Housing search activity and quantiles-based predictability of housing price movements in the USA

Rangan Gupta and Damien Moodley
Department of Economics, University of Pretoria, Pretoria, South Africa

Abstract

Purpose – Recent evidence from a linear econometric framework infers that housing search activity, captured from Google Trends data, can predict housing returns for the USA at a national and regional (metropolitan statistical area [MSA]) level. Based on search theory, the authors, however, postulate that search activity can also predict housing returns volatility. This study aims to explore the possibility of using online search activity to predict both housing returns and volatility.

Design/methodology/approach – Using a k-th order non-parametric causality-in-quantiles test allows us to test for predictability in a robust manner over the entire conditional distribution of both housing price returns and its volatility (i.e. squared returns) by controlling for nonlinearity and structural breaks that exist in the data.

Findings – The analysis over the monthly period of 2004:01 to 2021:01 produces results indicating that while housing search activity continues to predict aggregate US house price returns, barring the extreme ends of the conditional distribution, volatility is relatively strongly predicted over the entire quantile range considered. The results carry over to an alternative (the generalized autoregressive conditional heteroskedasticity-based) metric of volatility, higher (weekly)-frequency data (over January 2018–March 2021) and to over 84% of the 77 MSAs considered.

Originality/value – To the best of the authors' knowledge, this is the first study regarding predictability of overall and regional US housing price returns and volatility using search activity, based on a non-parametric higher-order causality-in-quantiles framework, which is insightful to investors, policymakers and academics.

Keywords International housing markets, Housing prices, Housing market analysis, Higher-order non-parametric causality in quantiles test, Housing returns and volatility, Housing search activity

Paper type Research paper

1. Introduction

In a recent paper, Møller *et al.* (2023) provide statistical evidence in favour of the hypothesis that online housing search activity (measured by a housing search index [HSI] obtained from Google Trends data) [1], which captures peoples' intentions of buying a house and

hence, proxies for housing demand, contain predictive information for housing price returns for the overall USA, and its regions. This is not surprising, since an increase in search activity is propagated into future periods, which, given various frictions in the housing market, would imply sluggish price adjustment in response to an increase in demand, so that search activity should hold predictive power for future variation in house prices. In this regard, the reader is referred to the conclusions drawn from the theoretical search-based models or Berkovec and Goodman (1996), Genesove and Han (2012) and Carrillo et al. (2015).

In this regard, Ngai and Sheedy (2022), extending the earlier works of Díaz and Jerez (2013), Ngai and Sheedy (2020) and Smith (2020), used a calibrated search and matching model with both endogenous inflows (new listings) and outflows (sales), to show that a single persistent housing demand shock induces more moving and increases the supply of houses on the market and hence, can quantitatively match the data on volatility of various housing market variables, including housing price returns variability. In other words, we can postulate that the HSI of Møller et al. (2023) should not only contain predictive information for house price returns but also its volatility.

To test our proposition, we use the k-th order non-parametric causality-in-quantiles framework of Balcilar et al. (2018). This econometric model allows us to test the predictability of the entire conditional distributions (capturing regimes) through quantiles of both housing price returns and squared returns, i.e. volatility simultaneously, by controlling for misspecification due to uncaptured non-linearity and regime changes with the HSI in a non-parametric manner – both of which we show to exist in our data set via formal statistical tests. As our focus is on volatility in this paper, being an extension to the work of Møller et al. (2023), to check for the robustness of our results, we also apply the first-order of the test to the conditional volatility as captured by the generalized autoregressive conditional heteroskedasticity (GARCH) model of Bollerslev (1986), which is a well-established approach to obtain an estimate for model-based volatility. While the primary focus is the aggregate US housing price returns and its volatility, just as in Møller et al. (2023), we also analyse the predictive impact of the HSI for the first and second moment of housing prices of 77 metropolitan statistical areas (MSAs), as it is well-known that the US housing market is highly segmented (Gupta et al., 2023). Based on data availability, we conduct these predictive experiments over the monthly period of 2004:01 to 2021:01.

Statistically speaking, US residential real estate represents about 85.00% of total household non-financial assets, 32.56% of total household net worth and 35.10% of US net wealth (financial accounts of the USA, Second Quarter, 2023) [2]. Hence, it is not surprising that housing price movements have been historically associated with aggregate and regional business cycles [Balcilar et al. (2014), Apergis et al. (2015), Nyakabawo et al. (2015), Emirmahmutoglu et al. (2016) and Payne and Sun (2023)]. Naturally, predicting the future path of housing price returns and its volatility contingent on the information content of the HSI in our current context is of immense value, to not only real estate consumers and investors but also to the policymaker. Understandably, information on the evolution of house price movements at a higher frequency would be immense value in making timely portfolio decisions (Bollerslev et al., 2016; Nyakabawo et al., 2018), and in particular to policy authorities from the perspective of nowcasting (Bańbura et al., 2011), which will assist in the designing of monetary and fiscal responses ahead of time to prevent possible recessions (Balcilar et al., 2020, 2021; Bouri et al., 2021). Hence, we also conduct our analysis at a weekly-frequency over the period of January 2018 to March 2021.

To the best of our knowledge, this is the first paper that evaluates the predictive power of HSI for overall and regional US housing price returns and volatility based on a non-parametric higher-order causality-in-quantiles framework, with the hypothesis derived from

Ngai and Sheedy (2022), who show that a single persistent housing demand shock induces more moving and increases the supply of houses on the market, and hence, can cause housing price returns variability. In the process, we add to the large existing literature on predicting the first and second moment of US house prices using various types of econometric models and predictors, the review of which is not only beyond the scope of this paper but also not its objective, with the reader referred to the recent works of Bork and Møller (2015), Bork et al. (2020), Segnon et al. (2021) and Gupta et al. (2022) for this purpose.

The remainder of the paper is structured as follows: Section 2 describes the data used for our analysis, as well as outlines the methodology. Section 3 presents the findings, and Section 4 concludes the paper.

2. Data and methodology

2.1 Data sets

This sub-section provides specific details concerning the data set used in the main analysis. Furthermore, it presents an overview of the econometric methodology implemented to perform our investigation.

As mentioned above, we make use of a newly developed HSI introduced by Møller et al. (2023) to test the possibility of using online search activity to predict housing returns and volatility of the aggregate USA and that, for 77 MSAs [3]. HSI is constructed using Google trends data, to quantify internet search activity related to housing demand. Google Trends data are available from 2004 onwards, resulting in a sample period of 2004:01 to 2021:01 at the monthly frequency. To obtain a measure of housing demand, Møller et al. (2023) initially used "buying a house" as their main search term and subsequently used a list of 22 related terms, namely: "when buying a house", "buying a home", "buy a house", "mortgage", "buying a new house", "before buying a house", "how to buy a house", "real estate", "steps to buying a house", "buying a house calculator", "first time buying a house", "buying a house process", "house buying process", "homes for sale", "building a house", "buying a house with bad credit", "cost of buying a house", "buying a house to rent", "mortgage calculator", "houses for sale", "buying a house tips" and "buying a foreclosure house".

To filter out the noise and more accurately estimate latent demand, Møller et al. (2023) use the elastic net estimator to select the ten most relevant search indexes and then apply principal component analysis (PCA) to summarize the most important information from these indices into one common component, which is interpreted as a summary measure for housing search and referred to as the HSI [4]. Note that the same approach is followed for the overall USA, but by now specifying the MSA for which the search is conducted [5]. Recall that the elastic net is an estimation method for linear regressions with many predictors, which allows the econometrician to choose the most relevant predictors in explaining the variability of the dependent variable. While, PCA is a data analysis technique, which is particularly useful for reducing the dimensionality of data sets, but ensuring the preservation crucial information.

We use the log returns (HR), i.e. the first difference of the natural logarithm of the seasonally adjusted monthly Federal Housing Finance Agency purchase-only house price index for the USA and the MSAs to capture housing price returns, with the corresponding squared values measuring volatility [6]. As indicated earlier, a GARCH model was estimated on the log returns also provide an alternative conditional (model-based) estimate of volatility. As part of our high-frequency analysis, we also use housing log returns at weekly frequency using the smoothed, seasonally adjusted weekly median sale prices from Zillow [7] for the overall USA over the 1st week of January, 2018 to the 4th week of March, 2021. Given the availability of the HSI data at a high frequency, this

analysis will depict to us if the results are dependent on the frequency of the data and, hence, the speed of information transmission from housing search to being reflected in the moments of house prices.

