# Re-thinking the housing-macro nexus: housing market churn in aggregate demand formation and stability, a wavelets based analysis.

Bazil A. Sansom

**Abstract**

I take a data driven approach to investigating the relationship between housing and macroeconomic cycles for the US economy. I use wavelet power and coherence spectra for housing and GDP series to identify the spectral ranges of and relationship between housing and GDP cycles. To unpack these relationships I then employ a 'network discovery' procedure, deploying bivariate and partial wavelet based statistics combined with graph theoretic methods to identify a network projection of nontrivial relationships between a detailed set of housing, financial and macro variables over an identified cycle frequency range. Finally I make a detailed wavelet based time-frequency domain study of those links identified as key relationships by the network-discovery procedure. My results confirm the importance of housing in US macroeconomic cycles at low frequencies, and identify secondary housing market volumes (existing homes) as a global "source", thus candidate 'root cause' of low frequency cycles, propagating most directly to the wider economy via residential investment and durable-consumption. This raises the unexpected possibility that secondary housing market volumes - a variable unconsidered in the theoretical or empirical literature - may be an autonomous source of fluctuations, and lead a significant low-frequency macroeconomic cycle.

## 1 Introduction

Study of the macroeconomic cycle has long emphasised the 'business cycle'. However not only was the US housing market conspicuously implicated in the 2007-8 crisis, but a growing body of mostly statistically driven empirical work seems to implicate housing much more widely in economic cycles and instability. There is a now well established long running association between housing market boom-bust and systemic banking crises (Borio and Drehmann 2009; Crowe et al. 2013; IMF 2012; IMF et al. 2012; Kindleberger and Aliber 2005). Credit growth is a powerful predictor of financial crises in advanced economies (Reinhart and Rogoff 2011; Schularick and Taylor 2009) and house prices and household credit in particular emerge as one of the best leading indicators (Alessi et al. 2014; Alessi and Detken 2009; Borio 2014; Borio and Drehmann

2009; Borio and Lowe 2002; Büyükkarabacak and Valev 2010; Davis and Zhu 2004; Dufrénot and Malik 2012; Szemere, Scatigna, and Tsatsaronis 2014) with mortgage debt driven credit booms more likely to 'go bad' (Bezemer and Zhang 2014; Jorda, Schularick, and Taylor 2014). There is evidence that in many advanced economies household debt (Barbosa-filho et al. 2008) house prices (Borio 2014; Goodhart and Hofmann 2008; IMF 2008; IMF et al. 2009) and residential investment in particular (Álvarez and Cabrero 2010; Davis and Heathcote 2005; Ferrara and Vigna 2009; Igan et al. 2011; IMF 2008; IMF et al. 2009; Kydland, Rupert, and Sustek 2016; Leamer 2007)[1] appear to lead the "business cycle". What is more a large and growing body of empirical work suggests that recessions preceded by larger increases in household debt (principally mortgage credit) tend to be more severe and protracted (Adalid and Detken 2007; IMF et al. 2012; Jorda, Schularick, and Taylor 2012; Jorda et al. 2014; Juselius and Drehmann 2015; King 1994; Mian, Sufi, and Verner 2015) and that recessions associated with house price busts tend to be deeper and longer (Claessens et al. 2012). Taken together, these various strands in the empirical literature underscore a need to clarify and deepen our understanding of the relationship between housing cycles and macroeconomic cycles.

Though long studied in economics, understanding cycles and instability continues to pose a considerable and key challenge. Economic cycles potentially arise for a variety of reasons - the propagation of shocks or the emergence of endogenous cyclicality through the interactions between economic processes. The highly interconnected and complex nature of the system means that once an oscillation is generated somewhere in the economy, it may easily propagate to other parts, making the root causes of macroeconomic cycles difficult to identify. The complex spectral content, and very often non-stationary character of macroeconomic time-series, pose further important challenges in both the time and frequency domains. However adequately identifying "root" oscillations in the system in particular may be an essential first step in an applied approach towards understanding and modelling cyclical dynamics in the economy. Strategies are thus needed by which to diagnose the *sources* of economic fluctuations, and the paths by which they propagate, that are both able to parse the relationships between variables at different frequencies, as well as robust to non-stationarity and able to shed light on transient relationships and structural change.

In this paper I propose a novel data driven strategy through the deployment of wavelet-based time-frequency domain analysis suitable for the study of non-stationary time-series, combined with graph-theoretic methods. I employ four different statistics: wavelet based *coherence*, *partial coherence*, *phase-difference* and *partial phase-difference.* Wavelet *coherence* statistics provide a spectral measure of correlation between variables; and *phase-difference* statistics provide a measure of the temporal ordering between variables. The time and frequency localised information yielded by these wavelets methods allows me to study the relationship between variables for particular cycle periodicity bands and how

[1] That housing investment peaks and troughs before business investment (which lags the cycle) has long been recognized going back to (Burns and Mitchell 1946; Long 1940) and others, but little addressed or explained by macroeconomics.

these spectral relationships have or have not varied or changed over a time. I deploy an algorithmic procedure for 'network discovery' based on summary versions of these statistics, leading to a coherence weighted phase directed network based representation of the relationship between variables at specific periodicity bands of interest. The partial version of these statistics allow me to study relationships controlling for other variables, thus helping to further clarify transmission paths in the network by helping to distinguish direct vs. indirect dependence between variables[2].

I first apply this methodology to the study, for the United States, of the relationship between key housing, financial and macroeconomic variables over relevant periodicity band. A number of interesting results emerge: in particular I identify strong coherence between low frequency cycles in housing and macro cycles with a clear spectral peak at roughly 8 year periodicity. The important and leading role of housing market volumes for the US is confirmed, however strikingly secondary housing market volumes (exiting home sales) emerge as leading residential investment, and as the ultimate "root cause" or *source*-variable for real and financial variables in the housing, finance, macro network under analysis over the low frequency band of interest. Meanwhile and unexpectedly house-prices emerge as a largely *sink*-variable as do mortgage debt and mortgage rates. Phase-lag based temporal orderings of bivariate statistics, suggest a transmission path from existing home sales, to residential investment, to durable consumption, to non-residential investment, to non-durable consumption. Meanwhile partial coherence and phase lag analysis suggest some non-trivial direct transmission from existing home sales to durable consumption.

These results are striking considering that existing home sales volumes feature neither in macroeconomic theory nor empirics. What is more the picture that emerges does not seem to provide much if any support for any of the theoretical perspectives established in the literature, considering the different but key roles attributed to prices and interest rates in e.g. both New Keynesian and Post-Keynesian theories of consumption, debt and demand cycles. Overall the work raises the possibility that cycles in secondary housing market volumes may be an autonomous source of cyclicality, propagating to the wider economy by driving a house building cycle, and durable consumption cycle, which in turn further propagate to other investment and consumption categories resulting in system level or aggregate cycle. This possibility and the results presented here require further and detailed confirmation and investigation but raise a number of important questions and suggest new empirical and theoretical directions for research.

The rest of this paper is organized as follows. Section 2 outlines my method and procedure in more detail, first introducing and defining the wavelet tools I

---

[2] In a further extension and refinement of this methodology I will deploy these partial statistics also via an unsupervised/data driven approach in order to derive a 'sparser' network of theoretically *direct* (in the sense of not dependent on any other node in the network under consideration) edges between variables providing a simple and intuitive representation of the direct dependence and temporal flow between variables at key cycle frequencies. For now I present provisional results based on bivariate wavelet statistics based network and detailed partial analysis of some key edges.

employ then my procedure for network discovery; section 0 presents my application of this methodology to the case of the US with an emphasis on unpicking the relationship between housing cycles, financial cycles, and macroeconomic cycles. In 3.1 I initially identify and characterise the spectral and temporal relationship between key housing variables and GDP; in 3.2 I implement my 'network discovery' methodology in order to unpack or unpick the relationship between housing variables and economic aggregates (studied in 3.1) in terms of more detailed variables; in 3.3 I then make further more detailed study of the key dependencies between variables identified by this 'network discovery' procedure making partial analyses informed by the topology of the network (which contains both direct and indirect links). Section 0 discusses the results of this application, their implications for our understanding of economic cycles and macroeconomic policy, implications and further work, and concludes.

## 2 Methodology and procedure

Both a key challenge and key point of interest in identifying the relationships between housing, financial and macro variables/cycles, is that these relationships may vary across frequencies (e.g. commonly different theoretical explanations are attributed to different business cycle periods) and over time (both variation over the cycle as well as structural breaks). The complex spectral content and non-stationarity often present in relevant time series pose a considerable challenge for conventional econometric and spectral methods. Traditional time-domain analysis/econometrics is not well suited to understanding the spectral structure or relationships between variables, and an interest in cycles leads naturally to spectral analysis (that is the study of time series and their relationships in the frequency domain). However economic data generally requires various treatments in order to achieve, or may fail to meet, stationarity requirements - besides, as above, transient or changing relationship may be a key point of interest - and Fourier based approaches to recovering time information and handling non-stationarity (the short windowed Fourier decomposition (SWFD)) are inefficient (Daubechies 1992). The wavelets transform was developed as a solution to these drawbacks and forms the basis for a growing set of powerful tools able to capture rich well-resolved time-frequency domain information about multivariate non-stationary time series. I apply wavelets based methods in order to study the spectral relationship between housing, financial and macro variables. Concretely, I employ wavelet *power spectrum*, and bivariate and conditional wavelet *coherence* and *phase-difference* statistics based on the continuous wavelet transform using the Morlet wavelet. Wavelet *power spectra* (*auto-spectra*) provide time-frequency information about the cyclical characteristics of individual variables of interest over my sample period (e.g. are there cycles of different periodicity in GDP and of what periodicities?); wavelet *coherence spectra* provide quantitative time-frequency information about the co-cyclicality of different variables, helping me to identify strong associations between cycles in different variables (e.g. which GDP cycles are shared with housing variables?); *phase-differences* provide time-frequency specific delays/lead-lags between different variables thus potentially some indication of the dominant causal flow and of phase-synchrony between variables (e.g. Do housing variables lead or lag specific GDP cycles?). Conditional versions of these different statistics help me to unpick independent from dependent (direct from indirect) relationships (i.e. $x \rightarrow z$ or $x \rightarrow y \rightarrow z$). (These statistics are introduced and defined in some more detail in section 2.1).

Another key challenge, is that the highly interconnected and complex nature of the system means that once an oscillation is generated somewhere in the economy, it may easily propagate to other parts, making the root causes of macroeconomic cycles and their propagation chains, difficult to identify or trace. This is made harder by our limited understanding of the housing-macro nexus and the large number of potentially relevant variables. Reflecting these characteristics of the problem under study, I take a data driven approach, deploying non-parametric wavelets statistics via an unsupervised algorithmic procedure for 'network discovery' in order to identify key relationships and the direction of flow between a relatively detailed set of housing, financial and macroeconomic variables. This approach has various advantages including not restricting the variables included for study (computing power/time the only limitation on this), and not imposing any restrictions on the relationships under study.

My project and general procedure in this study is to:

(1) Characterise the spectral relationship between key housing variables and aggregate demand over time in order to identify possible co-cyclicality between the housing market and aggregate demand, the relevant periodicity ranges for this, and lead-lag relationship for any shared cycles, as well as whether these relationships have changed or varied over the historical period under study/

(2) Unpack the relationships identified in stage (1) at the aggregate demand level by taking a data driven approach to discovering the network topology for a relatively detailed set of housing, financial and macroeconomic variables over specific periodicity ranges of common cycles identified at stage (1) as a way to try and identify candidate 'root cause' oscillations and their propagation paths.

(3) Make a detailed time-frequency study of key bivariate relationships identified by the network discovery procedure at stage (2) and attempt to unpick direct vs. indirect (independent vs. dependent) links in this network through the use of partial analysis informed by the topology of the network discovered at stage (2).

Section 2.1 makes a general introduction to the nature and usefulness of the wavelets statistics employed and provides detailed technical references for the interested reader. Section 2.2 sets out the 'network discovery' procedure.

## 2.1 Wavelets coherence and phase-difference statistics

### 2.1.1 Wavelets and the continuous wavelet transform

Given a nonstationary time series, analysing it through a transform that captures events locally in both time and frequency is appealing. The Short Time Fourier Transform (STFT) (Gabor 1946) was developed as one solution to this, where the discrete Fourier transform is applied to subsets of a time series via an arbitrary fixed-length moving window for analysis resulting in a grid like partition of the time-frequency plane with fixed time windows over which the series is assumed to be locally approximately stationary. The choice of window length is thus an important problem for STFT. A Heisenberg's uncertainty type principle (Daubechies 1992) applies to time-frequency domain analysis since exact frequency and exact time of occurrence in a signal cannot be both obtained. This is because a short window can enhance time information but implies a loss of frequency information; whereas a long window will enhance frequency resolution but implies a loss of time information (and introduction of nonstationarities). The fixed time-window length of the STFT is an inefficient solution to this problem.

The wavelet transform represents a more efficient response to this problem. As the Fourier transform decomposes/represents a signal in terms of combinations of sins and cosines, so a wavelet transform decomposes/represents a signal as a combination of "wavelets". However (by contrast with familiar trigonometric functions) wavelet functions are limited in time and frequency[3]. Mathematically speaking the wavelet transform is the *convolution* of the *scaled* and *translated* basic ("mother") wavelet function with the signal to be analysed. This means that the basic wavelet function is stretched or compressed (*scaled*) in order to capture different levels of detail (different scales i.e. different periodic components) and also moved through the signal (*translated*) in order to obtain time-localised correlation. In this way correlation between the wavelet and the signal in a scale-time plane is obtained generating and allowing representation of scale-time information about the spectral content of the time series.

