WORKING PAPER

13/2020

Modelling Returns in US Housing Prices – You're the One for Me, Fat Tails

Tamás Kiss, Hoang Nguyen and Pär Österholm

Economics/Statistics

ISSN 1403-0586

https://www.oru.se/institutioner/handelshogskolan/forskning/working-papers/
Örebro University School of Business
701 82 Örebro
SWEDEN

Modelling Returns in US Housing Prices – You're the One for Me, Fat Tails*

Tamás Kiss[#]
Division of Economics, School of Business, Örebro University

Hoang Nguyen[∀]
Division of Statistics, School of Business, Örebro University

Pär Österholm[◊]
Division of Economics, School of Business, Örebro University
National Institute of Economic Research

Abstract

In this paper, we analyse the heavy-tailed behaviour in the dynamics of housing-price returns in the United States. We investigate the sources of heavy tails by estimating autoregressive models in which innovations can be subject to GARCH effects and/or non-Gaussianity. Using monthly data ranging from January 1954 to September 2019, the properties of the models are assessed both within- and out-of-sample. We find strong evidence in favour of modelling both GARCH effects and non-Gaussianity. Accounting for these properties improves within-sample performance as well as point and density forecasts.

JEL Classification: C22, C52, E44, E47, G17

Keywords: Non-Gaussianity, GARCH, Density forecasts, Probability integral transform

---

* The authors gratefully acknowledge financial support from Jan Wallanders och Tom Hedelius stiftelse (grants number Bv18-0018, P18-0201 and W19-0021).

[#] Corresponding author. Örebro University, School of Business, 701 82 Örebro, Sweden
e-mail: tamas.kiss@oru.se

[∀] Örebro University, School of Business, 701 82 Örebro, Sweden
e-mail: hoang.nguyen@oru.se

[◊] Örebro University, School of Business, 701 82 Örebro, Sweden
e-mail: par.osterholm@oru.se

1. Introduction

In less than fifteen years, the world has now experienced both a global financial crisis and a virus pandemic – both of which have had dire economic consequences. These events have confirmed the fact that large swings in economic variables happen more frequently than what is implied by models based on the traditional assumption of normally distributed disturbances. Put differently, the distributions of many variables are characterized by fat tails (Fagiolo *et al.*, 2008; Ascari *et al.*, 2015), that is, they have a higher probability mass in the tails of the distribution than a normal distribution does.

The purpose of this paper is to study the topic of fat tails related to a key US variable, namely housing-price returns. Not only is this an important financial variable, it is also highly relevant from a macroeconomic perspective since housing prices have a close relation with the real economy (Catte *et al.*, 2004; Iacoviello, 2005; Campbell and Cocco, 2007; Iacoviello and Neri, 2010). There are a number of studies that document non-Gaussian behaviour in real-estate price changes; see, for example, Myer and Webb (1994), Young and Graff (1995), Graff *et al.* (1997), Maurer *et al.* (2004), Young *et al.* (2006), Pontines (2010) and Chiang *et al.*, (2012). However, these studies focus mainly on the unconditional distribution of the returns, which is important from a portfolio risk-management perspective.

In contrast, our analysis focuses on the dynamic properties of the process. We primarily aim to assess what the appropriate distributional features of the innovations are if we want to model the dynamics of the returns. This issue is more relevant from a macroeconomic perspective where we often want to model dynamic interactions; see, for example, Ahamada and Diaz Sanchez (2013). To evaluate this issue, we conduct within-sample analysis on five different models and validate our findings through an out-of-sample forecast exercise.

Our analysis is performed employing monthly data on US housing-price returns from January 1954 to September 2019. Looking at the data, we confirm that the return series is characterised by fat tails and accordingly investigate its sources. Relying on univariate autoregressive time series models, we allow for different sources of non-Gaussian tail behaviour. We assess the relevance of time-varying volatility – which can be a source of fat tails in the unconditional distribution of the variable – by estimating GARCH models. Whether error terms are drawn from distributions with fat tails is addressed by abandoning the assumption of a normal distribution for that of a Laplace distribution or a *t*-distribution. Finally, since the unconditional distribution of the data also appears to be somewhat skewed, we also estimate the model under the assumption that error terms are drawn from a Skew-*t* distribution.[1] Summarising our results, we find strong in-

[1] For a discussion regarding the importance of skewness, see, for example, Neftci (1984), Acemoglu and Scott (1997) and Bekaert and Popov (2019).

