An investigation into the existence of conditional volatility in U.S. house prices

Webster Zhou

Jan 2021

Data is taken from here. For this exercise, Stata is the software of choice.

GARCH/ARCH models can be employed to evaluate whether there is conditional volatility in U.S. house prices. GARCH type models are a generalisation of ARCH models, and they allow $E(\epsilon_t^2|\mathbf{\Psi}_{t-1}) = h(\mathbf{\Psi}_{t-1})$, where $\mathbf{\Psi}_{t-1}$ denotes information avalable up to time $t-1$. In other words, conditional variance of the ϵ_t is a function of observed information. This is contrasted to a standard ARMA model, which has a constant conditional variance because ϵ_t is assumed to have constant variance. GARCH models are useful in explaining volatility clustering in a series. The GARCH model assumes:

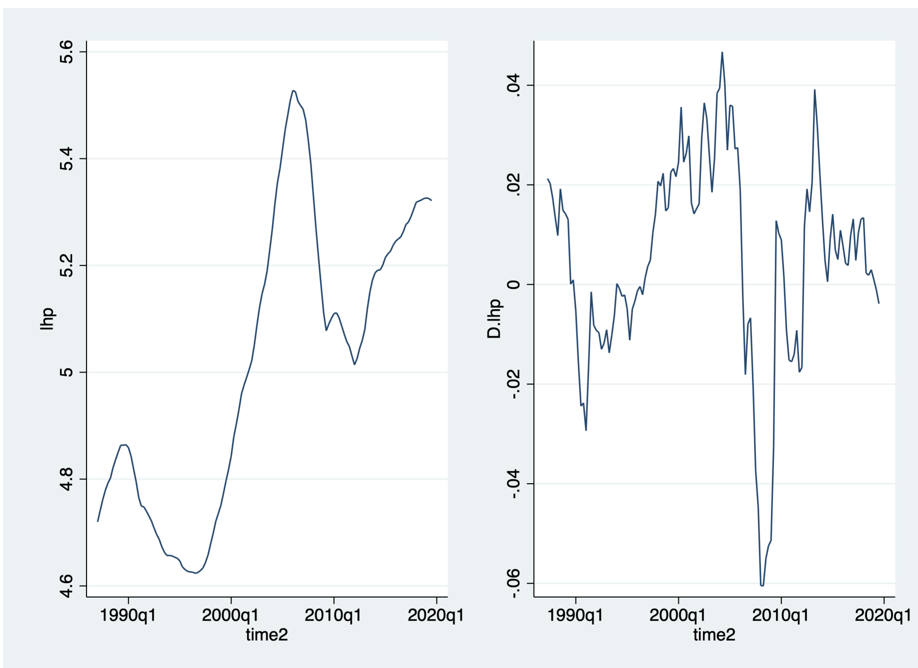
$$\epsilon_t^2|\mathbf{\Psi}_{t-1} \sim N(0, h_t)$$

$$h_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}$$

$$\alpha_0 > 0, \;\; \alpha_i \geq 0, \;\; \beta_j \geq 0, \;\; \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$$

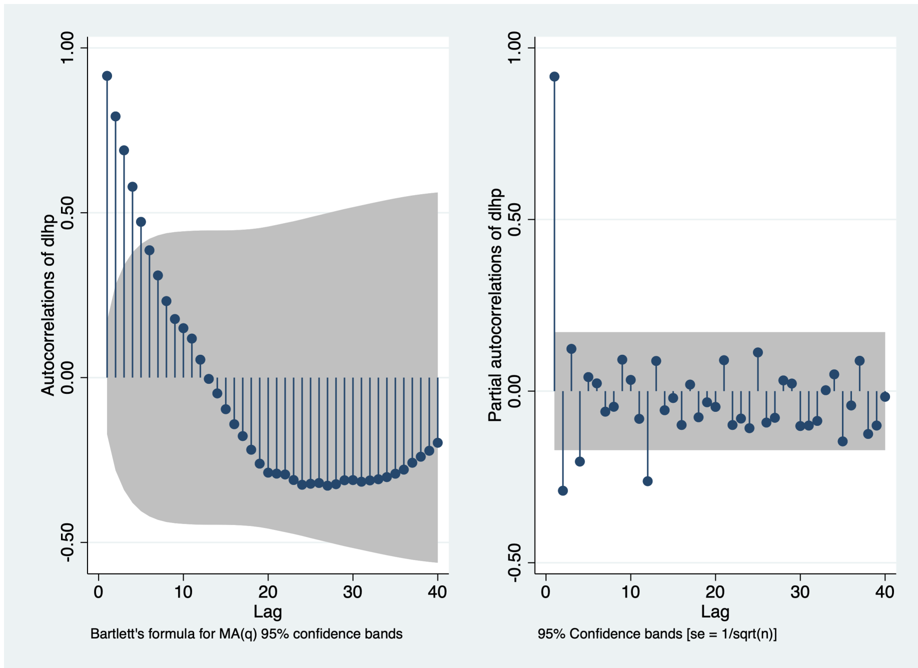
The restrictions on coefficients are necessary because the variance of a process is strictly positive. Additionally, $\sum_{i=1}^{max(p,q)} (\alpha_i + \beta_i) < 1$ is required for the model to be stable.