The width of a wavelet window is squeezed or stretched by scaling, while the number of oscillations remains constant, thus increasing or decreasing the frequency. Large wavelets filter out all but low frequencies, and small wavelets filter out all but high frequencies. The consequence is that, at low frequencies the wavelets are broad and poorly localized in time but well localized in frequency; whereas at high frequencies the wavelets are well localized in time, but the frequency band-width is large. This is sometimes referred to as "multi-resolution analysis" or "adaptive windowing" and makes the wavelet transform the time–frequency decomposition with the optimal trade-off between time and frequency resolution (Mallat 1999). In this study I adopt the Morlet wavelet[4] (Goupillaud, Grossmann, and Morlet 1984), a sine wave damped by a Gaussian envelope, expressed as $\psi(0)\eta = \pi^{-1/4}e^{i\omega_0\eta}e^{-\frac{1}{2}\eta^2}$, where $\psi(0)$ is the wavelet, $\omega_0$ is the dimensionless frequency, and $\eta = st$ where $s$ is the wavelet scale, and $t$ is time (Torrence and Compo 1998).

In the CWT, the analysing function is a "mother" wavelet, $\psi$, and the CWT compares the signal to translated (or *shifted*) – shifting its position in time - and

[3] See **Figure 16** in appendix 5.1 which illustrates the Morlet wavelet for an example.
[4] Illustrated in **Figure 16** Appendix 5.1.

scaled (or dilated) - compressed or stretched - versions of the mother wavelet, $\psi_{a,\tau}$. These analysing functions are sometimes referred to as a family of wavelet "daughters". By comparing the signal to the wavelet at various scales and positions a function of two variables is obtained.

Starting with a wavelet $\psi$ in $L^2(\mathbb{R})$, the analysing functions $\psi_{a,\tau}$ are generated by simply scaling $\psi$ by $a$ and translating it by $\tau$

$$\psi_{a,\tau}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-\tau}{a}\right), \quad a,\tau \in \mathbb{R}, a \neq 0 \qquad (1)$$

where $a$ is the *scale* (or *dilation*) parameter that controls the length of the wavelet; and $\tau$ is the *translation* (or *shift* or *position*) parameter, a location parameter that indicates where the wavelet is centred. The value of the scale parameter determines the level of *stretch* (if $|a| > 1$), or *compression* (if $|a| < 1$) of the wavelet. Shifting a wavelet simply means delaying (or advancing) its onset[5]. The coefficient $\frac{1}{\sqrt{|a|}}$ is an energy-normalized factor (the energy of the wavelet must be the same for different $a$ value of the scale).[6]

By continuously varying the values of the scale parameter, $a$, and the translation parameter, $\tau$, the CWT coefficients $W_{x;\psi}(a,\tau)$ are obtained.

Given a time series $x(t) \in L^2(\mathbb{R})$, its continuous wavelet transform (CWT), with respect to the wavelet $\psi$, is a function of two variables $W_{x;\psi}(a,\tau)$ defined as

$$W_{x;\psi}(a,\tau) = \frac{1}{\sqrt{|a|}}\int_{-\infty}^{\infty} x(t)\bar{\psi}\left(\frac{t-\tau}{a}\right)dt \qquad (2)$$

Where the over-bar denotes complex conjugation. By mapping the original series into a function of $\tau$ and $a$, information is simultaneously obtained on time and frequency.

If the wavelet is complex-valued, the CWT is a complex-valued function of scale and position. If the signal is real-valued, the CWT is a real-valued function of scale and position.

### 2.1.2 Wavelet power spectrum

The *wavelet power spectrum* describes the evolution of the variance of a time-series at different frequencies, thus the wavelet transform can be used to analyse time series that contain nonstationary power at many different frequencies

---

[5] Mathematically, delaying a function $f(t)$ by $k$ is represented by $f(t-k)$.

[6] Wavelet transforms commonly use $L2$ normalization of the wavelet. For the $L2$ norm, dilating a signal by $\frac{1}{a}$, where $a$ is greater than 0, is defined as follows:

$$\left\| x\left(\frac{t}{a}\right)\right\|_2^2 = a\|x(t)\|_2^2$$

The energy is $a$ times the original energy. To preserve the energy of the original signal, you must multiply the CWT by $\frac{1}{\sqrt{a}}$.

(Daubechies 1990). **Figure 17** (appendix section 5.1) borrowed from (Soares and Aguiar-Conraria 2011) illustrates how wavelet transform based local power spectrum is able to capture spectral information about transient cycles (here from a synthesized signal).

In some sense, the wavelet transform can be regarded as a generalization of the Fourier transform and by analogy with spectral approaches. One can compute the local *wavelet power spectrum* (WPS) by

$$(WPS)_x(a,\tau) = |W_x(a,\tau)|^2 \qquad (3)$$

We can interpret the wavelet power spectrum as depicting the *local variance.* The Fourier spectrum of a signal can be compared with the global wavelet power spectrum, which is defined as the average energy (average variance) contained in all wavelet coefficients of the same scale (period) $a$ over all time.

$$(GWPS)_x(a) = \int_{-\infty}^{\infty} |W_x(a,\tau)|^2 \, d\tau \qquad (4)$$

### 2.1.3 Local and global wavelets coherence

It is often desirable to quantify statistical relationships between two non-stationary signals. In Fourier analysis, the coherency is used to determine the association between two square-integrable signals $x$ and $y$. The coherence function is a direct measure of the correlation between the spectra of two time-series ((Chatfield 1989)). To quantify the association between two *non-stationary* signals, and by analogy with Fourier analysis, the *wavelet cross-spectrum* and the *wavelet coherence* may be calculated. As in the Fourier spectral approaches, the cross wavelet coherence can be defined as the ratio of the cross-wavelet spectrum to the product of the spectrum of each series, and can be thought of as the local correlation between two CWTs in the time-frequency plane. The following sections cover this in some detail.

The cross-wavelet transform of two time series, $x(t)$ and $y(t)$, first introduced by (Hudgins, Friehe, and Mayer 1993), is defined as

$$W_{x,y}(a,\tau) = W_x(a,\tau)\overline{W_y}(a,\tau) \qquad (5)$$

Where $W_x(a,\tau)$ and $W_y(a,\tau)$ are the wavelet transforms of $x$ and $y$ respectively.

When $y = x$, we obtain the wavelet power spectrum $W_{xx}(a,\tau) = |W_x(a,\tau)|^2 = (WPS)_x$ (see Eq. 3). The *cross-wavelet power* (XWP) is defined by

$$(XWPS)_{xy}(a,\tau) = \left|W_{xy}(a,\tau)\right| \qquad (6)$$

While we can interpret the *wavelet power spectrum* as depicting the *local variance*, the *cross-wavelet power* of two time series depicts their *local covariance* at each time and frequency. The cross-wavelet power thus gives a quantified indication of the similarity of power between two time series.

In analogy with the concept of coherency used in Fourier analysis, given $x(t)$ and $y(t)$ one can define their *complex wavelet coherency*, $\varrho_{x,y}$, by

$$\varrho_{x,y} = \frac{S(W_{xy}(a,\tau))}{[S(|W_x(a,\tau)|^2)S(|W_y(a,\tau)|^2)]^{1/2}} \qquad (7)$$

where $S$ denotes a smoothing operator in both time and scale. As in the Fourier case, smoothing is necessary, otherwise coherency would be identically one. In Fourier analysis this problem is overcome by smoothing the cross-spectrum before normalization. For wavelet analysis [and in particular with the Morlet wavelet,] smoothing can be achieved by a convolution in time and scale ((Torrence and Compo 1998)). Time and scale smoothing can be achieved by convolution with appropriate windows (time convolution is done with a Gaussian window and the scale convolution is performed by a rectangular window); see (Cazelles et al. 2007) or (Grinsted, Moore, and Jevrejeva 2004), for details.

The absolute value of the *complex wavelet coherency* denoted $R_{xy}(a,\tau)$ is called *wavelet coherency* and gives a measure of the *local correlation* between signals $x$ and $y$ in the time-frequency plane

$$R_{xy}(a,\tau) = \frac{|S(W_{xy}(a,\tau))|}{[S(|W_x(a,\tau)|^2)S(|W_y(a,\tau)|^2)]^{1/2}} \qquad (8)$$

with $0 \leq R_{xy}(a,\tau) \leq 1$. At points $(a,\tau)$ for which the denominator is zero, we define $R_{xy}(a,\tau) = 0$. Eq. 8 resembles the definition of the correlation coefficient. A coherence value of 0 signifies that the two time series are unrelated, whereas a coherence value of 1 indicates the two time series are linearly related at the given frequency and time.

The time-averaged or "*global wavelet coherency*" could simply be to take the time average of the coherence $R_{xy}(a,\tau)$ between signals $x$ and $y$ over the time index (as for global wavelet power spectrum Eq. 4). Following Schulte et al. (2016)) I define *global wavelet coherency* as

$$GWC_{xy}(a) = \frac{|S(W_{xy}(a))|}{[S(|W_x(a)|^2)S(|W_y(a)|^2)]^{1/2}} \qquad (9)$$

where $W_{xy}(a) = W_x(a)\overline{W_y}(a)$.

### 2.1.4 Phase-difference

The phase of a given time-series $x(t)$ can be viewed as the position in the pseudo-cycle of the series and it is parameterized in radian ranging from $-\pi$ to

$\pi$. The *phase-difference* statistic indicates the temporal order of two time-series. *Phase-difference* (and in particular *mean phase-difference*) has been widely applied as a measure and test of synchronisation between variables[7]. A stable phase relationship between variables over a period may indicate both partial synchrony (Pikovsky, Rosenblum, and Kurths 2001; Rosenblum et al. 2001); but also the temporal ordering of this relationship offers potentially relevant information for understanding the driver-response relationship between signals (see e.g. (Grinsted et al. 2004; Nolte et al. 2008)), or implies a differential in their rate of response to some third causal influence – in either case a matter of economic interest and an ambiguity that may itself be further unpicked by the application of partial phase-difference introduced by (Aguiar-Conraria and Soares 2014; Ng and Chan 2012) based on the continuous wavelet transform. Of course precedence is not the same thing as causality, what is more given bidirectional or unknown coupling, a finding that x leads y does not imply the absence of any influence from y to x and we can make no statement about this reverse direction based on phase-difference. Nevertheless we might expect the phase-difference to capture the dominant direction of flow between coupled series. If the phases of two signals are fully locked, their *phase-difference* - indicating a *phase-shift* - should be constant. Thus periods of synchronization can be detected simply by plotting the phase difference against time and looking for horizontal plateaux (Blasius, Huppert, and Stone 1999; Rosenblum et al. 2001). This sort of time plot of phase-difference may also reveal interesting information about slow trends and/or abrupt transitions in phase relationships at particular frequencies - wavelets based phase and partial phase lags are able to capture not only structural breaks and transient relations, but are also able to distinguish between different relations that occur at the same time but at distinct frequencies. It is helpful to consider the phase-relationship between variables over regions of high coherence in time-frequency space, and by averaging the phase-difference between variables over a specific scale band, it is possible to characterise their phase relationship for differed frequency ranges and track the evolution of this over time. Even where a phase-relationship is not entirely stable over time, looking at the distribution may reveal a tendency (unimodal distribution), or regime switching (bimodal distribution).

When the wavelet $\psi$ is complex-valued, the corresponding wavelet transform $W_x(a,\tau)$ is also complex-valued and may be separated into its real part $\Re\{W_x(a,\tau)\}$ and imaginary part, $\Im\{W_x(a,\tau)\}$. That is into its amplitude $|W_x(a,\tau)|$, and its phase (or phase-angle) $\phi_x(a,\tau)$: $W_x(a,\tau) = |W_x(a,\tau)|e^{-i\phi_x(a,\tau)}$. The phase-angle $\phi_x(a,\tau)$ of the complex number $W_x(a,\tau)$ can be obtained from

$$\phi_x(a,\tau) = tan^{-1}\left(\frac{\Im\{W_x(a,\tau)\}}{\Re\{W_x(a,\tau)\}}\right) \qquad (10)$$

where $tan^{-1}$ denotes the extension of the usual principal component of the arctan function to four quadrants[8].

[7] Phase-synchronisation, first described for chaotic oscillators (Rosenblum, Pikovsky, and Kurths 1996), is defined as the global entrainment of the phases while the amplitudes may remain uncorrelated. Different indices of phase synchronization such as the mean phase coherence (Kuramoto 1984; Mormann et al. 2000) have been introduced.

[8] Detail this

The complex wavelet coherency can be written in polar form, as $\varrho_{x,y} = |\varrho_{x,y}|e^{i\phi_{x,y}}$. The angle $\phi_{x,y}$ of the complex coherency is called the phase-difference (phase lead of $x$ over $y$) $\phi_{x,y}(a,\tau) = \phi_x(a,\tau) - \phi_y(a,\tau)$ or

$$\phi_{x,y}(a,\tau) = tan^{-1}\left(\frac{\Im(S(W_{x,y}(a,\tau))}{\Re(S(W_{x,y}(a,\tau))}\right) \qquad (11)$$

- A phase-difference of zero indicates that the time series $x(t)$ and $y(t)$ move together at the specified time-frequency;
- If $\phi_{x,y}(a,\tau) \in (0, \frac{\pi}{2})$, then the series move *in phase*, but the time series $x$ leads $y$; if $\phi_{x,y}(a,\tau) \in (-\frac{\pi}{2}, 0)$, then it is $y$ that is leading $x$;

A phase-difference of $\pi$ (or $-\pi$) indicates an *anti-phase* relation; if $\phi_{x,y}(a,\tau) \in (\frac{\pi}{2}, \pi)$, then $y$ is leading; time series x is leading if $\phi_{x,y}(a,\tau) \in (-\pi, -\frac{\pi}{2})$
$\phi_{x,y} \in (-\pi, -\frac{\pi}{2})$ (See **Figure 18** in appendix section 5.1).

### 2.1.5 Partial coherence and phase-difference

To estimate the interdependence, in the time-frequency domain, between two variables after eliminating the effect of other variables, we will rely on the concept of *partial coherency* (introduced by Aguiar-Conraria and Soares (2014) as a generalisation of the concept of simple wavelet coherency in the existing literature). If we find that, after controlling for a third variable, the (partial) coherency between two variables decreases in some region of the time-frequency space, we conclude that part of their interdependence was due to that third variable. On the other hand, if the opposite happens, one concludes that the third variable was clouding the relationship."

