchases are financed via a mortgage, the value for households with mortgages is used. The price-to-income ratio is measured simply by the ratio of the MSA-house price average and MSA per capita income (specifically from the BEA).

College-education levels for adults over the age of 25 are included as a proxy for educational attainment. Finally, to account for the role of investors, the owner-occupier rate serves as a proxy. According to data from the US Department of Housing and Urban Development, social housing in the USA accounts for approximately 1% of total stock and so public ownership is immaterial.

Table 26 presents descriptive statistics for the independent variables. A negative mean rate of real house price growth can be considered in the context of covering a substantial house market crash, and also that the MSAs are not weighted, so that stagnant housing markets in some second-tier cities may skew the results compared to a national measure. Construction is the most volatile series, with some extreme collapses in permit volumes seen during the housing crisis, and in many MSAs the pre-GFC peaks have never been recovered.

Variable	Mean	Median	Min	Max	SD	Skew	Kurt
Economic Growth	1.70	1.93	-27.05	10.89	3.27	-1.67	14.23
Population Growth	1.09	1.04	-28.73	5.26	1.60	-11.64	222.37
House Price Growth	-0.11	-0.08	-4.81	3.78	0.87	-0.40	5.08
Housing Costs	24.56	24.00	18.80	34.40	3.05	0.80	3.24
Land Elasticity	1.66	1.50	0.60	4.00	0.85	0.72	2.60
Permits	0.15	-0.74	-68.33	286.80	33.05	1.73	13.55
Price/Income	5.25	4.56	2.21	15.53	2.35	1.69	6.13
Owner-Occupier	65.14	65.80	48.68	79.00	5.78	-0.48	2.95
Education	31.88	30.86	18.15	49.35	6.02	0.67	3.30

Table 26: Descriptive Statistics for Determinants
Descriptive statistics for the independent variables used to estimate equation 4, as described in "Economics, Housing and Social Data". All growth values are annual, other than House Price Growth which is monthly

7.4. Empirical Results

If the estimated response is statistically insignificant, then the market is assumed to be operating rationally. If the response is significantly negative, then it suggests a reduced CSAD and herding, and likewise a significantly positive response suggests the CSAD increased non-linearly and reverse herding may exist. Having categorised the dependent variable as three distinct outcomes (herding, rational responses and reverse herding), a discrete choice model estimates the determinants of behavioural responses to large absolute increases in house prices.

	EcG	PopG	HPG	HouC	LSE	Per	P/I	OO	Edu
Economic Growth	1.00								
Population Growth	0.26	1.00							
House Price Growth	0.43	0.05	1.00						
Housing Costs	-0.15	-0.02	-0.33	1.00					
Land Elasticity	0.07	0.20	0.02	-0.62	1.00				
Permits	0.36	-0.02	0.36	-0.17	-0.02	1.00			
Price/Income	0.20	0.05	0.08	0.67	-0.51	-0.08	1.00		
Owner-Occupier	-0.14	-0.05	-0.10	-0.35	0.28	-0.23	-0.41	1.00	
Education	0.23	0.08	0.10	-0.02	-0.09	0.18	0.30	-0.20	1.00

Table 27: Correlations of Determinants
Pairwise correlations of the independent variables used to estimate equation 14, as described in "“Economics, Housing and Social Data”

Table 27 shows predicted levels of correlation which are not expected to pose an issue with econometric modelling.

7.4.1. Econometric Modelling

Monthly rationality responses are estimated from 5 year rolling regressions with give us monthly measures. Clearly, it cannot be assumed that responses to the regressors are contemporaneous and so the estimated response function from 2005 to 2009 is regressed on independent variables for 2005 to ascertain the leading nature of the determinants. This is repeated up until herding behaviour in 2019 using independent data up to 2015.

A series of linear probability models (LPM) are estimated. As these models are likely to encounter heteroskedastic errors, robust standard errors are employed in addition to time and location fixed effects. The LPMs are estimated via an ordinary least squares regression following the equation;

$$\mathbf{P(y = 1|x) = \propto_t + \beta_t x_t + \varepsilon_t} \qquad \textbf{(12)}$$

In this situation, the rational outcome is the base case and so the results measure what moves markets into irrational states. The use of LPM rests on an assumption of linear effects by the explanatory variables which may be unrealistic as the effects are expected to diminish as they move away from the average, and therefore non-linear estimations are also used to ensure the robustness of the results, specifically a multinomial logistic regression. As a set of 50 MSAs is available, then a panel dataset is used to estimate the model.

When the model is estimated with both linear and non-linear methods, the dependent variable derives from the direction and significance of the estimated non-linear coefficient. Essentially, when the response variable in Equation (7) is statistically significant then an irrational response has been identified. If the response variable is significantly negative, then that period is categorised under herding, and likewise is the estimated coefficient is significantly positive then the period is categorised as experiencing reverse herding. Therefore, the dependent variable is discrete, rather than continuous.

7.4.2. Results

Overall, when comparing coefficients from Table 28 with relative risk ratios from Table 29, the same pattern emerges. Widely, it seems that economic growth leads to more rationality as herding is less likely, with some weaker evidence for lower reverse herding. Whilst it may be assumed that economic growth would lead to overconfidence, this does not appear to be the case. House price appreciation was also expected to activate some speculative assumption of further price appreciation and therefore trigger a behavioural response. However the estimated coefficients for house price growth do not support this hypothesis.

	Panel A - Herding		**Panel B – Reverse Herding**	
Economic Growth	-0.005***	(0.001)	-0.001	(0.002)
Population Growth	0.000	(0.001)	0.003	(0.004)
House Price Growth	-0.002	(0.005)	0.002	(0.013)
Housing Costs	-0.026***	(0.004)	0.070***	(0.009)
Land Elasticity	-0.364***	(0.068)	0.111	(0.148)
Permits	-0.000	(0.000)	-0.000	(0.000)
Price/Income	0.006*	(0.003)	0.028***	(0.009)
Owner-Occupier	0.003**	(0.001)	0.012***	(0.003)
Education	-0.024***	(0.004)	-0.010	(0.010)
R^2	0.149		0.239	
Obs.	6468		6468	

Table 28: Linear Probability Estimates of Herding and Reverse Herding Outcomes
Each outcome is coded as 1 in the binary dependent variable. Figures in parentheses are heteroskedastic corrected standard errors. * denotes 10%, ** denotes 5%, *** denotes 1% significance

The failure of economic or house price growth to determine irrational responses suggests that the motivator may not simply be derived from headline macroeconomic figures. The lack of a clear speculative or overconfident trigger may in fact be in line with the argument of Daniel et al (1997) that individuals are likely to underreact to public information, instead overreacting to private information. Likewise, if overconfidence is present with strong appreciation and under inefficient conditions (Bao and Li 2020), then more work is required to separate these market conditions

