Working Papers

www.cesifo.org/wp

On the Cyclicity of Regional House Prices: New Evidence for U.S. Metropolitan Statistical Areas

Michael A. Flor
Torben A. Klarl

CESIFO WORKING PAPER NO. 5471

CATEGORY 6: FISCAL POLICY, MACROECONOMICS AND GROWTH
AUGUST 2015

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: www.CESifo-group.org/wp

ISSN 2364-1428

On the Cyclicity of Regional House Prices: New Evidence for U.S. Metropolitan Statistical Areas

Abstract

This paper is mainly concerned with the analysis of regional house price cycles. Based on a newly available data set consisting of the 40 largest U.S. Metropolitan Statistical Areas (MSAs), we introduce a wavelet transform based metric to study the housing cycle synchronization across MSAs. We derive several conclusions: (i) We show that regional housing cycle dissimilarities are significantly and strongly connected to geography. (ii) We show that U.S. regional housing cycles are considerably shorter compared to business cycles. (iii) By employing statistical methods, we also show that regional housing prices significantly converge after the bursting of the bubble only at higher business cycle frequencies, whereas for lower business cycle frequencies this is not the case. This is a notable result, because it directly implies that housing cycles behavior is different even at the business cycle.

JEL-Code: C320, C330, E320, R110, R310.

Keywords: continuous wavelet transform, housing cycles, regional house price synchronization, wavelet distance matrix.

Michael A. Flor
University of Augsburg
Department of Economics
Universitätsstraße 16
Germany – 86159 Augsburg
michael.flor@wiwi.uni-augsburg.de

Torben A. Klarl
University of Augsburg
Department of Economics
Universitätsstraße 16
Germany – 86159 Augsburg
torben.alexander.klarl@wiwi.uni-augsburg.de

August 4, 2015

We would like to thank Monika Bredow and Franziska Kruse for invaluable editorial assistance in the preparation of this manuscript.

1 Introduction and Motivation

Kindleberger and Aliber (2005) point out that financial crises emerge more frequently than are usually recognized and, further, their unpleasant consequences are forgotten with the beginning of the next upswing (MacLennan and O’Sullivan (2011)). The recent financial crisis with its origin lay at the heart of the collapse of the subprime mortgage boom and the U.S. house price bubble can be seen as an excellent example for *glocalization* in the way that the appearance of a local shock has global consequences but also feedback effects to local areas (Martin (2011)): There is recent evidence that both, the severity and magnitude of house price irruption varied across regions (Chowdhury and MacLennan (2014)). This article aims to improve our understanding of U.S. house price cycles to obtain a better assessment of the future consequences of boom-bust episodes.

House prices besides other macroeconomic time series can be generally seen as the combination of components working at different frequencies. This can be justified with the idea that economic processes are the result of agents’ decisions, from which some base their decisions on short-run movements, whilst others are long-term orientated. Further, policy endeavors, such as the Home Affordable Modification Program (HAMP) or the Home Affordable Refinancing Program (HARP)² to support and stabilize the U.S. housing market could have short- as well as long-term impacts, thus affecting the economy in different ways at different frequencies. Finally, it could be possible that policy induced effects evolve over time, because of the initialization of a new policy directive or changes of the institutional settings. To sum up, a comprehensive investigation of housing cycles requires a method which allows us to analyze the behavior of the housing price evolution at different frequencies over time. Of course, applying standard time series or purely frequency domain-orientated methods fail to accomplish such an analysis, which implies only a limited understanding of house price cycles. Further, another unpleasant attribute of the standard time series toolbox is directly associated with the fact that it entirely neglects the spatial interrelationship of regional house price evolution. These two shortcomings motivate us to employ a new method - the wavelet analysis - which, as it will turns out, elegantly overcomes those shortcomings.

Therefore, the aim of this paper is to uncover the dynamics of housing prices for urban areas of the U.S. with the help of regional data and ask the following questions: First, what can we say about the geography of housing price changes? Second, what

²For more details on each program, please refer to <http://www.makinghomeaffordable.gov>.

clusters of U.S. MSAs emerge in terms of synchronicity of housing price cycles, both with the national U.S. housing price cycle as well as among the 40 largest MSAs? And finally, do we observe a MSA wide convergence of housing cycles after the bursting of the U.S. housing bubble at *every* frequency?

In particular, our paper makes the following major points: First, we show that regional housing cycle dissimilarities are significantly and strongly connected to geography. Second, we show that U.S. regional housing cycles are considerably shorter compared to business cycles. This significant result is obtained, because the wavelet analysis allows us to uncover transient relations. And third, by employing statistical methods, we also show that regional housing prices significantly converges after the bursting of the bubble only at higher business cycle frequencies, whereas for lower business cycle frequencies this is not the case. This is a notable result, as it directly implies that housing cycles behavior is different even at the business cycle, and, thus, a policy maker should be keenly aware of this fact.

The reminder of this article is organized as follows. The next section depicts the related literature, section three explains the used data, and section four illustrates the empirical methodology. Section five describes our results and section six concludes. Finally, an appendix contains additional explanations and results.

2 Related Literature

2.1 Wavelets

Until a few years ago, the wavelet transform was mainly used in several scientific disciplines, such as acoustics and engineering, but its application in economics, until today is limited.

For an early contribution to the literature in economics, see Ramsey and Lampart (1998a, 1998b), who employ wavelets for the decomposition of economic linkage between expenditure and income on the one hand, and money and income on the other. Another paper using wavelet analysis in an economic context is Kim and In (2003), who study the relationship between financial variables and the U.S. industrial production. Crowley (2007) presents a comprehensive guide to wavelets and their applications for economists. Thereafter Aguiar-Conraria et al. (2008) show within a cross wavelet analysis that the relation between monetary policy and macroeconomic variables has changed and evolved with time and that these changes are not homogeneous across

the different frequencies. Further, Aguiar-Conraria and Soares (2011a) focus on the link between oil price increases and inflation and Aguiar-Conraria and Soares (2011b) analyze business cycle synchronization across the EU-15 and the Euro-12 countries with wavelet analysis tools. Also with wavelet analysis, Aguiar-Conraria et al. (2012b) assess the relation between the yield curve and the macroeconomy in the U.S. between 1961 and 2011. A recent contribution to this literature in economics is Mandler and Scharnagl (2014), who investigate the relationship between the M3 money growth and the Harmonized Index of Consumer Prices (HICP) inflation in the Euro area within a wavelet framework. *Inter alia*, they find strong co-movements close to a one-to-one relationship between the very long-run fluctuations (24-32 years) with money growth leading inflation by about two to three years. Another recent paper belonging to this growing literature is Ko and Lee (2015). They examine the link between economic policy uncertainty and stock prices with the help of wavelet analysis. They find that this relationship is generally negative but changes over time regarding the considered frequency.

2.2 House Price Dynamics

Among the first studies in focusing on U.S. housing regions, Clark and Coggin (2009) consider regional house price trends between 1975:I and 2005:II. They employ a unobserved components, structural time series model in the first stage and a bivariate unit root procedure to test for absolute versus relative convergence in the second stage. In this regard they find mixed evidence for regional convergence in the U.S. house price indices (HPIs).³ Furthermore, they argue that the estimated regional housing cycles are between 1.37 and 2.25 years, in contrast to the average postwar business cycle in the U.S. with about 5.5 years. The authors conclude that the mixed evidence for convergence leads to little support for their simple and linear vector error correction (VEC) model of regional housing prices because of the obviously contained non-linear development of regional housing prices.

In contrast to Clark and Coggin (2009), Corradin and Fontana (2013) take the non-linearities of housing prices appropriately into account. They estimate an error correction model with Markov-switching to uncover the house price dynamics for thirteen European countries. Corradin and Fontana (2013) find that (i) the growth rates within regimes differ largely across countries, (ii) for some European countries the

³In the paper at hand we underpin their result, because the subprime bubble drives the convergence.

observed high growth phases are associated with a stable regime, and (iii) European housing markets have been more synchronized since 2000, following a growing trend in the time span 2002-2006 and a dramatic downturn after the Lehman collapse in 2008 and during the Euro area sovereign debt crisis. Regarding the first result, the authors mention that for some European countries, such as the UK and Ireland, the high growth phases might be associated with booms in the housing markets, while for others the high growth phases are characterized by modest growth in housing prices.

Furthermore, Chowdhury and MacLennan (2014) examine the regional U.K. house price dynamics at *Nomenclature des unités territoriales statistiques* (NUTS) 1 level with Markov-switching vector auto regression (MSVAR) methods. To reduce the dimensionality of the data and to check for common factors, they perform a principal component factor analysis, with communalities and a varimax rotation. They find that the U.K. is divided by two large groups of regions with marked differences in the amplitude and duration of the cyclical regimes between the two groups. Findings from their study also show that during economic downturns, housing prices in regions belonging to group one fall faster than for regions in group two. However, during positive growth regimes, the house price growth rates in group one are higher than in group two. Moreover, analyzing the duration of the regimes, it becomes evident that the duration of positive growth regimes is longer in group one than in group two. Finally, the findings also suggest that the response of housing prices to growth is faster in group one compared to group two. Chowdhury and MacLennan (2014) argue that these significant differences in both growth rates and average regime durations for the groups are attributed to the existence of regional heterogeneities in housing markets, financial sectors, and labor markets.