Figure 1: Time series of U.S. house prices in natural logs (left) and first differences (right)



Natural logs ($ln(HP)$) and first differences ($\Delta ln(HP)$) of the house price index are plotted in Figure 1. There is an upward trend over time in levels, but clustering is not immediately obvious just through visual inspection. Since GARCH modelling requires the underlying series to be stationary, $\Delta ln(HP)$ is used as the variable for analysis; an ADF test with 3 lags (based on AIC), a constant, and no trend was ran to make sure that it was indeed I(0).

Figure 2: ACF and PACF plots of $ln(HP)$



$$\Delta ln(HP)_t = \phi_0 + \phi_1 \Delta ln(HP)_{t-1} + \theta \upsilon_{t-1} + \upsilon_t \tag{1}$$

Looking at Figure 2, the ACF plot of $\Delta ln(HP)$ series shows a geometric decay over time, indicating AR structure in the series. The PACF plot has a large spike in lag 1 and tapers down quickly afterwards, suggesting that the AR and MA structure (if any) is of low lag length. Ultimately, based on visual inspection and model selection based on AIC, an ARMA(1,1) model (Equation 1) yielded the lowest IC. From Figure 3, the ARMA(1,1) specification appear to have adequately accounted for serial correlation in the data; to confirm this, Breusch–Godfrey tests were also conducted on $\hat{\upsilon}_t$. However, skewness and kurtosis of the residuals indicate non-normality (Table 1). This result provides some evidence that the $\Delta ln(HP)$ series may suffer from heteroscedasticity.

Then, squared residuals $\hat{\upsilon}_t^2$ is analysed. It is not immediately clear just through looking at the evolution of this series in Figure 4 that there is clear heteroscedasticity for the resulting residuals. A formal ARCH LM test is conducted and the results are reported in Table 2. The results points towards the presence of ARCH effects of lag 1, hence it may be appropriate to incorporate that in the model.