	Panel A - Herding		Panel B – Reverse Herding	
Economic Growth	0.899***	(0.020)	0.963*	(0.020)
Population Growth	0.943**	(0.028)	1.253	(0.344)
House Price Growth	0.974	(0.168)	1.025	(0.083)
Housing Costs	0.591***	(0.075)	1.548***	(0.107)
Land Elasticity	0.000***	(0.000)	0.834	(0.816)
Permits	0.996	(0.004)	0.998	(0.002)
Price/Income	1.146	(0.189)	1.219***	(0.063)
Owner-Occupier	1.094**	(0.049)	1.049***	(0.020)
Education	0.535***	(0.077)	0.906	(0.059)
Pseudo R^2	0.2405		0.2405	
Obs.	6468		6468	

Table 29: Logistic Estimates of Herding and Reverse Herding Outcomes
Both Panels A and B are estimated from the same multinomial model and present relative risk ratios. Rational behaviour is the base case. Figures in parentheses are heteroskedastic corrected standard errors. * denotes 10%, ** denotes 5%, *** denotes 1% significance

Higher housing costs relative to take home pay reverse herding to be more likely, a finding which is consistent with the price-income ratio. In parallel with the former finding, markets with lower housing costs appear less likely to experience herding. When considering the economic significance of the estimated effects, these findings suggest that the overall "cost" of the market is a very strong determinant of behaviour. This is challenging to reconcile with existing literature on herding in effectively costless financial markets. If these costs are considered as constraints to liquidity, then this may illustrate potential paths for further analysis (Galariotis et al. 2016 and Campajola et al. 2020).

Higher rates of homeownership make irrational behaviour generally more likely, which was contrary to expectations as it was anticipated that homeownership caused more conservatism in investment decisions. However, it may align with findings of general overconfidence in homeowners (Hwang et al. 2020). Being an indication of greater irrationality overall, where the direction of impact is the same for herding and reverse herding, is not necessarily inconsistent as there is a rational versus irrational argument as well as the herding versus reverse herding argument.

From another homeowner perspective, higher rates of education are strongly indicative of lower herding. If these homeowners feel more informed, or more confident in their own abilities, they would have weaker justification for rational herding. However, it would be expected that these same factors may trigger reverse herding, for which there is no statistical or economic support.

The broad conclusion is that housing-specific factors are the main determinants of herding and reverse herding behaviour, which in turn motivates further investigation into housing market structures and local ownership characteristics and demographics.

Finally, in line with expectations, markets with more elastic land supply have significantly lower likelihood of experiencing herding. This suggests that in the context of the market structure the supply response does have some function, although further work could measure the speed of responses.

7.5. Conclusion

The estimated results are broadly robust across linear and non-linear modelling, and show consistently that higher housing costs and price-income ratios are associated with an increased chance of reverse herding. This may simply come from investors in unaffordable conditions hoping to invest and trade their way into more sustainable housing tenure. More elastic land supply may predict lower herding as, if construction is more feasible, investors observe significant investment and development opportunities, and be less motivated to act irrationally.

The relative weakness of economic and house price growth may be argued from the perspective that homeowners, or any active individuals in the housing market, are constrained more by their own household or corporate budget constraints, as seen via the importance of housing costs, and so be relatively less motivated by wider market price signals. It is interesting to consider that home-ownership has some consistently irrational impact on behaviour.

The use of a panel estimation has also illustrated the wide spatial and temporal variation in behaviour between MSAs, in line with existing research on housing and urban economics. In addition to pricing and growth, behavioural dynamics can also demonstrate considerable variation. This motivates continued analysis of local variation in price dynamics and investor behaviour.

The main proposed avenues for further work largely relate to data and measurement. The housing characteristics of affordability and cost require more nuanced and robust measurement.

Specifically, the economic significance of the costs variable suggests this requires more investigation. Whilst herding and reverse herding is highly conditional on overall market price dynamics such as volatility, there is some underlying evidence that variation in irrationality may be driven by the social factors rather than the economic ones (Baddeley 2010).

Another consideration is estimating which characteristics contribute to market readjustment to a rational equilibrium in behaviour. One obvious factor is the supply function and the speed of responses to irrational pricing. This will relate to the persistence of irrationality, and the conditions that determine the overall market adjustment.

Finally, an attempt at deriving the causality of rationality would deepen the theoretical understanding of the factors considered and would also inform any practical exercise to construct leading indicators.

8. Conclusion

8.1. A Review of the Motivation

This thesis has approached the question of herding in housing markets via a series of connected research questions. Broadly speaking, an empirical analysis of herding behaviour in US housing markets was conducted, identifying irrational behaviour, providing a theoretical structure for the observed behaviour in the context of a property market, testing for the relationship between herding and price bubbles, and estimating determinants of herding behaviour.

Several main themes were drawn from the initial literature review, which in turn supported the main research questions and informed the hypotheses within them.

Firstly, herding was a recurrent phenomenon across different geographies and asset classes. However, this pattern was not consistent, and herding appeared as a fleeting rather than persistent phenomenon. For example, Klein (2015) showed that there had been long term changes in herding and there had been less evidence of herding after the GFC (Zhou and Anderson 2013). Even within the same asset class, the empirical evidence was sometimes contradictory and was often conditional on other factors.

Secondly, although there was some discussion in the literature (and in grey literature) about herding being a motivator for price bubbles, limited empirical work existed to estimate the impact of herding on other market measures. If, as suggested, price bubbles will be followed by price collapses, then further work on the practical and policy implications is clearly motivated. The scale of housing as an asset class and its wider systemic economic importance further demonstrate the relevance of these specific research questions, even before considering housing's role as shelter. Investigating the connection between herding and price bubbles also supports the construction of leading indicators as a tool for forecasting and risk management.

Thirdly, much of the observed variation in herding behaviour is conditional upon other market metrics. However, whilst factors such as recessions, volatility and banking crises were clear explanations for the presence of significant herding, scope exists to estimate the underlying determinants. Estimating the structural characteristics of rational markets would illustrate why spatial variation is observed in herding, in addition to the stylised assumptions about information efficiency.

Lastly, although the field of behavioural real estate goes back to the work of Shiller at least, a significant portion of behavioural investment research relates to highly liquid and efficient securitised asset markets, and therefore there remains a gap in the literature around more inefficient and decentralised asset classes, although there is a burgeoning literature on cryptocurrencies. Even the consideration that houses cannot be shorted reveals that mis-pricing is expected to persist more than in other asset classes. Therefore, it is quite possible that inefficient behavioural dynamics, such as herding and bubbles, can exist on a prevalence and scale that is greater than in markets with short selling mechanisms.

Rather than being constrained by the classical definition of rationality, this thesis has discussed how an understanding of information asymmetries, especially in periods of market volatility, demonstrate that herding often has a very rational explanation. One of the subsidiary interests was disentangling rational and irrational motivations, especially as it became clear that the structure of housing markets is so distinct from equities that it may make a priori assumptions invalid, and this topic also motivates future research.