#### *2.1.5.1 Partial coherence*

As in the case of ordinary wavelet coherency (and Fourier analysis), to compute partial wavelet coherence it is necessary to perform a smoothing operation on the cross-spectra. Denoting the smoothed version of $W_{ij}(a,\tau)$ (see section 2.1.1 and Eq. 2), with $S_{ij}(a,\tau)$ we have that

$$S_{ij}(a,\tau) = S(W_{ij}(a,\tau)) \qquad (12)$$

where $S$ as before is a certain smoothing operator (see section and 2.1.3 equation 7).

Let $\mathcal{S}$ denote the $p{\times}p$ matrix of all the smoothed cross-wavelet spectra $S_{ij}(a,\tau)$ and dropping $(a,\tau)$ for convenience of notation

$$\mathcal{S} = \begin{bmatrix} S_{11} & \cdots & S_{1p} \\ \vdots & \ddots & \vdots \\ S_{p1} & \cdots & S_{pp} \end{bmatrix} \qquad (13)$$

This matrix (13) is an Hermitian matrix.

The *complex partial wavelet coherency* of $x_1$ and $x_j$ $(2 \leq j \leq p)$ allowing all the other series will be denoted by $\varrho_{1,j.\boldsymbol{q}_j}$ and is given by

$$\varrho_{1,j.\boldsymbol{q}_j} = -\frac{\mathcal{S}_{j1}^d}{\sqrt{\mathcal{S}_{11}^d \mathcal{S}_{jj}^d}} \qquad (14)$$

where $\mathcal{S}^d$ is the determinant of the matrix $\mathcal{S}$ and $\mathcal{S}_{j1}^d$ is the cofactor of the element in position (j, 1) of $\mathcal{S}$, and $\boldsymbol{q}_j$ is short notation for all the indexes in $\boldsymbol{q}$ excluding the index $j$, i.e. $\boldsymbol{q}_j = \{2, \ldots, p\}\{j\}$.

The *partial wavelet coherency* of $x_1$ and $x_j$ allowing for all the other series, denoted $r_{1j.\boldsymbol{q}_j}$, is given by

$$r_{1j.\boldsymbol{q}_j} = -\frac{\left|\mathcal{S}_{j1}^d\right|}{\sqrt{\mathcal{S}_{11}^d \mathcal{S}_{jj}^d}} \qquad (15)$$

whilst the squared partial wavelet coherency of $x_1$ and $x_j$ allowing for all the other series is given by

$$r_{1j.\boldsymbol{q}_j}^2 = -\frac{\left|\mathcal{S}_{j1}^d\right|^2}{\mathcal{S}_{11}^d \mathcal{S}_{jj}^d} \qquad (16)$$

See Aguiar-Conraria and Soares (2014) for a more detailed exposition.

#### 2.1.5.2 Partial phase difference

Having defined the *complex partial wavelet coherency* $\varrho_{1,j.\boldsymbol{q}_j}$ between the series $x_1$ and $x_j$, after removing the influence of all the other remaining series, Aguiar-Conraria and Soares then further define the *partial phase-difference* of $x_1$ over $x_j$, given all the other series, as the angle of $\varrho_{1,j.\boldsymbol{q}_j}$. This phase-difference is denoted by $\phi_{1,j.\boldsymbol{q}_j}$, i.e.

$$\phi_{1,j.\boldsymbol{q}_j} = tan^{-1}\left(\frac{\Im(\varrho_{1,j.\boldsymbol{q}_j})}{\Re(\varrho_{1,j.\boldsymbol{q}_j})}\right) \qquad (17)$$

### 2.1.6 Statistical significance

Using Monte Carlo methods, the statistical significance of wavelet coherence was found by generating a large number of sets of synthetic data series with the same lag-1 autocorrelation coefficients as the input time series, calculating the wavelet coherence for these sets, and then estimating the significance level at each scale using values outside the cone of influence (Aguiar-Conraria and Soares 2014; Grinsted et al. 2004; Schulte et al. 2016; Soares and Aguiar-Conraria 2011; Torrence and Compo 1998). While some empirical distributions have been derived, as Aguiar-Conraria and Soares (2014), if computer time is not a constraint, given that these distributions were derived by Monte Carlo simulations, then one might as well just do the Monte Carlo simulations directly, and this is the approach I have taken.

## 2.2 Network-discovery procedure

Direct inspection of the detailed time-frequency plane information provided by these wavelet spectra, can provide a rich picture of the (conditional) bivariate spectral relationship between variables. However given our limited understanding of the economic cycle or housing-macro nexus, an initial empirical problem is to identify which relationships are important. What is more, in order to understand the relative and overall synchrony between cycles in different variables and to identify "root cause" cycles and transmission paths through the economy, it is essential to look not only at pairwise relationships, but to derive a more global picture of the wider topology of the network of relations between a relatively detailed set of housing, financial and macro variables. Outside of economics various "reverse-engineering" approaches have been taken to uncovering network relationships among different elements (genes, proteins, metabolites, neurons, brain areas, ecological populations etc.), and to the identification of sources of oscillations within and their propagation paths on a network from time series data (in neuroscience (Shao et al. 2015; Zou et al. 2010), in engineering (Yuan and Qin 2014), in ecology (Damos 2016; Detto et al. 2013; Zhang 2011), in chemistry (Bauer et al. 2007) and so on). These studies often rely on correlation and on Granger statistics. I develop and apply a novel data driven procedure for 'network discovery' exploiting *wavelet coherence* and *phase-difference* as well as relatively newly available *partial-wavelet coherence* and *phase-difference* statistics (Aguiar-Conraria and Soares 2014) (see methodology section 2.1.5 above). A *coherence* based similarity-network provides a system wide picture of the relative synchronisation between variables; while phase-difference statistics provide empirical information on the temporal flow or transmission paths between variables on this network. The general procedure is summarised by the flow chart in **Figure 1.**

I first construct an initial network by considering all possible links by computing bivariate pair-wise global wavelet coherence spectra (as in section 2.1.3 and equation 9) for all possible bivariate pairs among the full set of variables under analysis. I then select the periodicity band of interest. These cuts should account for the spectral structure of relationships between variables in the network. In this application, I use a preliminary study and inspection of the spectral structure of the global coherence between housing variables and GDP (section 3.1) to inform this decision. I select this spectral range based on local minima for the global coherence curves, either side of the spectral peak in housing/GDP coherence that I wish to unpack with my network discovery procedure. However other and potentially unsupervised approaches might be taken and may be worth exploring[9].

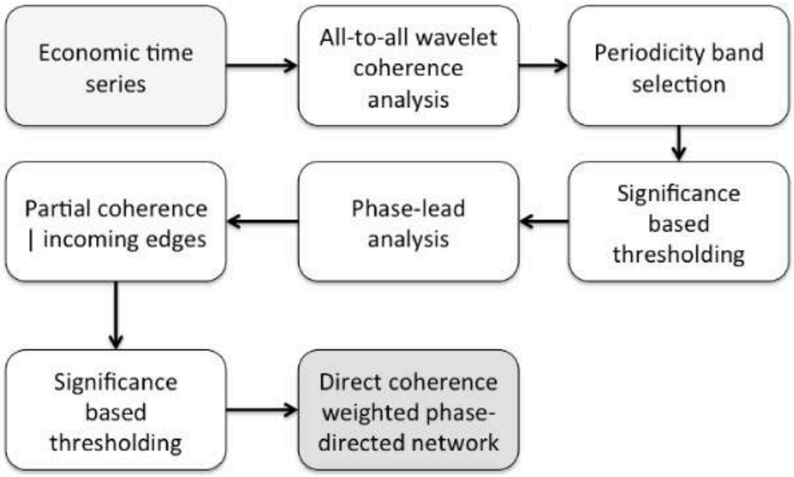



**Figure 1:** Procedure for network discovery.

Having selected the periodicity band of interest, I then obtain the spectral peak coherence magnitude within this band for each pair producing an all-to-all undirected global-wavelet-coherence weighted network for the periodicity range of interest. In order to thin out trivial links I then threshold this network by statistical significance discarding any edges (any links between variables) that are insignificant at the 5% level (this is done based on a simple bootstrapping procedure based on ar1 coefficients and 500 series as in section 2.1.6. Other confidence levels might of course be applied).

I then obtain the period band specific phase difference (see section 2.1.4) for all non-trivial edges and apply kernel density smoothing in order to derive a continuous distribution allowing me to obtain a peak phase-difference statistic that I use to quantify the overall or dominant tendency in the phase relationship[10] (see methodology section 2.1.4 for more detail), thus a temporal ordering between variables (Pikovsky et al. (2001) and Rosenblum (2001) argue that a unimodal distribution of the phase difference between two signals (for the chosen range of frequencies and/or periods) indicates a statistical tendency for the two time-series to be phase locked (conversely, the lack of association is characterized by a broad and uniform distribution) and that this may be used to identify coupling even in noisy signals.). I treat the direction of flow to be from *lead* to *lag* variables (i.e. if X leads Y this generates the directed edge X-Y).

By this procedure I derived a coherence weighted phase-directed network of non-trivial links between the variables under study for the periodicity band of

[9] For example obtaining local maxima for all pairwise global coherence curves, then applying a classical k-means clustering analysis does not yield a stable partitioning for less than k=3 clusters (see appendix section 5.6), suggesting a possible 3 periodicity bands based on the clustering of local spectral peaks in the relationship between variables.
[10] This is reasonable for a uni-modal distribution. A more rigorous refinement might test and account for material bi-modality of distributions as an indication of regime switching or an important phase shift.

interest. In principal this is likely to capture all significant direct links between variables but also indirect links between variables (where direct means a relationship is not significantly dependent on other variables in the network under study). Since I am interested in identifying important *sources* of cyclicality and transmission paths through the economy/network, I now consider the *out-degree* and *in-degree* of individual nodes; as well as the temporal ordering not just for single edges, but within larger groups of connected variables. High out-degree (the number of outgoing edges) variables may be characterised as *sources* (thus potential drivers of the system level cycle we want to understand) and high in-degree (incoming edges) variables as *sinks* within the system. Meanwhile the full temporal ordering of edges between sources and sinks may (though need not necessarily) provide some indication of the direct transmission chain. Thus a path between a source and a sink defined by a rule of always taking the shortest phase-path may be one way of thinking about the path from *source* to *sink* variables on this network.

Having used an algorithmic procedure to identify all non-trivial links, I then use the topology of the resulting network in order to (1) identify key relationships for more detailed study (i.e. important source nodes and their edges) and (2) in order to inform my conditional analysis, testing specific edges in the network, conditional on other incoming edges, in order to obtain information about independent vs. dependent aspects of these links and investigate the direct edge structure of the network. Conditional analysis is conducted using the partial coherence and phase statistics introduced by Aguiar-Conaria and Soares (Aguiar-Conraria and Soares 2014) (see methodology section 2.1.5)[11].

This procedure has a number of obvious advantages: it is data driven, spectrally focussed thus able to parse complex spectral relationships; robust to non-stationarity as a characterisation of relationships over a given period of reference; non-parametric thus not sensitive to model specification. What is more since the only limit on the size of the variable set that may be considered by this procedure is computing power, it has the advantage of allowing all relationships of potential interest to be considered, thus the study of a wider and/or a more detailed network is possible. By explicitly taking more variables and relationships into account, it not only has the potential to provide a more comprehensive and/or detailed structure, but may also provide a network structure and system analysis less biased by omitted variables and researchers' deductive or prior assumptions making it a potentially useful starting point for empirical or applied modelling work. The results can then guide and inform theoretical hypotheses and further empirical work to investigate particular sub-sets of relationships identified as important.

[11] In a further iteration and refinement of this methodology, this procedure might be implemented via a further algorithmic procedure to iteratively test each edge conditional on each other variable in the network, discarding any edge found to be insignificant conditional on any other variable in the network in order to systematically uncover a direct network such that no edge indicates (conditional) independence between two variables, whereas a directed edge (X – Y) signifies coherence with X leading Y.

# 3 Application: The housing-macro nexus in the US economy

## 3.1 The housing cycle and the economic cycle

In this section I will initially study the spectral structure of GDP and of key housing variables – existing home sales, house building, and house prices by obtaining both the wavelet power spectrum for each series, as well as looking at the wavelet coherence and phase-relationship between GDP and housing variables.

### 3.1.1 Power spectrum analysis

Looking at the power spectrum for GDP and housing variables will allow me to identify and characterise the existence of cyclicality in these individual series as a first step in studying housing and macro cycles and their relationship.

**US GDP (real)**

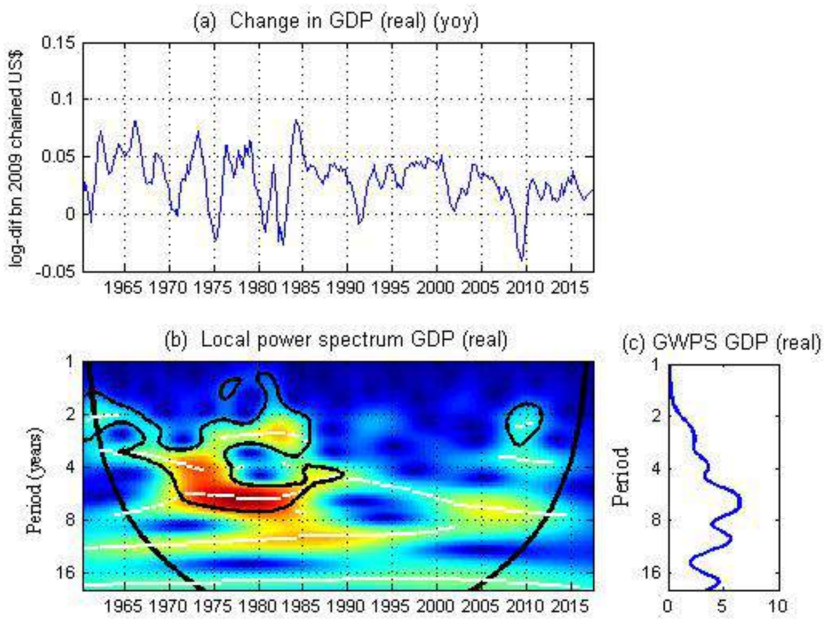



**Figure 2:** These charts present yoy change in quarterly real GDP data from q1 1960 to q2 2017[12]. (a) Time-series plot of the data. (b) The local wavelet power spectrum for this time series. Dark contours show contours of statistically significant regions and the parabola marks the cone of influence – areas outside of this cone may be subject to edge effects and should not be trusted. (c) The time averaged wavelet power spectrum for this series. over the same period.