Table A1 and Figure A1 in the Appendix of the paper summarizes the HR and HSI variables for the overall USA over 2004:01 to 2021:01. As can be seen from Table A1, HR is negatively skewed and has excess kurtosis, resulting in a non-normal distribution as indicated by the overwhelming rejection of the null of normality under the Jarque–Bera test. This evidence of heavy-tails provides preliminary justification for using a quantiles-based approach, rather than a conditional mean-reliant method, to predictability, as using the latter is likely to miss important information at various parts of the conditional distribution of housing returns.

2.2 Econometric model

In this sub-section, we briefly present the methodology for testing nonlinear causality via a hybrid approach as developed by Balcilar *et al.* (2018), who combine the higher-order conditional mean-based non-parametric approach of Nishiyama *et al.* (2011), with the quantiles-based first-moment framework of Jeong *et al.* (2012). Let y_t denote housing returns and x_t the HSI. Further, let.

$Y_{t-1} \equiv (Y_{t-1}, \cdots, Y_{t-p})$, $X_{t-1} \equiv (x_{t-1}, \cdots, x_{t-p})$, $Z_t = (X_t, Y_t)$ and $F_{y_t}(y_t | \bullet)$ denote the conditional distribution of y_t given $\bullet$. Defining $Q_\theta(Z_{t-1}) \equiv Q_\theta(y_t | Z_{t-1})$ and $Q_\theta(Y_{t-1}) \equiv Q_\theta(y_t | Y_{t-1})$, we have $F_{y_t | Z_{t-1}} \{ Q_\theta(Z_{t-1}) | Z_{t-1} \} = \theta$ with probability one. The (non)causality in the θ-th quantile hypotheses to be tested are:

$$H_0: \quad P\Big\{ F_{y_t | Z_{t-1}} \big\{ Q_\theta(Y_{t-1}) | Z_{t-1} \big\} = \theta \Big\} = 1 \tag{1}$$

$$H_1: \quad P\Big\{ F_{y_t | Z_{t-1}} \big\{ Q_\theta(Y_{t-1}) | Z_{t-1} \big\} = \theta \Big\} < 1 \tag{2}$$

Jeong *et al.* (2012) show that the feasible kernel-based test statistics has the following format:

$$\hat{J}_T = \frac{1}{T(T-1)h^{2p}} \sum_{t=p+1}^{T} \sum_{s=p+1, s \neq t}^{T} K\left(\frac{Z_{t-1} - Z_{s-1}}{h} \right) \hat{\varepsilon}_t \hat{\varepsilon}_s \tag{3}$$

where $K(\bullet)$ is the kernel function with bandwidth h, T is the sample size, p is the lag order and $\hat{\varepsilon}_t = \mathbf{1}\{y_t \leq \hat{Q}_\theta(Y_{t-1})\} - \theta$ is the regression error, where $\hat{Q}_\theta(Y_{t-1})$ is an estimate of the θ-th conditional quantile and $\mathbf{1}\{\bullet\}$ is the indicator function. The *Nadarya-Watson* kernel estimator of $\hat{Q}_\theta(Y_{t-1})$ is given by:

$$\hat{Q}_\theta(Y_{t-1}) = \frac{\sum_{s=p+1, s \neq t}^{T} L\left(\frac{Y_{t-1} - Y_{s-1}}{h} \right) \mathbf{1}\{y_s \leq y_t\}}{\sum_{s=p+1, s \neq t}^{T} L\left(\frac{Y_{t-1} - Y_{s-1}}{h} \right)} \tag{4}$$

with $L(\bullet)$ denoting the kernel function.

Balcilar *et al.* (2018) extend the model of Jeong *et al.* (2012) framework, based on the work of Nishiyama *et al.* (2011), to the *second* (or higher) moment which allows us to test the

causality between the HSI and housing returns volatility. In this case, the null and
alternative hypotheses are given by:

$$H_0: \quad P\left\{F_{y_t^k|Z_{t-1}}\left\{Q_\theta(Y_{t-1})|Z_{t-1}\right\} = \theta\right\} = 1, \quad k = 1, 2, \ldots, \quad K \tag{5}$$

$$H_1: \quad P\left\{F_{y_t^k|Z_{t-1}}\left\{Q_\theta(Y_{t-1})|Z_{t-1}\right\} = \theta\right\} < 1, \quad k = 1, 2, \ldots, \quad K \tag{6}$$

The causality-in-variance test can then be calculated by replacing y_t in equations (3) and (4)
with y_t^2. As pointed out by Balcilar *et al.* (2018) a rescaled version of the $\hat{J}_T$ has the standard
normal distribution. The testing approach is sequential and failing to reject the test for $k = 1$
does not automatically lead to no causality in the *second* moment; one can still construct the
test for $k = 2$.

The empirical implementation of causality testing via quantiles entails specifying three
key parameters: the bandwidth (h), the lag order (p) and the kernel types for $K(\cdot)$ and $L(\cdot)$.
We use a lag order of one based on the Schwarz Information Criterion (SIC). We determine h
by the leave-one-out least-squares cross-validation. Finally, for $K(\cdot)$ and $L(\cdot)$, we use
Gaussian kernels.

3. Empirical findings

Before we discuss the findings from the causality-in-quantiles test, for the sake of
completeness and comparability we conduct the standard linear Granger causality test, with
a lag-length of 1, as determined by the SIC. The resulting $\chi^2(1)$ test statistic associated with
the causality running from HSI to HR is 61.5012 with a p-value of 0.0000, i.e. the null
hypothesis that housing search activity does not Granger cause housing returns, in line with
Møller *et al.* (2023), is strongly rejected. However, the linear framework is unable to provide
information on regime-specific, i.e. quantiles-based, predictability, besides being silent about
the causal influence on volatility, i.e. squared returns. Naturally, we turn to the k-th order
non-parametric causality-in-quantiles test next. But to econometrically motivate this
framework, we statistically examine the presence of nonlinearity and structural breaks in
the relationship between the HSI and HR. Nonlinearity and regime changes, if present,
would warrant the use of the non-parametric quantiles-in-causality approach, since this
data-driven test would formally address the issues of nonlinearity and structural breaks in
the relationship between the variables under investigation.

For this purpose, we first apply the Brock *et al.* (1996, BDS) test on the residual derived
from the HR equation involving one lag each of HR and HSI. Table A2 in the Appendix
presents the results of the BDS test of non-linearity. As the table shows, we find evidence, at
least at the 5% level of significance, for the rejection of the null hypothesis of *i.i.d.* residuals
at various embedded dimensions (m), which, in turn, is indicative of non-linearity in the
relationship between housing search activity and housing price returns. To further motivate
the causality-in-quantiles approach, we next use the powerful *UDmax* and *WDmax* tests of
Bai and Perron (2003), to detect 1 to M structural breaks in the relationship between HR and
HSI, allowing for heterogeneous error distributions across the breaks. When we apply these
tests to the HR equation involving one lag each of HR and HSI, we detect four breaks on:
2008:12, 2012:03, 2014:09 and 2017:03 associated with the downturns and lower search
activity in the housing market during the global financial and the European sovereign debt
reflecting weak economic conditions, but then sustained economic recovery and improved
HSI since 2014 (see, Figure A1).

Given the strong evidence of non-linearity and structural breaks in the relationship between HR and HSI, we now turn our attention to the causality-in-quantiles test, which is robust to misspecification in the linear model due to its non-parametric nature, besides allowing us to test for predictability over the entire conditional distributions of both returns and volatility. The results are reported in Figure 1, whereby we test the regime-specific null hypothesis of no – Granger causality running from HSI to HR and HR^2 over the quantile range of 0.10 to 0.90 based on the standard normal test statistic. As can be seen from the figure, predictability for housing returns from HSI holds over the range of 0.20 to 0.80 at least at the 5% level of significance, with the strongest causal influence observed at the median. Interestingly, there is no evidence of predictability at the extreme quantiles of 0.10 and 0.90. In other words, allowing for a quantiles-based model, we provide more nuanced evidence of predictability as detected by Møller et al. (2023) from a linear (conditional mean-based) predictive regression framework, as we are able to detect varied strength of causality conditional on the regimes of the market. Put alternatively, we can now say that the impact of HSI on HR increases as we move from a bearish regime to a bullish regime, with a peak at the median, but there is no evidence of causal influence at the two extreme ends of the market [8]. These findings tend to support the idea that for exceptionally weak and strong phases of the real estate market, i.e. at the quantiles of 0.10 and 0.90, participants tend to herd (Babalos et al., 2015; Ngene et al., 2017) and, hence, does not require information of a predictor like HSI to gauge the future path of HR. Note that, housing markets are also known to commove during booms and busts (Cotter et al., 2015) corresponding to the unconditional upper and lower quantiles of HR, so the role of a fundamental, like HSI, might become invalid here. Furthermore, the lack of predictability at the upper quantiles could also be
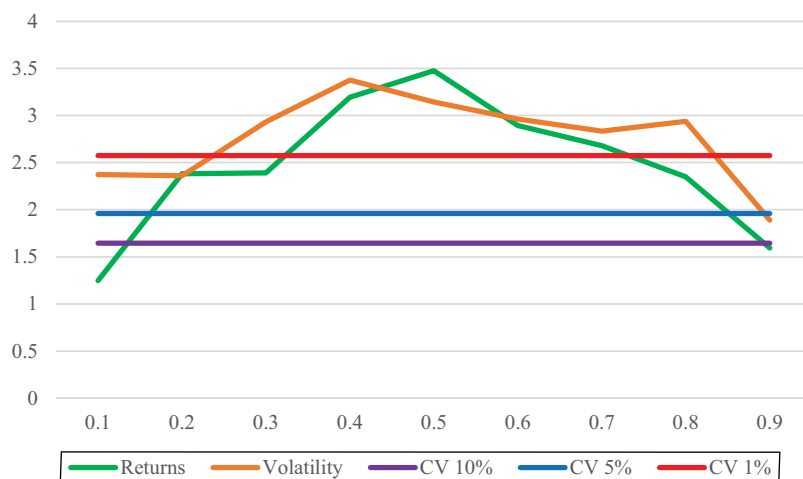
Figure 1.
k-th order causality-in-quantiles test results for housing price returns and volatility for the USA using monthly data: 2004:01–2021:01