Figure 3: ACF and PACF plots of ARMA(1,1) residuals

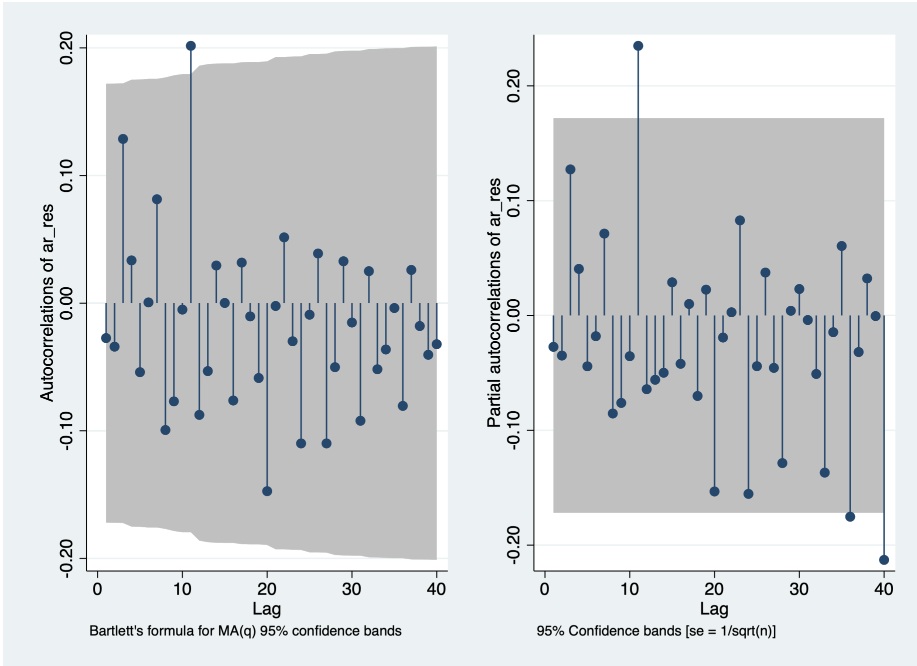


Table 1: Normality Test on $\hat{v}_t$ series

H_0 : Series is normally distributed	
H_1 : Series is not normally distributed	
Observations	130
Skewness	0.75
Kurtosis	5.26
Jarque-Brera Test statistic $\sim \chi^2_2$	39.99
Probability	0.00 (Reject H_0)

Table 2: ARCH LM test

H_0 : no ARCH effects at lag p

H_1 : ARCH(p) disturbance

Lags	Chi sq	P value	Decision (10% significance)
1	3.23	0.07	Reject H_0
2	3.51	0.17	Do not reject H_0
3	3.42	0.33	Do not reject H_0
4	3.35	0.50	Do not reject H_0

Specification of an ARCH(1) process $h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$ gives an estimate of $\alpha_1 = 0.37$ with a p-value of 0.01. The graph of the estimated conditional variance is shown in Figure 5. This is a plot of the one-step ahead variance for each observation in the sample. The value at time t is the forecast based on information available at t-1. Comparing this graph with that of the first differences of house price in Figure 1, it is evident that the increase in conditional variance is associated with a clustering of large positive and negative observations.

Overall, U.S. house prices seem to exhibit volatility clustering behaviour, and an ARCH(1) model seems to model the conditional volatility adequately. A TARCH model was further estimated to test for any potential asymmetry in the process but there was insufficient evidence to suggest so.

Figure 4: Time series of squared residuals $\hat{\upsilon}_t^2$

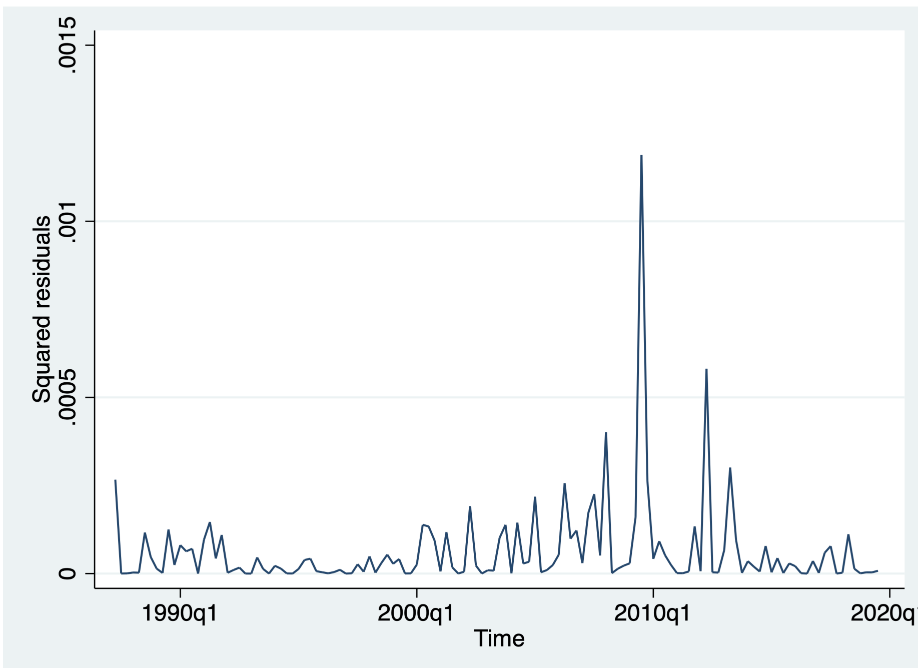


Figure 5: Estimated conditional variance based on an ARCH(1) model