The recently published paper by Ghent and Owyang (2010), which is closely related to ours, analyzes the relationship between housing and the business cycle at the MSA-level. Based on a time-varying transition probabilities (TVTPs) MSVAR, *inter alia*, they conclude that an increase in housing prices may have different short and long-term effects on the economy, given there is a significant frequency-dependence. However, these authors do not explicitly focus on the frequency domain of housing cycles.

2.3 Added Value of the Paper to the Literature

Following the majority of the above mentioned literature, it is rather obvious to employ the MSVAR method and run the following (naive) regression for the i th MSA region

which controls for a state dependent variance in order to investigate the housing cycle behavior:

$$\Delta p_{it} = \mu_{i,s_{l,t}} + \sum_{j=1}^{\tilde{p}} \phi_j(\Delta p_{it-j} - \mu_{i,s_{l,t-j}}) + \sigma_{i,t} \epsilon_{l,t},$$

where $\mu_{i,s_{l,t}}$ is the mean rate of house price change for region i in state l at time t .⁴ Based upon this regression, we can show whether housing cycles are symmetric or not.

As an example, we have run the regression for the U.S. and for two states, like Ghent and Owyang (2010): a recession and a boom state.⁵ *Inter alia*, we observe that housing cycles are asymmetric: The duration in the boom state is considerably larger compared to the recession state, with a significantly lower variance of the real house price growth rate in the high growth state. Moreover, as suggested by the estimated transition probabilities, both states are highly persistent.

Although the contributions mentioned above have enriched our understanding of house cycle behavior, nevertheless, several questions remain unanswered. For instance, we cannot satisfactorily investigate whether or not cycles at *different* frequencies contribute equally to the total variance of the house price change. Even if we focus on business cycle frequencies, in this paper, we show that higher business cycle frequencies contribute in a significantly *different* way the total variance compared to lower business cycle frequencies. It is impossible to derive this result based on the MSVAR methodology, which only focuses solely on the time dimension but leaves out the frequency domain. Further, we cannot give an answer to the question whether or not house price cycles at different frequencies are more synchronized after the burst of the housing bubble in 2007. It is straightforward that those questions are highly relevant for policy. Hence, we need a method which combines both, the frequency as well as the time domain to investigate satisfactorily the cyclical behavior of house prices.

Many papers conclude that the wavelet transform is better adapted for analyzing localized variations of power within a time series than the Fourier transform because a wavelet function is indexed by two parameters, time and frequency, and the Fourier

⁴Clearly, in order to control for spatial auto-correlation in the residuals, a multivariate MSVAR of order $\tilde{p}$ should be estimated. However, in this contribution we cannot estimate such a model due to the degree of freedom (DOF) set of problems.

⁵The results of the regression exercise can be found in appendix 7.2. Further, it would be desirable to estimate a three-state model, however, the probability of non-convergence of the employed optimization routine, such as quasi-Newton rises with the number of states. This is even the case here, mainly due to the fact of the relatively short time series.

function has the frequency as the only parameter.⁶ Therefore, we will use the wavelet analysis to uncover the U.S. housing cycle and its dynamics within a time-frequency domain. Furthermore, the wavelet transform handles frequencies in a logarithmic rather than linear way and thus, it is also preferable to a windowed Fourier transform, which was first introduced by Gabor (1946).⁷ Further, Bloomfield et al. (2004) argue that "wavelet analysis is an improved tool over classical Fourier analysis for studying temporal relations".

To the best of our knowledge, the paper at hand is the first which uses wavelet analysis within an urban context to analyze time series in both the time and frequency domain.

3 Data

As a real estate price index we use the Federal Housing Finance Agency (FHFA), which is the successor of the OFHEO, expanded HPI for the U.S. as a whole and for 40 Metropolitan Statistical Areas (MSAs).⁸ The considered MSAs are standardized county-based areas or rather geographic entities that have at least one core urbanized area with a population of 50,000 or more and include adjacent territory that has a high degree of social and economic integration (measured by commuting to work).⁹ If specified criteria are met and an MSA contains a single core population greater than 2.5 million, the MSA is divided by the FHFA into Metropolitan Statistical Area Divisions (MSADs). Nevertheless, for an MSA evaluation, we have to construct the MSA index as a weighted average of the indices for the component divisions, where the weights are the relative shares of the division's population estimate from the U.S. Census Bureau in 2010. For example, the MSA population file shows that the Anaheim MSAD (11244) had a population estimate of 3,010,232, while the Los Angeles MSAD (31084) had a population estimate of 9,818,605. In this situation, the Los Angeles-Long Beach-Anaheim MSA (31080) index is composed of 23.46 % of the Anaheim division index and 76.54 % of the Los Angeles division index with an overall population estimate of 12,828,837.

⁶For example see Farge (1992).

⁷See Daubechies (1990).

⁸The FHFA reports three different HPIs: the purchase-only HPI, the all-transactions HPI, and the expanded HPI. For a more detailed consideration of these HPIs see the description in the appendix 7.1.

⁹See Goetz et al. (2014). For the considered MSAs see table (7.1) in the appendix.

Compared with the popular S&P/Case-Shiller Home Price Index, the FHFA expanded HPI (i) is broader in terms of sampling¹⁰, (ii) is also preferable with respect to geographic coverage and frequency¹¹, and (iii) is reported quarterly, seasonally adjusted, and unadjusted.¹²

We use the U.S. Consumer Price Index (CPI) "All Items Less Shelter" taken from the Federal Reserve Economic Data (FRED) to adjust the HPI for inflation. Since we also want to look at the house price dynamics before and after the housing bubble in 2007, all considered time series are restricted to the first quarter in 2001 till the fourth quarter in 2013 to ensure a balanced panel.

To sum up, the FHFA expanded HPI by MSA as well as the FHFA expanded U.S. HPI are set to 100 in 2009, are quarterly and seasonally adjusted, and are deflated by the just mentioned CPI.

4 Empirical Methodology

4.1 Some Basic Ingredients

The methodology used is based on the so-called continuous wavelet transform (CWT). A superb introduction into this methodology in economics is given by Aguiar-Conraria and Soares (2011a, 2011b, and 2014).¹³ For an easier interpretation we often refer to the notation of the former mentioned authors. Intuitively, a wavelet transform of a discrete time series can be described as a simultaneously mapping of the time series into a time

¹⁰The latter is built on the FHFA purchase-only index, the FHFA all-transactions index, and includes county recorder data licensed from CoreLogic as well as from DataQuick Information Systems; thus this HPI depicts also the high end of the market. These data, like the S&P/Case-Shiller data, reflect selling prices recorded at county assessor as well as recorder offices and therefore include not only the Enterprise data from the Federal Home Loan Mortgage Corporation (Freddie Mac) and the Federal National Mortgage Association (Fannie Mae). The underlying homes include those with sub-prime loans, jumbo mortgages, U.S. Department of Veterans Affairs (VA) home loans, Federal Housing Administration (FHA) endorsed loans, and other types of financing arrangements.

¹¹The FHFA expanded HPI is available for the 50 largest MSAs and their regarding MSADs, in contrast to the S&P/Case-Shiller Home Price Index, where only 20 MSA indices are available, and thus the former mentioned index represents the rural areas better than the latter mentioned index.

¹²The U.S. Census Bureau's X-12-ARIMA procedure is the method used by the FHFA for seasonal adjustment. They employ the automated ARIMA model-selection algorithm in X-12, which searches through a series of seasonality structures and selects the first that satisfies the Ljung-Box test for serial correlation. See the FHFA FAQs.

¹³For a detailed description of the properties of wavelets see also Chiann and Morettin (1998).

and frequency function. Being more precise, for a discrete time series $z(t) \in \mathcal{L}^2(\mathbb{R})$, the CWT with respect to the wavelet ψ is given by the following function of two variables τ and s :

$$W_z(\tau, s) = \int_{-\infty}^{+\infty} z(t) \frac{1}{\sqrt{|s|}} \hat{\psi}^* \left(\frac{t - \tau}{s} \right) dt. \quad (4.1)$$

The time domain position of the wavelet is denoted by τ , whereas its position in the frequency domain is given by s . τ controls the location of the wavelet. If $|s| > 1$, the wavelet is stretched, whereas for $|s| < 1$, the wavelet is compressed. Contrary to the Fourier transform, the CWT provides us with a time localization of the time series. Finally, the $\hat{\psi}^*$ denotes the use of complex conjugation.

The choice of the so-called mother wavelet $\psi(t)$ is an important task. The relevant literature suggests several wavelet functions which offer different characteristics (e.g. Mexican Hat, Haar, Morlet etc.).¹⁴ The choice of the "correct" wavelet depends on the application the researcher has in mind. In the present study, we are particularly concerned with the task to study the synchronism of regional real estate prices. Hence, we have to search for a complex-valued wavelet as we are primarily interested in both, the phase and the amplitude of the considered time series. Being the most popular complex-valued wavelet, we choose the Morlet wavelet as advised by Aguiar-Conraria and Soares (2011b) because this features some pleasant attributes, which considerably simplifies the interpretation of the results which we will derive below.¹⁵ For instance, the Morlet wavelet has an optimal joint time-frequency concentration in the sense that the Heisenberg uncertainty attains the minimum possible value, so that this complex-valued wavelet is quite well localized in both time and frequency space.¹⁶ Moreover, the Morlet wavelet consists of a complex exponential modulated by a Gaussian, where we set the central angular frequency ω_0 equal to 6, so that the terms scale and period can be used interchangeably.¹⁷ To test for significance, a Monte Carlo simulation with 10,000 draws of white noise time series was used.¹⁸

¹⁴See for example Farge (1992), Weng and Lau (1994, 1995), and Torrence and Compo (1998).