Notes: Vertical axis reports the standard normal test statistic for the hypothesis that there is no Granger causality for a particular quantile on the horizontal axis running from housing search activity index (HSI) to housing price returns (HR) and squared returns (volatility; HR2); CV 10%, CV 5% and CV 1 % correspond to the critical values of 1.645, 1.96 and 2.575 respectively
Source: Figure created by authors

signalling market efficiency related to HSI in line with the quantiles-based test of efficiency of Tiwari *et al.* (2020) for the overall USA, as well as at the MSA-level.

Interestingly however, the predictability of HSI for squared returns, i.e. volatility is observed over its entire conditional distribution at least at the 5% level of significance for majority of the quantiles (barring the 90th quantile, where causality holds at the 10% level), with a peak at the quantile of 0.40. In other words, we provide strong evidence in favour of our hypothesis that housing search activity can lead to house price volatility over and above housing returns, with the effect holding irrespective of the size this price variability, unlike returns.

Although robust predictive inference is derived based on the non-parametric causality-in-quantiles test, it is also interesting to estimate the sign of the effect of the HSI on HR and HR^2 at various quantiles, especially to validate the theoretical positive relationship outlined in the introduction. But, in a non-parametric framework, this is not straightforward, as we need to use the first-order partial derivatives. Estimation of the partial derivatives for non-parametric models can give rise to complications, because non-parametric methods exhibit slow convergence rates, due to the dimensionality and smoothness of the underlying conditional expectation function. However, one can look at a statistic that summarizes the overall effect or the global curvature (i.e. the global sign and magnitude), but not the entire derivative curve. In this regard, a natural measure of the global curvature is the average derivative (AD) using the conditional pivotal quantile, based on approximation or the coupling approach of Belloni *et al.* (2019), which allows us to estimate the partial ADs of HR or HR^2 with respected to HSI. Based on these ADs reported in Table 1, we find consistent evidence of a positive predictive effect of HSI on housing price returns and its volatility.

When we rely on a GARCH-based metric of volatility [9], our findings, as reported in Figure 2, continue to be robust in the sense that predictability is again observed over the entire conditional distribution, with a peak at the median, in a quite strong manner at the 1% level of significance, except at the two ends where the same holds at the 10% level. Barring the highest quantile of 0.90, a similar result is also obtained for weekly squared returns, as seen from Figure 3. At the same time from Figure 3, it must be noted that causality of HSI on HR is restricted now over the quantile range of 0.30–0.70, i.e. compared to the monthly data, though sample periods are different, the lack of predictability at the two ends of the conditional distribution of house price returns gets extended.

At the regional-level, as can be seen from Table 2, 65 of the 77 MSAs considered, i.e. in 84.42% of the cases considered, there is evidence of predictability running from HSI to HR

Quantile	HR	HR^2
0.10	0.8934	0.0637
0.20	0.8667	0.1696
0.30	0.7906	0.1680
0.40	0.7151	0.2372
0.50	0.7775	0.2351
0.60	0.7213	0.2635
0.70	0.6685	0.2758
0.80	0.6361	0.2767
0.90	0.5778	0.2175

Notes: Entries correspond to average derivative (AD) estimates of the sign of the effect of HSI on to housing price returns (HR) and its volatility (HR^2) at a particular quantile
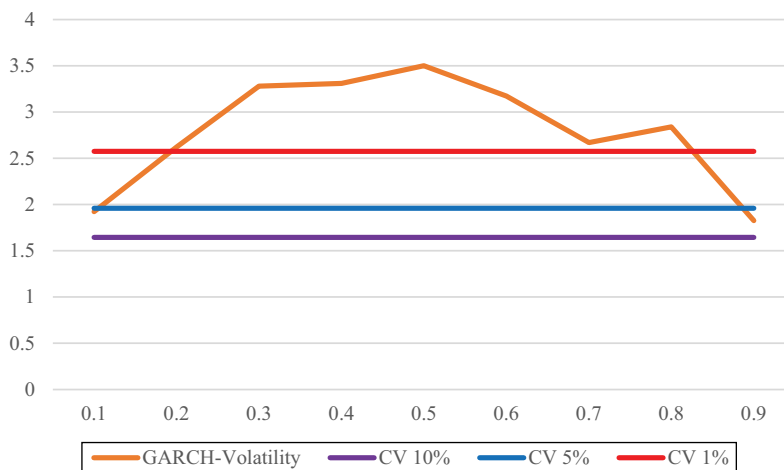**Source:** Table created by authors

Table 1.
Average derivative estimates for the effect of HSI on housing price returns and volatility for the USA using monthly data: 2004:01–2021:01

Notes: Vertical axis reports the standard normal test statistic for the hypothesis that there is no Granger causality for a particular quantile on the horizontal axis running from housing search activity index (HSI) to GARCH-based housing price returns volatility; CV 10%, CV 5% and CV 1 % correspond to the critical values of 1.645, 1.96 and 2.575 respectively
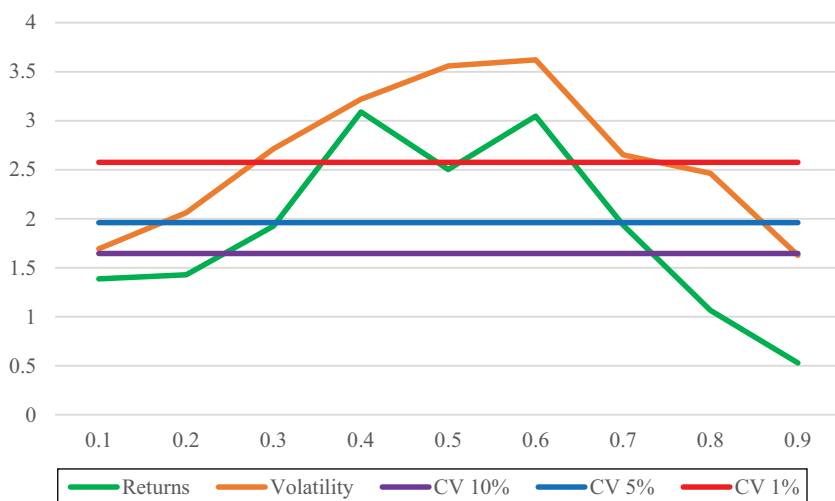**Source:** Figure created by authors

(for at least one quantile of the conditional distribution at the 1% to 10% level of significance). In line with the results for the overall USA, for these instances, predictability peaks at quantiles closer to the median and fades away at the extreme ends. Furthermore, as reported in Table 3, predictability for housing price returns volatility is detected for 68 of the 77 MSAs, i.e. in 88.31% of instances, again with an inverted u-shaped pattern of the test statistic registering its highest value close to the median. But in this case, just as for the aggregate USA, the coverage of causality over the conditional distribution of volatility is relatively higher compared to that of HR, in terms of the number of quantiles for which predictability is observed.

In sum, we tend to conclude that the predictability of HSI for house price volatility, unlike housing returns is, in general, not regime-specific and tends to be stronger in the sense of its coverage of the entire conditional distribution of the former, with these observations tending to hold both at the aggregate and MSA-level of the USA housing market. In other words, we provide for the first time, robust empirical validation of the theoretical proposition found in the housing search-theoretic models that housing search activity is likely to predict not only housing price returns but also its volatility.