**Figure 2** presents the wavelet power spectrum for GDP which shows significant cycles at a high and low frequencies during the 1970s and 1980s. While these become less strong after 1990 and high frequency cycles seem to die out (consistent with the "Great Moderation"), there is nevertheless a clear persistent

[12] For the interested reader the power spectrum of GDP data from q1 1947 to q2 2017 is presented in Figure 19 in Appendix section 5.2.

spectral peak at low frequencies (7 to 9 year periodicity) over the entire nearly 60 year sample period including through the Great Moderation, 1990s and 2000s.

### US Existing Home Sales

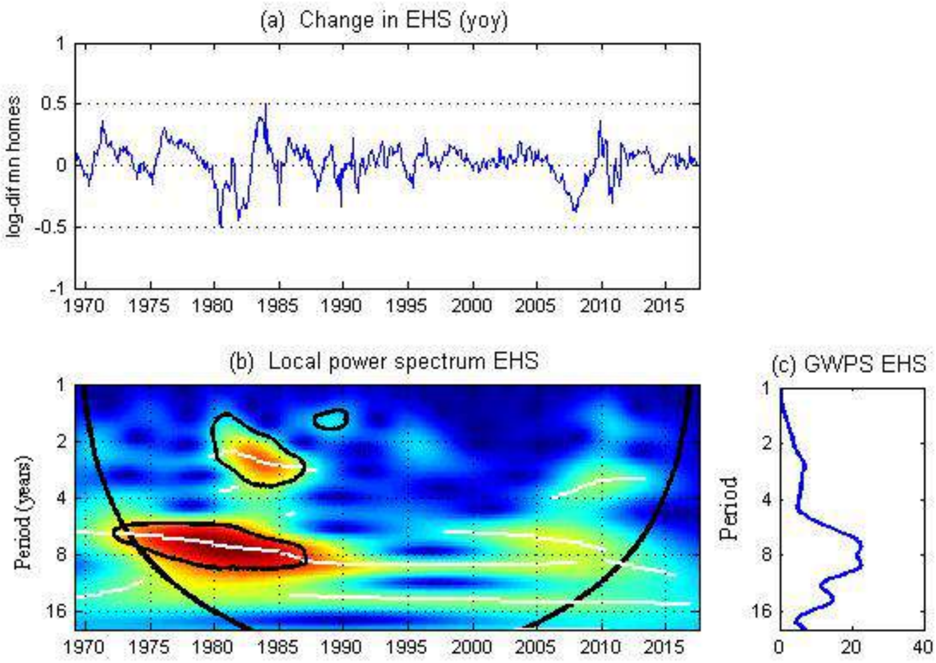



**Figure 3:** These charts present yoy change in monthly *existing* home sales (NAR monthly existing home sales data) from Jan 1969 to July 2017. (a) Time-series plot of the data. (b) The local wavelet power spectrum for this time series. Dark contours show contours of statistically significant regions. (c) The time averaged wavelet power spectrum for this series.

**Figure 3** presents the wavelet power spectrum for *existing home sales* which is simpler and more stable than that for GDP but shows some clear similarity especially at low frequencies where there is a particularly strong and stable spectral peak close to 8 year periodicity. Although like GDP housing sales volumes cyclicality also appears to weaken after 1990 (a period over which volumes rose steadily up to 2005).

**Figure 4** presents the wavelet power spectrum for *new home sales* volumes which seems to suggest that new home sales share a roughly 8 year cycle with *existing home sales* although this cycle appears to have "slowed down" over the sample period moving from roughly 7 years during the 1970s and 1980s, to around 9 years from 1990. Like existing home sales this cycle is significant in the 1970s and 1980s and moderates in the 1990s, however unlike for existing home sales volumes, it is also significant over the 2000s.

**US New Home Sales**[13]

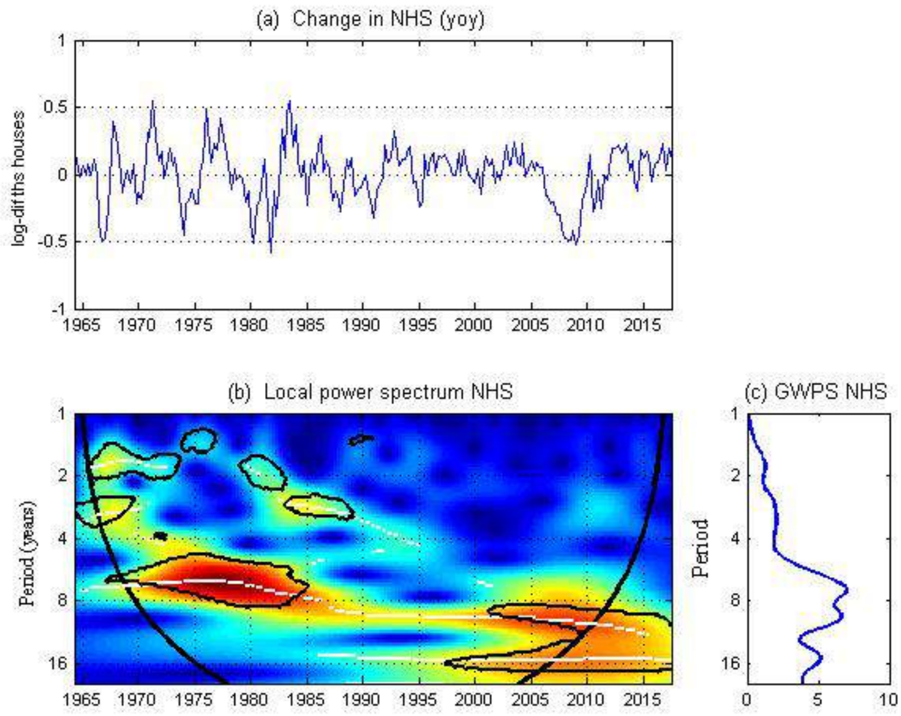



**Figure 4:** These charts present yoy change in quarterly *new* home sales (NHS) data from q1 1964 to q2 2017. (a) Time-series plot of the data. (b) The local wavelet power spectrum for this time series. Dark contours show contours of statistically significant regions. (c) The time averaged wavelet power spectrum for this series.

**Figure 5** presents the power spectrum for the BIS long house price index series, which exhibits rather stable 8 and 15 year cycles across the sample period. However while GDP and housing volumes exhibit on the whole reduced power after 1990, house price cycles appear to have become stronger since 1990. What is more, the other series all exhibited a roughly 8 year and a lower frequency spectral peak, however while for the other series power is concentrated at the 8 year cycle, the main spectral peak for prices is at about 15 years. It is not clear whether the lack of high frequency content reflects underlying economic processes, or some filtering of high frequency bands by the index calculation procedure.

[13] The same analysis is conducted for new house starts data (see Figure 20 Appendix section 5.2) as an alternative indicator of new building and a variable closer to the investment decision (since new home sales involves simultaneously both new house supply and demand). Starts and sales have similar spectral structure.

US House Price Index

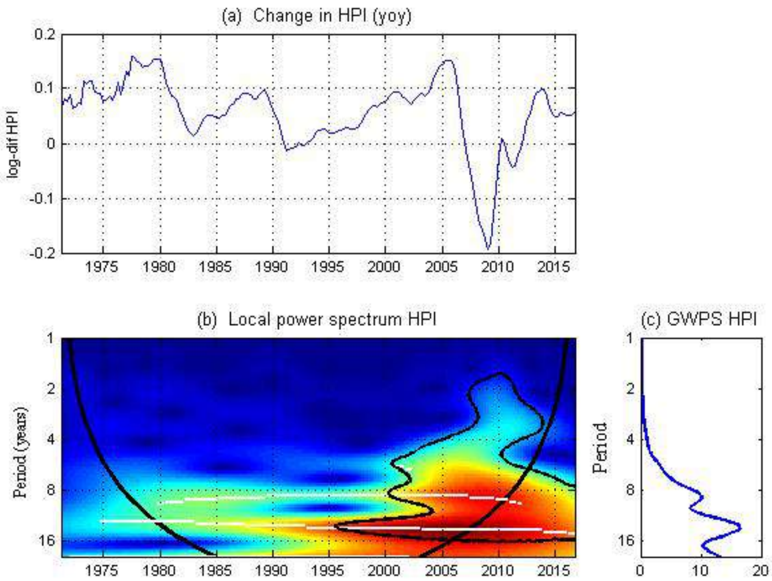



**Figure 5:** These charts present yoy change in quarterly house prices (BIS qtly US house price index data) from q1 1970 to q4 2016. (a) Time-series plot of the data. (b) The local wavelet power spectrum for this time series. Dark contours show contours of statistically significant regions and the parabola marks the cone of influence outside of which edge effects may distort results (c) The time averaged wavelet power spectrum for this series.

In summary: there seems to have been a general reduction in the cyclicality of GDP growth after the 1980s, and in particular higher frequency cyclicality present in the 1970s and 1980s is no longer apparent over the 1990s and 2000s, however some low frequency cyclicality remains clearly apparent in the power spectrum across the entire sample period. Housing market volumes appear to exhibit a similar reduction in power at high frequencies but persistent low frequency cycles. However while high frequencies weaken there appears to have been a significant low frequency cycle in new homes beginning in the mid 1990s. Meanwhile house prices show little cyclicality before the 1990s, but a clear low frequency price cycle appears to develop starting in the early 1990s. Overall there seem to be some visible similarities between power spectrums, and this is something I will now test rigorously by looking at coherence spectra between GDP and these housing variables in order to make characterise the association between overall aggregate demand and the housing market.

### 3.1.2 Coherence and phase-difference analysis

Looking at the coherence spectrum and phase relationship between GDP and housing variables will allow me to identify and characterise the strength and significance of synchronisation between aggregate demand and housing cycles, as well as the lead-lag relationships between them at relevant frequencies. This will help me to make a top-level answer to the question under study: what is the relationship between housing and macro cycles? This will motivate and inform further more detailed unpacking of this relationship with my network discovery procedure. The results of this analysis are presented in Figure 6, Figure 7 and Figure 8.

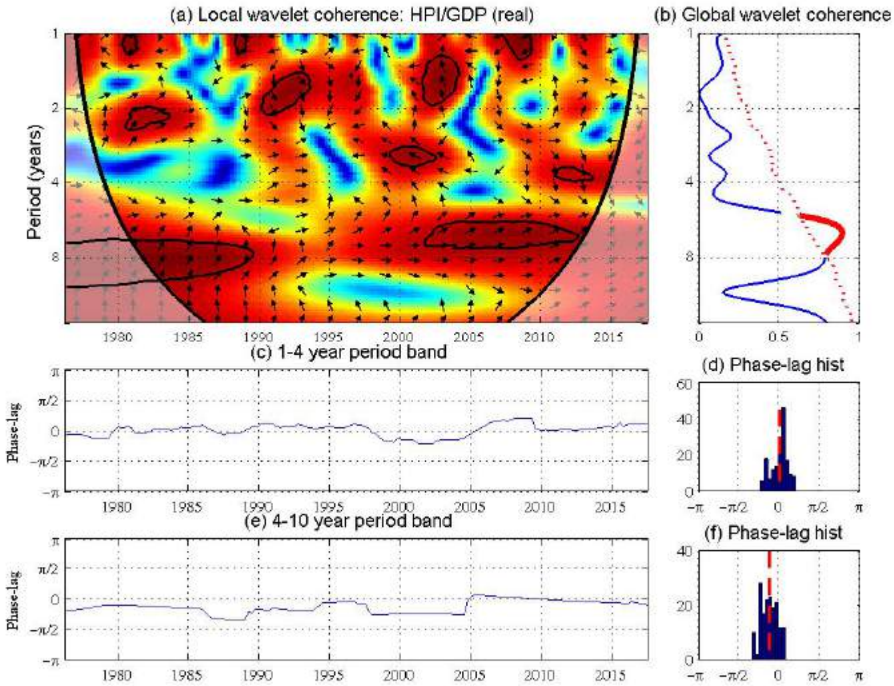



**Figure 6:** These charts present analysis of qtly US real GDP and FHFA HPI (all transactions) data from 1975-2017. (a) Local wavelet coherency; (b) Global wavelet coherency; (c) and (e) Phase-differences averaged across 1-4 and 4-10 year periodicity bands; (d) and (f) Histograms of the phase-difference statistics presented in (c) and (e). Significance is tested at 5% level based on 500 bootstrapped series.

**Reading and interpreting the figures**: (a) Local wavelet coherence: the heat map shows coherence strength between 0-1; dark contours show the 5% significance level estimated based on 500 bootstrapped series; and the parabola marks the cone of influence - the area outside of which may be affected by edge effects. (b) Global wavelet coherence: the blue line is the global coherence spectrum, while sections of the curve significant at the 5% level (above the thin dashed red line) are bin old red. (c) and (e) Phase-difference averages across periodicity bands: positive values indicate x (here HPI) leads y (here GDP), and visa versa. A phase-difference between 0 and $\pm\pi/2$ indicates the variables are in-phase (or positively related) while a phase-difference of absolute magnitude greater than $\pi/2$ indicates an out of phase (or negative) relationship (see methodology section 2.1.4).

A number of striking results arise in this analysis: looking first at coherency, both housing market prices, and volumes clearly show a strong coherence with GDP at low frequencies as may be seen from the strong and statistically significant spectral peaks in the global wavelet coherence spectrums between these housing variables and GDP at an approximately 8 year periodicity (Figure 6-Figure 21 (b)). The local-wavelet coherence spectrum indicates this relationship has been relatively stable especially for volumes variables, which show strong coherence not only more consistently over the historical period under study, but also over a broader spectral range than do prices. Nevertheless we can observe that prices especially but also secondary market volumes appear to exhibit stronger spectral coherence in the 1980s and in the 2000s (consistent with strong housing cycle episodes over these decades), with weaker coherence during the 1990s. Meanwhile new home sales are strongly significant at this periodicity band right across the entire more than 40 year sample range.