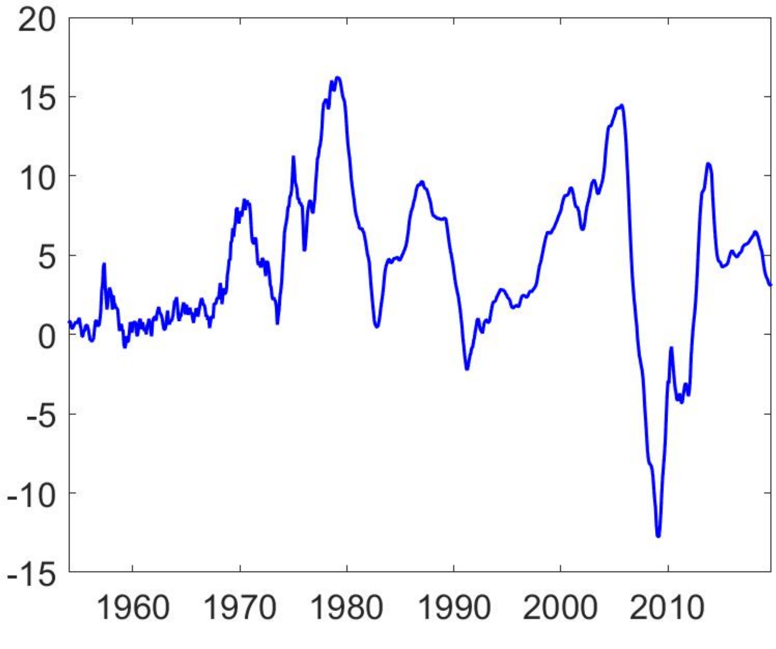
sample evidence for highly persistent, time-varying volatility and heavier-than-Gaussian tails. This translates into improved short-term point and density forecasts, with GARCH effects being the most important feature. Skewness, on the other hand, appears to be a less relevant aspect to consider in modelling.

The rest of this paper is organised as follows: In Section 2, we present the data. The methodological framework is discussed in Section 3 and our empirical findings are shown in Section 4. Finally, Section 5 concludes.

2. Data

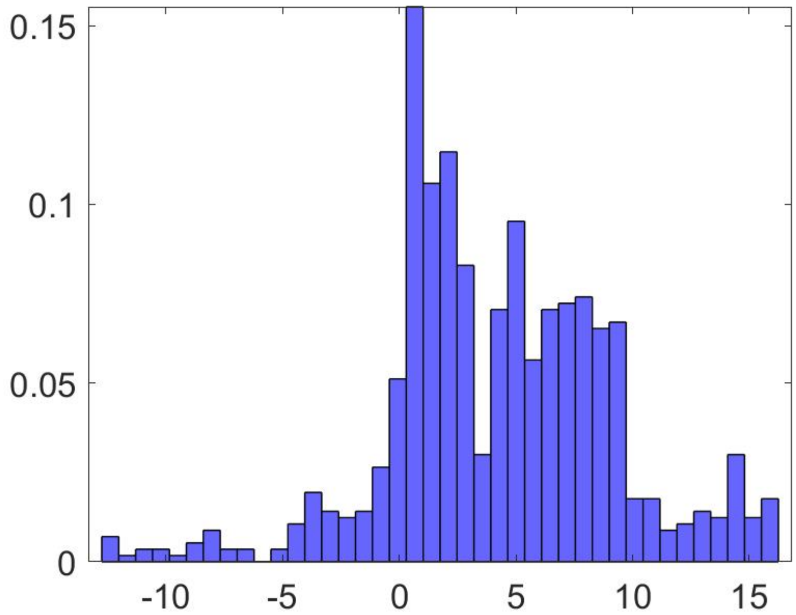
We base our analysis on Shiller's (2015) nominal house-price index. Data are monthly and span the period January 1953 to September 2019. Returns are calculated as $r_t = \frac{P_t - P_{t-12}}{P_{t-12}}$, where P_t is the price index at time *t*. A time-series plot of the return data is given in Figure 1 and a histogram showing the unconditional distribution is provided in Figure 2. Some key descriptive statistics are given in Table 1.

Figure 1. Housing-price return data.



Note: Twelve-month change in percent on vertical axis.

Figure 2. Histogram of housing-price return data.



Note: Twelve-month change in percent on horizontal. Relative frequency on vertical axis.

Table 1. Descriptive statistics and Jarque-Bera test statistic.

Mean	Variance	Skewness	Kurtosis	Jarque-Bera
4.278	24.390	-0.150	3.813	24.464

Note: The critical value at the five percent level of the Jarque-Bera test is 5.99.

As both the histogram in Figure 2 and the descriptive statistics in Table 1 suggest, the tails of the distribution appear heavier than what is implied by a normal distribution; the kurtosis of the unconditional distribution is around 3.8. There is also a slight negative skewness. Taken together, this means that the Jarque-Bera test strongly rejects an assumption of normality.

In Table 2, we compare the unconditional house-price returns and a normally distributed variable using the kurtosis ($K_{\alpha,\tau}$), peakedness ($P_{\eta,\tau}$) and tailness ($T_{\alpha,\eta}$) measures proposed by Liu (2019). These measures are based on the ratio of interquantile ranges of the variables, using various quantiles of the distribution denoted by α, η, and τ (for example , $\alpha = 0.05$ means that we use the 5th percentile of the distribution when calculating kurtosis and tailness).[2] Of particular importance are the tails of the return distribution, which are heavier than the normal distribution at all levels. There is also some evidence for asymmetry in the unconditional distribution further out in the tails (α=0.01), but skewness is not a robust feature if we use different quantiles (α).

Table 2. Comparison of quantile-based kurtosis, peakedness and tailness.