4. Conclusions

In a recent study, Møller *et al.* (2023) developed a Google-based online search volume index of housing activity as a measure of underlying housing demand to show that the metric can predict housing price returns of the USA and its MSAs. Based on recent models of housing search theory, in particular of Ngai and Sheedy (2022), we can also postulate that this HSI should also be able to predict volatility in house prices. To test our

Notes: Vertical axis reports the standard normal test statistic for the hypothesis that
there is no Granger causality for a particular quantile on the horizontal axis running
from housing search activity index (HSI) to housing price returns (HR) and squared
returns (volatility; HR2); CV 10%, CV 5% and CV 1 % correspond to the critical
values of 1.645, 1.96 and 2.575 respectively
Source: Figure created by authors

hypothesis, we use the k-th order non-parametric causality-in-quantiles framework of
Balcilar *et al.* (2018), which is capable of capturing predictability of the entire conditional
distributions of both housing price returns and squared returns, i.e. volatility
simultaneously, by controlling for misspecification due to uncaptured non-linearity and
regime changes with the HSI, which we statistically show to exist in our data set over the
monthly period of 2004:01 to 2021:01.

We show that while housing search activity continues to predict aggregate USA house
price returns under the misspecified linear Granger causality model, as in Møller *et al.* (2023),
the same, in general, also holds true for the quantiles-causality framework, barring at the
extreme ends of the conditional distribution of returns. Our results thus provide a more
nuanced evidence of causality running from HSI to housing price returns, with an inverted
u-shape of the strength of the underlying standard normal test statistic of the k-th order non-
parametric test of causality-in-quantiles, which reached its peak at the median. In other
words, the strongest evidence of causality from HSI to HR is obtained around the
conditionally normal state of housing returns. Comparatively, volatility is found to
relatively strongly predicted over the entire quantile range considered of squared returns,
with the highest value test statistic again registered close to the conditional median, i.e.
normal volatility-regime. Our results tend to carry over to an alternative (the GARCH-based)
metric of volatility, as well as for higher-frequency, i.e. weekly data over January, 2018 to
March, 2021. When we take a regional perspective by delving into 77 MSAs, we find that the
predictive impact of HSI is detected for 65 and 67 of the cases for the first and second
moment of house prices, respectively. In other words, the causal influence of HSI is
dominant not only for the overall USA but also at the local-level, with strong evidence in

MSA	Quantiles								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Miami, FL	1.0315	1.4009	1.4803	1.1712	0.7886	0.8352	1.1120	1.0232	0.9610
Los Angeles–Long Beach	1.6172	2.4281**	3.1131***	2.6601***	2.9570***	2.4529**	2.5215**	1.8816*	1.7244*
San Francisco, CA	0.3846	0.5856	0.5890	0.7001	0.8839	0.6040	0.9766	0.8529	0.8039
San Diego, CA	0.0297	0.0125	0.0049	0.0011	0.0069	0.0026	0.0049	0.0056	0.0019
Salt Lake City–Ogden, UT	0.8457	1.1103	1.6508*	2.1129**	1.6261	1.4383	1.0168	0.7117	0.4247
New York, NY	0.5398	0.6002	0.9871	0.8503	1.0702	0.6045	0.7042	0.4508	0.1450
New Orleans, LA	0.2549	0.9730	0.9229	0.8069	0.4172	0.4403	0.4563	0.1943	0.0415
Chicago, IL	0.0783	0.0204	0.0424	0.0151	0.0150	0.0142	0.0296	0.0314	0.0470
West Palm Beach–Boca Raton, FL	0.7922	0.5847	0.5895	0.4146	0.5078	1.1368	1.1028	1.0023	0.1704
Boston–Worcester–Lawrence–Lowell–MA-NH	0.5223	0.7183	1.1918	1.2904	1.1493	1.5974	2.0438**	1.1216	0.6838
Seattle–Bellevue–Everett, WA	1.3265	2.1514**	1.3000	2.2435**	2.0830**	1.7651*	0.9728	0.9178	0.6126
Sarasota–Bradenton, FL	0.5239	0.7290	0.8896	1.3928	1.3280	0.9660	0.9136	0.8324	0.4598
Milwaukee–Waukesha, WI	0.9720	1.5021	1.2144	1.7936*	1.9767***	1.9644***	1.5705	1.1497	0.8825
Jacksonville, FL	1.0741	1.8793*	2.1737**	1.4697	1.4369	1.5129	1.4428	1.1978	1.0618
Portland–Vancouver, OR–WA	1.5964	1.5876	3.2543***	3.1052***	2.8820***	2.6871***	3.1567***	2.1055**	1.2170
Orlando, FL	0.2242	0.1142	0.1436	0.1765	0.2541	0.2259	0.1878	0.1155	0.0102
Charleston–North Charleston, SC	1.1628	1.4145	1.3444	1.5315	1.8511*	2.4513**	1.6873*	1.6276	0.7773
Pittsburgh, PA	0.0297	0.0125	0.0109	0.0011	0.0069	0.0026	0.0003	0.0003	0.0227
Baltimore, MD	0.8807	1.5390	1.7551*	1.1636	1.5181	1.7940*	1.7453*	1.4437	0.8052
Detroit, MI	1.3360	1.9523*	2.9844***	3.2718***	2.4783***	1.2232	0.8939	1.1375	0.5979
Las Vegas, NV–AZ	1.2482	2.4415**	1.8425*	2.4518**	1.8768*	1.1186	1.0021	0.8975	0.7407
Rochester, NY	1.2886	0.9997	1.5033	2.1694**	2.3091**	3.0241***	2.1271**	1.2446	0.8146
Tucson, AZ	1.3234	1.0723	2.1754**	2.1017**	1.8669*	2.0634**	1.5311	1.5100	0.9449
Knoxville, TN	0.9294	1.3284	1.1705	1.4031	1.8139*	2.0568**	2.2482**	2.2348**	1.1416
Minneapolis–St. Paul, MN–WI	0.8279	1.4193	1.1155	1.4068	1.7771*	2.5054***	2.6168***	2.8726***	1.9257*
Hartford, CT	0.9639	2.0299**	2.4521**	2.3619**	1.5036	1.1521	1.4457	1.1273	0.6974
Springfield, MA	1.7069*	2.1698***	1.8684*	2.0740**	2.1631**	1.2705	0.8792	1.0704	1.1143
Denver, CO	1.2080	2.4052**	3.1959***	2.6348***	1.7198*	1.5048	1.0974	1.3067	0.8487
Providence–Warwick–Pawtucket, RI	1.1515	1.4914	1.9584*	1.9702**	1.9663**	2.5125**	2.7372***	1.6985**	1.2235
Washington, DC–MD–VA–WV	1.1079	1.8825*	2.0805**	1.9628**	2.3268***	2.6011***	3.1993***	2.5861***	1.6609**
Phoenix–Mesa, AZ	1.0222	1.6476*	2.6507***	2.5610***	2.0493***	2.2587***	2.1026***	1.7802*	0.5833