8.2. A Comment on Spatial Scale

There is extensive literature that demonstrates regional disparities in economics and housing markets (Hortas-Rico and Gómez-Antonio 2020, Lerbs and Oberst 2014, Oikarinen et al. 2015), and there is variation in herding behaviour between interconnected equity markets (Chang et al. 2000) in addition to the temporal variation in dynamics. This suggests that a national-level analysis may not accurately capture the full dynamics of herding.

One of the conclusions from the literature review was that there was limited research on herding in local markets without a central clearing place. Whilst listed securities may be liquid and traded without consideration of location, housing is fixed in space. Therefore, more idiosyncratic markets, such as urban housing sectors, may have even greater variation, presumably conditional on city-level characteristics. This requires careful consideration of the appropriate definition of the markets before undertaking analysis of the drivers of between-market variation.

This thesis uses contiguous urban areas as sensible units for measuring herding in integrated markets. Housing in other urban markets does not provide a substitutable good and therefore herding should be measured in the same city or metropolitan area. Therefore, using MSA-level analysis estimates herding at a sensible level and measures how localised property market structures will impact herding behaviour. This also allows a panel framework to be exploited to determine how different characteristics impact herding and bubbles.

8.3. Research Questions, Aims and Objectives

The first discrete research question aimed to estimate the prevalence of herding behaviour across MSAs, to identify whether in fact reverse herding would be common in light of the innate homeowner overconfidence that had been identified, and to assess what other factors irrational market outcomes are conditional upon.

The spatial and temporal prevalence of herding was identified at the metropolitan level, which provides an appropriate spatial scale to understand the intra-market conditions. A non-linear asset pricing model estimated the scale of herding and reverse herding responses, which also provided a framework to understand the pattern of herding and the factors it is conditional upon.

There is large spatial and temporal variation in behaviour, whilst there is a largely equal split between rational and irrational behaviour, both conclusions that align with existing research on herding. With regards to the specific assumptions about real estate markets, and in line with expectations around confident overtrading, reverse herding is found to be more prevalent than herding. Both behaviours are found to be conditional, as herding is more prevalent when markets are down, volatile (in line with rational herding) or less confident, and reverse herding is more prevalent when markets price are appreciating, stable or overconfident. Also, temporal changes were identified as herding behaviour declined noticeably after the GFC whilst reverse herding became far more common.

The second research question addressed the limited empirical investigation into the commonly cited argument that price bubbles were the result of herding. The size and importance of housing markets justifies research on triggers for price bubbles, especially due to the link between the latter and severe price declines. In addition, much research on housing markets pre-GFC stated that price appreciation was a response to fundamentals and there was no evidence of a bubble. However, the validity of this conclusion was questioned after the event, and there remains a requirement for further understanding of linkages between herding and price bubbles.

A set of bubble proxies was used in a vector autoregression framework to estimate the connection between herding and price bubbles and found, somewhat counter to expectations, evidence that herding seemed to be a result of excess growth. The conclusion provides an area for further research that considers both the theoretical underpinning of the transmission mechanism and also robust understanding of the market characteristics that drive these connections.

As with the previous research question, the GFC caused a permanent change in market dynamics. It appears that there is probably quite substantial further scope for research in temporal and structural changes in market behaviour, especially when considering the market structure and other institutional factors. The spatial element was also apparent when considering some of the potential contextual factors for price bubbles, as there is strong evidence of some size effect, perhaps resulting from the role of gateway cities in attracting property investment.

A metric of expected returns was introduced to measure the potential speculation which would trigger the irrational behavioural motivation for herding. Indeed, the role of rational herding seems limited which may suggest that the herding is speculative, further supporting previous evidence that these behavioural outcomes are highly conditional on sentiment.

Having established that herding and reverse herding are conditional on various market factors, the third research question seeks to understand some of the underlying property, economic and social factors that determine the presence of rational market outcomes. This is especially interesting in the context of the spatial variation which informs the practical perspective of investment in diversified markets.

Both a linear probability regression and a multinomial logistic regression are used to estimate a discrete choice model of rational, herding and reverse herding behaviour. Economic growth was not a major determinant of rational behaviour, and indeed the important factors were non-economic. The strongest predictors of irrational behaviour generally were house price appreciation, land supply elasticity and college-level education, the latter of which may indicate overconfidence in abilities and tally with conclusions from the first research question. High homeownership rates are significant determinants of more rational outcomes, as they potentially result in more conservative decision making.

This illustrates how the three initial conclusions drawn from the literature review form the main research questions. These in turn formed the three empirical chapters that are at the core of the thesis. The final research question, about the decentralised and spatially fixed market structure and the relevance of real estate's unique characteristics in determining rational market dynamics, runs through all three sections and provides what may be considered as the institutional setting for the thesis.

To summarise; herding in housing markets exhibits a clear partnership with reverse herding, both of which differ across space and time, and are highly dependent on market conditions. Some unique real estate characteristics can explain much of the observed behaviour, and there are associations between herding and bubble formation that provide useful frameworks for forecasting and risk modelling.

8.4. Significance and Implications

The main significance of the findings relates to the context of the research. Housing represents around $40 trillion of total household wealth of $140 trillion, demonstrating both the absolute and relative importance of this asset to both individuals and the wider economy. In addition, as housing has a primary consumption function as shelter, and also functions as some form of positional good, then this complicates the social, psychological and economic factors at play, suggesting that more robust consideration of herding behaviour is required.

Some results align with findings in other asset classes, as housing does clearly have very important functions as an investment good and store of wealth for individuals and households. Namely; rationality is conditional and varies widely over time, there is some connection between herding and bubbles, and there are some structural factors that determine the level of rational outcomes.

However, some findings were significantly different enough to warrant deeper analysis. Whilst reverse herding was expected, the scale of it compared to herding was perhaps not anticipated, and the significance of this is important. The theory on herding gives little consideration to the concept of reverse herding, and both theoretical and practical conclusions that can be drawn from these results need further development. Most importantly, the potential outcomes of reverse herding need to be considered, especially their implications for market stability.

The findings from the first research question found that herding and reverse herding are fairly symmetric in conditionality, in that there was a clear opposing pattern of when these dynamics occurred. However, this conditionality, whilst fairly clear, was not always overwhelming, and the determinants drawn from the third question were not symmetric. The significance of this is that a substantial asset class may be predisposed to an irrational dynamic with limited understanding of that dynamic's linked outcomes.

Also significant is that the connection between herding and price bubbles may not be as clear as existing theory would suggest. If herding is not causing bubbles, then there may be some other behavioural factor that needs to be considered. If rational herding is not the motivator, it may be speculation. If markets are not self-correcting because they are inefficient and lack a short selling mechanism, then some considerations should be taken about alternative policy interventions. These major differences that appeared should impact considerations of stylised facts and risk.