	α	$K_{\alpha,\tau}$	$P_{\eta,\tau}$	$T_{\eta,\tau}$	$RT_{\eta,\tau}$	$LT_{\eta,\tau}$
Normal distribution	0.010	3.449	1.706	2.022	1.011	1.011
	0.025	2.906	1.706	1.704	0.852	0.852
	0.050	2.439	1.706	1.430	0.715	0.715
Unconditional returns	0.010	3.979	1.404	2.834	1.256	1.579
	0.025	3.317	1.404	2.363	1.141	1.222
	0.050	2.615	1.404	1.862	1.013	0.850

Note: Comparison of the quantile-based kurtosis, peakedness and tailness among a normal random variable and the unconditional house-price returns. The measurements are calculated based on ratios of interquantile ranges with α = (0.01, 0.025, 0.05) and η = 0.125, τ = 0.25 quantile levels (see Footnote 2 and Liu, 2019).

3. Methodological framework

We use univariate autoregressive (AR) models to model house return series. Our baseline is the homoscedastic AR(p) model:

$$r_t = \gamma_0 + \sum_{j=1}^{p} \gamma_j r_{t-j} + \sigma_\nu \nu_t \tag{1}$$

with $\nu_t \sim N(0,1)$. The number of lags is determined using the Schwarz (1978) information criterion, which suggests a lag length of p =5; for comparability we employ the same number of lags across all model specifications.

We then allow for a more flexible error term in the model. First, we let the second moment of the series vary over time. Doing that, we choose a robust approach – relying on the GARCH(1,1) specification (Bollerslev, 1986). Hence, the second model is the AR(p)-GARCH(1,1):

---

[2] The measures are related to each other and the quantiles of the distribution by the formula

$$K_{\alpha,\tau} = P_{\eta,\tau} T_{\alpha,\eta} = \frac{Q_{1-\eta} - Q_\eta}{Q_{1-\tau} - Q_\tau} \frac{Q_{1-\alpha} - Q_\alpha}{Q_{1-\eta} - Q_\eta},$$

where Q_i is the i-th quantile of the distribution and the quantiles used satisfy $0 < \alpha < \eta < \tau < 0.5$. Further, the tail ($T$) can be decomposed to right (RT) and left (LT) tails according to

$$T_{\alpha,\eta} = RT_{\alpha,\eta} + LT_{\alpha,\eta} = \frac{Q_{1-\alpha} - Q_{0.5}}{Q_{1-\eta} - Q_\eta} + \frac{Q_{0.5} - Q_\alpha}{Q_{1-\eta} - Q_\eta}.$$

$$r_t = \gamma_0 + \sum_{j=1}^{p} \gamma_j r_{t-j} + \epsilon_t \tag{2}$$

$$\epsilon_t = \sigma_t v_t \tag{3}$$

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2 \tag{4}$$

where v_t is assumed to be an *iid* error term distributed according $v_t \sim N(0,1)$, as above. In the third and fourth models, we relax the assumption of normality of the error term, while maintaining symmetry. This is done by using $v_t \sim Laplace$ and $v_t \sim Student - t(\kappa)$, where κ is the degrees of freedom. Of specific interest here is the fact that both the Laplace distribution and the Student-*t* distribution allow for heavier-than-Gaussian tails.[3] In the fifth model, we finally assume that the error terms are drawn from the Skew-*t* distribution (Hansen, 1994) – that is, $v_t \sim Skew - t(\lambda, \kappa)$, where κ is the degrees of freedom and λ is the skewness parameter. This means that we allow for both heavy tails and asymmetry in the error term.[4]

Our primary focus is to evaluate which distributional properties are important for modelling the dynamics of the housing-price return series. This we evaluate based on both in-sample estimates (using maximum likelihood) and an out-of-sample forecasting exercise. We consider both point and density forecasts. The point forecasts are evaluated based on the root mean square error (RMSE) of the forecast, accompanied by the Diebold-Mariano test (Diebold and Mariano, 1995). We formulate the test such that a positive test value implies that the given model outperforms the benchmark.[5]

In terms of density nowcast evaluation we make use of the probability integral transform (PIT) proposed by Diebold *et al.* (1998). As a first check, we look at the histograms of the PIT and assess visually how close they are to the uniform distribution. We also employ more rigorous statistical techniques such as the Kullbach-Leibler (KL) divergence (Kullback and Leibler, 1951) of the PIT series from an *iid* uniform distribution and formal statistical tests of uniformity, in particular the Kolmogorov-Smirnov (KS) and the Anderson-Darling (AD) test (Anderson and Darling, 1952).[6] For all these measures, a lower value implies that the given model produces a better density forecast.

---

[3] Fagiolo *et al.* (2008) found the Laplace distribution useful when modelling the fat tails of GDP growth rates. The Student-*t* distribution has been used more widely in the empirical literature; see, for example, Cúrdia *et. al.* (2014), Clark and Ravazzolo (2015), Cross and Poon (2016) and Kiss and Österholm (2020).

[4] All distributions are parameterized to have a variance of one. The density function of the Skew-*t* distribution can be found in Hansen (1994).

[5] As pointed out by Diebold (2015), the Diebold and Mariano (1995) tests in its standard form is an appropriate choice even if we test for nested models, for which the original assumptions of the test do not formally hold.

[6] The KLIC has been used in a number of applications in the context of density forecasts; see, for example, Cogley *et al.* (2005), Robertson *et al.* (2005), Hall and Mitchell (2007), Diks *et al.* (2010), and Mitchell and Wallis (2011).