| MSA | Quantiles | | | | | | | | |
---	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Scranton–Wilkes-Barre–Hazleton, PA	0.7570	1.1224	1.3310	2.3950**	2.4429**	2.9612***	2.3274**	2.4546***	1.6816*
Harrisburg–Lebanon–Carlisle, PA	0.9990	1.5114	1.7468*	1.9619**	2.8813***	1.8169*	1.7417*	1.6672**	0.9664
Bakersfield, CA	1.0639	0.8970	2.5892***	2.2546**	1.8009*	1.7610*	2.5116**	1.4811	0.7470
Philadelphia, PA–NJ	1.3665	1.8516*	1.7397*	2.2293**	3.0995***	2.2630**	1.4608	1.4072	0.7484
Colorado Springs, CO	1.2947	1.9661**	2.7020***	2.4998**	2.7360***	2.1308**	1.9941**	1.7832*	1.2118
Albany–Schenectady–Troy, NY	1.4288	1.8400*	1.6962*	1.8392*	2.0853**	2.9321***	2.2955***	1.9511**	1.0322
Baton Rouge, LA	1.5405	1.8065*	1.9067*	1.7699*	2.2405**	2.0414**	1.5779	1.5663	0.7470
Memphis, TN–AR–MS	1.9584*	2.8592***	3.8258***	3.9451***	3.4101***	2.1399**	1.9956**	2.2495***	1.6166
Buffalo–Niagara Falls, NY	1.7342**	2.8795***	3.0687***	3.4654***	2.7187***	1.8934*	2.0187**	1.8165*	1.2763
Fresno, CA	1.5792	1.5859	2.0167**	2.5058**	3.1755***	3.7592***	2.9497***	1.9863**	1.3931
Mobile, AL	1.5873	2.0574**	2.5170**	2.6588***	2.7147***	2.4968**	3.0013***	2.2005**	1.6771*
Stockton–Lodi, CA	1.2421	1.7719*	2.3989**	3.0228***	2.4810*	1.9310*	2.2088***	2.0231***	1.1430
Raleigh–Durham–Chapel Hill, NC	1.7304**	1.7121*	1.5564	1.9191*	1.9940**	1.7604*	2.0654***	2.4008***	1.7610*
Albuquerque, NM	1.1578	1.7769*	2.3672**	2.7646***	3.8512***	2.3503**	2.1771**	1.8785*	1.1242
Birmingham, AL	0.9247	1.7948*	1.9787**	2.3515**	2.5025**	3.3346***	2.5357**	2.2238**	1.3168
Dallas, TX	2.1536**	3.2999***	4.0244***	3.6804***	2.8397***	2.6609***	2.4549**	2.3625**	1.7889**
Syracuse, NY	1.4190	1.8243*	2.2039**	2.2401**	2.3091**	2.7562***	2.8785***	1.9178**	1.4464
Toledo, OH	1.2193	2.0333**	2.2469**	2.0911**	2.5175**	2.7884***	3.3375***	2.7302***	1.7709**
Nashville, TN	1.8309*	2.7301***	3.4952***	3.4010***	2.9874***	2.3399**	2.6290***	2.9454***	2.0077**
Houston, TX	1.4731	1.5189	2.0764**	2.2869**	2.6411***	2.7310***	2.9190***	2.4201***	1.2746
Louisville, Kentucky–in	1.7590*	2.0245**	2.2745**	2.1711**	1.9693***	1.6940*	1.4193	1.2770	1.1261
El Paso, TX	1.6071	1.6035	3.2674***	3.1187***	2.9011***	2.7018***	3.1655***	2.1168***	1.2268
St. Louis, MO–IL	1.6979**	2.1327**	2.0580**	2.3715**	3.1767***	2.5749**	1.6855**	1.6462**	0.9160
Grand Rapids–Muskegon–Holland, MI	1.3830	2.1189**	2.8493***	2.6240***	2.8924***	2.7741**	2.1277**	1.8781**	1.2852
Cincinnati, OH–KY–IN	1.4276	1.7081*	1.6916*	3.2130**	2.6618***	2.7388***	2.0568**	1.7322*	1.0176
Atlanta, GA	1.1597	2.0861**	2.2544**	1.9151*	2.1358**	2.5020**	2.5864***	2.0286**	1.2088
Akron, OH	1.3681	1.4483	1.5130	1.9876**	2.0251**	1.8975*	1.7117*	1.9233**	1.1918
Richmond–Petersburg, VA	1.1380	1.6732*	2.5305**	2.9522***	2.1733**	1.0307	0.7218	0.8730	0.3825
Youngstown–Warren, OH	1.0673	0.8120	1.2339	1.8985*	1.9496*	2.5551**	1.7199*	0.9278	0.5679
Columbia, SC	1.0899	0.8096	1.7444*	1.6330	1.4198	1.7113*	1.3049	1.2619	0.7193
Columbus, OH	0.5767	1.1656	1.4594	2.0776**	1.4285	1.4978	1.0568	0.9371	0.6818
Greenville–Spartanburg–Anderson, SC	1.3893	2.0681**	2.5085***	2.6938***	2.3514**	2.0502**	1.7485**	1.2987	0.7215
Little Rock–North Little Rock, AR	0.4724	0.6745	1.0939	0.7762	1.1838	0.6972	0.8240	0.8068	0.3496

(continued)

Table 2.

Table 2.

MSA	Quantiles								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
San Antonio, TX	0.5114	1.0929	1.2055	1.2786	1.6673*	2.7928***	1.8211*	1.5456	0.7899
Austin–San Marcos, TX	1.0708	1.5119	1.1402	1.2476	1.3066	0.5248	0.2587	0.4986	0.7598
Charlotte–Gastonia–Rock Hill, NC–SC	0.6165	0.8297	1.3283	1.4147	1.2774	1.7696*	2.1862**	1.2142	0.7585
Greensboro–Winston–Salem–High Point, NC	0.4463	1.0402	1.1558	1.3421	0.9448	1.3012	2.0395**	1.7061*	1.0543
Kansas City, MO–KS	1.7648*	2.5646**	1.9925**	2.8205***	2.7256***	1.9966**	1.3909	1.4963	1.0217
Oklahoma City, OK	2.4923**	3.1768***	3.9544***	4.4959***	4.2997***	4.0092***	3.6874***	3.0245***	2.1822**
Tulsa, OK	1.2549	1.8172*	2.1911**	2.2828**	3.1208***	2.0177**	2.0014**	2.0320**	1.1488
Omaha, NE–IA	1.0514	1.8565*	2.1434**	1.4386	1.3987	1.4674	1.4104	1.1597	1.0404
McAllen–Edinburg–Mission, TX	1.5781	1.5605	3.2318***	3.0817***	2.8491***	2.6618***	3.1415***	2.0862**	1.2003
Dayton–Springfield, OH	0.7699	0.9734	1.5553	1.6369	1.8718*	1.3761	1.3163	1.2357	0.7780
Indianapolis, IN	1.2560	1.5460	1.4584	1.6395	1.9304*	2.6180***	1.8842**	1.7453*	0.8653
Fort Wayne, IN	0.4695	0.7598	1.0543	1.1525	0.9726	1.7609*	1.5190	0.9813	0.7564
Wichita, KS	0.7452	0.7934	0.9622	1.5843	1.6864*	1.4958	1.0378	1.4877	0.8458

Notes: ***, ** and ***indicate rejection of the null hypothesis of no Granger causality running from the housing search activity index (HSI) to housing price returns (HR) for a particular quantile at 1, 5 and 10% levels of significance (i.e. critical values of 2.575, 1.96 and 1.645 of the standard normal test statistic), respectively
Source: Table created by authors

MSA					Quantiles				
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Miami, FL	0.8703	0.8950	1.3486	1.8427*	2.2876**	2.2929**	2.1602**	1.9359*	1.2383
Los Angeles–Long Beach, CA	0.7213	1.2580	1.1667	1.2354	1.2994	0.8677	0.8762	1.0332	0.3263
San Francisco, CA	0.7275	1.2885	1.5128	1.5848	1.8735*	2.4123**	2.8068***	1.8178*	1.3707
San Diego, CA	0.0747	0.3316	0.4756	0.5282	0.2470	0.4307	0.3147	0.1035	0.1510
Salt Lake City–Ogden, UT	0.5626	0.7968	0.8965	0.6792	1.8796*	1.4745	1.7500*	2.0929**	0.7326
New York, NY	1.9700**	2.5368**	3.0200***	2.8922***	2.8510***	2.6424***	2.5665***	2.8133***	2.0357**
New Orleans, LA	2.0347**	2.8877***	3.1719***	3.7951***	3.4485***	3.3966***	3.1575***	2.7593***	1.8943*
Chicago, IL	2.5074**	3.0688***	3.2208***	3.0966***	3.3560***	3.3225***	3.3806***	2.5228**	1.8023*
West Palm Beach–Boca Raton, FL	2.3858**	3.5280***	3.6042***	3.2835***	3.2161***	3.1981***	3.0464***	2.9806***	1.9629**
Boston–Worcester–Lawrence–Lowell–MA-NH	2.3321**	3.6856***	3.8455***	3.7483***	3.7849***	3.9532***	4.0359***	3.4730***	2.4074**
Seattle–Bellevue–Everett, WA	1.7174*	3.2167***	3.4201***	4.0392***	3.5926***	3.3423***	3.0351***	2.9114***	1.3935
Sarasota–Bradenton, FL	0.0042	0.0125	0.0049	0.0026	0.0069	0.0011	0.0003	0.0056	0.0019
Milwaukee–Waukesha, WI	0.4985	0.5995	0.5232	1.2423	1.2790	1.3837	1.5230	0.7894	0.3720
Jacksonville, FL	1.2901	3.1035***	3.2118***	2.8427***	2.5329**	2.2116**	1.4379	0.7276	0.5150
Portland–Vancouver, OR–WA	0.4586	0.6138	0.9899	1.3995	1.2190	1.2726	2.0073**	0.8690	0.3902
Orlando, FL	2.4296***	3.9410***	3.4785***	3.6268***	3.7040***	3.4718***	3.0547***	2.9759***	1.8225*
Charleston–North Charleston, SC	0.3375	0.3577	0.7429	0.7065	1.2109	1.2095	0.7105	1.0772	0.5840
Pittsburgh, PA	2.5278***	3.0941***	3.3136***	3.6568***	4.0745***	3.9911***	3.3000***	2.2853***	1.7135*
Baltimore, MD	1.9005*	3.0557***	4.1613***	4.3945***	4.3343***	4.7444***	5.2918***	4.2823***	2.2293**
Detroit, MI	0.9201	0.9385	1.4495	1.9332*	2.3648**	2.3560**	2.2123**	1.9660**	1.2604
Las Vegas, NV–AZ	1.4486	1.6593*	1.4347	1.6429	2.2149**	2.0583**	1.7868*	1.5392	0.8722
Rochester, NY	0.6384	1.1238	1.3760	1.4089	1.7111*	2.2279**	2.5964***	1.5997	1.1879
Tucson, AZ	0.9062	1.9164*	2.5067**	2.1125**	2.1777**	2.0804**	1.4190	1.2101	1.2126
Knoxville, TN	0.9554	1.8263*	2.0436**	1.7790*	2.4685**	2.2650**	2.3140**	2.5703**	1.2816
Minneapolis–St. Paul, MN–WI	1.3702	1.8024*	2.2160**	1.8775*	1.8316*	1.6450*	1.8495*	2.2751**	1.6257
Hartford, CT	1.7256*	2.5832***	3.0645***	2.6259***	1.7448*	1.4598	1.8168*	1.5229	0.8309
Springfield, MA	0.7719	1.2192	1.2653	1.3262	1.4975	1.6951*	2.4270**	1.6050	0.6698
Denver, CO	2.0919**	2.5273**	2.6077***	2.9588***	3.0687***	3.2438***	2.7256***	2.1361**	1.7162*
Providence–Warwick–Pawtucket, RI	1.1044	2.3733**	2.0048**	1.8707*	2.2137**	2.0510**	1.7915*	2.0273**	1.3533
Washington, DC–MD–VA–WV	1.2803	2.2880**	2.5206**	2.1444**	2.0815**	2.4187**	2.6160***	2.3878**	1.7671*