8.5. Contribution

This thesis makes several contributions to the existing understanding of herding in real estate markets, and in housing in particular.

Firstly, it adds to fairly limited research on herding in direct real estate and housing. The value of this particular market focus is twofold. Firstly, housing represents a significance asset class in size and systemic importance. Secondly, housing is a heavily owner-occupied market with a core consumption function operated in mainly by unsophisticated investors, demonstrating a clear difference from herding in other financial instruments.

Secondly, herding is analysed at a spatial level that provides a robust framework for understanding the practical housing markets. This follows from a consideration of the physical immovability of real estate, the practical restrictions on substitutable goods and the lack of a central clearing place.

Thirdly, reverse herding, which is relatively under-researched, is considered as an equally important dynamic in herding as a broader concept. Both theoretical and empirical arguments are presented for the prevalence of reverse herding in real estate and a unique measure of overconfidence is proposed to examine its role as a driver of reverse herding.

Fourthly, the significance and scale of the connection between herding and price bubbles is estimated, an association which has limited empirical analysis in general and especially in real estate and housing. The impact of size effects on spatial variation in the connection is proposed to be as a result of investor-focussed markets, which is itself especially relevant to investors in light of the narrative on “gateway cities”.

Finally, the determinants of rational and irrational outcomes in housing markets are assessed so that local variation in behaviour can be explained. Following this, the relative importance of real estate characteristics and the general economic conditions is discussed.

Overall, in line with the fourth research question that runs throughout the three empirical chapters, a discussion is continued on the characteristics of real estate markets that influence the findings.

If acquiring information on housing has high associated costs, due to the time and costs of physical travel, then alternative word-of-mouth sources and information cascades may provide a more efficient alternative. Therefore, as the cost of acquiring information has reduced in recent years due to the development of online listings platforms, then this cost based motivation for herding has also reduced. This may provide an additional explanation for the temporal reduction in herding.

From a liquidity perspective, if herding is more prevalent in high-liquidity assets (Galariotis 2016), then this may explain why herding is relatively less common in illiquid property markets, and may even go as far as to provide an explanation for greater observations of reverse herding.

8.6. Limitations

There are some empirical limitations to the thesis which should be considered as directions for further research. For the data assessed via the adjusted asset pricing model, the original approach proposed by CCK (2000) used homogeneous securities, specifically exchange-traded equities, which are regularly and frequently priced. However, individual houses are not regularly or frequently traded, and so the ZIP-level neighbourhoods provide a proxy for individual behaviour due to the data limitations. Clearly, if individual property transaction data and characteristics were available then micro-level behaviour could be assessed in a different empirical framework.

Likewise, when estimating the determinants of rational outcomes, better quality rental data may have helped understand if there is an affordability aspect. This could also have applied to considerations of imputed rent to better capture the consumption function of housing. Deeper analysis of the latter would allow for its separation from the investment driven behaviour that has formed the core of the empirical analysis.

Throughout this thesis, there is a reliance on the notion of highly correlated behaviour as an indicator of herding. In fact, the claim can only strictly be made that these responses show evidence of herding, rather than that they are proof of herding. Therefore, one conceptual challenge faced in this thesis is that, whilst a strong argument can be built for herding in housing markets, herding can be definitively stated as the cause. There are additional rationales for the presence of correlated behaviour in any given situation, which clearly demonstrates the need for further theoretical and empirical work on the causality of herding behaviour.

However, the empirical framework does match the behavioural concepts presented, and the theoretical framework does align with the results, suggesting that it provides a suitable explanation for the findings. For example, incorporation of volatility follows the logic of market efficiency and rational herding, and the use of overconfidence proxies matches previous understanding of memory formation.

A further issue may be the use of potentially subjective market definitions such as "up" and "down" markets, and the concept of market maturity. Whilst a binary interpretation may not fully capture all the price-driven dynamics, it does allow an initial analysis of conditional herding which then informs the subsequent research.

The empirical aspect of this thesis has heavily leaned upon a derivation from the CAPM. Whilst this is appropriate for the market and data employed in this thesis, it does raise some queries as to whether it fully isolates herding rather than other potential explanations for highly correlated price movements.

Conceptually, framing herding as being a mechanism that is simply the opposite of the rational base case may not allow for a fully independent and comprehensive investigation of herding as a distinct market dynamic. There should be future work that aims to find appropriate metrics for detecting herding as a phenomena in its own right, rather than a deviation from the classical assumptions of mainstream finance.

For example, any analysis of herding in securitised real estate could employ a volume-based approach to herding, or even potentially match transactions between parties. There could also be some consideration of the use of CSAD-based models for risk management, as unexpected levels of price dispersions may have implications for value loss and forecasting, regardless of the specific behavioural motivation.

On a deeper level, the application of asset pricing models to real estate may be limited by frequency, heterogeneity and spatial level. Practically, direct property can be valued no more frequently than monthly, and so there may need to be some consideration when applying analytical methods taken directly from equity research. However, this lack of frequency is an intrinsic part of property valuations and so perhaps should not cause undue concern.

Remaining with the issue of asset pricing models, part of the initial discussion considered the definitions of rationality and how that informed both measurement and interpretation. The empirical sections of this thesis are built around adjusted asset pricing models that test for significant non-linear responses to price movements. This is well established as the most commonly accepted approach to testing for the presence of herding behaviour, derives from mainstream methodologies in asset pricing, and has been applied successfully to housing markets.

Therefore, one long-term goal is to develop this basis to consider the non-securitised aspect of real estate and incorporate recent trends in behavioural asset pricing. Overall, the limitations of the study provide avenues for the future development of some aspects of behavioural real estate, specifically herding in housing markets.

8.7. Further Research

Beyond the limitations of the thesis, which themselves indicate direction for further work, several conclusions from the research questions require deeper investigation. Some reconciliation is needed between the conditionality of herding upon down markets in the first research question with the dependency of herding on excess returns in the second question. Whilst this seems counterintuitive, it may be that the results are sensitive to herding or bubbles measure which suggests that more robust modelling could help disentangle the effects. Similarly, separating rational and irrational herding would allow disentanglement of the relative importance of information asymmetry and speculation.

Overall, further work would ensure the robustness of results, from the perspective of data, measurement and pricing models. For example, housing-specific measurements of sentiment and overconfidence require more calibration to ensure their applicability in leading indicators.

More broadly, the second research question's use of different bubble measures to estimate the robustness of findings has clearly demonstrated the scope for better measures of price bubbles, including scale as well as existence. This would support simulated models of herding and bubbles to be incorporated into forecasting and risk modelling. An additional task would be deeper analysis on the persistence of returns.