4. Empirical analysis

We start by presenting estimation results based on the full sample for all five specifications.[7] Then we turn to an out-of-sample forecast exercise where we use recursive estimation to obtain both point and density forecasts. We focus on short horizon forecasts ($h = 1$) for two reasons. First, the one-period-ahead forecasting results are most closely related to the in-sample findings, and hence it can be considered as a validation tool. Second, from a forecasting perspective, the simple univariate forecast is likely to produce a reasonable conditional mean forecast in the short run.[8]

4.1 Estimation results

Table 3 presents the estimation results for the five specifications. Note that the housing-price return series is fairly persistent (the sum of the autoregressive coefficients are close to unity). Engle's (1982) ARCH test shows a strong presence of conditional heteroscedasticity for the residuals in the baseline model in column 1. Also, the Jarque-Bera test suggests that residuals are non-Gaussian in this specification. This is further supported by Figure 3, where it is clear that a *t*-distribution fits the residuals of the baseline specification better than a normal distribution.

The GARCH(1,1) specification (column 2) seems to take care of conditional heteroscedasticity quite well. However, the variance is very persistent and its parameters are not precisely estimated.[9] For the specification with Laplace distribution, where no extra parameter is added to capture heavy-tails, the persistence of the variance is somewhat lower. However, it seems that the Laplace distribution is not flexible enough to capture the behaviour of the standardized residuals, as their transformation imposing a normal distribution still shows signs of non-Gaussianity according to the Jarque-Bera test.[10] In contrast, the more flexible approaches, the *t* and the Skew-*t* distribution, seem appropriate to capture non-Gaussianity. The skewness parameter however is low and non-significant (column 5).

---

[7] Figure 1 suggests that the time-series behaviour of the return series may have changed around 1975. In fact, the Shiller (2015) dataset changes source in January 1975. Therefore we also estimate the model on the shorter subsample, namely January 1975 to September 2019. Results, collected in Table A1 in the Appendix, are qualitatively very similar to full sample estimates.

[8] We would like to stress that our aim is to assess how important it is to allow for flexibility in the higher order moments of the disturbances when modelling housing-price returns – not to generate the best possible model for the conditional mean from a forecasting perspective.

[9] In fact, the parameters α_1 and α_2 sum up to unity for the normal-GARCH, Student-*t*-GARCH and Skew-*t*-GARCH specifications. However, looking at the results using the Laplace distribution and the shorter sample, integrated volatility does not seem to be a robust feature of the data, therefore we do not impose it in any of the specifications with conditional heteroscedasticity.

[10] The Jarque-Bera tests are based on z_t, the PIT series of the standardized residuals, for which $\Phi^{-1}(z_t)$ is standard normally distributed (where Φ^{-1} is the inverse cumulative distribution function of the standard normal distribution). We test this by applying the Jarque-Bera test on $\Phi^{-1}(z_t)$.

Table 3. Within-sample estimation results

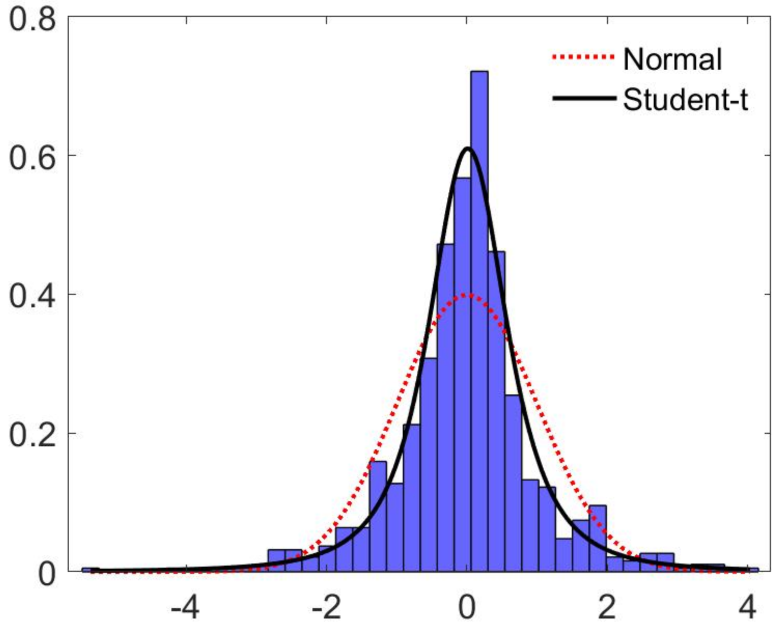
		(1)	(2)	(3)	(4)	(5)
Mean equation [AR(5)]	γ_0	0.038	0.033	0.035	0.033	0.032
		(0.023)	(0.185)	(0.000)	(0.467)	(0.013)
	γ_1	1.564	1.816	1.823	1.820	1.819
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	γ_2	-0.509	-0.756	-0.743	-0.741	-0.740
		(0.000)	(0.000)	(0.000)	(0.023)	(0.000)
	γ_3	-0.115	-0.271	-0.321	-0.307	-0.303
		(0.103)	(0.011)	(0.000)	(0.147)	(0.020)
	γ_4	0.203	0.396	0.415	0.400	0.397
		(0.003)	(0.209)	(0.000)	(0.000)	(0.006)
	γ_5	-0.152	-0.189	-0.180	-0.178	-0.178
		(0.000)	(0.263)	(0.000)	(0.001)	(0.003)
GARCH equation	ω		0.000	0.000	0.001	0.001
			(0.745)	(0.260)	(0.895)	(0.702)
	α_1		0.172	0.123	0.175	0.175
			(0.459)	(0.001)	(0.824)	(0.671)
	α_2		0.828	0.804	0.825	0.825
			(0.000)	(0.000)	(0.000)	(0.001)
Degrees of freedom	ν				7.870	7.941
					[26.854]	[3.297]
Skewness parameter	λ					-0.031
						(0.854)
ARCH-test		131.592	13.661	14.431	13.440	13.494
		(0.000)	(0.322)	(0.274)	(0.338)	(0.334)
JB-test		295.989	51.528	15.288[a]	0.616[a]	0.070[a]
		(0.000)	(0.000)	(0.000)	(0.735)	(0.966)
T		789	789	789	789	789