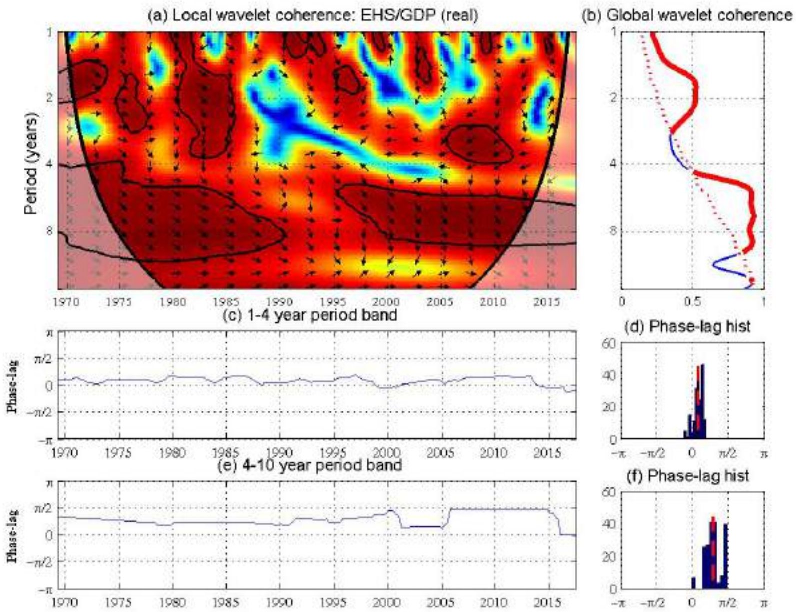



**Figure 7:** These charts present analysis of qtly US real GDP and existing home sale (EHS) data from 1968-2017. (a) Local wavelet coherency; (b) Global wavelet coherency; (c) and (e) Phase-differences averaged across 1-4 and 4-10 year periodicity bands; (d) and (f) Histograms of the phase-difference statistics presented in (c) and (e). Significance is tested at 5% level based on 500 bootstrapped series. For a detailed explanation of how to read the figure see **Figure 6**.

Phase information is also striking. Not only is the coherence between volumes and GDP stronger than is the coherence between prices and GDP, but also the phase-difference statistics indicate that housing market volumes clearly *lead* GDP, while house prices clearly *lag* GDP. This tendency is clearly visible both in the time-plot of phase relationship summarised for the 4-10 year periodicity band (Figure 6-Figure 8 (e)) and in the histogram summaries of these time series (Figure 6-Figure 8 (f)). The time evolution of the phase relationship between these variables however, also reveals further interesting detail. For example we see that from around 1998 the lag of house prices over GDP at the 4-8 period band grew (Figure 6 (e)) i.e. house prices fell increasingly behind; meanwhile existing home sales over the same periodicity band had been increasing their lead over GDP up to 2000, after which they rapidly moved into phase-synchrony remaining rather simultaneous with GDP up to 2005, which coincides with the peak in housing market volumes, at which time the phase-lead of EHS over GDP jumped back up, since when EHS have led GDP particularly strongly since (Figure 7 (e)). This may suggest some critical slowing down[14] ahead of the housing market peak in 2005 (and deep subsequent downturn). What is more while house prices lag GDP, they may nevertheless afford the most leading indicator if this phase-information is considered. What is more, it is interesting to note the very strong

[14] Critical slowing down before a transition may show up as a phase-shift. Williamson et al (Williamson, Bathiany, and Lenton 2016) show that increasing the system timescale as it approaches a local bifurcation (that is, critical slowing down before a turning point) shows up as an increasing phase-lag in the system response relative to the forcing.

and particularly stable lead EHS show over GDP since the market peaked in 2005 (i.e. during a period of falling home sales).

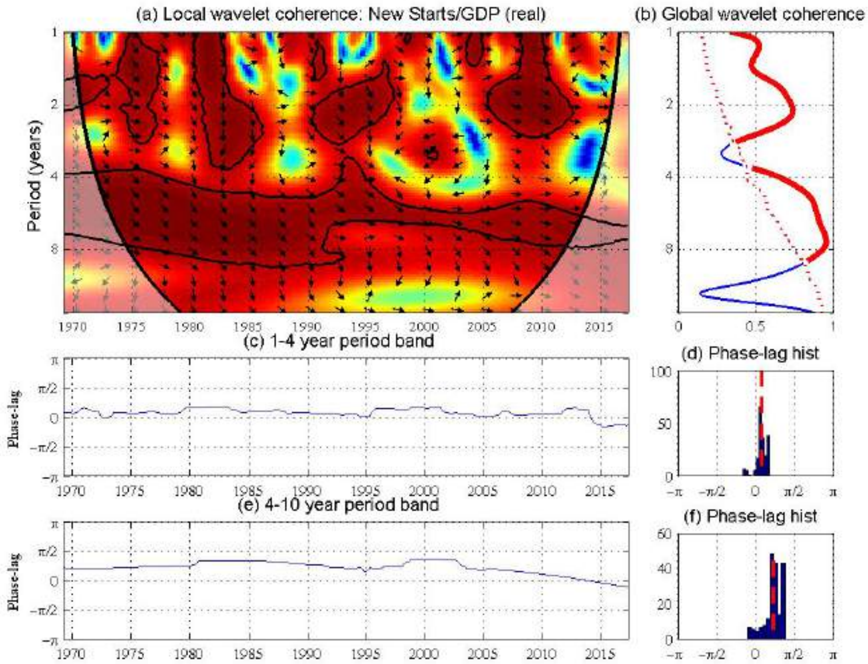



**Figure 8:** These charts present analysis of qtly US real GDP and new housing starts data from 1968-2017. (a) Local wavelet coherency; (b) Global wavelet coherency; (c) and (e) Phase-differences averaged across 1-4 and 4-10 year periodicity bands; (d) and (f) Histograms of the phase-difference statistics presented in (c) and (e). Significance is tested at 5% level based on 500 bootstrapped series. For a detailed explanation of how to read the figure see **Figure 6**.

It is interesting to note the existence of a clear statistically significant common 2 year, housing volume led cycle (both existing home sales and house building) (Figure 7 and Figure 8), suggesting housing may be important not only in low frequency demand and financial cycles (an association that remains poorly understood but that has been made in the business and financial cycle literature), but also for the shorter "business cycle". That this shorter cycle is shared only with the neglected housing market volumes variables, but not present for the house price data analysed (which shares only the long cycle), may explain why an association between housing and shorter demand cycles has not been previously made.

Overall the housing volumes cycle (in EHS and house building – see Figure 7 and Figure 8 but also Figure 21 presenting the same analysis for new home sales rather than starts) emerge as strongly associated with and leading the economic cycle at low frequencies in particular, while house price cycles not so. Nevertheless phase-relation information suggests the possibility that prices may provide a leading indicator, and potentially play a role in the volumes cycle. Overall this preliminary analysis is consistent with the possibility that a housing cycle may propagate to the wider economy, but surprisingly it is housing market volumes not prices, which emerge as the leading variable.

### 3.2 Network discovery and ‘root cause’ analysis

Having identified a strong relationship and common cycle between housing variables and GDP it would be interesting to better identify the transmission from housing variables to aggregate demand as well as the relationships between housing variables. I also wish to consider financial variables as potential sources of cyclicality or important mediating variables in the housing-macro nexus. In order to begin to unpack these relationships, I study the housing variables *house prices* (HPI), and *existing home sales* (EHS) data, and *private residential fixed investment* (PRFI) and further decompose GDP into *non-residential investment* (PNRFI) and *durable* (Durables) and *non-durable* personal consumption expenditure (Non-durables) components. I study these detailed housing and macro variables alongside important associated financial variables *home mortgage credit* (M) and *consumer credit* (CC) as well as *mortgage rates* (Mi) and *long-run rates* (LTi). In all cases I study year on year change in the variable under study using quarterly data over the sample period 1975-2017. I focus this analysis over the 4 to 10 year periodicity range over which I found GDP coherence with housing variables (section 3.1) exhibits a clear spectral peak since this significant and time persistent cycle seems a key feature to explain. In this first iteration I do not study the roughly 2 year cycle apparently shared by GDP and housing volumes, which is left to a further iteration/extension of the initial work presented here.

**Figure 9** presents the results from the 5% significance level thresholded bivariate global coherence weighted and phase-lag directed network resulting from the unconditional analysis (see methodology section 2.2). A number of striking results emerge from this initial network discovery procedure. The important and leading role of the housing market in the economic cycle is confirmed, however and perhaps most strikingly, existing home sales (EHS) emerge as the key “source” node in the network thus as a candidate ‘root cause’ of the economic cycle at this frequency. EHS have zero in-degree, and residential investment is led only by the existing home sales cycle (**Figure 10**). Both lead and are strongly associated with a number of other variables in the network. Indeed from EHS it is possible to reach every other node in the network except for long term rates (which is the only other variable with zero in-degree, however from long term rates it is only possible to reach financial variables with no path to aggregate demand). This gives EHS a unique position in the network and implies cycles in EHS may propagate to all variables in the network (besides LTi). Meanwhile house prices emerge as a sink node from which you can only reach financial (and no real) variables. This perhaps comes as a surprise given the growing literature identifying house prices as a leading indicator, as well as on account of the important role given to collateral values and wealth in macroeconomic theory; while housing market volumes have not been studied or theorised in the macroeconomic literature. Financial variables themselves emerge as pure sinks (with no path to aggregate demand). What is more credit quantities lead mortgage rates rather than the other way around (indeed mortgage rates emerge – all be it among a limited set of financial variables – as a global financial sink).

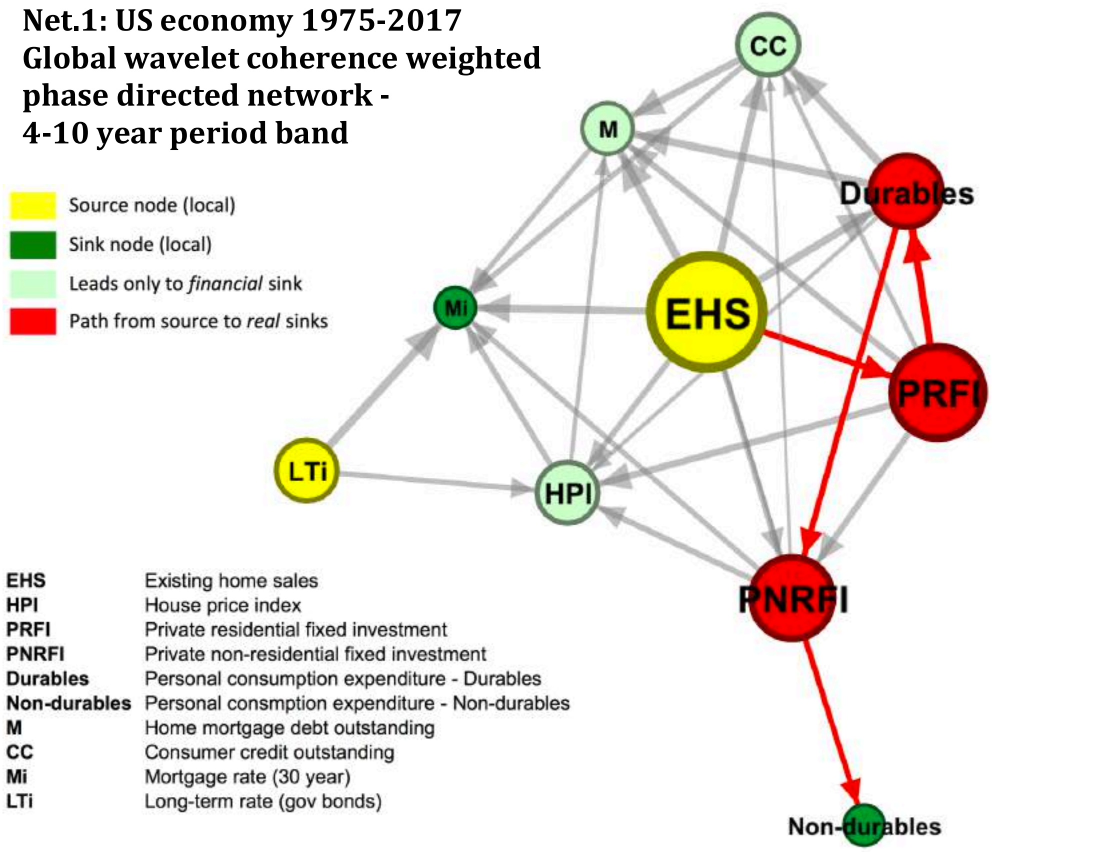



**Figure 9:** A directed edge (X –Y) signifies coherence significant at the 95% confidence level with X leading Y. Edges are *phase-lag* directed from lead to lag variable; edges are *global coherence* weighted; node size is proportional to *out-degree*; *source* nodes are coloured yellow and *sink* nodes are coloured dark green; nodes from which you can only reach financial variable sinks are coloured pale green. Red arrows and nodes trace a path from the real source EHS to the real sink Non-durables by always following the edge with the smallest phase-lag i.e. this path describes the temporal ordering between these variables. Analysis is based on year-on-year change for quarterly data from q1 1975 to q2 2017 (about 40 years of data).

Since the bivariate global wavelet coherence employed in estimating these edges is not a conditional statistic, it is very likely that these non-trivial links capture both direct and indirect relationships between variables. While these phase-lead statistics are not able to unpick independent from dependent lead/lags, they do provide a complete temporal ordering between the variables in the network. For example starting from the *source* existing home sales (EHS) (from where as already noted it is possible to reach any other node except long-term rates) and following the shortest phase-lead edge at each step (i.e. the nearest direct neighbour in a temporal sense) generates a path from *existing home sales* to the local sink *non-durable consumption* that runs *existing home sales →residential investment → durable consumption → non-residential investment → non-durable consumption* (this path is highlighted in red in **Figure 9**). While this temporal ordering does not by any means rule out other direct links, it may give indicate a transmission chain.