(continued)

Table 3.

MSA				Quantiles					
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Phoenix–Mesa, AZ	1.3113	3.0041***	2.8881***	3.5670****	3.0388****	2.1941**	2.0342**	1.9509*	0.7595
Scranton–Wilkes-Barre–Hazleton, PA	1.0561	1.2866	1.1985	1.8976*	2.0250**	2.4707**	2.6621***	1.7903**	0.9243
Harrisburg–Lebanon–Carlisle, PA	1.2589	3.0361****	3.1269****	2.7738***	2.4950**	2.1588**	1.3991	0.6818	0.5005
Bakersfield, CA	1.1184	1.8107*	2.1036**	2.6435***	2.6641***	2.0956**	2.6565****	1.6240	1.1635
Philadelphia, PA–NJ	1.3496	1.9497*	2.1062**	2.3734**	2.5730**	2.9948***	3.7178****	2.5190***	1.7351*
Colorado Springs, CO	1.0634	1.8171*	1.9610**	2.5025**	2.7436***	2.4228**	2.2589**	2.5706***	1.2080
Albany–Schenectady–Troy, NY	0.9097	1.5745	3.0043****	3.9817****	3.7736****	3.0433***	3.3014****	1.5944	0.8832
Baton Rouge, LA	0.9090	2.0467**	2.8309****	2.6395****	2.6998****	3.1491****	3.5969****	2.7533****	1.3631
Memphis, TN–AR–MS	1.7452*	2.9557****	3.0763****	3.8195****	3.3128****	2.8290****	2.7686****	2.7780****	1.9302**
Buffalo–Niagara Falls, NY	2.5312**	2.8628****	3.1145****	3.3351****	3.7601****	3.5848****	3.6605****	2.9605****	2.2033**
Fresno, CA	2.1116**	2.5803****	3.0367****	2.8128***	3.1856****	2.9397****	2.9705****	2.2250**	1.9952**
Mobile, AL	2.0718**	2.9141****	2.9610****	3.0311****	3.0133****	3.1211****	2.6828***	2.7084****	1.9030*
Stockton–Lodi, CA	2.3866**	3.5564****	3.9375****	3.8241****	4.1971****	4.1110****	3.8094****	3.2976****	2.3427**
Raleigh–Durham–Chapel Hill, NC	2.5448**	3.3876****	4.1297****	4.0599****	4.0583****	3.9497****	3.5219****	3.2509****	2.4664**
Albuquerque, NM	2.6521****	3.6986****	3.8698****	4.4678****	4.3920****	4.2729****	3.9686****	3.4709****	2.5719**
Birmingham, AL	2.3044**	3.0809****	3.3656****	3.4531****	3.5504****	3.2768****	2.9780****	2.8053****	2.1962**
Dallas, TX	2.9449****	4.0323****	4.6233****	4.8084****	4.5618****	4.5782****	4.2095****	3.6711****	2.9286***
Syracuse, NY	3.0918****	4.0467****	4.6168****	4.8342****	4.3974****	3.7560****	3.6350****	3.1527****	2.5518**
Toledo, OH	2.3007**	3.6586****	3.8203****	3.7104****	3.7419****	3.9179****	4.0102****	3.4479****	2.3914**
Nashville, TN	2.2370**	3.0924****	3.2461****	3.8869****	3.7642****	3.5127****	3.2286****	3.3857****	2.1694**
Houston, TX	2.7515****	3.9438****	4.4109****	4.2427****	4.3691****	4.2894****	3.9229****	3.5593****	2.5118**
Louisville, KY–IN	2.5604**	3.5874****	4.3698****	3.6192****	4.0795****	3.9993****	3.6115****	3.1413****	2.0755**
El Paso, TX	2.2442**	3.0897****	3.0920****	3.3278****	3.1480***	3.1096****	2.9105****	2.9025****	2.1632**
St. Louis, MO–IL	2.2421**	2.9522****	3.3401****	3.2736****	3.5053****	3.5653****	3.8323****	3.2400****	2.3885**
Grand Rapids–Muskegon–Holland, MI	2.5279**	4.0167****	3.6860****	3.9030****	3.9668****	3.7672****	3.3539****	3.2432****	2.0050**
Cincinnati, OH–KY–IN	2.2388**	2.9998****	3.4880****	3.7923****	4.1827****	4.3543****	3.7351****	3.4423****	2.5692**
Atlanta, GA	2.5835****	3.1895****	3.4345****	3.7882****	4.1923****	4.1291****	3.4081****	2.4040**	1.8084*
Akron, OH	2.2533***	3.4198****	3.9638****	3.8596****	3.6327****	3.7273****	3.7135****	3.2775****	2.3498**
Richmond–Petersburg, VA	0.7017	0.7832	1.0523	1.5680	2.0568**	2.1281**	1.9971**	1.7976*	1.1402
Youngstown–Warren, OH	1.9397*	2.8534****	3.1442****	2.9138***	2.8365****	2.5522**	2.6584***	2.5106***	1.7579*
Columbia, SC	0.2409	0.8537	1.2956	0.9853	0.6724	0.8433	0.5656	0.3082	0.3845
Columbus, OH	0.5614	0.7676	1.3391	1.1282	1.6215	1.4030	1.5389	1.2441	1.1088

(continued)

MSA	Quantiles								
---	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Greenville–Spartanburg–Anderson, SC	0.7214	1.2265	1.6382	1.6722*	1.2023	0.7185	0.8307	1.4257	0.5427
Little Rock–North Little Rock, AR	0.5026	0.6365	0.8494	1.6220	1.3867	0.8672	1.1021	0.8013	0.4811
San Antonio, TX	0.9044	0.8797	0.6511	0.7969	0.7757	0.6941	1.2459	0.7842	0.3383
Austin–San Marcos, TX	0.7190	0.9101	1.2932	1.1730	1.7133*	2.4632**	2.1453**	1.3625	0.8940
Charlotte–Gastonia–Rock Hill, NC–SC	1.5769	2.4250**	1.6712*	1.6665*	2.0390**	2.0765**	0.9206	0.8035	0.8019
Greensboro–Winston–Salem–High Point, NC	1.8372*	0.8819	0.6037	0.7192	0.7967	0.8120	0.8495	1.0102	0.4916
Kansas City, MO–KS	1.6603*	3.1778***	3.3138***	3.9360***	3.3434***	3.1102***	2.8301***	2.7626***	1.2862
Oklahoma City, OK	0.7622	0.8718	0.8250	1.5776	1.7039*	2.0880**	2.2510**	1.3679	0.6632
Tulsa, OK	0.7513	0.9374	1.3036	1.3077	1.7797*	2.0324**	1.7535*	1.1327	0.4420
Omaha, NE–IA	1.2067	2.9158***	2.9786***	2.6564***	2.4341**	2.0698**	1.3378	0.6098	0.4794
McAllen–Edinburg–Mission, TX	0.6717	1.2159	1.7774*	0.7030	1.0303	0.9283	0.7171	0.6253	0.3363
Dayton–Springfield, OH	1.0504	1.7978*	1.9333*	2.4781**	2.7191***	2.4009**	2.2392**	2.5539**	1.1952
Indianapolis, IN	1.2693	1.5350	1.5973	2.2893**	2.3702**	2.0141**	1.4383	1.1400	0.7896
Fort Wayne, IN	0.5610	1.0066	1.5165	1.9027*	1.7622*	1.5717	0.8940	0.8619	0.6646
Wichita, KS	0.3596	1.0858	0.7555	1.5275	1.3643	0.9991	0.8174	0.7549	0.4577

Notes: ***, ** and *indicate rejection of the null hypothesis of no Granger causality running from the housing price returns (HR^2) for a particular quantile at 1, 5 and 10% levels of significance (i.e. critical values of 2.575, 1.96 and 1.645 of the standard normal test statistic), respectively to squared housing search activity index (HSI)
Source: Table created by authors

Table 3.

favour of our hypothesis that housing search activity tends to predict housing returns volatility, over and above returns.