Note: Model (1) is a homoscedastic model with Gaussian error terms. Model (2) employs a GARCH(1,1) specification with Gaussian error terms. Models (3), (4) and (5) employ a GARCH(1,1) specification together with Laplace, *t*-, and Skew-*t*-distributed error terms, respectively. The ARCH-test is Engle's test for conditional heteroscedasticity (Engle, 1982) conducted with twelve lags. The JB-test is the Jarque-Bera test for conditional heteroscedasticity. "a" indicates that the test is based on the probability integral transform of the (standardized) residuals (see footnote 10). *p*-values are in parentheses (); standard errors for the degrees of freedom of the *t*-distribution are in brackets [].“T” is the number of observations.

Figure 4 shows the estimated conditional volatility of the return series. All GARCH specifications look quite similar; the biggest discrepancy is found for the Laplace which – due to its excessively heavy tails – allows for a generally lower variance. Volatility also clusters; the series starts out with a higher general level of volatility, which falls considerably during between 1980 and 2005, and peaks again during the financial crisis.

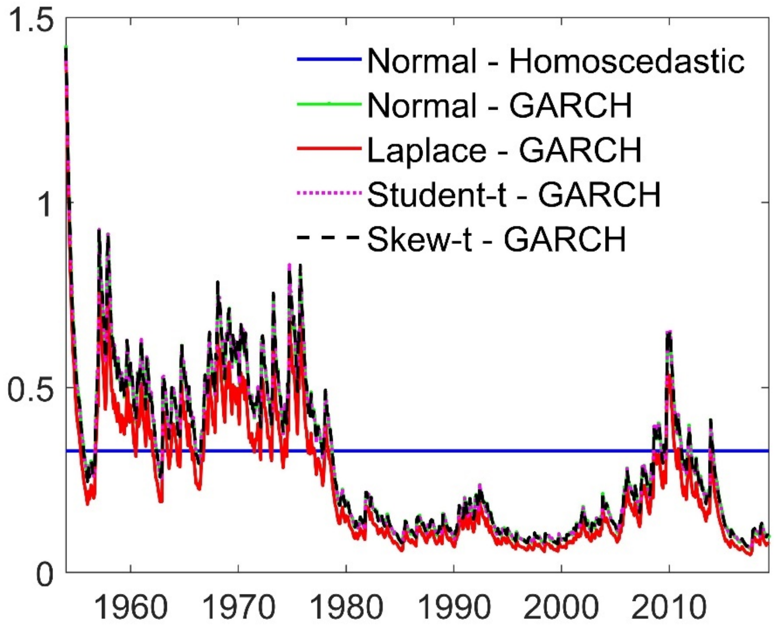
Having assessed the different specifications within-sample, we next conduct a recursive out-of-sample forecast exercise to validate our results. If we with this exercise establish that a more complex specification also generates better out-of-sample forecasts, we can conclude that that our results are not due to over-fitting the data in-sample.

Figure 3. Residuals from homoscedastic model.



Note: Density on vertical axis.

Figure 4. Conditional volatility under different assumptions for the error term.



Note: Conditional standard deviation on vertical axis.

4.2 Out-of-sample analysis

We evaluate the one-period-ahead forecasting performance of the models using an expanding window technique starting in January 1974; this yields a total of 548 forecasts to evaluate.[11] Results are collected in Table 4 for both point and density forecast measures.

Table 4. Forecast evaluation results.

	RMSE	DM	KS	AD	KL
Normal - Homoscedastic	0.279	-	3.534	25.498	0.179
Normal - GARCH	0.259	3.491	1.019	1.743	0.013
Laplace - GARCH	0.258	3.394	2.366	13.141	0.066
Student-t - GARCH	0.258	3.495	0.844	0.590	0.008
Skew-t - GARCH	0.258	3.468	0.814	0.521	0.006

Note: "RMSE" is the root mean squared error of the point forecasts. "DM" is the Diebold-Mariano test statistic where the model with homoscedastic, normally distributed errors is chosen to be the benchmark. "KS" is the Kolmogorov-Smirnov test statistic. "AD" is the Anderson-Darling test statistic. "KL" is the Kullback-Leibler divergence of the PIT of the residuals from the uniform distribution on the [0,1] interval. All three density nowcast measures are based on the probability integral transform of the density forecasts produced by each model. The results are based on 548 out-of-sample forecasts and an expanding window for parameter estimation. The critical value at the five percent level for the Kolmogorov-Smirnov test is 0.058 for N=548 and for the Anderson-Darling test it is 2.492.