Having identified spatial variation via the first research question, one major area to extend the modelling approach would be to incorporate local effects. Whilst variation between MSAs has been discussed, closer consideration of sub-markets within MSAs would measure if herding is MSA-wide or clustered in certain parts of an MSA. Spatial analysis could provide some understanding of spatial dependency or spillovers in herding (Hyun and Milcheva 2018), helping to measure if neighbourhoods are impacted both by their neighbours as well as wider MSA and national effects.

Likewise, future work could investigate if herding and bubbles exist on the same spatial scale, which would also build upon the concept of spatial links. If this allowed for more robust estimation of connections between herding and price bubbles, then it would allow for the construction of useful leading indicators. In turn, leading indicators could have some practical use in forecasting, risk management and portfolio construction. The construction of these would also be improved by the incorporation of asymmetric responses and non-linear effects, as well as MSA-specific measures of dissimilarity.

The results on rational herding need further reconciliation, which requires more work on volatility and the context of information transmission. A model of information costs in housing markets based as asymmetric availability and the costs for private acquisition may assist more comprehensive analysis of rational herding. Another key factor in information efficiency would be analysing changes in behaviour resulting from the transition to online listings platforms.

The parallel motivation to rational herding is the speculative assumption. Isolating the latter effect requires further work on excess returns, which may require measurements of fundamental returns in the market, perhaps through an imputed rents approach or more mainstream forecasting techniques.

As mentioned, housing retains a primary consumption function as shelter and therefore another issue to disentangle would be the investment and consumption functions. This is further complicated by differing tax treatments between administrative geographies, and links with the need for further work on institutions and market structure.

Whilst market states have been modelled in an essentially binary "up" and "down" context, this does raise the issue of subjective definitions for conditional analysis. More sophisticated and nuanced treatment of price dynamics may yield more insight into the mechanisms that drive herding. Whilst asymmetric effects have been estimated, and indeed much of the interpretation of results is dependent on market conditions, there is clearly scope for more robust analysis, including perhaps finer sub-samples of market conditions. For example, "up" markets could be separated into accelerating or decelerating growth, and the persistence of the market condition could be incorporated. Another aspect may again be the spatial and temporal variation in these results.

Lastly, all these factors feed into estimating the determinants of rational behaviour via the choice model illustrated in the third research question. Better incorporating local characteristics, the unique property factors, and disentangling speculative and consumption effects may also assist in modelling the time lags in the various herding and bubble mechanisms, and estimating the contemporaneous effects. If institutional structures can also be factored in, then ultimately it may lay the groundwork for estimating causal factors.

8.8. Final Comments

This thesis represents a comprehensive assessment of herding in housing markets, through identification, impacts and determinants, with a consistent focus on the real estate context.

Real estate, specifically US metropolitan housing markets, shares some similarities with all asset classes as they all demonstrate inconsistency and conditionality in patterns of herding. However, these patterns also show significant evidence of spatial variation due to the unique localised aspect of real estate as an investment asset. Equally significant is that reverse herding is far more prevalent than herding. Both these conclusions have implications for future research.

The initial literature review identified several gaps in knowledge regarding herding in housing markets, and this thesis has made some effort toward filling these gaps in an innovative way. A dataset that offers relatively small scale price data was used to identify behaviour across naturally-defined markets, rather than arbitrary administrative boundaries. Two commonly discussed irrational behaviours were brought together in an empirical setting. Lastly, a choice modelling approach estimated what local socio-economic and real estate characteristics determine the rationality of market outcomes.

As with previous literature, broad spatial and temporal variation is identified in herding behaviour. Reverse herding is actually far more prevalent, which is attributed to the costs of acquiring private information in an inefficient market such as housing. There is limited evidence that herding creates price bubbles, and in fact price bubbles may trigger herding, which can also follow the direction of logic. Lastly, some characteristics are identified that lead to herding and reverse herding behaviour. It can be concluded that herding in housing markets shows some distinctly different behaviour from securitised markets and from theory, and the market structure may inform much of this.

The significance of these findings is conditional itself upon the context that little research was found on herding in housing markets, and it has been established that the size and importance of this asset class surely motivates in depth understanding of dynamics that can be associated with sub-optimal market outcomes. Property markets are very local due to the physical immovability of the asset plus the lack of a central clearing place, so national-level herding has limited practical use unless considering market convergence and integration. As with many aspects of real estate investment, there has been limited discussion of why herding may take place in real estate markets relative to equity markets, and asymmetric access to information could represent an important factor in understanding the significant findings about reverse herding. Largely there is a requirement for further work on understanding the variation between exchange-traded securities and real estate in terms of the pricing dynamics, and relating this to the theoretical underpinnings and empirical analysis.

9. References

1. Alpizar, F., Carlsson, F. and Johansson-Stenman, O., 2005. How much do we care about absolute versus relative income and consumption? *Journal of Economic Behavior & Organization,* 56 (3), 405-421.
2. An, X., Cordell, L. and Nichols, J.B., 2020. Reputation, information, and herding in credit ratings: Evidence from CMBS. *The Journal of Real Estate Finance and Economics,* 61, 476-504.
3. Anenberg, E., 2011. Loss aversion, equity constraints and seller behavior in the real estate market. *Regional Science and Urban Economics,* 41 (1), 67-76.
4. Arrow, K., et al., 2004. Are we consuming too much? *Journal of Economic Perspectives,* 18 (3), 147-172.
5. Attaran, M., 1986. Industrial diversity and economic performance in US areas. *The Annals of Regional Science,* 20 (2), 44-54.
6. Avery, C.N. and Chevalier, J.A., 1999. Herding over the career. *Economics Letters,* 63 (3), 327-333.
7. Baddeley, M., 2010. Herding, social influence and economic decision-making: socio-psychological and neuroscientific analyses. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 365 (1538), 281-290.
8. Bakas, D., Magkonis, G. and Oh, E.Y., 2022. What drives volatility in Bitcoin market? *Finance Research Letters,* 50, 103237.
9. Baker, M. and Stein, J.C., 2004. Market liquidity as a sentiment indicator. *Journal of Financial Markets,* 7 (3), 271-299.
10. Baker, M. and Wurgler, J., 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives,* 21 (2), 129-151.
11. Baldauf, M., Garlappi, L. and Yannelis, C., 2020. Does climate change affect real estate prices? Only if you believe in it. *The Review of Financial Studies,* 33 (3), 1256-1295.
12. Ball, L.M. and Romer, D.H., 1993. No title. *Inflation and the Informativeness of Prices,* .
13. Banerjee, A.V., 1992. A simple model of herd behavior. *The Quarterly Journal of Economics,* 107 (3), 797-817.
14. Bao, H. and Li, S.H., 2020. Investor overconfidence and trading activity in the Asia Pacific REIT markets. *Journal of Risk and Financial Management,* 13 (10), 232.
15. Barber, B.M. and Odean, T., 2013, The behavior of individual investors. *In:* The behavior of individual investors. *Handbook of the Economics of Finance.* Elsevier, 2013, pp. 1533-1570.
16. Barber, B.M. and Odean, T., 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance,* 55 (2), 773-806.