¹⁵The introduction of the Morlet wavelet can be traced back to the work accomplished by Goupillaud et al. (1984), who used the wavelet transform to represent a one-dimensional seismic data set as a two-dimensional function of both time and frequency. See also Weng and Lau (1994).

¹⁶For instance see Jevrejeva et al. (2003), Aguiar-Conraria et al. (2008, 2012a, 2012b), and Aguiar-Conraria and Soares (2014). See also Goupillaud et al. (1984), Debauchies (1990), Farge (1992), and Weng and Lau (1995).

¹⁷See Torrence and Webster (1999) and Cazelles et al. (2007).

¹⁸See also Torrence and Webster (1999).

For a time series $\{z_k, k = 0, 1, \dots, K-1\}$ with K observations, we obtain the discrete version of equation (4.1) with time step δt as:

$$W_l^z(s) = \frac{\delta t}{\sqrt{s}} \sum_{k=0}^{K-1} z_k \hat{\psi} \left((k-l) \frac{\delta t}{s} \right), \text{ for } l = \{0, 1, \dots, K-1\}. \quad (4.2)$$

The wavelet power spectrum can be computed as $|W_k^z|^2$. For two time series, z_k and y_k , the cross wavelet transform can be analogously represented as:

$$W_k^{zy} = W_k^z \hat{W}_k^y,$$

whereas the cross wavelet power is defined as $|W_k^{zy}|$. Intuitively, for an univariate time series environment, economic turbulences, such as the recent financial and economic crisis or the bursting of the housing bubble in the U.S. in 2007 should be directly associated with higher wavelet power as the wavelet power spectrum reflects the variance of these time series. In other words, the wavelet transform is able to capture such local events in time.¹⁹ Nevertheless we are dealing with finite time series, so that there will be so-called edge effects at the beginning and end of the wavelet power spectrum. This region is designated by Torrence and Compo (1998) as the cone of influence (COI). One possible solution is to pad zeros to the considered time series before doing the wavelet transform.²⁰

For our analysis, we require an instrument which quantifies and detects appropriately the relationship between two variables, namely the coherence between the FHFA expanded HPI for every considered MSA and the national counterpart. The wavelet coherence, which is defined as

$$R_k(s) = \frac{|S(s^{-1}W_k^{zy}(s))|}{S(s^{-1}|W_k^z|)^{0.5}S(s^{-1}|W_k^y|)^{0.5}}, \quad (4.3)$$

helps us to shed light into the frequency-time dependencies between those two variables. S represents a smoothing parameter, both in scale and time (see Torrence and Webster (1999), Cazelles et al. (2007), and Aguiar-Conraria et al. (2008)).

Finally, to obtain an answer to our question whether an MSA specific HPI leads the U.S. HPI or vice versa, we employ the concept of phase difference as this provides us with the delay of the oscillations of our two time series as a function of frequency.²¹

¹⁹See Fan and Gençay (2010).

²⁰For a more detailed contemplation of these errors see for example Weng and Lau (1994) as well as Torrence and Compo (1998).

²¹See Torrence and Webster (1999), Bloomfield et al. (2004), and Cazelles et al. (2007).

The phase difference reads as

$$\phi_{z,y} = \arctan \left[\frac{\mathcal{I}(W_k^{zy})}{\mathcal{R}(W_k^{zy})} \right], \quad (4.4)$$

with $\phi_{z,y} \in [-\pi, \pi]$. $\mathcal{R}(\cdot)$ denotes the real part and $\mathcal{I}(\cdot)$ represents the imaginary part of a complex number. Intuitively, if two time series move together at a given frequency, the phase difference turns out to be zero. For $\phi_{z,y} \in [0, \frac{\pi}{2}]$, the two time series move in phase, whereas z leads y . If $\phi_{z,y} \in [-\frac{\pi}{2}, 0]$, y leads z . For $\phi_{z,y} \in [\frac{\pi}{2}, \pi]$, the series are negatively correlated, whereas y leads z . For $\phi_{z,y} \in [-\pi, -\frac{\pi}{2}]$, both series are also negatively correlated, with z leading y .²² Intuitively, the phase difference helps us to clearly identify the delay between two time series' oscillations.

However, as our main impetus is the analysis of regional house price convergence, our introduced methodology so far suffers from the fact that it cannot detect possible regional clusters, i.e. MSAs which can be grouped because they behave similar in terms of their underlying house price evolution. In turn, to have that kind of information would be very attractive because this helps us to answer the question how clusters behave before and after the bursting of the U.S. housing bubble and whether or not all clusters exhibit the same house price dynamics as embodied by the i.e. national HPI. For that reason we have to introduce a method which allows us to compare the wavelet spectra between regions. We introduce this tool in the next subsection.

4.2 Dissimilarities between MSAs: a Wavelet Interpretation

As we draw our attention towards the investigation of regional house price convergence in the U.S., we need a method which allows us to exactly quantify the evolution of the dissimilarity of regional house price dynamics between MSAs during time, frequency and space. In technical words, we require a procedure which provides us with a pairwise similarity measure by comparing the wavelet transform of two time series which embody the house price dynamics. It is important to note that a direct comparison of the wavelets is not suitable, because the existence of regions of low power may lead to a biased dissimilarity measure (see Aguiar-Conraria et al. (2013)). Hence, we follow the suggestion made by Aguiar-Conraria and Soares (2011b) and perform a singular value decomposition (SVD) of the matrix $W_x W_y^H$, where W_y^H represents the conjugate transpose of W_y . This enables us to focus on high-power time-frequency regions only. With the SVD, we extract leading patterns $\mathbf{l}_x^k$ and $\mathbf{l}_y^k$ for $k = \{1, \dots, K\}$ as well as

²²For instance see Aguiar-Conraria et al. (2012a, 2012b, and 2014).

singular vectors $\mathbf{u}_x^k$ and $\mathbf{u}_y^k$ for $k = \{1, \dots, K\}$ which can be used to approximate the original matrices as

$$W_x \approx \sum_{k=1}^K \mathbf{u}_x^k \mathbf{l}_x^k \quad (4.5)$$

and

$$W_y \approx \sum_{k=1}^K \mathbf{u}_y^k \mathbf{l}_y^k. \quad (4.6)$$

Finally, we compute the distance between two MSAs, x and y , as

$$\Gamma(W_x, W_y) = \frac{\sum_{k=1}^K \xi_k^2 [d(\mathbf{l}_x^k, \mathbf{l}_y^k) + d(\mathbf{u}_x^k, \mathbf{u}_y^k)]}{\sum_{k=1}^K \xi_k^2}, \quad (4.7)$$

where $d(\mathbf{a}, \mathbf{b})$ is the distance between two vectors $\mathbf{a}$ and $\mathbf{b}$ computed by calculating the Hermitian angle between the consecutive points of the two vectors and taking the mean of those values.²³ The Hermitian angle Ψ_H can be calculated from $\cos(\Psi_h) = \frac{|\langle \mathbf{a}, \mathbf{b} \rangle_{\mathbb{C}}|}{\|\mathbf{a}\| \|\mathbf{b}\|}$ with the Hermitian inner product $\langle \mathbf{a}, \mathbf{b} \rangle_{\mathbb{C}} = \mathbf{a}^H \mathbf{b}$ and the corresponding norm $\|\mathbf{a}\| = \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle_{\mathbb{C}}}$ and $\|\mathbf{b}\| = \sqrt{\langle \mathbf{b}, \mathbf{b} \rangle_{\mathbb{C}}}$, respectively. Finally, ξ_k for $k = \{1, \dots, K\}$ are obtained from the SVD.

As we use the Hermitian angle, based on equation (4.7), we obtain distance measures between $100 \times [0, \frac{\pi}{2}]$. A value close to zero implies that for two MSAs, their housing cycles are highly synchronized, which in turn means that their individual contribution of housing cycles at every observed frequency to the total variance of house price development is similar. Moreover, the contribution is also synchronized over time, and, finally, the up- and downswing of the housing market occur simultaneously for two observed MSAs.

Next, we use the introduced tools to (i) examine the cyclicity of regional HPIs based on the wavelet power spectra, (ii) conduct a MSA to MSA comparison with the wavelet spectral distances and define MSA specific clusters which exhibit most similar cyclicity, and (iii) employ the cross wavelet and phase difference methodology to investigate whether there is evidence for a more pronounced synchronization pattern among MSAs after the bursting of the U.S. housing bubble at the beginning of 2007.

²³We also use the Euclidian angle which deliver similar results.

5 Results

In this section we present our results. As mentioned in the introduction to this paper, our main impetus of the study is to revisit the cyclicity of U.S. housing prices with a focus towards geography. In particular, we want to give an answer to the following questions: First, what can we say about the geography of housing price changes? Second, which clusters of U.S. MSAs emerge in terms of synchronicity of housing price cycles, both with the national U.S. housing price cycle as well as among the 40 largest MSAs? And finally, do we observe a MSA wide convergence of housing cycles after the bursting of the U.S. housing bubble? To give an answer to these questions, we employ the above introduced wavelet tools. In this context, we want to stress that this study does not attempt to answer the question what factors can explain the cyclicity of the housing cycle we observe, but rather we want to discover whether regional entities became more or less synchronized in terms of their housing price development.