For point forecasts, the homoscedastic model clearly performs worst. Accounting for time-varying volatility statistically significantly improves the forecast, as can be seen from the Diebold-Mariano test. However, allowing for non-Gaussian error terms does not provide additional improvement in point forecasts over the Gaussian GARCH specification.

For density forecasts, all features – that is, conditional heteroscedasticity, fat-tails and skewness – seem to have some relevance. Comparing results in Table 4, allowing for GARCH effect improves quite a lot on the density forecasts; further improvements are generated by employing a *t*-distributed or a skewed error term, with the latter being less important. However, using the Laplace distribution to capture non-Gaussianity seems to produce the least favourable results.

To gain further insight for the out-of-sample results, the histograms of the PIT for each specification are presented in Figure A1 in the Appendix. The graph of the homoscedastic model (top left panel in Figure A1) is peaked, implying that the homoscedastic, Gaussian model has a too wide predictive density.[12] Allowing for GARCH effects – while still assuming a Gaussian error term – eliminates this problem and produces

[11] That is, we first estimate the models on the sample January 1953 to January 1974 and make predictions for February 1974. We then expand the sample to January 1953 to February 1974, re-estimate the models and predict March 1974. We continue in this manner until we reach the end of the sample, where we estimate the models using data from January 1953 to August 2019 and make predictions for September 2019.

[12] This finding is concordant with the fact that our within-sample analysis indicates that the unconditional volatility overestimates the conditional one in the larger part of the out-of-sample evaluation period; see Figure 4.

PITs that are fairly closely uniformly distributed (top right panel in Figure A1). The Laplace distribution is less successful at capturing the tail observations (second row, left panel in Figure A1), while the Student-*t*, and Skew-*t* distributed error terms appear most successful in bringing PITs close to uniformity (bottom left and right panels in Figure A1 respectively).

5. Conclusions

Fat tails in housing-price returns have been established in several studies. We contribute to this literature by analysing the heavy-tailed behaviour in the context of modelling and forecasting the return series. We find evidence of both conditional heteroscedasticity and non-Gaussian behaviour. Non-Gaussianity mostly takes the form of excess kurtosis, while the support for skewed behaviour is weaker. We also find that accounting for these features help improve both point and density forecasts.

Our results highlight the importance of underlying distributional assumptions when modelling housing-price series. In addition, our findings point to the relevance of considering time-varying volatility when modelling macroeconomic time series; after all, housing-price returns are not only a financial variable but also an important macroeconomic variable. While this aspect of macroeconomic modelling has grown in popularity over the last fifteen years – after having been made popular by Cogley and Sargent (2005) and Primiceri (2005) – we nevertheless believe that macroeconomists have not quite given it sufficient attention.

References

Acemoglu, D. and Scott, A. (1997). "Asymmetric Business Cycles: Theory and Time-Series Evidence", *Journal of Monetary Economics* 40, 501-533.

Ahamada, I. and Diaz Sanchez, J. L. (2013), "A Retrospective Analysis of the House Price Macro-Relationship in the United States", *International Journal of Central Banking* 9, 153-174.

Anderson, T. W. and Darling, D. A. (1952), "Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes", *Annals of Mathematical Statistics* 23, 193-212.

Ascari, G., Fagiolo, G. and Roventini, A. (2015), "Fat-Tail Distributions and Business-Cycle Models", *Macroeconomic Dynamics* 19, 465-476.

Bekaert, G. and Popov, A. (2019), "On the Link between the Volatility and Skewness of Growth." *IMF Economic Review* 67, 746-790.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics* 31, 307-327.

Campbell, J. Y. and Cocco, J. F. (2007), "How Do House Prices Affect Consumption? Evidence from Micro Data", *Journal of Monetary Economics* 54, 591-621.

Catte, P., Girouard, N., Price, R. and André, C. (2004), "Housing Markets, Wealth and the Business Cycle." OECD Economics Working Paper No. 394.

Chiang, M. C., Tsai, I. C. and Lee, C. F. (2012), "The Downside Risk of US and UK Housing Markets." *Journal of Real Estate Portfolio Management* 18, 257-272.

Clark, T. E. and Ravazzolo, F. (2015), "Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility", *Journal of Applied Econometrics* 30, 551-575.

Cogley, T., Morozov, S. and Sargent, T. J. (2005), "Bayesian Fan Charts for U.K. Inflation: Forecasting and Sources of Uncertainty in an Evolving Monetary System", J*ournal of Economic Dynamics and Control* 29, 1893-1925.

Cogley, T., and Sargent, T. J. (2005), "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US", *Review of Economic Dynamics* 8, 262-302.

Cross, J. and Poon, A. (2016), "Forecasting Structural Change and Fat-Tailed Events in Australian Macroeconomic Variables", *Economic Modelling* 58, 34-51.