17. Barber, B.M., Odean, T. and Zhu, N., 2009. Systematic noise. *Journal of Financial Markets,* 12 (4), 547-569.

18. Barnes, M.L. and Hughes, A.T.W., 2002. A quantile regression analysis of the cross section of stock market returns.

19. Bekaert, G. and Wu, G., 2000. Asymmetric volatility and risk in equity markets. *The Review of Financial Studies,* 13 (1), 1-42.

20. Bekiros, S., et al., 2017. Herding behavior, market sentiment and volatility: will the bubble resume? *The North American Journal of Economics and Finance,* 42, 107-131.

21. Bernstein, A., Gustafson, M.T. and Lewis, R., 2019. Disaster on the horizon: The price effect of sea level rise. *Journal of Financial Economics,* 134 (2), 253-272.

22. Bikhchandani, S., Hirshleifer, D. and Welch, I., 1998. Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives,* 12 (3), 151-170.

23. Bikhchandani, S., Hirshleifer, D. and Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy,* 100 (5), 992-1026.

24. Blasco, N., Corredor, P. and Ferreruela, S., 2012. Market sentiment: a key factor of investors' imitative behaviour. *Accounting & Finance,* 52 (3), 663-689.

25. Brown, P.H., Bulte, E. and Zhang, X., 2011. Positional spending and status seeking in rural China. *Journal of Development Economics,* 96 (1), 139-149.

26. Campajola, C., Lillo, F. and Tantari, D., 2020. Unveiling the relation between herding and liquidity with trader lead-lag networks. *Quantitative Finance,* 20 (11), 1765-1778.

27. Carlino, G. and DeFina, R., 1999. The differential regional effects of monetary policy: Evidence from the US states. *Journal of Regional Science,* 39 (2), 339-358.

28. Carlino, G. and DeFina, R., 1998. The differential regional effects of monetary policy. *Review of Economics and Statistics,* 80 (4), 572-587.

29. Carlsson, F., Johansson-Stenman, O. and Martinsson, P., 2007. Do you enjoy having more than others? Survey evidence of positional goods. *Economica,* 74 (296), 586-598.

30. Chang, C. and Lin, S., 2015. The effects of national culture and behavioral pitfalls on investors' decision-making: Herding behavior in international stock markets. *International Review of Economics & Finance,* 37, 380-392.

31. Chang, E.C., Cheng, J.W. and Khorana, A., 2000. An examination of herd behavior in equity markets: An international perspective. *Journal of Banking & Finance,* 24 (10), 1651-1679.

32. Chiang, T.C. and Zheng, D., 2010. An empirical analysis of herd behavior in global stock markets. *Journal of Banking & Finance,* 34 (8), 1911-1921.

33. Choi, K. and Yoon, S., 2020. Investor sentiment and herding behavior in the Korean stock market. *International Journal of Financial Studies,* 8 (2), 34.

34. Choi, N. and Skiba, H., 2015. Institutional herding in international markets. *Journal of Banking & Finance,* 55, 246-259.

35. Christie, W.G. and Huang, R.D., 1995. Following the pied piper: do individual returns herd around the market? *Financial Analysts Journal,* 51 (4), 31-37.

36. Chuang, W., Lee, B. and Wang, K., 2014. US and domestic market gains and Asian investors' overconfident trading behavior. *Financial Management,* 43 (1), 113-148.

37. Conrad, J., Gultekin, M.N. and Kaul, G., 1991. Asymmetric predictability of conditional variances. *The Review of Financial Studies,* 4 (4), 597-622.

38. Coskun, E.A., Lau, C.K.M. and Kahyaoglu, H., 2020. Uncertainty and herding behavior: evidence from cryptocurrencies. *Research in International Business and Finance,* 54, 101284.

39. Coulson, N.E., McCoy, S.J. and McDonough, I.K., 2020. Economic diversification and the resiliency hypothesis: Evidence from the impact of natural disasters on regional housing values. *Regional Science and Urban Economics,* 85, 103581.

40. Cui, Y., Gebka, B. and Kallinterakis, V., 2019. Do closed-end fund investors herd? *Journal of Banking & Finance,* 105, 194-206.

41. Damianov, D.S. and Escobari, D., 2016. Long-run equilibrium shift and short-run dynamics of US home price tiers during the housing bubble. *The Journal of Real Estate Finance and Economics,* 53, 1-28.

42. Daniel, K.D., Hirshleifer, D.A. and Subrahmanyam, A., 1997. A theory of overconfidence, self-attribution, and security market under-and over-reactions. *Self-Attribution, and Security Market Under-and Over-Reactions (February 19, 1997),* .

43. DeCoster, G.P. and Strange, W.C., 2012. Developers, herding, and overbuilding. *The Journal of Real Estate Finance and Economics,* 44 (1-2), 7-35.

44. Deng, X., Hung, S. and Qiao, Z., 2018. Mutual fund herding and stock price crashes. *Journal of Banking & Finance,* 94, 166-184.

45. Dermisi, S. and McDonald, J., 2010. Selling prices/sq. ft. of office buildings in down town Chicago—How much is it worth to be an old but Class A building? *Journal of Real Estate Research,* 32 (1), 1-22.

46. Devenow, A. and Welch, I., 1996. Rational herding in financial economics. *European Economic Review,* 40 (3-5), 603-615.

47. Duffee, G.R., 2001. Asymmetric cross-sectional dispersion in stock returns: evidence and implications. *In: Federal Reserve Bank of San Francisco,* Citeseer, .

48. Economou, F., Kostakis, A. and Philippas, N., 2011. Cross-country effects in herding behaviour: Evidence from four south European markets. *Journal of International Financial Markets, Institutions and Money,* 21 (3), 443-460.

49. Effinger, M.R. and Polborn, M.K., 2001. Herding and anti-herding: A model of reputational differentiation. *European Economic Review,* 45 (3), 385-403.

50. Ekholm, A. and Pasternack, D., 2008. Overconfidence and investor size. *European Financial Management,* 14 (1), 82-98.

51. Flynn, S.M., 2012. Noise-trading, costly arbitrage, and asset prices: Evidence from US closed-end funds. *Journal of Financial Markets,* 15 (1), 108-125.

52. Frank, R.H., 2008. Should public policy respond to positional externalities? *Journal of Public Economics,* 92 (8-9), 1777-1786.

53. Frank, R.H., 2005. Positional externalities cause large and preventable welfare losses. *American Economic Review,* 95 (2), 137-141.

54. Freybote, J. and Seagraves, P.A., 2017. Heterogeneous investor sentiment and institutional real estate investments. *Real Estate Economics,* 45 (1), 154-176.

55. Galariotis, E.C., Krokida, S. and Spyrou, S.I., 2016. Herd behavior and equity market liquidity: Evidence from major markets. *International Review of Financial Analysis,* 48, 140-149.