5.1 Hierarchical Tree Clustering and Scaling Map Analysis

As mentioned in section 3, we employ data for 40 U.S. MSAs from the first quarter of 2001 until the last quarter of 2013. Hence, our data set comprises the complete housing boom-bust phase as well as the recovery phase of the housing market. Our empirical strategy is as follows: First, we compute the dissimilarity between the house price cycle in the 40 MSAs based upon the dissimilarity index introduced with equation (4.7). Hence, we arrive at a 40×40 dissimilarity matrix, which is hard to interpret.²⁴ Therefore, we try to re-express the information contained in the dissimilarity index with a cluster analysis. We follow Aguiar-Conraria et al. (2013) and perform a hierarchical tree clustering. The idea we follow is to cluster MSAs which exhibits similar housing cycles. Our starting point is to employ the entire set of states and group them stepwise round by round, whereas in every round a new dissimilarity matrix has to be computed. The computation stops when on the last step only one cluster appears, which by construction contains the entire sample as members. Therefore, the computation is tremendous and is increasing non-linearly with the number of regions one considers in the analysis. Figure (5.1) shows the result of the clustering exercise.

The difficulty of this kind of analysis is that there is a certain DOF, how many clusters the researcher wants to define. One can define four clusters which are visualized with different colours in figure (5.1). We follow this classification, but however, we have

²⁴The dissimilarity index matrix can be found in tables (7.3)-(7.5) stored the appendix.

Figure 5.1: Hierarchical Cluster Tree

further decided to split the big red cluster into two sub-clusters. As it will turn out, this strategy allows us to distill interesting details regarding the geographic distribution of house price cycles. The big two red clusters contain MSAs which are significantly synchronized with the national housing market. These two clusters are followed by a green, blue, and black cluster which each contain regions which are not synchronized with the national housing price cycle. It is interesting to observe which MSAs the blue, green, and black clusters each contain. For instance, the blue and black cluster exclusively consist of so-called rust belt regions, such as Indianapolis-Carmel-Anderson, Kansas City or Philadelphia-Camden-Wilmington. Also interesting is the composition of MSAs which define the green cluster. This cluster consists of all MSAs in Texas except Austin-Round Rock. That is, however, not surprising. Beside others, Glaeser and Gyourko (2002) have shown that there is a strong causal relationship between higher housing prices and prescriptive land use policies. This also confirms the findings by Huang and Tang (2012), who point out that there is a significant link between supply inelasticity and price declines during the recent bust of the U.S. housing market. For instance, before the crash of the housing market, the U.S. averaged median housing price was \$ 417,800, which is remarkably higher compared to less regulated areas,

such as Houston-The Woodlands-Sugar Land and Dallas-Forth Worth-Arlington with an average housing price of \$159,300. However, this does not hold true for every Texan region, as shown by our analysis. Austin-Round Rock belongs to the first red sub-cluster, because its house price cycle is highly more synchronized with Charlotte-Concord-Gastonia than with the remaining MSAs in Texas, which can be seen by consulting figure (5.2). Here we try to reduce the information of the dissimilarity matrix to a two-column matrix, which contains the position of each MSA on two orthogonal axes. This allows us to graphically represent each MSA on a two-dimensional map. However, one has to keep in mind that due to the fact that we do not use Euclidian distances, these measures are not perfectly accurate. Hence, we interpret the results in an ordinal manner. Nevertheless, figure (5.2) impressively confirms the representation of the dendrogram by arguing that the sample of U.S. MSAs can be well represented by four clusters. We marked each of them with colored shaded areas which corresponds to the dendrogram representation.

Figure 5.2: Scaling Map

So far our analyses regarding housing cycles are rather descriptive. In order to

assess whether housing cycles synchronization is statistically significant, we employ a Monte-Carlo analysis. In other words, per draw we fit an $ARMA(1, 1)$ -model under the null and construct new samples which are drawn from a Gaussian distribution with the same variance as the estimated error terms. We re-sample 10,000 times the distance for each pair of MSAs. We further calculate 1%, 5%, and 10% critical values. Our analysis shows that for 26 out of 40 MSAs we can reject the null hypothesis that these MSAs are not synchronized at the 5% significance level, whereas the number grows to 38 MSAs at the 10% significance level. Putting it differently, we conclude that only 8 MSAs are not synchronized with the national housing cycle during the first quarter of 2001 until the last quarter of 2013. Those 8 remaining MSAs belong either to the rust belt area, such as Columbus, Indianapolis-Carmel-Anderson, Kansas City, Philadelphia-Camden-Wilmington, Pittsburgh or St. Louis or are Texan MSAs, such as Dallas-Fort Worth-Arlington and Houston-The Woodlands-Sugar Land.

What we can thus observe is that the national U.S. housing cycle is indeed not independent of geography which is graphically fleshed out with figure (5.3), where we have shaded the different MSAs regarding the significance of synchronization with the U.S. housing cycles²⁵.

Figure 5.3: Map

In a nutshell, we can summarize our findings so far as follows: The great majority of U.S. MSAs are synchronized with the national housing cycle. From a technical time series perspective this implies that, first, the contribution of cycles at each considered

²⁵We have delegated the computation of the dissimilarity index to the appendix (see table (7.6)).

frequency to the total variance is similar for those MSAs. Second, this contribution takes place at the same time, and finally, the up- and downswings happen simultaneously. Nevertheless, we also show that some MSAs are not synchronized with the national housing cycles, and those MSAs can also be categorized geographically, if we focus simultaneously on the time as well as on the frequency domain.

5.2 A Wavelet Evaluation for MSA Cluster

In this subsection we make use of our cluster results derived from the dendrogram analysis in the preceding subsection to give an answer to the question whether the clustered house price cycles are synchronized at every business cycle frequency. We used cluster-specific, MSA-averaged house price growth rates.²⁶ Hence, we remove short- as well as long-run noise based upon a wavelet-based filter and estimate the power spectra between 1 and 8 years frequency. Further, we do not only report the wavelet power spectrum but also calculate Monte-Carlo based, 5% significance levels. The results of the exercise can be found summarized with figure (5.4). Note that group x stands for cluster x , for $x = \{1, 2, \dots, 5\}$, whereas USA refers to the national house cycle. Table (7.1) provides information to which cluster a MSA belongs. For the purpose of interpretation, let us first focus on the upper-right graph, which shows the estimated power at each time-frequency region. Regions with warm colors (yellow to red (extreme)) show high power, whereas regions associated with cold colors (green to blue (extreme)) represent regions with low power. The white lines show the maxima of a wavelet power spectrum's undulations, which hence, provides us with a direct estimate of the cycle period. The COI is represented with a thick black line. From this graph we observe that 4 year cycles are important to explain the total variance of the house price growth rate, which is plotted in the upper-left panel.

5.2.1 Estimated Wavelet Power Spectra for MSA Cluster

It is rather apparent that a similar pattern can be found for the second and fifth group, which together form the red cluster (see dendrogram analysis). However, the second, third and fourth group instead behave different. For those groups it seems that not only 4 year cycles are important but also the 2 year cycles at the end of the sample period after 2009 to explain the total variance of the house price growth rate.²⁷

²⁶Therein we follow Rappaport (2007). See also figure (7.2) in the appendix.

²⁷For the third group, the importance of 2 year cycles starts later, from 2011.

Figure 5.4: HPI Growth Rate and Wavelet Power Spectrum

Notes: *On the left:* Quarterly HPI growth rate. *On the right:* Wavelet Power Spectrum. The white lines show the maxima of a wavelet power spectrum's undulations. The COI is represented with a thick black line. Power is low for cold colors (starting with blue) and high for warm colors (ending with red).

Although interesting, the analysis so far is silent regarding the question whether or not our identified clusters share each others high regions of coherency. In other words, we should conjecture that right after the bursting of the housing bubble in 2007, which hit nearly all regions in the U.S. simultaneously, all considered MSAs behave as if they exhibit a high degree of synchronization. Moreover, our analysis conducted so far is silent to the question whether or not the national house cycle leads the regional house cycle. In particular, this kind of question cannot be satisfactorily answered with standard time series tools at hand. To overcome this obstacle is one of the appealing attributes of the wavelet analysis.

5.2.2 Phase Difference and Coherency Cross Wavelets for MSA Cluster

Broadly spoken, the coherency cross wavelets provides us with information on the correlation between two time series, whereas the phase difference tell us whether one time series leads another, or in other words, gives us information about the delay between oscillations of two specific time series. In the following we synonymously say that two specific regions are highly coherent given they are strongly locally correlated. To test whether the local correlation is statistically significant, we employ the same method as for calculating critical values for the MSA-specific house cycle distances.

Figure 5.5: Wavelet Coherency between the U.S. and Cluster x , where $x \in \{1, 2, 3, 4, 5\}$, Phases and Phase Difference at the 1-4 and 4-8 Frequency Bands.

Notes: *On the left:* Cross wavelet. The black (grey) contour shows the 5% (10%) significance level. Significance levels have been obtained by conducted Monte-Carlo-simulations with 10,000 replications each. The COI is represented with a green, dashed line. Coherency is low for cold colors (starting with blue) and high for warm colors (ending with red). *On the right:* The green line represents the U.S. phase, the blue line the cluster x 's phase. The red line represents the phase difference.