Cúrdia, V., Del Negro, M. and Greenwald, D. L. (2014), "Rare Shocks, Great Recessions", *Journal of Applied Econometrics* 29, 1031-1052.

Diebold, F. X. (2015), "Comparing Predictive Accuracy, Twenty Years Later: A Personal Perspective on the Use and Abuse of Diebold–Mariano Tests", *Journal of Business and Economic Statistics* 33, 1-9.

Diebold, F. X., Gunther, A. and Tay, K. (1998), "Evaluating Density Forecasts with Application to Financial Risk Management", *International Economic Review* 39, 863-883.

Diebold, F. X. and Mariano, R. S. (1995). "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics* 13, 253-263.

Diks, C., Panchenko, V. and van Dijk, D. (2010), "Out-of-Sample Comparison of Copula Specifications in Multivariate Density Forecasts", *Journal of Economic Dynamics and Control* 34, 1596-1609.

Engle, R. F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica.* 50, 987-1007.

Fagiolo, G., Napoletano, M. and Roventini, A. (2008), "Are Output Growth-Rate Distributions Fat-Tailed?", *Journal of Applied Econometrics* 23, 639-669.

Graff, R., Harrington, A., and Young, M. (1997), "The Shape of Australian Real Estate Return Distributions and Comparisons to the United States", *Journal of Real Estate Research* 14, 291-308.

Hall, S. G. and Mitchell, J. (2007), "Combining Density Forecasts." *International Journal of Forecasting* 23, 1-13.

Hansen, B. E. (1994), "Autoregressive Conditional Density Estimation." *International Economic Review* 35, 705-730.

Iacoviello, M. (2005), "House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle," *American Economic Review* 95, 739-764.

Iacoviello, M. and Neri, S. (2010), "Housing Market Spillovers: Evidence from an Estimated DSGE Model", *American Economic Journal: Macroeconomics* 2, 125-164.

Kiss, T. and Österholm, P. (2020), "Fat Tails in Leading Indicators", *Economics Letters* 193, 109317.

Kullback, S. and Leibler, R. A. (1951), "On Information and Sufficiency", *Annals of Mathematical Statistics* 22, 79-86.

Liu, X. (2019). "On Tail Fatness of Macroeconomic Dynamics", *Journal of Macroeconomics* 62, 103154.

Maurer, R., Reiner, F. and Sebastian, S. P. (2004), "Financial Characteristics of International Real Estate Returns: Evidence from the UK, US, and Germany", *Journal of Real Estate Portfolio Management* 10, 59-76.

Mitchell, J. and Wallis, K. F. (2011), "Evaluating Density Forecasts: Forecast Combinations, Model Mixtures, Calibration and Sharpness." *Journal of Applied Econometrics* 26, 1023-1040.

Myer, F. N. and Webb, J. R. (1994). "Statistical Properties of Returns: Financial Assets versus Commercial Real Estate", *Journal of Real Estate Finance and Economics* 8, 267-282.

Neftci, S. N. (1984). "Are Economic Time Series Asymmetric over the Business Cycle?", *Journal of Political Economy* 92, 307-328.

Pontines, V. (2010), "Fat-Tails and House Prices in OECD Countries", *Applied Economics Letters* 17, 1373-1377.

Primiceri, G. (2005), "Time Varying Structural Vector Autoregressions and Monetary Policy", *Review of Economic Studies* 72, 821-852.

Robertson, J. C., Tallman, E. W. and Whiteman, C. H. (2005), "Forecasting Using Relative Entropy", *Journal of Money, Credit and Banking* 37, 383-401.

Schwarz, G. (1978), "Estimating the Dimension of a Model", *Annals of Statistics*, 6, 461-464.