56. Galariotis, E.C., Rong, W. and Spyrou, S.I., 2015. Herding on fundamental information: A comparative study. *Journal of Banking & Finance,* 50, 589-598.

57. Gallimore, P. and McAllister, P., 2004. Expert judgement in the processes of commercial property market forecasting. *Journal of Property Research,* 21 (4), 337-360.

58. Garber, P., 2000. *Famous First Bubbles: The Fundamentals of Early Manias.* The MIT Press.

59. Gębka, B. and Wohar, M.E., 2013. International herding: Does it differ across sectors? *Journal of International Financial Markets, Institutions and Money,* 23, 55-84.

60. Giannakis, E. and Bruggeman, A., 2017. Economic crisis and regional resilience: Evidence from Greece. *Papers in Regional Science,* 96 (3), 451-476.

61. Giglio, S., et al., 2021. Climate change and long-run discount rates: Evidence from real estate. *The Review of Financial Studies,* 34 (8), 3527-3571.

62. Gleason, K.C., Mathur, I. and Peterson, M.A., 2004. Analysis of intraday herding behavior among the sector ETFs. *Journal of Empirical Finance,* 11 (5), 681-694.

63. Glossner, S., et al., 2021. Do institutional investors stabilize equity markets in crisis periods? Evidence from COVID-19. *Evidence from COVID-19 (December 8, 2021).Swiss Finance Institute Research Paper,* (20-56).

64. Goodfellow, C., Bohl, M.T. and Gebka, B., 2009. Together we invest? Individual and institutional investors' trading behaviour in Poland. *International Review of Financial Analysis,* 18 (4), 212-221.

65. Gray, D., 2018. Convergence and divergence in British housing space. *Regional Studies,* 52 (7), 901-910.

66. Griffin, J.M., Harris, J.H. and Topaloglu, S., 2003. The dynamics of institutional and individual trading. *The Journal of Finance,* 58 (6), 2285-2320.

67. Griffin, J.M., Nardari, F. and Stulz, R.M., 2007. Do investors trade more when stocks have performed well? Evidence from 46 countries. *The Review of Financial Studies,* 20 (3), 905-951.

68. Gupta, R. and Kabundi, A., 2010. The effect of monetary policy on house price inflation: A factor augmented vector autoregression (FAVAR) approach. *Journal of Economic Studies,* 37 (6), 616-626.

69. Gyourko, J. and Saiz, A., 2006. Construction costs and the supply of housing structure. *Journal of Regional Science,* 46 (4), 661-680.

70. Harras, G. and Sornette, D., 2011. How to grow a bubble: A model of myopic adapting agents. *Journal of Economic Behavior & Organization,* 80 (1), 137-152.

71. Henrich, J., et al., 2001. In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review,* 91 (2), 73-78.

72. Holt, J.R. and Borsuk, M.E., 2020. Using Zillow data to value green space amenities at the neighborhood scale. *Urban Forestry & Urban Greening,* 56, 126794.

73. Hong, Y., Tu, J. and Zhou, G., 2007. Asymmetries in stock returns: Statistical tests and economic evaluation. *The Review of Financial Studies,* 20 (5), 1547-1581.

74. Hopkins, E. and Kornienko, T., 2004. Running to keep in the same place: Consumer choice as a game of status. *American Economic Review,* 94 (4), 1085-1107.

75. Hortas-Rico, M. and Gómez-Antonio, M., 2020. Expansionary zoning and the strategic behaviour of local governments. *Regional Studies,* 54 (3), 388-402.

76. Hott, C., 2012. The influence of herding behaviour on house prices. *Journal of European Real Estate Research,* 5 (3), 177-198.

77. Hott, C., 2009. Herding behavior in asset markets. *Journal of Financial Stability,* 5 (1), 35-56.

78. Huang, E.J., 2015. The role of institutional investors and individual investors in financial markets: Evidence from closed-end funds. *Review of Financial Economics,* 26, 1-11.

79. Hwang, M. and Quigley, J.M., 2006. Economic fundamentals in local housing markets: evidence from US metropolitan regions. *Journal of Regional Science,* 46 (3), 425-453.

80. Hwang, S., Cho, Y. and Shin, J., 2020. The impact of UK household overconfidence in public information on house prices. *Journal of Property Research,* 37 (4), 360-389.

81. Hwang, S. and Salmon, M., 2004. Market stress and herding. *Journal of Empirical Finance,* 11 (4), 585-616.

82. Hyun, D. and Milcheva, S., 2018. Spatial dependence in apartment transaction prices during boom and bust. *Regional Science and Urban Economics,* 68, 36-45.

83. Joshi, N.K., 2016. Local house prices and mental health. *International Journal of Health Economics and Management,* 16, 89-102.

84. Kahneman, D. and Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica,* 47 (2), 363-391.

85. Kindleberger, C., 1978. *Manias, Panics and Crashes.*

86. Klein, A.C., 2013. Time-variations in herding behavior: Evidence from a Markov switching SUR model. *Journal of International Financial Markets, Institutions and Money,* 26, 291-304.

87. Kort, J.R., 1981. Regional economic instability and industrial diversification in the US. *Land Economics,* 57 (4), 596-608.

88. Lam, K.S. and Qiao, Z., 2015. Herding and fundamental factors: The Hong Kong experience. *Pacific-Basin Finance Journal,* 32, 160-188.

89. Lan, T., 2014. Herding behavior in China housing market. *International Journal of Economics and Finance,* 6 (2), 115-124.

90. Lantushenko, V. and Nelling, E., 2017. Institutional property-type herding in real estate investment trusts. *The Journal of Real Estate Finance and Economics,* 54 (4), 459-481.

91. Leguizamon, S.J. and Ross, J.M., 2012. Revealed preference for relative status: Evidence from the housing market. *Journal of Housing Economics,* 21 (1), 55-65.

92. Lei, V., Noussair, C.N. and Plott, C.R., 2001. Nonspeculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs. actual irrationality. *Econometrica,* 69 (4), 831-859.

93. Lerbs, O.W. and Oberst, C.A., 2014. Explaining the spatial variation in homeownership rates: Results for German regions. *Regional Studies,* 48 (5), 844-865.

94. LeRoy, S.F. and Porter, R.D., 1981. The present-value relation: Tests based on implied variance bounds. *Econometrica,* , 555-574.

95. Levy, G., 2004. Anti-herding and strategic consultation. *European Economic Review,* 48 (3), 503-525.

96. Liao, T., Huang, C. and Wu, C., 2011. Do fund managers herd to counter investor sentiment? *Journal of Business Research,* 64 (2), 207-212.

97. Litimi, H., BenSaïda, A. and Bouraoui, O., 2016. Herding and excessive risk in the American stock market: A sectoral analysis. *Research in International Business and Finance,* 38, 6-21.