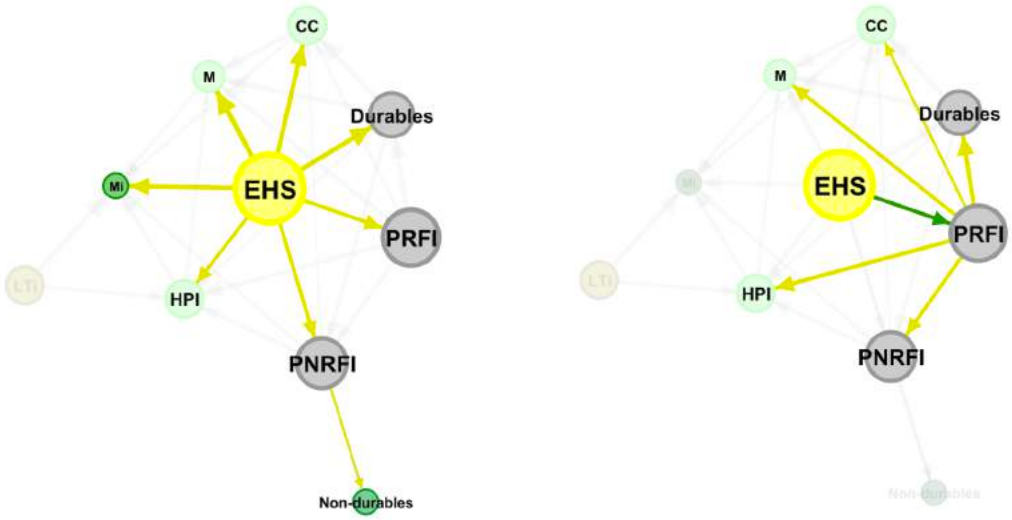



**Figure 10:** This highlights important aspects of the topology of the network presented in **Figure 9**. Yellow arrows indicate out-edges and green arrows indicate in-edges. We see that EHS has a zero in-degree, and that the only incoming edge received by residential investment is from existing home sales.

Overall the results of the initial network discovery exercise seems to raise questions regarding the precedence given to house prices and financial variables in the wider macroeconomic literature; confirm the importance of house building identified by a small existing literature; and suggest the interesting and unexpected possibility that a cycle in secondary housing market volumes drives the aggregate demand cycle. Although the possibility that the phase-lead existing sales have over residential investment may represent a faster response by existing home sales compared to residential investment to some omitted common driver must also be considered. Either way the important question arises of why housing market sales volumes cycle. This may be the result of two-way causality with other variables, something not captured by phase-direction which is by definition uni-directional.

## 3.3 Detailed conditional study of EHS links to demand variables

While network 1 (**Figure 9**) is already informative, it would be very helpful if we were able to disentangle independent vs. dependent (or direct vs. indirect) edges in this graph. For example EHS is PRI's only incoming edge, what is more since house sales volumes are widely read in the real estate industry as a key indicator of market demand, it seems relatively easy to imagine that house building might follow a cycle in existing home sales; however is there a direct channel from existing home sales to non-residential investment? Or is this non-trivial link mediated by residential investment and/or durable consumption etc.? Partial coherence and phase-lag statistics provide a tool with which we may be able to gain some further insight into these questions by helping us to unpick independent from dependent or direct from indirect edges in the network presented in **Figure 9.**

**Unpicking direct vs. indirect channels from EHS to GDP**

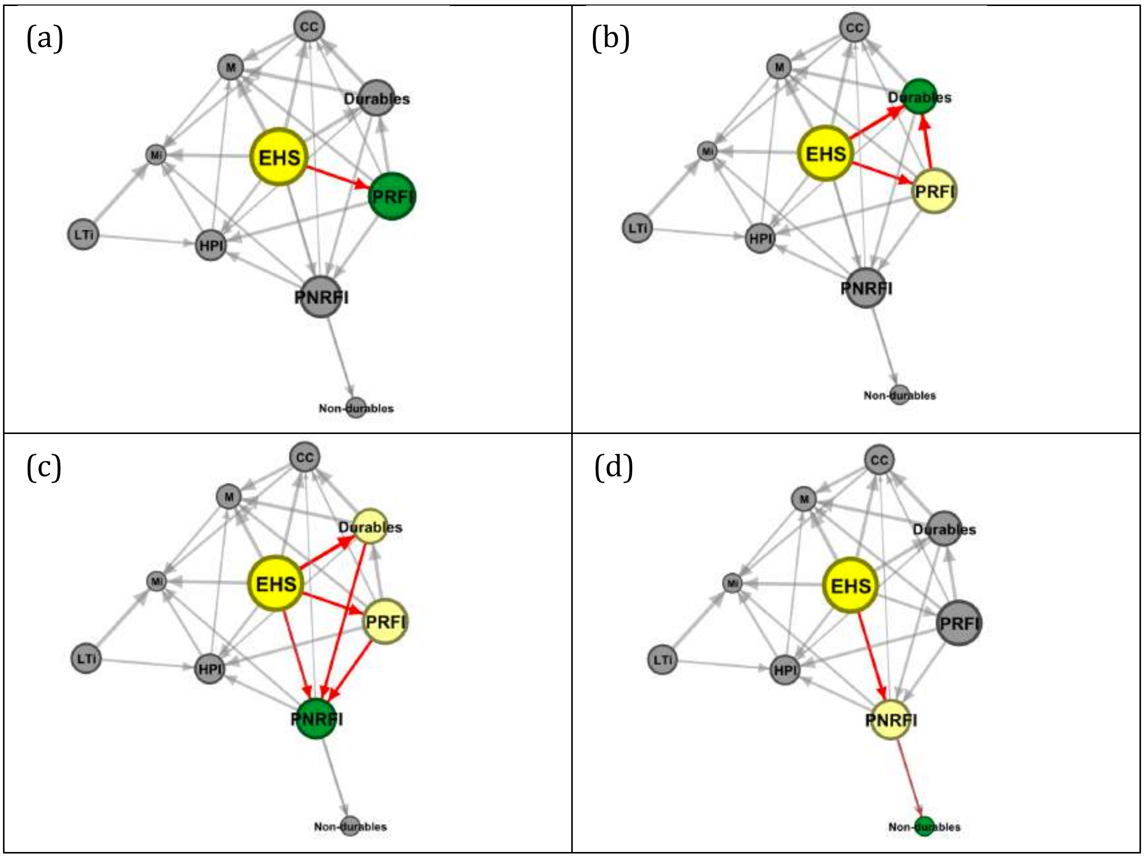



**Figure 11:** The graph resulting from network-discovery process in (**Figure 9**) suggests statistically significant coherence between existing home sales (EHS) and all four aggregate demand components included in the analysis (durable and non-durable consumption (Durables, Non-Durables), residential and non-residential investment (PRFI, PNRFI)) with EHS leading demand variables. The graph also suggests the possibility however, that some of these may be indirect or at least have indirect components. For example whilst EHS is the only leading edge for PRFI, both EHS and PRFI lead and exhibit significant coherence with durables. This raises the question whether EHS have any direct impact on durable consumption, or whether this relationship is mostly mediated by housing investment etc.

Since existing home sales emerge from my network discovery procedure as the key source variable and potential "root cause" of aggregate demand cycles at low frequencies in network 1, I will focus my further analysis on the links from EHS to demand variables. I will therefore look at the coherence and phase relationship for all edges between existing home sales and demand variables, comparing bivariate and partial results. For the partial analysis I will, more specifically, look at each of these edges conditional on the potential indirect links suggested by the topology of network 1 as presented in **Figure 9** (this idea is illustrated in **Figure 11**). Partial wavelet coherence and phase-difference conditional on other incoming edges should provide some insight into which demand components are directly, and which only indirectly impacted by housing transactions in order to try and achieve more detailed insight into the transmission path from housing to wider demand.

**EHS to PRI**

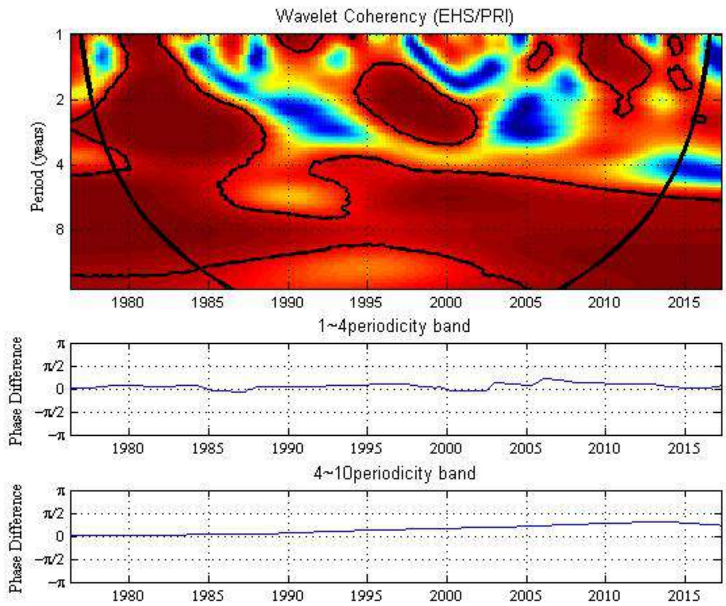



**Figure 12:** These charts present analysis of qtly existing home sales (EHS) and private residential fixed investment (PRFI) data 1975-2017.

Since private residential investment is led only by existing home sales (**Figure 11** (a)), I do not make a conditional analysis for this relationship, and look only at the bivariate coherence and phase-difference (**Figure 12**). This relationship exhibits unbroken extremely strong and significant coherence at low frequencies across the entire sample period (1975-2017) as well as across high and low frequency bands during the 1980s. Existing home sales seem to lead residential investment over high and low frequency bands, and at low frequencies this lead appears to have increased over time. These results seem consistent with the possibility that a cycle in existing house sales drives a cycle in new house building (consistent for example with the plausible argument that house builders respond to sales volumes as an indicator of market demand), but might also be consistent with existing sales responding faster than residential investment to some omitted common driver. It certainly confirms the strength of this association.

Since the network topology suggests the possibility that the link between existing home sales and durable consumption may be mediated by residential investment (**Figure 11** (b)), I estimate coherence and phase-difference for EHS and durables, not only bivariate but also conditional on residential investment (**Figure 13**). While the simple bivariate analysis shows strong and significant coherence at low frequencies and a strong lead of EHS over durables, the partial statistics suggest a more complex story. Conditional on PRI, EHS appear to have lagged durable consumption up until 1995 after which they led, this lead growing rapidly until they move out of phase, around 2004 with EHS now in the lead. However, it is striking that it is only since 2000, and especially since EHS peaked in 2005, that there is any strong and significant coherence between EHS and durable consumption, once the common cycle with PRI is partialed out.

**EHS to Durable Consumption**

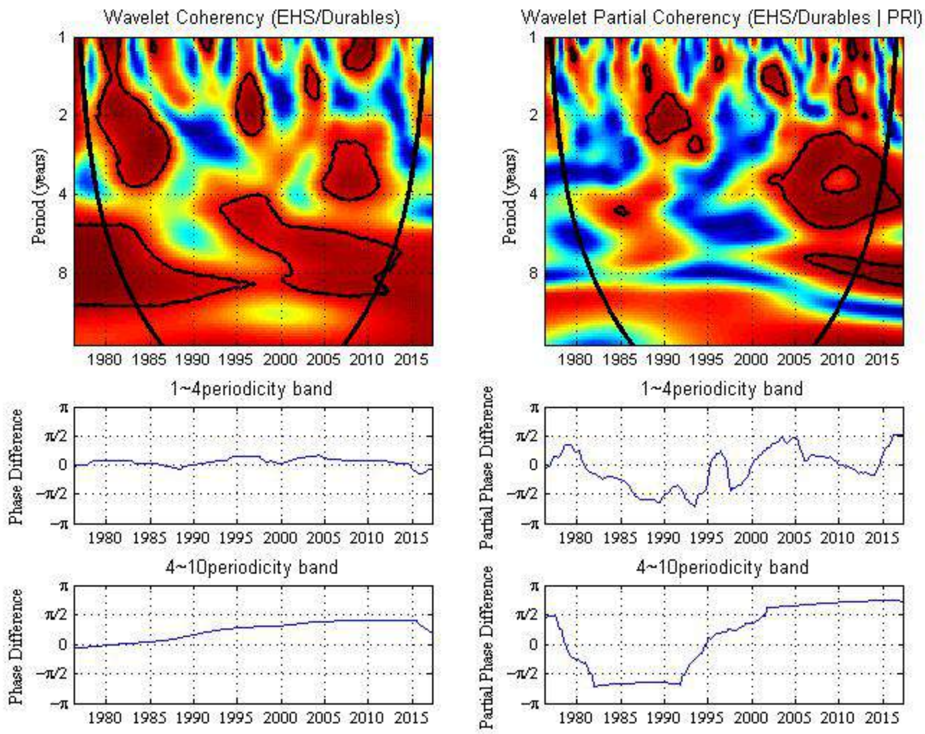



**Figure 13:** These charts present analysis of qtly existing home sales (EHS) and durable consumption (Durables) data 1975-2017. The left hand panel presents bivariate analysis, and the right hand panel partial analysis conditional on private residential fixed investment (PRI) – durable consumption’s only other leading edge (**Figure 9** and see **Figure 11** (b)). See figures in 3.1.2 for how to interpret the graphs.

Since the network topology suggests the possibility that the link between existing home sales and non-residential investment may be mediated by both residential investment and durables (**Figure 11** (c)), I estimate coherence and phase-difference for EHS and non-residential investment, not only bivariate but also conditional on residential investment and durables (**Figure 14**). Significant coherence shows up in this relationship from the 1990s with EHS leading strongly suggesting an important role for housing in this investment cycle, however this coherence seems to be eliminated in the conditional case suggesting the impact of EHS on non-residential investment may be principally operating via residential investment and durable consumption[15].

[15] While partial coherence at low frequencies seems mostly weak, partial phase lag shifts from strongly leading to strongly lagging up to 1995, after which EHS lead non-residential investment. This may simply reflect the strong acceleration in EHS. However we had better focus our phase analysis on areas of significant coherence.