In figure (5.5) in the upper-left panel we see the wavelet coherency between each MSA HPI and the national HPI. On the upper-right hand we present two graphs. On

the left top, we plot the phase difference between the two house price growth rates for the 1-4 years frequency-band, whereas in the right top, we focus on the longer run, the 4-8 years confidence band. A phase difference - which is drawn in red - between zero and $\frac{\pi}{2}$ implies that the U.S. is leading; between $-\frac{\pi}{2}$ and zero, we have that the considered MSA is leading. As mentioned above, the series are out of phase (or negatively correlated) between $\frac{\pi}{2}$ and π , with a leading MSA, and, finally, between $-\pi$ and $-\frac{\pi}{2}$, where the national house cycle is leading. The reason we focus on both, the 1-4 year and 4-8 year frequency bands is borne due to the fact that first, we focus on the business cycle, and, second, we have found that the wavelet power marked concentration of energy at 2 and 4 year cycles, respectively.

Figure 5.6: Wavelet Coherency between Cluster 1 and Cluster x , where $x \in \{2, 3, 4, 5\}$, Phases and Phase Difference at the 1-4 and 4-8 Frequency Bands.

Notes: *On the left:* Cross wavelet. The black (grey) contour shows the 5% (10%) significance level. Significance levels have been obtained by conducted Monte-Carlo-simulations with 10,000 replications each. The COI is represented with a green, dashed line. Coherency is low for cold colors (starting with blue) and high for warm colors (ending with red). *On the right:* The green line represents the Cluster 1's phase, the blue line the Cluster x 's phase. The red line represents the phase difference.

Next, we turn to the econometric interpretation of our results. From a global perspective, figure (5.5) shows that there are obviously no significant episodes of inverse

co-movements, or in other words, the series move in phase for all episodes of significant coherencies. In particular, one common feature to every observed cluster is that after the bursting of the housing bubble in the end of 2006, we observe a high coherency region for 1-4 year as well as for 4-8 year frequency bands, because the downturn in the housing market affects all observed regions simultaneously, and, hence, all group specific cycles are highly synchronized with the national cycle.

A second level of indications which can be observed from figure (5.5) is that lead-lag episodes not only differ between clusters but also are time and frequency dependent. For instance, the second and third cluster exhibit a similar lead-lag pattern through the entire sample period and for all observed frequencies, whereas we conclude that before the housing boom each cluster cycle led the national cycle at the shorter-run (1-4 year frequency bands), whereas after 2007 there is evidence that the national cycle leads the cluster cycles. It is further interesting that this pattern is totally different in the longer-run (4-8 year frequency bands), where the national cycle always leads the cluster cycles. The fourth cluster instead behaves differently. At the high coherency regions, starting from 2006 onwards, we observe that the cluster cycles lead the national cycle in the short, as well as in the long-run. Cluster 1 and 5 behave rather similar in the long-run, whereas it seems that in the short-run the first cluster is slightly more synchronized than the fifth cluster.

In a nutshell, from this exercise we can take on board that the first and fifth cluster are more synchronized with the national cycle than the remaining clusters. This is particularly true in the longer-run. Hence, this conclusion further motivates us to conduct a wavelet coherency analysis comparing the first and fifth cluster with the remaining clusters. The results are presented in figure (5.6) and figure (5.7).

The overall conclusion we have drawn in the first exercise still holds: after the bursting of the housing bubble the clusters are highly synchronized at the shorter-run, but not in the longer run at 4-8 year frequency bands. If we draw our attention towards the phase differences, we observe that both in the short, as well as in the long-run, the fifth as well as the first cluster lead the cycles of the second and third cluster after 2007.

Hence, this exercise robustifies our analysis carried out at the national level. Finally, we directly compare the second and third cluster. Intuitively, we should expect a high region of coherency at the shorter-run. This is indeed the case, as shown by figure (5.8). We further observe that the second cluster leads the cycle of the third cluster until 2011 in the short-run, whereas in the longer-run this conclusion is reversed.

Figure 5.7: Wavelet Coherence between Cluster 5 and Cluster x , where $x \in \{1, 2, 3, 4\}$, Phases and Phase Difference at the 1-4 and 4-8 Frequency Bands.

Notes: *On the left:* Cross wavelet. The black (grey) contour shows the 5% (10%) significance level. Significance levels have been obtained by conducted Monte-Carlo-simulations with 10,000 replications each. The COI is represented with a green, dashed line. Coherency is low for cold colors (starting with blue) and high for warm colors (ending with red). *On the right:* The green line represents the Cluster 5's phase, the blue line the Cluster x 's phase. The red line represents the phase difference.

If we draw our results together, we observe that the clustered cycles are highly synchronized with each other at the shorter-run, whereas in the longer-run the coherency is rather low. The same conclusion can be made for a direct comparison between clusters and the national housing cycle. Hence, our results also confirm the result by Clark and Coggin (2009), who state that regional housing cycles are significantly shorter than business cycles.

6 Conclusion

In this contribution we employed wavelet tools to study the time and frequency-varying patterns of synchronization of housing cycles for U.S. MSAs. By employing both a novel method as well as a new data set we derive several interesting results regarding

Figure 5.8: Wavelet Coherency between Cluster 2 and Cluster 3, Phases and Phase Difference at the 1-4 and 4-8 Frequency Bands.

Notes: *On the left:* Cross wavelet. The black (grey) contour shows the 5% (10%) significance level. Significance levels have been obtained by conducted Monte-Carlo-simulations with 10,000 replications each. The COI is represented with a green, dashed line. Coherency is low for cold colors (starting with blue) and high for warm colors (ending with red). *On the right:* The green line represents the Cluster 2's phase, the blue line the Cluster 3's phase. The red line represents the phase difference.

the cyclicity of regional housing prices, from which we want to highlight the most important. First, we have shown that regional house price cycles and geography are closely related. In particular, we have identified several clusters which also have a clear geographic interpretation. Two clusters comprise regions located at the West and East coast of the U.S., whereas another two clusters comprise MSAs which belong to the rust belt or to Texas. Second, our analysis further provides comprehensive and significant evidence for a pronounced convergence pattern of U.S. housing prices after the bursting of the housing bubble in 2007 at the shorter-run (1-4 year frequency bands). This general conclusion, however, does not hold true for the longer-run, where the results are mixed. It seems that only the West and East Coast MSAs exhibit a strong coherence between each other and the national cycle even in the long-run. In a nutshell, we find that housing shocks cannot be exclusively seen to be a local phenomena. If this would be the case, shocks should operate locally, and hence, we shouldn't observe clusters as we do in our analysis, not for all but for some business cycle frequencies. Hence, this result stand in some contrast with the findings made by Ghent and Owyang (2010), but can be explained with the fact that these authors leave out the frequency domain heterogeneity of house cycles in their analysis.

Although our results enriches our understanding of the complex pattern of housing price cyclicity, our analysis also has some drawbacks. First of all, our sample only covers large MSAs. However, there is evidence that smaller MSAs have a direct and non-neglectful influence on house price development. However, due to data-limitation problems we cannot control for this issue. Further, we know that our analysis as well

as any other mentioned contribution in this paper neglects the existence of spatial correlation, which may influence our overall conclusions. Nevertheless, our methodology, in contrast to existing papers, allows a deeper and more diverse look on housing cycles coherence than standard economic time series tools, as we do not only focus on the time domain but also on the frequency domain.

7 Appendix

7.1 Data

Two traditional HPIs are, on the one hand the FHFA purchase-only HPI and on the other the S&P/Case-Shiller Home Price Index. The former is based on transactions involving conforming, conventional mortgages purchased or securitized by Fannie Mae or Freddie Mac, where only mortgage transactions on single-family properties are included, and is a weighted, repeat-sales index, meaning that it measures average price changes in repeat sales or refinancings on the same properties. Both indices employ the same fundamental repeat-valuations approach and are restricted to purchase prices in index calibration, but nevertheless there are a number of data and methodology differences.²⁸ Among the dissimilarities are:²⁹

- The S&P/Case-Shiller indices use information obtained from county assessor and recorder offices.
- The S&P/Case-Shiller indices are value-weighted, meaning that price trends for more expensive homes have greater influence on estimated price changes than other homes, whereas the FHFA’s purchase-only index weights price trends equally for all properties.
- The S&P/Case-Shiller National Home Price Index, for example, does not have valuation data from 13 states, while the FHFA’s U.S. index is calculated using data from all states.

To work around the first two problems, the FHFA constructed an expanded HPI. In general, the methodology to construct the FHFA expanded HPI is the same as is used in the construction of the standard purchase-only HPI, except that a supplemented data set is used for estimation.³⁰ Thus, the augmented data include sales price information from Fannie Mae and Freddie Mac mortgages. Furthermore, the augmented data also include two new information sources:³¹ (i) transactions records for houses with mortgages endorsed by FHA and (ii) county recorder data licensed from CoreLogic and

²⁸The repeat sales method was first proposed by Bailey, Muth, and Nourse (1963), and is extended later by Case and Shiller (1987, 1989). The FHFA all-transaction HPI also includes refinance appraisals in index calibration.

²⁹See also the FHFA FAQs.

³⁰For a detailed technical description of the OFHEO HPI, see for instance Calhoun (1996).