98. Lux, T., 1995. Herd behaviour, bubbles and crashes. *The Economic Journal,* 105 (431), 881-896.

99. Mcallister, P., Newell, G. and Matysiak, G., 2008. Agreement and accuracy in consensus forecasts of the UK commercial property market. *Journal of Property Research,* 25 (1), 1-22.

100. Miller, N. and Pandher, G., 2008. Idiosyncratic volatility and the housing market. *Journal of Housing Research,* 17 (1), 13-32.

101. Ngene, G.M., Sohn, D.P. and Hassan, M.K., 2017. Time-varying and spatial herding behavior in the US housing market: Evidence from direct housing prices. *The Journal of Real Estate Finance and Economics,* 54, 482-514.

102. Odean, T., 1999. Do investors trade too much? *American Economic Review,* 89 (5), 1279-1298.

103. Oikarinen, E., Peltola, R. and Valtonen, E., 2015. Regional variation in the elasticity of supply of housing, and its determinants: The case of a small sparsely populated country. *Regional Science and Urban Economics,* 50, 18-30.

104. Palomares-Linares, I. and Van Ham, M., 2020. Understanding the effects of homeownership and regional unemployment levels on internal migration during the economic crisis in Spain. *Regional Studies,* 54 (4), 515-526.

105. Papastamos, D., Matysiak, G. and Stevenson, S., 2015. Assessing the accuracy and dispersion of real estate investment forecasts. *International Review of Financial Analysis,* 42, 141-152.

106. Parr, J.B., 1965. Specialization, diversification and regional development. *The Professional Geographer,* 17 (6), 21-25.

107. Philippas, N., et al., 2013. Herding behavior in REITs: Novel tests and the role of financial crisis. *International Review of Financial Analysis,* 29, 166-174.

108. Pierdzioch, C., Rülke, J. and Stadtmann, G., 2012. Housing starts in Canada, Japan, and the United States: Do forecasters herd? *The Journal of Real Estate Finance and Economics,* 45, 754-773.

109. Ritter, J.R., 2003. Behavioral finance. *Pacific-Basin Finance Journal,* 11 (4), 429-437.

110. Rivas, R., et al., 2019. The impact of colleges and hospitals to local real estate markets. *Journal of Big Data,* 6 (1), 1-24.

111. Ro, S. and Gallimore, P., 2014. Real estate mutual funds: Herding, momentum trading and performance. *Real Estate Economics,* 42 (1), 190-222.

112. Ro, S., et al., 2019. Herding behavior among residential developers. *The Journal of Real Estate Finance and Economics,* 59, 272-294.

113. Saiz, A., 2008. On local housing supply elasticity. *Available at SSRN 1193422,* .

114. Scharfstein, D.S. and Stein, J.C., 1990. Herd behavior and investment. *The American Economic Review,* , 465-479.

115. Seiler, M.J., Lane, M.A. and Harrison, D.M., 2014. Mimetic herding behavior and the decision to strategically default. *The Journal of Real Estate Finance and Economics,* 49, 621-653.

116. Shapiro, A.H., Sudhof, M. and Wilson, D.J., 2022. Measuring news sentiment. *Journal of Econometrics,* 228 (2), 221-243.

117. Shiller, R.J., 1982. Consumption, asset markets and macroeconomic fluctuations. *In: Carnegie-Rochester Conference Series on Public Policy,* Elsevier, pp. 203-238.

118. Sias, R.W., 2004. Institutional herding. *The Review of Financial Studies,* 17 (1), 165-206.

119. Sibande, X., et al., 2023. Investor Sentiment and (Anti) Herding in the Currency Market: Evidence from Twitter Feed Data. *Journal of Behavioral Finance,* 24 (1), 56-72.

120. Siegel, P.B., Alwang, J. and Johnson, T.G., 1995. A structural decomposition of regional economic instability: a conceptual framework. *Journal of Regional Science,* 35 (3), 457-470.

121. Siegel, P.B., Alwang, J. and Johnson, T.G., 1994. Toward an improved portfolio variance measure of regional economic stability. *Review of Regional Studies,* 24 (1), 71-86.

122. Solnick, S.J. and Hemenway, D., 2005. Are positional concerns stronger in some domains than in others? *American Economic Review,* 95 (2), 147-151.

123. Solnick, S.J. and Hemenway, D., 1998. Is more always better?: A survey on positional concerns. *Journal of Economic Behavior & Organization,* 37 (3), 373-383.

124. Solnick, S.J., Hong, L. and Hemenway, D., 2007. Positional goods in the United States and China. *The Journal of Socio-Economics,* 36 (4), 537-545.

125. Thoma, M., 2013. Bad advice, herding and bubbles. *Journal of Economic Methodology,* 20 (1), 45-55.

126. Trueman, B., 1994. Analyst forecasts and herding behavior. *The Review of Financial Studies,* 7 (1), 97-124.

127. Tsai, I., 2015. Spillover effect between the regional and the national housing markets in the UK. *Regional Studies,* 49 (12), 1957-1976.

128. Veblen, T., 1899. *The Theory of the Leisure Class.* Routledge.

129. Vieira, E.F.S. and Pereira, M.S.V., 2015. Herding behaviour and sentiment: Evidence in a small European market. *Revista De Contabilidad,* 18 (1), 78-86.

130. Vogiazas, S. and Alexiou, C., 2017. Determinants of housing prices and bubble detection: Evidence from seven advanced economies. *Atlantic Economic Journal,* 45 (1), 119-131.

131. Wagner, J.E. and Deller, S.C., 1998. Measuring the effects of economic diversity on growth and stability. *Land Economics,* , 541-556.

132. Wei, S., Zhang, X. and Liu, Y., 2012. Status competition and housing prices.

133. Welch, I., 1992. Sequential sales, learning, and cascades. *The Journal of Finance,* 47 (2), 695-732.

134. West, K.D., 1988. Bubbles, fads and stock price volatility tests: a partial evaluation. *The Journal of Finance,* 43 (3), 639-656.

135. Wiley, J.A., 2017. Leverage, liquidity and information in commercial property prices. *Journal of Property Research,* 34 (2), 77-107.

136. Yao, J., Ma, C. and He, W.P., 2014. Investor herding behaviour of Chinese stock market. *International Review of Economics & Finance,* 29, 12-29.

137. Zahirovic-Herbert, V. and Chatterjee, S., 2011. What is the value of a name? Conspicuous consumption and house prices. *Journal of Real Estate Research,* 33 (1), 105-126.

138. Zhang, D. and Fan, G., 2019. Regional spillover and rising connectedness in China's urban housing prices. *Regional Studies,* 53 (6), 861-873.

139. Zhou, J. and Anderson, R.I., 2013. An empirical investigation of herding behavior in the US REIT market. *The Journal of Real Estate Finance and Economics,* 47, 83-108.