**EHS to PNRI**

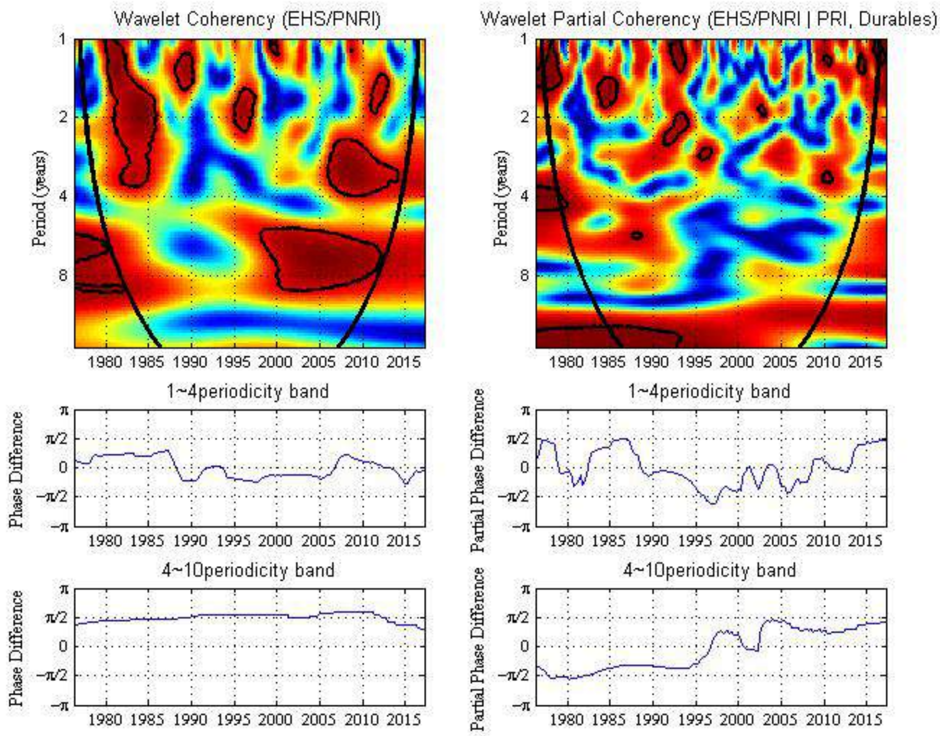



**Figure 14:** Bivariate and partial wavelet coherence and phase-difference for existing home sales (EHS) and private non-residential fixed investment (PNRFI). Partial statistics are conditional on private residential fixed investment (PRFI) and durable consumption (Durables) – PNRFI's other two leading edges in the network represented in **Figure 9** (see **Figure 11** (c)).

The much weaker coherence between existing home sales and non-durable consumption, since non-durable consumption is a local sink variable, is potentially mediated by a number of other variables (**Figure 11** (d)). I estimate coherence and phase-difference for EHS and non-durable consumption, not only bivariate but also conditional on non-residential investment (non-durable consumptions only other significant incoming edge) (**Figure 15**). While some areas of strong coherence show up - specifically from 2000 - and EHS lead consumption strongly in the bivariate analysis, lower frequency coherence seems to be substantially eliminated when studied conditional on non-residential investment suggesting that for low frequencies at least, this relationship is indirect and operating via non-residential investment, itself driven by residential investment and durables cycles[16].

[16] Interestingly some areas of significant coherence do show up at higher frequencies.

## EHS to Non-durables

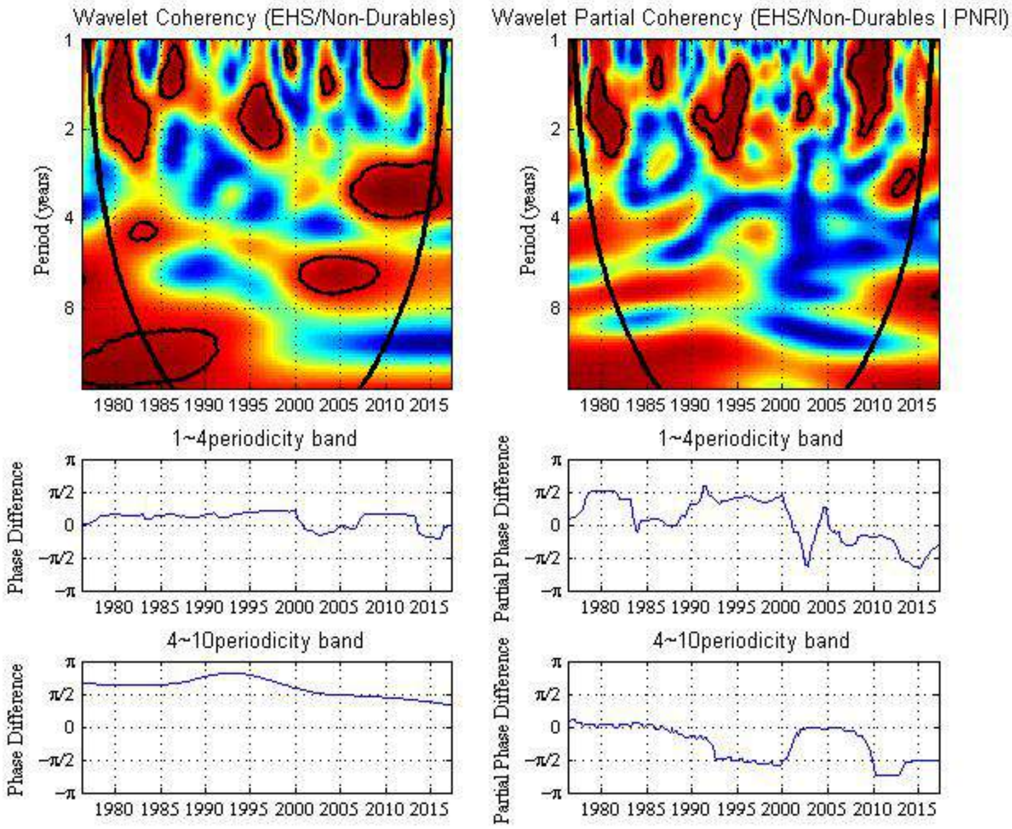



**Figure 15:** Bivariate and partial wavelet coherence and phase-difference for existing home sales (EHS) and non-durable consumption (Non-Durables). Partial statistics are conditional on private non-residential fixed investment - non-durables consumptions only other edge in the network represented in **Figure 9**. (see **Figure 11** (d)).

Overall a very significant and consistent relationship between existing home sales and residential investment is identified. There is also evidence of a non-trivial link direct path from market volumes to durable consumption operating independently of residential investment. While there is strong and significant coherence between housing market volumes and non-residential investment over the entire post 1995 period, partial analysis suggests this relationship may be largely indirect and substantially mediated by residential investment and durable consumption patterns. Similarly the relationship between housing market volumes and non-durable consumption seems to be mediated by non-residential investment, itself driven by residential investment and durable consumption.

## 4 Discussion and conclusion

While overall this data driven empirical work broadly confirms the increasingly recognised importance of housing cycles in the business cycle, a number of striking and novel results emerge. The importance widely attributed to house prices in existing macroeconomic literatures is not supported, with house prices emerging as a 'sink' variable not a 'source', lagging both aggregate demand and other housing variables. Meanwhile the importance and leading character of house building cycles previously identified in a small existing literature is very much confirmed with residential investment showing high coherence with and a temporal lead over other demand components. However the analysis also unexpectedly identifies non-trivial low frequency cycles in existing home sales (secondary housing market volumes) as leading residential investment cycles. Existing home sales thus emerge as the most leading variable. Since house sales volumes are widely regarded as an important indicator of market demand in the real estate industry, it seems entirely plausible that a cycle in existing home sales could transmit to house building, although the possibility that the phase-lead of EHS over residential investment could represent a faster response to some third common driver omitted from my analysis must also be considered. Either way the question of why secondary housing market volumes cycle emerges as an important issue. Partial analysis suggests non-trivial transmission from existing home sale volumes cycles not only to residential investment but also directly to durable consumption, with residential investment and durable consumption cycles transmitting to non-residential investment, and this "business cycle" in non-residential investment, further transmitting in turn to wider non-durable consumption. The possibility thus emerges that a low frequency cycle in housing market volumes (a variable and a cycle unstudied in the macroeconomic literature), propagates through the economy, and is the root cause of non-trivial cycles in aggregate demand. Further work is required to test and interpret these results, but they seem to suggest the basis for an alternative and novel theory of long-run economic cycles and beg further empirical and theoretical effort. Further research might: consider a larger set of variables; take a more unsupervised and systematic approach to deploying partial analysis for network thinning (and cleaning out indirect edges); also apply bi-directional directed statistics to uncover the causal structure of the network; pursue empirical and theoretical work on both why home sales cycle, as well as on studying and explaining the links between home sales and durables and house building (since these appear to be the most important and direct links). Perhaps some of the most obvious questions emerging from this study are: (1) What is the link between existing home sales and residential investment? (2) What is the link between existing home sales and durable consumption? But also and crucially, (3) Why do existing home sales volumes cycle? Seriously addressing these questions is outside of the scope of this particular study and something I take up in further work and papers.

**Acknowledgements**
I am happy and grateful to benefitting from code for wavelets analysis developed and made available by A. Grinsted and by Maria Joana Soares and Luís Aguiar-Conraria (available at below referenced web-links) to which I have referred extensively and from which I have benefited from very much in developing the analysis here presented. My thanks also to Engelbert Stockhammer and to Robert Jump for discussion and interest.

Grinsted: https://uk.mathworks.com/matlabcentral/fileexchange/47985-cross-wavelet-and-wavelet-coherence
Aguiar-Conrria and Soares: https://sites.google.com/site/aguiarconraria/-joanasoares-wavelets

## References

Adalid, Ramon, and Carsten Detken. 2007. *Liquidity Shocks and Asset Boom/bust Cycles*.

Aguiar-Conraria, Luís, and Maria Joana Soares. 2014. “The Continuous Wavelet Transform: Moving beyond Uni and Bivariate Analysis.” *Journal of Economic Surveys* 28(2):344–75.

Alessi, Lucia et al. 2014. “Comparing Different Early Warning Systems: Results from a Horse Race Competition among Members of the Macro-Prudential Research Network.” 1–27.

Alessi, Lucia, and Carsten Detken. 2009. *“Real Time” Early Warning Indicators for Costly Asset Price Boom/Bust Cycles: A Role for Global Liquidity*.

Álvarez, Luis J., and Alberto Cabrero. 2010. *Does Housing Really Lead the Business Cycle?*

Barbosa-filho, Nelson H., Codrina Rada Von Arnim, Lance Taylor, and Luca Zamparelli. 2008. “Cycles and Trends in U.S. Net Borrowing Flows.” *Journal of Post Keynesian Economics* 30(4):623–48.

Bauer, Margret et al. 2007. “Chemical Process Using Transfer Entropy.” 15(1):12–21.

Bezemer, Dirk, and Lu Zhang. 2014. “From Boom to Bust in the Credit Cycle: The Role of Mortgage Credit.”

Blasius, Bernd, Amit Huppert, and Lewi Stone. 1999. “Complex Dynamics and Phase Synchronization in Spatially Extended Ecological Systems.” *Nature* 399(May):354–59.

Borio, Claudio. 2014. “The Financial Cycle and Macroeconomics: What Have We Learnt?” *Journal of Banking and Finance* 45:182–98.

Borio, Claudio, and Mathias Drehmann. 2009. “Assessing the Risk of Banking Crises – Revisited.” *BIS Quarterly Review* (March):29–46.

Borio, Claudio, and Philip Lowe. 2002. *Asset Prices, Financial and Monetary Stability: Exploring the Nexus*.

Burns, Arthur F., and Wesley C. Mitchell. 1946. “Measuring Business Cycles.” in *Studies in Business Cycles no. 2*. New York: NBER.

Büyükkarabacak, B., and N. T. Valev. 2010. “The Role of Household and Business Credit in Banking Crises.” *Journal of Banking & Finance* 34(6):1247–56.

Cazelles, Bernard, Mario Chavez, Guillaume Constantin De Magny, Simon Hales, and Jean-francois Gue. 2007. “Time-Dependent Spectral Analysis of

Epidemiological Time-Series with Wavelets." *Journal of the Royal Society, Interface* 4:625–36.
Chatfield, C. 1989. *The Analysis of Time Series*. edited by Hall and Chapman.
Claessens, Stijn, Ayhan M. Kose, Marco E. Terrones, and IMF. 2012. "What Happens During Recessions, Crunches and Busts?" *Economic Policy* 60:653–700.
Crowe, Christopher, Giovanni Dell Ariccia, Deniz Igan, and Pau Rabanal. 2013. "How to Deal with Real Estate Booms: Lessons from Country Experiences." *Journal of Financial Stability* 9(3):300–319.
Damos, Petros. 2016. "Using Multivariate Cross Correlations, Granger Causality and Graphical Models to Quantify Spatiotemporal Synchronization and Causality between Pest Populations." *BMC Ecology* 1–17.
Daubechies, I. 1990. "The Wavelet Transform Time-Frequency Localization and Signal Analysis." *IEEE Transactions on Information Theory* 36:961–1004.
Daubechies, Ingrid. 1992. *Ten Lectures on Wavelets*. Philadelphia, USA: SIAM.
Davis, E.Philip, and Haibin Zhu. 2004. *Bank Lending and Commercial Property Cycles: Some Cross-Country Evidence*.
Davis, Morris, and Jonathan Heathcote. 2005. "Housing and the Business Cycle." *International Economic Review* 46:751–84.
Detto, Matteo et al. 2013. "Multivariate Conditional Granger Causality Analysis for Lagged Response of Soil Respiration in a Temperate Forest." 4266–84.
Dufrénot, Gilles, and Sheheryar Malik. 2012. "The Changing Role of House Price Dynamics over the Business Cycle." *Economic Modelling* 29(5):1960–67.
Ferrara, Laurent, and Olivier Vigna. 2009. *Cyclical Relationships between GDP and Housing Market in France: Facts and Factors at Play*.
Gabor, D. 1946. "Theory of Communication." *Journal Institute Electrical Engineers* 93:429–457.
Goodhart, Charles, and Boris Hofmann. 2008. *House Prices, Money, Credit and the Macroeconomy*.
Goupillaud, P., A. Grossmann, and J. Morlet. 1984. "Cycle Octave and Related Transforms in Seismic Signal Analysis." *Geoexploration* 23:85–102.
Grinsted, A., J. C. Moore, and S. Jevrejeva. 2004. "Application of the Cross Wavelet Transform and Wavelet Coherence to Geophysical Time Series." *Nonlinear Processes in Geophysics* 11:561–66.
Hudgins, L., C. Friehe, and M. Mayer. 1993. "Wavelet Transforms and Atmospheric Turbulence." *Physics Review Letters* 71:3279–3282.
Igan, Deniz, Alain Kabundi, Francisco Nadal, De Simone, and Marcelo Pinheiro. 2011. "Housing, Credit, and Real Activity Cycles: Characteristics and Comovement." *Journal of Housing Economics* 20(3):210–31.
IMF. 2008. "The Changing Housing Cycle and the Implications for Monetary Policy." Pp. 103–32 in *World Economic Outlook: Housing and the Business Cycle*.
IMF. 2012. *World Economic Outlook, Chapter 3*.
IMF, Deniz Igan, A. Kabundi, and Nd Simone. 2009. *Three Cycles: Housing, Credit, and Real Activity*.
IMF, Daniel Leigh, Deniz Igan, John Simon, and Petia Topalova. 2012. "Dealing With Household Debt." Pp. 1–36 in *IMF World Economic Outlook*.
Jorda, Oscar, Moritz Schularick, and Alan M. Taylor. 2012. "When Credit Bites Back: Leverage, Business Cycles, and Crises." *Federal Reserve Bank of San Francisco Working Paper Series* (October):1–42.