³¹See the FHFA FAQs.

from DataQuick Information Systems. The FHFA mentions that the licensed county recorder data do not include records in many U.S. counties, particularly rural ones, but to ensure that the addition of these data to the estimation sample does not unduly bias index estimates toward price trends in urban areas, the expanded data HPI for certain states is estimated by weighting price trends in areas with CoreLogic and DataQuick Information Systems coverage and other areas.

FIPS	MSAs	Abbreviations	Region	Group
12060	Atlanta-Sandy Springs-Roswell (GA)	At	3	1
12420	Austin-Round Rock (TX)	Au	3	5
12580	Baltimore-Columbia-Towson (MD)	Ba	3	1
14460	Boston-Cambridge-Newton (MA-NH)	Bo	1	1
16740	Charlotte-Concord-Gastonia (NC-SC)	Cha	3	5
16980	Chicago-Naperville-Elgin (IL-IN-WI)	Chi	2	1
17140	Cincinnati (OH-KY-IN)	Ci	2	1
17460	Cleveland-Elyria (OH)	Cl	2	1
18140	Columbus (OH)	Co	2	1
19100	Dallas-Fort Worth-Arlington (TX)	Da	3	2
19740	Denver-Aurora-Lakewood (CO)	Den	4	1
19820	Detroit-Warren-Dearborn (MI)	Det	2	5
26420	Houston-The Woodlands-Sugar Land (TX)	Ho	3	2
26900	Indianapolis-Carmel-Anderson (IN)	In	2	3
27260	Jacksonville (FL)	Ja	3	5
28140	Kansas City (MO-KS)	Ka	2	3
29820	Las Vegas-Henderson-Paradise (NV)	LV	4	5
31080	Los Angeles-Long Beach-Anaheim (CA)	LA	4	5
33100	Miami-Fort Lauderdale-West Palm Beach (FL)	Mia	3	5
33340	Milwaukee-Waukesha-West Allis (WI)	Mil	2	1
33460	Minneapolis-St. Paul-Bloomington (MN-WI)	Min	2	1
34980	Nashville-Davidson-Murfreesboro-Franklin (TN)	Na	3	5
35620	New York-Newark-Jersey City (NY-NJ-PA)	NYJ	1	1
36740	Orlando-Kissimmee-Sanford (FL)	Or	3	5
37980	Philadelphia-Camden-Wilmington (PA-NJ-DE-MD)	Phi	1	1
38060	Phoenix-Mesa-Scottsdale (AZ)	Pho	4	5
38300	Pittsburgh (PA)	Pi	1	4
38900	Portland-Vancouver-Hillsboro (OR-WA)	Po	4	5
39300	Providence-Warwick (RI-MA)	Pr	1	1
40140	Riverside-San Bernardino-Ontario (CA)	Ri	4	5
40900	Sacramento-Roseville-Arden-Arcade (CA)	Sa	4	5
41180	St. Louis (MO-IL)	SL	2	1
41700	San Antonio-New Braunfels (TX)	SA	3	2
41740	San Diego-Carlsbad (CA)	SD	4	5
41860	San Francisco-Oakland-Hayward (CA)	SF	4	5
41940	San Jose-Sunnyvale-Santa Clara (CA)	SJ	4	5
42660	Seattle-Tacoma-Bellevue (WA)	Se	4	5
45300	Tampa-St. Petersburg-Clearwater (FL)	Ta	3	5
47260	Virginia Beach-Norfolk-Newport News (VA-NC)	Vi	3	1
47900	Washington-Arlington-Alexandria (DC-VA-MD-WV)	Wa	3	5

Table 7.1: MSAs and Classifications.

Notes: Regions are indicated as 1 (Northeast), 2 (Midwest), 3 (South), and 4 (West). Group $\hat{=}$ Cluster.

7.2 Real house prices dynamics: MSVAR results for the U.S.

Table (7.2) reports the results for a MSVAR(1) regression for the U.S. with two states, s_{lt} , $l = \{1, 2\}$: State s_{1t} corresponds to the recession state, whereas state s_{2t} denotes the high growth state. Newey-West HAC standard errors are used. The number of AR-terms is selected on sequential t and F -Tests together with a focus on maximizing the likelihood function. It turns out that a MSVAR(1) fits the data best. 1%, 5%, and 10% significance levels are reported as *, **, and ***, respectively.

Variable	Coefficient
$\mu_{i,s_{1t}}$	-1.70*
$\mu_{i,s_{2t}}$	1.39*
σ_1	1.48*
σ_2	0.56*
π_{11}	0.94*
π_{21}	0.03*
duration s_{1t}	17.17
duration s_{2t}	35.27
No. of AR-terms	1
Log-lik.	-71.68
Obs.	50

Table 7.2: MSVAR(1) regression for the U.S. with two states.

Notes: Newey-West HAC standard errors are used. 1%, 5%, and 10% significance levels are indicated as *, **, and ***, respectively.

We observe that the average duration in state 1 is about 17 quarters, with a mean real house price growth rate of -1.7%. The average duration in the high growth state is considerably higher, with about 35 quarters. The mean real house price growth rate in the high growth state is estimated with 1.39%, which is significantly and considerably larger compared to the mean growth rate in the low growth state.

From the estimation results obtained above, we can easily derive the matrix of transition probabilities as:

$$\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} = \begin{pmatrix} 0.94 & 0.06 \\ 0.03 & 0.97 \end{pmatrix}. \quad (7.1)$$

From this matrix we can derive that both the low growth state as well as the high

growth state are highly persistent.

Figure 7.1: MSVAR(1) for the U.S.

We further compute the smoothed transition probabilities by applying a smoothing algorithm by Kim (1994), using the entire sample information for the complete time span and contrast them with the NBER recession indicator as well as with the real house price growth rate. We see that the standard deviation of the real house price growth rate in recession times (smoothed probabilities close to one) is significantly larger compared to high growth times, as also supported by our regression analysis. In other words, the low growth state is more volatile compared to the high growth state. Despite its simplicity, the MSVAR(1) model matches fairly well the bust of the house price bubble in 2006 as the associated smoothed probability indicating a low growth state is close to 1 in this period, as can be seen from figure (7.1).

7.3 Estimated pairwise dissimilarity indices for MSAs

The following tables show the estimated pairwise dissimilarity indices for each MSA.

	At	Au	Ba	Bo	Cha	Chi	Ci	Cl	Co	Da	Den	Det	Ho	In	Ja
At	0.00	25.44	14.41***	18.54	14.26**	9.82**	18.64**	15.13**	18.12**	24.31	14.56***	11.29**	23.01	30.39	14.66***
Au		0.00	25.81	35.13	13.83*	25.60	40.82	33.57	41.25	22.43***	37.18	20.88***	21.75***	46.63	18.43**
Ba			0.00	9.89*	17.28***	6.13*	20.13***	14.01**	21.94***	32.89	17.03***	9.82**	29.39	37.15	9.59**
Bo				0.00	25.30	12.04**	15.02**	12.39*	18.03**	32.69	14.17**	17.77***	32.58	32.47	18.22***
Cha					0.00	15.15**	27.40	22.16***	27.71	22.47	24.44	13.08**	17.23**	34.25	13.27**
Chi						0.00	19.23***	14.39*	21.48***	30.48	14.64***	6.87*	28.03	36.13	8.66*
Ci							0.00	9.56*	8.35*	19.56***	10.35*	24.60	25.98	23.60**	27.20
Cl								0.00	12.79*	21.44***	9.40*	18.58***	25.49	25.22***	20.84***
Co									0.00	24.43	14.31**	26.69	26.82	20.16**	29.18
Da										0.00	18.52***	32.76	12.32*	25.61***	33.45
Den											0.00	19.94	25.57	23.26**	22.32
Det												0.00	29.84	41.29	4.96*
Ho													0.00	25.29***	30.79
In														0.00	43.74
Ja															0.00
Ka															
LV															
LA															
Mia															
Mil															
Min															
Na															
NYJ															
Or															
Phi															
Pho															
Pi															
Po															
Pr															
Ri															
Sa															
SL															
SA															
SD															
SF															
SJ															
Se															
Ta															
Vi															
Wa															

Table 7.3: Estimated pairwise dissimilarity index between housing cycles based on equation (4.7) and multiplied by 100.

Notes: As we use the Hermitian angle, we obtain distance measures between $100 \times [0, \frac{\pi}{2}]$. A value close to zero implies that two MSAs imply that their housing cycles are highly synchronized. 1%, 5%, and 10% significance levels are indicated as *, **, and ***, respectively.