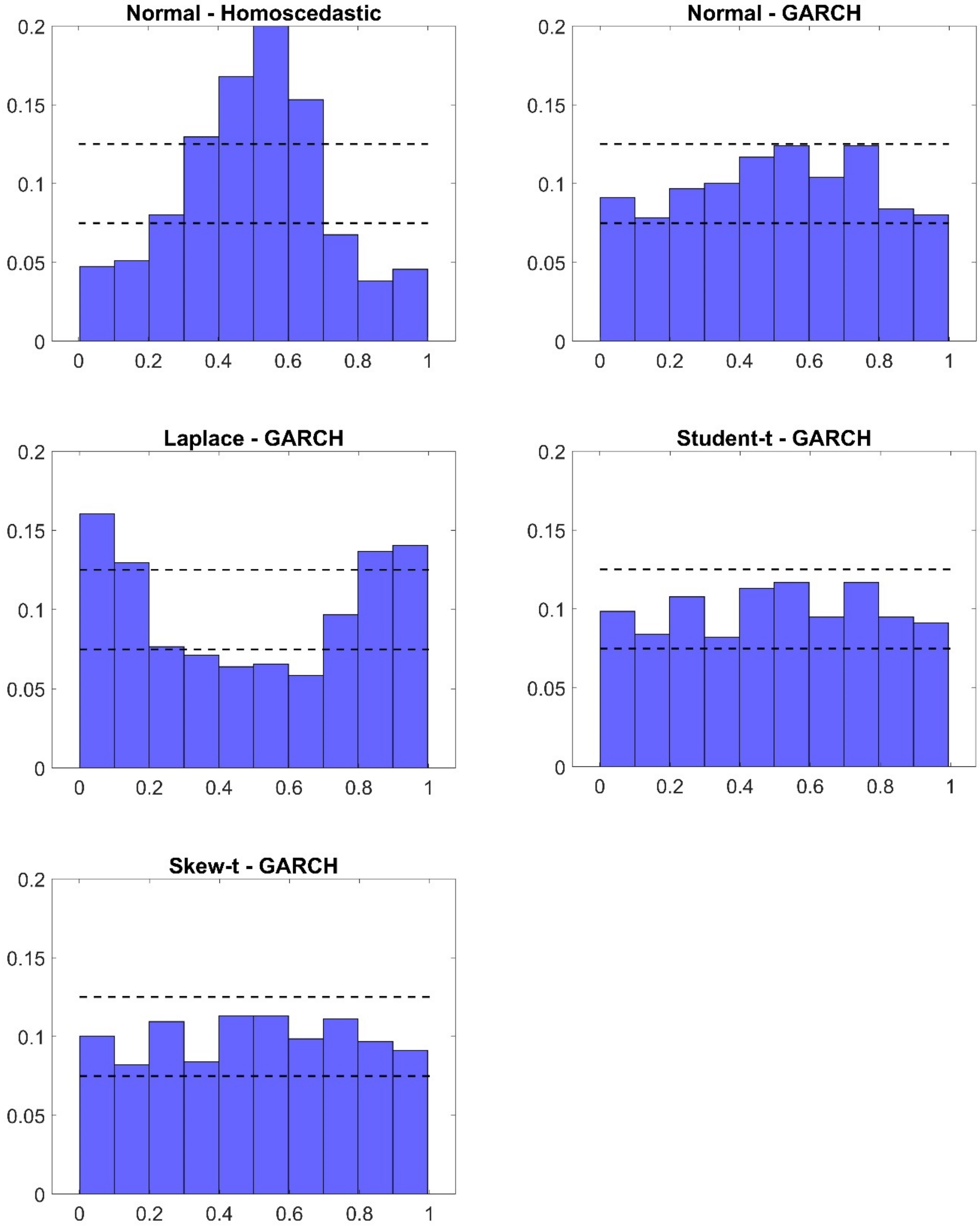
Shiller, R. J. (2015), *Irrational Exuberance: Revised and Expanded Third Edition*. Princeton University Press.

Young, M. S. and Graff, R. A. (1995). "Real Estate Is Not Normal: A Fresh Look at Real Estate Return Distributions", *Journal of Real Estate Finance and Economics* 10, 225-259.

Young, M. S., Lee, S. L. and Devaney, S. P. (2006), "Non-Normal Real Estate Return Distributions by Property Type in the UK", *Journal of Property Research* 23, 109-133.

Appendix

Figure A1. Histograms of probability integral transform for out-of-sample forecasts.



Note: Relative frequency on vertical axis.

Table A1. Within-sample estimation results, January 1975 to September 2019.

		(1)	(2)	(3)	(4)	(5)
Mean equation [AR(5)]	γ_0	0.028	0.031	0.033	0.028	0.026
		(0.032)	(0.003)	(0.000)	(0.004)	(0.010)
	γ_1	1.909	1.970	1.937	1.970	1.970
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	γ_2	-0.919	-0.897	-0.838	-0.909	-0.915
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	γ_3	-0.234	-0.469	-0.538	-0.454	-0.438
		(0.018)	(0.000)	(0.000)	(0.000)	(0.000)
	γ_4	0.430	0.640	0.717	0.638	0.627
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	γ_5	-0.192	-0.249	-0.284	-0.250	-0.248
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
GARCH equation	ω		0.001	0.001	0.001	0.001
			(0.003)	(0.010)	(0.008)	(0.006)
	α_1		0.215	0.124	0.192	0.188
			(0.000)	(0.001)	(0.000)	(0.000)
	α_2		0.742	0.724	0.759	0.762
			(0.000)	(0.000)	(0.000)	(0.000)
Degrees of freedom	ν				6.734	6.958
					[1.948]	[2.020]
Skewness parameter	λ					-0.065
						(0.283)
ARCH-test		115.610	8.634	8.082	8.094	8.032
		(0.000)	(0.734)	(0.778)	(0.777)	(0.783)
JB-test		1817.200	49.303	9.748[a]	1.758[a]	0.201[a]
		(0.000)	(0.000)	(0.008)	(0.415)	(0.904)
T		537	537	537	537	537

Note: Model (1) is a homoscedastic model with Gaussian error terms. Model (2) employs a GARCH(1,1) specification with Gaussian error terms. Models (3), (4) and (5) employ a GARCH(1,1) specification together with Laplace, *t*-, and Skew-*t*-distributed error terms, respectively. The ARCH-test is Engle's test for conditional heteroscedasticity (Engle, 1982) conducted with twelve lags. The JB-test is the Jarque-Bera test for conditional heteroscedasticity. "a" indicates that the test is based on the probability integral transform of the (standardized) residuals (see footnote 10). *p*-values are in parentheses (); standard errors for the degrees of freedom of the *t*-distribution are in brackets [].“T” is the number of observations.