In sum, using a novel methodology involving the k-th order non-parametric causality-in-quantiles framework, we are able to test the predictability of the entire conditional distributions of both housing price returns and squared returns, i.e. volatility simultaneously by controlling for misspecification due to uncaptured non-linearity and regime changes with the HSI – both of which we show to exist in our data set via formal statistical tests. In the process, we are able to provide robust evidence of causality running from HSI to not only returns but also volatility at various states or regimes of the two moments of housing prices, which has indeed implications for various market agents. Overall, our paper provides an empirical test of the theoretical proposition that housing search should produce housing market volatility for the first time.

As our predictive analysis is performed at the monthly as well as weekly frequencies associated with housing returns, our results can be used by policymakers to obtain high-frequency information about where the housing market is headed due to changes in housing search activity and predict the future path of low-frequency, i.e. quarterly, economic activity variables, such as growth of gross domestic product, at monthly and weekly-levels, given that house price movements are known to lead US business cycles. At the same time, since higher search activities result in higher future housing returns and volatility, the monetary authorities might need to undertake contractionary monetary policies to ensure the inflation target is met, as higher housing returns (and volatility depicting more trading) is likely to be associated with increased aggregate demand via the wealth-effect channel. Moreover, monthly and weekly predictions of housing returns and volatility contingent on online housing search activity, capturing latent demand, would also help investors to make optimal portfolio allocation decisions in a timely manner, as clearly with an increase in the HSI, both returns and risk in investing the housing market increases, producing a risk-return trade-off, which, in turn, needs to be compared with other asset market the agent might be investing in at the same point in time. Finally, from the perspective of a researcher, our results suggest that the housing market is, in fact, inefficient in the semi-strong sense [10], given the predictive role of search activity, but this result is also contingent on the phase of the housing returns, which excludes bearish- and bullish-regimes. This finding also implies that theoretical models relating housing price movements with search activities, should involve consideration of the underlying initial regime the market is in. In other words, our results have important implications for policy authorities, investors and academics.

Since in-sample predictability does not necessarily translate into out-of-sample gains, as part of future research, it would be interesting to extend our analysis to a full-fledged forecasting exercise using the k-th order non-parametric causality-in-quantiles test, as outlined in Bonaccolto *et al*. (2018). One could also go beyond a bivariate set-up when forecasting by incorporating a large array of other predictors when forecasting housing returns volatility in particular in a machine learning set-up, and hence compare the relative importance of HSI with the other control variables over the forecasting sample.

Notes

1. A recent report by the National Association of Realtors (NAR, 2023) shows that home buyers use the internet as their main source of information about the housing market, with as many as 96% of home buyers using the internet to search for a home.

2. www.federalreserve.gov/releases/z1/20230908/z1.pdf

3. The data is available for download from the research segment of the website of Professor Christian Montes Schütte at: https://sites.google.com/view/christian-montes-schutte/research?authuser=0.

4. Before extracting the first principal component, the indexes are used in their logarithms, a sequential testing strategy is used to account for the possibility that the individual Google Trends series could follow different trends, and seasonality is removed by regressing each series on monthly dummy variables to study the residuals from this regression.

5. While search activity for individuals residing in a given MSA counts in the overall search volume for that particular MSA, some individuals may also be interested in buying a home in one of the neighbouring MSAs. To allow for such potential moves across MSA borders, Møller *et al.* (2023) also include search activity in the state in which the MSA is located.

6. The data is available for download at: www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index.aspx

7. The data can be accessed at: www.zillow.com/research/data/

8. From the perspective of unconditional HR, it is likely that the result could be reflecting noisy data, especially since lowest and highest returns correspond to the peak of the global financial crisis and the outbreak of the COVID-19 pandemic, respectively (see, Figure A1).

9. Complete details of the parameter estimates of the GARCH model are available upon request from the authors.

10. Note that, efficiency in asset markets are tested in the context of univariate models to highlight the random-walk nature of asset prices (i.e., current price movements explained only by its lagged price) via the non-rejection of the unit root hypothesis can be considered a weak test, while predictability of asset returns involving one and multiple predictors can be termed as semi-strong and strong, respectively.

References

Apergis, N., Simo-Kengne, B., D., Gupta, R. and Chang, T. (2015), "The dynamic relationship between house prices and output: evidence from US metropolitan areas", *International Journal of Strategic Property Management*, Vol. 19 No. 4, pp. 336-345.

Babalos, V., Balcilar, M. and Gupta, R. (2015), "Herding behavior in real estate markets: Novel evidence from a markov-switching model", *Journal of Behavioral and Experimental Finance*, Vol. 8 No. C, pp. 40-43.

Bai, J. and Perron, P. (2003), "Computation and analysis of multiple structural change models", *Journal of Applied Econometrics*, Vol. 18 No. 1, pp. 1-22.

Balcilar, M., Bouri, E., Gupta, R. and Kyei, C.K. (2021), "High-Frequency predictability of housing market movements of the United States: the role of economic sentiment", *Journal of Behavioral Finance*, Vol. 22 No. 4, pp. 490-498.

Balcilar, M., Bouri, E., Gupta, R. and Wohar, M.E. (2020), "Mortgage default risks and High-Frequency predictability of the U.S. Housing market: a reconsideration", *Journal of Real Estate Portfolio Management*, Vol. 26 No. 2, pp. 111-117.

Balcilar, M., Gupta, R., Nguyen, D.K. and Wohar, M.E. (2018), "Causal effects of the United States and Japan on Pacific-Rim stock markets: nonparametric quantile causality approach", *Applied Economics*, Vol. 50 No. 53, pp. 5712-5727.

Balcilar, M., Gupta, R. and Miller, S.M. (2014), "Housing and the great depression", *Applied Economics*, Vol. 46 No. 24, pp. 2966-2981.

Bańbura, M., Giannone, D. and Reichlin, L. (2011), "Nowcasting", In Clements, M.P. and Hendry, D.F. (Eds) *Oxford Handbook on Economic Forecasting*, Oxford University Press, Oxford, pp. 63-90.

Belloni, A., Chernozhukov, V., Chetverikov, D. and Fernandez-Val, I. (2019), "Conditional quantile processes based on series or many regressors", *Journal of Econometrics*, Vol. 213 No. 1, pp. 4-29.

Berkovec, J.A. and Goodman, J.L. (1996), "Turnover as a measure of demand for existing homes", *Real Estate Economics*, Vol. 24 No. 4, pp. 421-440.

Bollerslev, T. (1986), "Generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics*, Vol. 31 No. 3, pp. 307-327.

Bollerslev, T., Patton, A. and Wang, W. (2016), "Daily house price index: construction modelling and longer-run predictions", *Journal of Applied Econometrics*, Vol. 31 No. 6, pp. 1005-1025.

Bonaccolto, G., Caporin, M. and Gupta, R. (2018), "The dynamic impact of uncertainty in causing and forecasting the distribution of oil returns and risk", *Physica A: Statistical Mechanics and Its Applications*, Vol. 507 No. C, pp. 446-469.

Bork, L. and Møller, S.V. (2015), "Forecasting house prices in the 50 states using dynamic model averaging and dynamic model selection", *International Journal of Forecasting*, Vol. 31 No. 1, pp. 63-78, doi: 10.1016/j.ijforecast.2014.05.005.

Bork, L., Møller, S.V. and Pedersen, T.Q. (2020), "A new index of housing sentiment", *Management Science*, Vol. 66 No. 4, pp. 1563-1583, doi: 10.1287/mnsc.2018.3258.

Bouri, E., Gupta, R., Kyei, C.K. and Shivambu, R. (2021), "Uncertainty and daily predictability of housing returns and volatility of the United States: Evidence from a higher-order nonparametric causality-in-quantiles test", *The Quarterly Review of Economics and Finance*, Vol. 82 No. C, pp. 200-206.