	Ka	LV	LA	Mia	Mil	Min	N'a	NYJ	Or	Phi	Pho	Pi	Po	Pr	Ri
At	24.53	17.61	16.29	16.52	14.51**	11.26**	11.35*	15.79***	20.17	21.80	21.65	34.94	15.99**	17.81	16.88
Au	43.31	19.55**	19.94**	16.90**	35.94	28.94	16.71**	34.88	14.94**	43.70	13.51*	51.57	14.37*	32.91	18.88**
Ba	30.51	12.79***	10.23**	12.94**	11.98*	6.57*	12.55**	8.58*	17.33	20.15	19.17	46.42	14.40***	7.64*	12.07**
Bo	26.17	21.10	18.42	21.23	7.80*	9.94*	20.99	5.38*	25.53	12.52**	27.22	42.37	22.86	3.71*	20.37
Cha	31.69	15.60**	15.03**	14.07**	23.74	18.22**	6.35*	23.13	15.31**	31.36	15.82**	36.28	11.76*	23.40	15.02**
Chi	29.58	11.96**	9.71**	12.09**	11.79*	3.71*	10.05*	9.77**	16.50	20.46	18.24	44.55	12.85**	11.16**	11.38**
Ci	15.44**	32.05	27.25	30.22	8.81*	16.80**	25.11	14.69**	34.96	8.82*	35.76	26.11	29.05	16.24*	29.59
Cl	18.56**	25.65	20.79***	24.16	8.91*	12.98**	18.75***	12.84*	28.65	13.53**	29.82	28.96	23.06	12.11*	23.13
Co	13.10*	33.40	28.69	32.25	12.26*	19.87**	25.70	17.23**	36.73	11.07*	38.09	26.27	31.79	18.94**	30.89
Da	21.74***	37.67	33.06	33.26	28.22	28.87	25.75	33.16	36.02	30.57	33.62	32.12	30.09	33.02	34.10
Den	19.02**	26.77	23.08	25.08	9.62*	13.41*	21.17	13.17**	29.56	14.21**	30.58	35.63	24.92	15.72***	25.02
Det	34.88	7.17*	7.52*	7.36*	18.17***	9.21**	7.97*	16.12	11.11**	27.00	12.74***	47.95	7.31*	16.40***	7.71*
Ho	20.85***	33.32	30.12	31.17	27.70	27.69	22.06	30.35	33.70	31.65	32.47	27.00	28.79	30.91	30.95
In	13.29*	45.43	42.91	45.93	28.99	35.05	36.10	33.67	50.55	25.55***	51.17	26.55***	45.30	33.24	44.07
Ja	37.67	4.61*	3.92*	3.92*	19.95***	11.07**	8.27*	16.69***	8.25*	28.32	10.15**	52.63	5.27*	16.28***	4.00*
Ka	0.00	40.70	36.12	40.15	20.86***	27.40	31.65	25.89	45.13	18.14**	46.15	23.34***	39.30	26.35	37.78
LV	0.00	0.00	3.67*	4.38*	23.75	13.80***	11.17*	19.50	7.05*	32.46	8.92**	52.98	6.08*	18.66	2.60*
LA			0.00	5.00*	20.55***	11.33**	10.54*	17.05	8.88**	28.85	10.85**	48.77	6.99*	15.97***	2.23*
Mia				0.00	23.20	14.25***	9.84*	19.93	4.68*	31.58	6.59*	55.15	4.16*	19.34	3.46*
Mil				0.00	0.00	10.60*	20.15***	6.39*	27.71	9.57*	29.31	34.39	23.62	8.98*	22.60
Min						0.00	13.20**	7.55*	18.60	19.06***	20.43	39.82	15.26***	9.12*	13.11**
Na							0.00	19.08***	12.31**	28.48	13.31**	41.13	7.85*	19.23***	10.70*
NYJ								0.00	24.13	13.42**	26.00	41.39	21.51	5.49*	18.99
Or									0.00	36.03	2.17*	60.67	5.14*	23.65	6.93*
Phi										0.00	37.64	32.27	32.18	14.77**	31.03
Pho											0.00	60.20	6.41*	25.34	8.88**
Pi												0.00	50.30	43.46	51.10
Po													0.00	21.01	5.88*
Pr														0.00	17.92
Ri															0.00
Sa															
SL															
SA															
SD															
SF															
SJ															
Se															
Ta															
Ta															
Vi															
Wa															

Table 7.4: Continued: Estimated pairwise dissimilarity index between housing cycles based on equation (4.7) and multiplied by 100.

Notes: As we use the Hermitian angle, we obtain distance measures between $100 \times [0, \frac{\pi}{2}]$. A value close to zero implies that two MSAs imply that their housing cycles are highly synchronized. 1%, 5%, and 10% significance levels are indicated as *, **, and ***, respectively.

	Sa	SL	SA	SD	SF	SJ	Se	Ta	Vi	Wa
At	15.85	19.15***	22.70**	17.31	16.32	19.62	17.08	17.26	16.07***	18.28
Au	23.04***	40.46	24.41***	22.27***	18.56**	14.58**	13.92*	16.85**	31.31	17.56**
Ba	9.17**	20.32***	24.08**	8.27*	11.83***	14.94***	16.43	13.03***	5.74*	11.15**
Bo	15.90***	14.27**	25.46***	16.37***	19.72	23.38	25.03	21.14	5.82*	19.20
Cha	17.10**	28.68	19.13**	16.07**	15.10**	14.68**	12.34**	14.44**	21.04***	15.09**
Chi	8.07*	19.68***	24.87***	9.59**	11.04**	15.34***	14.78***	12.13**	8.68*	12.23**
Ci	27.10	4.99*	24.12**	27.26	29.13	31.71	31.60	30.00	16.31**	28.20
Cl	21.15***	10.68*	21.65**	20.99***	22.76	25.33	25.02	23.94	10.97*	22.25
Co	29.36	7.88*	24.46***	28.21	31.00	33.22	33.13	32.33	18.70**	30.79
Da	36.27	21.27***	17.98**	33.86	32.67	30.13	30.62	33.34	34.29	32.32
Den	21.87	11.09*	23.24**	22.83	23.61	26.78	27.11	24.80	15.33**	23.55
Det	7.27*	25.53	28.02***	9.70**	7.35*	11.88**	8.83**	7.67*	14.06**	9.65**
Ho	34.00	27.34	16.09*	29.11	31.32	28.75	28.66	31.57	30.67	31.64
In	42.54	21.18**	24.18***	38.00	41.30	40.64	45.99	46.41	34.25	43.88
Ja	5.31*	27.89	27.73***	6.66*	4.00*	7.97*	7.44*	3.86*	14.66**	5.78*
Ka	37.48	13.92*	24.26***	33.71	36.94	37.31	40.17	40.61	27.91	38.95
LV	5.37*	32.23	31.64	6.21*	2.87*	7.58*	7.17*	3.88*	18.09***	6.85*
LA	3.95*	27.63	29.04	4.08*	3.25*	7.54*	8.52**	4.72*	15.35***	5.27*
Mia	7.22*	30.89	29.04	8.34**	4.21*	5.82*	4.88*	1.55*	18.09***	4.56*
Mil	19.21***	9.45*	22.29**	19.31***	22.22	25.51	25.91	23.22	7.82*	21.58
Min	9.41**	17.43**	25.78***	10.60**	12.90**	17.45***	17.07	14.31***	8.03*	14.00***
Na	11.83**	25.99	22.23***	12.17**	10.45*	12.01*	9.17*	10.03*	16.68**	11.61*
NYJ	15.18***	14.49*	25.61***	14.92***	18.95	22.68	23.52	19.83	5.51*	18.21
Or	11.27***	35.59	31.35	11.95**	7.59*	6.06*	4.50*	4.83*	22.43	8.06*
Phi	27.47	6.44*	25.94***	27.63	30.73	33.86	34.46	31.46	15.33**	29.43
Pho	13.04***	36.66	30.85	13.97***	6.99*	5.52*	5.52*	6.63*	24.15	9.73**
Pi	53.61	28.36	37.02	45.21	52.27	51.23	49.85	54.79	42.88	53.34
Po	9.33**	30.56	27.57***	10.41**	5.86*	6.70*	2.47*	4.04*	19.32	6.79*
Pr	14.17***	15.62**	25.19***	13.37**	17.73	21.14	22.88	19.32	4.44*	17.27
Ri	5.33*	29.86	30.01	5.64*	3.17*	6.55*	6.94*	3.04*	17.32***	5.36*
Sa	0.00	27.37	29.99	4.11*	4.80*	10.09**	11.07**	7.05*	13.97**	7.49*
SL		0.00	23.24***	27.01	29.28	31.66	32.63	30.82	16.57**	29.02
SA			0.00	28.11***	26.41***	28.99	29.12	23.56**	26.61***	
SD				0.00	6.07*	9.72**	11.78**	8.20**	13.39**	7.15*
SF					0.00	5.91*	7.43*	3.76*	17.03***	5.36*
SJ						0.00	7.36*	5.57*	20.19	5.61*
Se							0.00	5.35*	21.32	8.48**
Ta								0.00	18.11***	4.40*
Vi									0.00	16.23***
Wa										0.00

Table 7.5: Continued: Estimated pairwise dissimilarity index between housing cycles based on equation (4.7) and multiplied by 100.

Notes: As we use the Hermitian angle, we obtain distance measures between $100 \times [0, \frac{\pi}{2}]$. A value close to zero implies that two MSAs imply that their housing cycles are highly synchronized. 1%, 5%, and 10% significance levels are indicated as *, **, and ***, respectively.

7.4 Cluster analysis

The next figure shows the house price growth rate averaged over the identified clusters. We have used this data for our cluster-related, wavelet analysis.