Brock, W., Dechert, D., Scheinkman, J. and LeBaron, B. (1996), "A test for independence based on the correlation dimension", *Econometric Reviews*, Vol. 15 No. 3, pp. 197-235.

Carrillo, P.E., de Wit, E.R. and Larson, W. (2015), "Can tightness in the housing market help predict subsequent home price appreciation? Evidence from the United States and the Netherlands", *Real Estate Economics*, Vol. 43 No. 3, pp. 609-651.

Cotter, J., Gabriel, S. and Roll, R. (2015), "Can housing risk be diversified? A cautionary tale from the housing boom and bust", *Review of Financial Studies*, Vol. 28 No. 3, pp. 913-936.

Díaz, A. and Jerez, B. (2013), "House prices, sales, and time on the market: a search-theoretic framework", *International Economic Review*, Vol. 54 No. 3, pp. 837-872.

Emirmahmutoglu, F., Balcilar, M., Apergis, N., Simo-Kengne, B.D., Chang, T. and Gupta, R. (2016), "Causal relationship between asset prices and output in the US: evidence from State-Level panel granger causality test", *Regional Studies*, Vol. 50 No. 10, pp. 1728-1741.

Genesove, D. and Han, L. (2012), "Search and matching in the housing market", *Journal of Urban Economics*, Vol. 72 No. 1, pp. 31-45.

Gupta, R., Marfatia, H.A., Pierdzioch, C. and Salisu, A.A. (2022), "Machine learning predictions of housing market synchronization across US states: the role of uncertainty", *The Journal of Real Estate Finance and Economics*, Vol. 64 No. 4, pp. 523-545.

Gupta, R., Ma, J., Theodoridis, K. and Wohar, M.E. (2023), "Is there a national housing market bubble brewing in the United States?", *Macroeconomic Dynamics*, Vol. 27 No. 8, pp. 2191-2228.

Jeong, K., Härdle, W.K. and Song, S. (2012), "A consistent nonparametric test for causality in quantile", *Econometric Theory*, Vol. 28 No. 4, pp. 861-887.

Møller, S.V., Pedersen, T.Q., Erik Christian Montes Schütte, E.C., S. and Timmermann, A. (2023), "Search and predictability of prices in the housing market", *Management Science*, Vol. 70 No. 1, doi: 10.1287/mnsc.2023.4672.

National Association of Realtors (NAR) (2023), "Home buyers and sellers generational trends report", Technical Report, National Association of Realtors, available at: www.nar.realtor/sites/default/files/documents/2023-home-buyers-and-sellers-generational-trends-report-03-28-2023.pdf.

Ngai, L.R. and Sheedy, K.D. (2020), "The decision to move house and aggregate housing-market dynamics", *Journal of the European Economic Association*, Vol. 18 No. 5, pp. 2487-2531.

Ngai, L.R. and Sheedy, K.D. (2022), "The ins and outs of selling houses: Understanding Housing-Market volatility. Centre for macroeconomics (CFM)", London School of Economics and Political Science, Discussion Papers No. 2213.

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*, Vol. 54 No. 4, pp. 482-514.

Nishiyama, Y., Hitomi, K., Kawasaki, Y. and Jeong, K. (2011), "A consistent nonparametric test for nonlinear causality - Specification in time series regression", *Journal of Econometrics*, Vol. 165 No. 1, pp. 112-127.

Nyakabawo, W.V., Gupta, R. and Marfatia, H.A. (2018), "High frequency impact of monetary policy and macroeconomic surprises on Us MSAS, aggregate Us housing returns and asymmetric volatility", *Advances in Decision Sciences*, Vol. 22 No. 1, pp. 204-229.

Nyakabawo, W.V., Miller, S.M., Balcilar, M., Das, S. and Gupta, R. (2015), "Temporal causality between house prices and output in the U.S.: a bootstrap rolling-window approach", *North American Journal of Economics and Finance*, Vol. 33 No. 1, pp. 55-73.

Payne, J.E. and Sun, X. (2023), "Time-varying connectedness of metropolitan housing markets", *Real Estate Economics*, Vol. 51 No. 2, pp. 470-502.

Segnon, M., Gupta, R., Lesame, K. and Wohar, M.E. (2021), "High-Frequency volatility forecasting of US housing markets", *The Journal of Real Estate Finance and Economics*, Vol. 62 No. 2, pp. 283-317.

Smith, E. (2020), "High and low activity spells in housing markets", *Review of Economic Dynamics*, Vol. 36 No. C, pp. 1-28.

Tiwari, A.K., Gupta, R. and Wohar, M.E. (2020), "Is the housing market in the United States really Weakly-Efficient?", *Applied Economics Letters*, Vol. 27 No. 14, pp. 1124-1134.
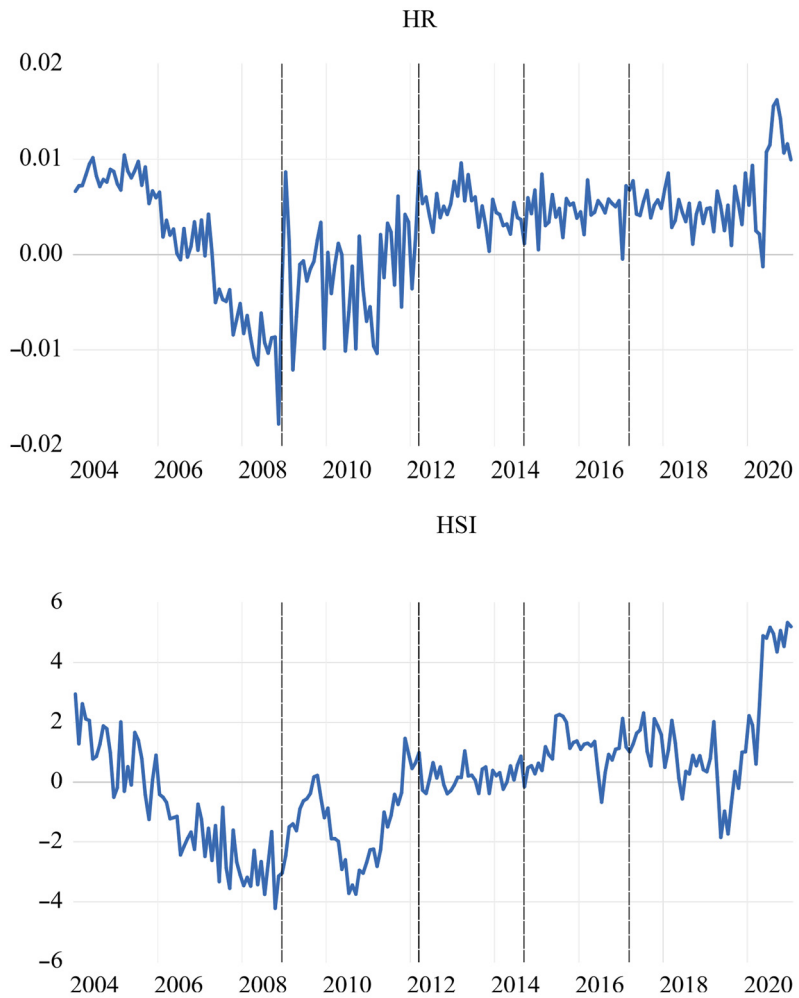
HR

HSI

Figure A1.
Data plots of housing
price returns (HR)
and housing search
activity index (HSI)

Notes: Dashed lines correspond to the four structural breaks detected on: 2008:
12, 2012:03, 2014:09 and 2017:03
Source: Figure created by authors

	Variable	
Statistic	Housing price returns (HR)	Housing search activity index (HSI)
Mean	0.0027	0.0000
Median	0.0041	0.2032
Maximum	0.0162	5.3304
Minimum	−0.0178	−4.2176
SD	0.0057	1.9094
Skewness	−0.8805	0.2334
Kurtosis	3.7467	3.4026
Jarque–Bera	31.2536***	3.2452
Observations	205	

Notes: SD = stands for standard deviation; The null hypotheses of the Jarque–Bera test correspond to the null of normality; ***indicates rejection of the null hypothesis at the 1% level of significance
Source: Table created by authors

	Dimension (m)				
Independent variable	2	3	4	5	6
HSI	2.0535**	4.1866***	4.5429***	5.1949***	5.6212***

Notes: Entries correspond to the z-statistic of the BDS test with the null of *i.i.d.* residuals, with the test applied to the residuals recovered from the housing returns equation with one lag each of housing price returns and housing search activity index (HSI); ** and ***indicates rejection of the null hypothesis at 5% and 1% levels of significance, respectively
Source: Table created by authors

Corresponding author
Damien Moodley can be contacted at: u18037667@tuks.co.za