Figure 7.2: Cluster-specific house price growth rates

7.5 Estimated pairwise dissimilarity indices between the U.S. housing cycles and MSAs

	At	Au	Ba	Bo	Cha	Chi	Ci	Cl	Co	Da
U.S.	11.51**	23.81***	5.47*	13.65**	14.80**	3.28*	20.81***	15.41**	23.40	30.21
	Den	Det	Ho	In	Ja	Ka	LV	LA	Mia	Mil
U.S.	16.50***	4.96*	28.63	37.76	6.06***	31.15	9.11**	7.02*	9.50**	14.37**
	Min	Na	NYJ	Or	Phi	Pho	Pi	Po	Pr	Ri
U.S.	5.20*	9.15*	11.97**	13.91***	23.03	15.74***	44.52	10.45**	12.00**	8.59***
	Sa	SL	SA	SD	SF	SJ	Se	Ta	Vi	Wa
U.S.	5.51***	21.31	25.80***	7.73*	8.17*	12.93**	12.38**	9.55**	10.13*	10.11**

Table 7.6: Estimated pairwise dissimilarity indices between the U.S. housing cycles and MSAs based on equation (4.7) and multiplied by 100.

Notes: As we use the Hermitian angle, we obtain distance measures between $100 \times [0, \frac{\pi}{2}]$. A value close to zero implies that housing cycles are highly synchronized. 1%, 5%, and 10% significance levels are indicated as *, **, and ***, respectively.

References

- [1] Aguiar-Conraria, Luís and Maria Joana Soares (2011a): Oil and the Macroeconomy: using wavelets to analyze old issues. *Empirical Economics*, 40, pp. 645-655.
- [2] Aguiar-Conraria, Luís and Maria Joana Soares (2011b): Business cycle synchronization and the Euro: A wavelet analysis. *Journal of Macroeconomics*, 33, pp. 477-489.
- [3] Aguiar-Conraria, Luís and Maria Joana Soares (2014): The Continuous Wavelet Transform: moving beyond uni- and bivariate analysis. *Journal of Economic Surveys*, 28, pp. 344-375.
- [4] Aguiar-Conraria, Luís, Nuno Azevedo, and Maria Joana Soares (2008): Using wavelets to decompose the time-frequency effects of monetary policy. *Physica A: Statistical Mechanics and its Applications*, 387, pp. 2863-2878.
- [5] Aguiar-Conraria, Luís, Pedro C. Magalhães, and Maria Joana Soares (2012a): Cycles in Politics: Wavelet Analysis of Political Time Series. *American Journal of Political Science*, 56, pp. 500-518.
- [6] Aguiar-Conraria, Luís, Manuel M. F. Martins, and Maria Joana Soares (2012b): The yield curve and the macro-economy across time and frequencies. *Journal of Economic Dynamics & Control*, 36, pp. 1950-1970.
- [7] Aguiar-Conraria, Luís, Pedro C. Magalhães, and Maria Joana Soares (2013): The nationalization of electoral cycles in the United States: a wavelet analysis. *Public Choice*, 156, pp. 387-408.
- [8] Bailey, Martin J., Richard F. Muth, and Hugh O. Nourse (1963): A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association*, 58, pp. 933-942.
- [9] Bloomfield, D. Shaun, R. T. James McAteer, Bruce W. Lites, Philip G. Judge, Mihalis Mathioudakis, and Francis P. Keenan (2004): Wavelet Phase Coherence Analysis: Application to a Quiet-Sun Magnetic Element. *The Astrophysical Journal*, 617, pp. 623-632.
- [10] Calhoun, Charles A. (1996): OFHEO House Price Indexes: HPI Technical Description. Office of Federal Housing Enterprise Oversight. Mimeo.
- [11] Case, Karl E. and Robert J. Shiller (1987): Prices of Single Family Homes Since 1970: New Indexes for Four Cities. *New England Economic Review*, September,

pp. 45-56.

- [12] Case, Karl E. and Robert J. Shiller (1989): The Efficiency of the Market for Single-Family Homes. *The American Economic Review*, 79, pp. 125-137.
- [13] Cazelles Bernard, Mario Chavez, Guillaume Constantin de Magny, Jean-Francois Guégan, and Simon Hales (2007): Time-dependent spectral analysis of epidemiological time-series with wavelets. *Journal of the Royal Society Interface*, 4, pp. 625-636.
- [14] Chiann, Chang and Pedro A. Morettin (1998): A wavelet analysis for time series. *Journal of Nonparametric Statistics*, 10, pp. 1-46.
- [15] Chowdhury, Rosen Azad and Duncan Maclellann (2014): Regional house price cycles in the UK, 1978-2012: a Markov switching VAR. *Journal of European Real Estate Research*, 7, pp. 345-366.
- [16] Clark, Steven P. and T. Daniel Coggin (2009): Trends, Cycles and Convergence in U.S. Regional House Prices. *Journal of Real Estate Finance and Economics*, 39, pp. 264-283.
- [17] Corradin, Stefano and Alessandro Fontana (2013): House Price Cycles in Europe. *ECB Working Paper Series* No 1613/November 2013.
- [18] Crowley, Patrick M. (2007): A Guide to Wavelets for Economists. *Journal of Economic Surveys*, 21, pp. 207-267.
- [19] Daubechies, Ingrid (1990): The Wavelet Transform, Time-Frequency Localization and Signal Analysis. *IEEE Transactions on Information Theory*, 36, pp. 961-1005.
- [20] Fan, Yanqin and Ramazan Gençay (2010): Unit root tests with wavelets. *Econometric Theory*, 26, pp. 1305-1331.
- [21] Farge, Marie (1992): Wavelet Transform and their Applications to Turbulence. *Annual Review of Fluid Mechanics*, 24, pp. 395-457.
- [22] Gabor, D. (1946): Theory of communication. *Journal of the Institution of Electrical Engineers*, 93, pp. 429-457.
- [23] Ghent, Andra C. and Michael T. Owyang (2010): Is housing the business cycle? Evidence from US cities. *Journal of Urban Economics*, 67, pp. 336-351.
- [24] Glaeser, Edward L. and Joseph Gyourko (2002): The Impact of Zoning on Housing Affordability. *NBER Working Paper* 8835.

- [25] Goetz, Martin, Luc Laeven, and Ross Levine (2014): Does the Geographic Expansion of Bank Assets Reduce Risk? *NBER Working Paper* 20758.
- [26] Goupillaud, P., A. Grossman, and J. Morlet (1984): Cycle-octave and related transforms in seismic signal analysis. *Geoexploration*, 23, pp. 85-102.
- [27] Huang, Haifang and Yao Tang (2012): Residential land use regulation and the US housing price cycle between 2000 and 2009. *Journal of Urban Economics*, 71, pp. 93-99.
- [28] Jevrejeva S., J. C. Moore, and A. Grinsted (2003): Influence of the Arctic oscillation and El Niño-Southern Oscillation (ENSO) on ice conditions in the Baltic Sea: The wavelet approach. *Journal of Geophysical Research*, 108, pp. 10/1-10/11.
- [29] Kim, Chang-Jin (1994): Dynamic linear models with Markov-switching. *Journal of Econometrics*, 60, pp. 1-22.
- [30] Kim, Sangbae and Francis Haeuck In (2003): The Relationship between Financial Variables and Real Economic Activity: Evidence From Spectral and Wavelet Analyses. *Studies in Nonlinear Dynamics & Econometrics*, 7, article 4.
- [31] Kindleberger, Charles P. and Robert Z. Aliber (2005): Manias, Panics, and Crashes: A History of Financial Crises, 5th Edition. John Wiley & Sons, New York.
- [32] Ko, Jun-Hyung and Chang-Min Lee (2015): International economic policy uncertainty and stock prices: Wavelet approach. *Economic Letters*, 134, pp. 118-122.
- [33] Maclellan, Duncan and Anthony O'Sullivan (2011): The global financial crisis: challenges for housing research and policies. *Journal of Housing and Build Environment*, 26, pp. 375-384.
- [34] Mandler, Martin and Michael Scharnagl (2014): Money growth and consumer price inflation in the euro area: a wavelet analysis. *Deutsche Bundesbank*, Discussion Paper No 33/2014.
- [35] Martin, Ron (2011): The local geographies of the financial crisis: from the housing bubble to economic recession and beyond. *Journal of Economic Geography*, 11, pp. 587-618.
- [36] Ramsey, James B. and Camille Lampart (1998a): The Decomposition of Economic Relationships by Time Scale Using Wavelets: Expenditure and Income. *Studies in Nonlinear Dynamics & Econometrics*, 3, pp. 23-42.

- [37] Ramsey, James B. and Camille Lampart (1998b): Decomposition of Economic Relationships by Timescale Using Wavelets: Money and Income. *Macroeconomic Dynamics*, 2, pp. 49-71.
- [38] Rappaport, Jordan (2007): A Guide to Aggregate House Price Measures. *Economic Review*, 92, pp. 41-71.
- [39] Torrence, Christopher and Gilbert P. Compo (1998): A Practical guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, 79, pp. 61-78.
- [40] Torrence, Christopher and Peter J. Webster (1999): Interdecadal Changes in the ENSO-Monsoon System. *Journal of Climate*, 12, pp. 2679-2690.
- [41] Weng, Hengyi and K.M. Lau (1994): Wavelets, Period Doubling, and Time-Frequency Localization with Application to Organization of Convection over the Tropical Western pacific. *American Meteorological Society*, September, pp. 2523-2541.
- [42] Weng, Hengyi and K.M. Lau (1995): Climate Signal Detection Using Wavelet Transform: How to Make a Time Series Sing. *Bulletin of the American Meteorological Society*, 76, pp. 2391-2402.