Jorda, Oscar, Moritz Schularick, and Alan M. Taylor. 2014. *The Great Mortgaging : Housing Finance, Crises, and Business Cycles*.

Juselius, Mikael, and Mathias Drehmann. 2015. "Leverage Dynamics and the Real Burden of Debt." (501).

Kindleberger, C., and R. Aliber. 2005. *Manias, Panics & Crashes: A History of Financial Crises*. edited by N J Hoboken. Wiley.

King, Mervyn. 1994. "Debt Deflation: Theory and Evidence." *European Economic Review* 38:419–45.

Kuramoto, Yoshiki. 1984. *Chemical Oscillations,Waves, and Turbulence*. Berlin: Springer.

Kydland, Finn E., Peter Rupert, and Roman Sustek. 2016. "Housing Dynamics over the Business Cycle." *International Economic Review* 57(4):1149–78.

Leamer, E. E. 2007. *Housing Is the Business Cycle*.

Long, Clarence Jr. 1940. *Building Cycles and the Theory of Investment*. Princeton and London: Princeton University Press and Oxford University Press.

Mallat, S. G. 1999. *A Wavelet Tour of Signal Processing*. Academic Press.

Mian, Atif, Amir Sufi, and Emil Verner. 2015. "Household Debt and Business Cycles Worldwide." *NBER Working Papers* (609).

Mormann, Florian, Klaus Lehnertz, Peter David, and Christian E. Elger. 2000. "Mean Phase Coherence as a Measure for Phase Synchronization and Its Application to the EEG of Epilepsy Patients." 144:358–69.

Ng, Eric K. W., and Johnny C. L. Chan. 2012. "Geophysical Applications of Partial Wavelet Coherence and Multiple Wavelet Coherence." *Journal of Atmospheric and Oceanic Technology* 29:1845–53.

Nolte, Guido et al. 2008. "Robustly Estimating the Flow Direction of Information in Complex Physical Systems." *Physical Review Letters* 100.

Pikovsky, A. S., M. G. Rosenblum, and J. Kurths. 2001. *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge, UK: Cambridge University Press.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2011. *This Time Is Different*. Princeton University Press.

Rosenblum, Michael G., Arkady S. Pikovsky, and Jürgen Kurths. 1996. "Phase Synchronization of Chaotic Oscillators." *Physical Review Letters* 76(11):1804–7.

Rosenblum, Michael, Arkady Pikovsky, Jurgen Kurths, Carsten Schafer, and Peter A. Tass. 2001. "Phase Synchronization: From Theory to Data Analysis." Pp. 279–321 in *Handbook of Biological Physics, Vol. 4, Neuro-informatics*, edited by F Moss and S Gielen.

Schularick, Moritz, and Alan M. Taylor. 2009. *Credit Booms Gone Bust: Monetary Policy, Leverage Cycles and Financial Crises, 1870-2008*.

Schulte, Justin A., Raymond G. Najjar, and Ming Li. 2016. "Journal of Hydrology: Regional Studies The Influence of Climate Modes on Streamflow in the Mid-Atlantic Region of the United States." *Journal of Hydrology: Regional Studies* 5:80–99.

Shao, Pei-chiang et al. 2015. "Granger Causality-Based Synaptic Weights Estimation for Analyzing Neuronal Networks." *Journal of Computational Neuroscience* 38(3):483–97.

Soares, Maria Joana, and Luís Aguiar-Conraria. 2011. *The Continuous Wavelet Transform: A Primer*.

Szemere, Robert, Michela Scatigna, and Kostas Tsatsaronis. 2014. "Residential Property Price Statistics across the Globe." *BIS Quarterly Review*

(September):61–76.
Torrence, Christopher, and Gilbert P. Compo. 1998. "A Practical Guide to Wavelet Analysis." *Bulletin of the American Meteorological Society* 79:61–78.
Williamson, Mark S., Sebastian Bathiany, and Timothy M. Lenton. 2016. "Early Warning Signals of Tipping Points in Periodically Forced Systems." *Earth System Dynamics* 7:313–26.
Yuan, Tao, and S.Joe Qin. 2014. "Root Cause Diagnosis of Plant-Wide Oscillations Using Granger Causality." *Journal of Process Control* 24(2):450–59.
Zhang, Wenjun. 2011. "Constructing Ecological Interaction Networks by Correlation Analysis: Hints from Community Sampling." *Network Biology* 1(2):81–98.
Zou, Cunlu, Christophe Ladroue, Shuixia Guo, and Jianfeng Feng. 2010. "Identifying Interactions in the Time and Frequency Domains in Local and Global Networks - A Granger Causality Approach." *Bioinformatics* 11(337).

# 5 Appendix

## 5.1 Wavelets

**Figure 16: Morlet wavelet**

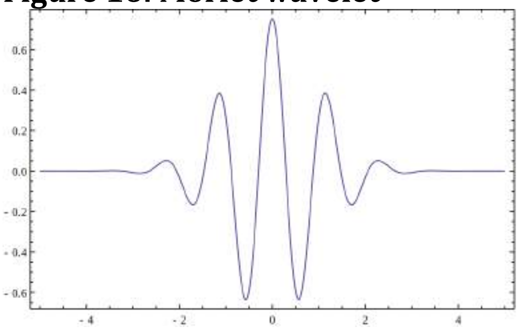
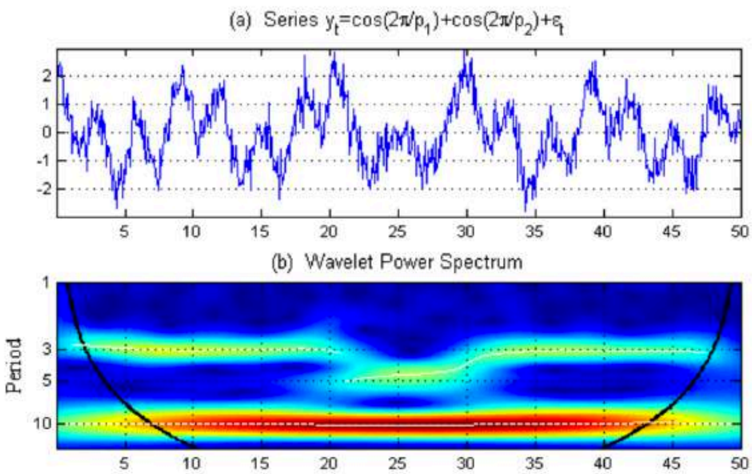
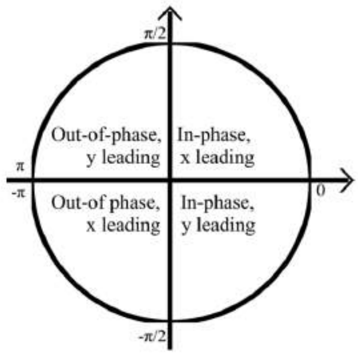


**Figure 17: Wavelets power spectrum plot**




*Source:* (Soares and Aguiar-Conraria 2011)

**Figure 18: Phase-difference circle**

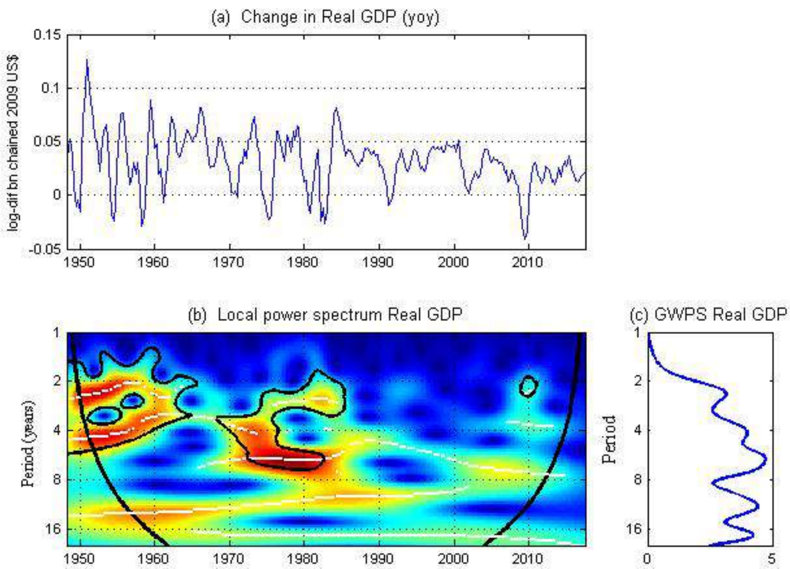



## 5.2 Data definitions and sources

| **Variable** | **Description** | **Source** |
|---|---|---|
| Existing Home Sales (EHS) | Number of units. Existing-home sales are based on closing transactions of single-family, town-homes, condominiums and cooperative homes. | National Association of Realtors (Release *Existing Home Sales*). |
| New Home Sales (NHS) | | U.S. Bureau of the Census (Release: *New Residential Sales*). |
| New housing starts | Total new privately owned housing units started. | U.S. Bureau of the Census (Release: *New Residential Construction*). |
| House price index (HPI) | All-Transactions House Price Index for the United States. Estimated using sales prices and appraisal data. | U.S. Federal Housing Finance Agency (Release: *House Price Index*). |
| House price index (HPI) | | National sources, BIS Residential Property Price database. |
| Long-term rates (LTi) | Long-term interest rates refer to government bonds maturing in ten years. Rates are implied by price at which bonds traded. | OECD Long-term interest rates, US. |
| Mortgage rates (Mi) | 30-Year Fixed Rate Mortgage Average in the United States. | Freddie Mac (Release: Primary Mortgage Market Survey). |
| Mortgage debt (M) | Households and Nonprofit Organizations; Home Mortgages; Liability, Level. | Board of Governors of the Federal Reserve System (US). (Release: Z.1 Financial Accounts of the United States. Series ID LA153165105.Q.) |
| Private residential fixed investment (PRFI) | | U.S. Bureau of Economic Analysis, Gross Domestic Product (BEA Account Code: A011RC1). |
| Private non-residential fixed investment (PNRFI) | Private investment, fixed, non-residential. | U.S. Bureau of Economic Analysis, Gross Domestic Product (BEA Account Code: A008RC1). |
| Durable consumption | Personal consumption expenditure – durables. | U.S. Bureau of Economic Analysis, Gross Domestic Product (BEA Account Code: DDURRC1). |
| Non-durable consumption | Personal consumption expenditure – non-durables. | U.S. Bureau of Economic Analysis, Gross Domestic Product (BEA Account Code: DNDGRC1). |

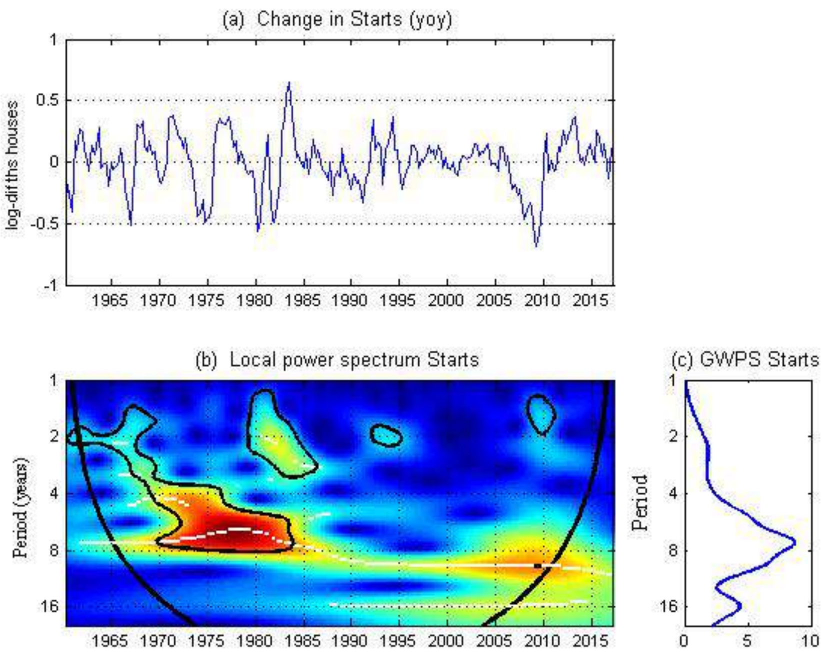
## 5.3 Power spectrums

**Figure 19: US real GDP power spectrum**




**Note**: This figure presents a the GWPS for US real GDP data over a longer period reflecting the greater availability of GDP data compared to housing variables.

**Figure 20: US New Housing Starts power spectrum**



## 5.4 Coherence analysis

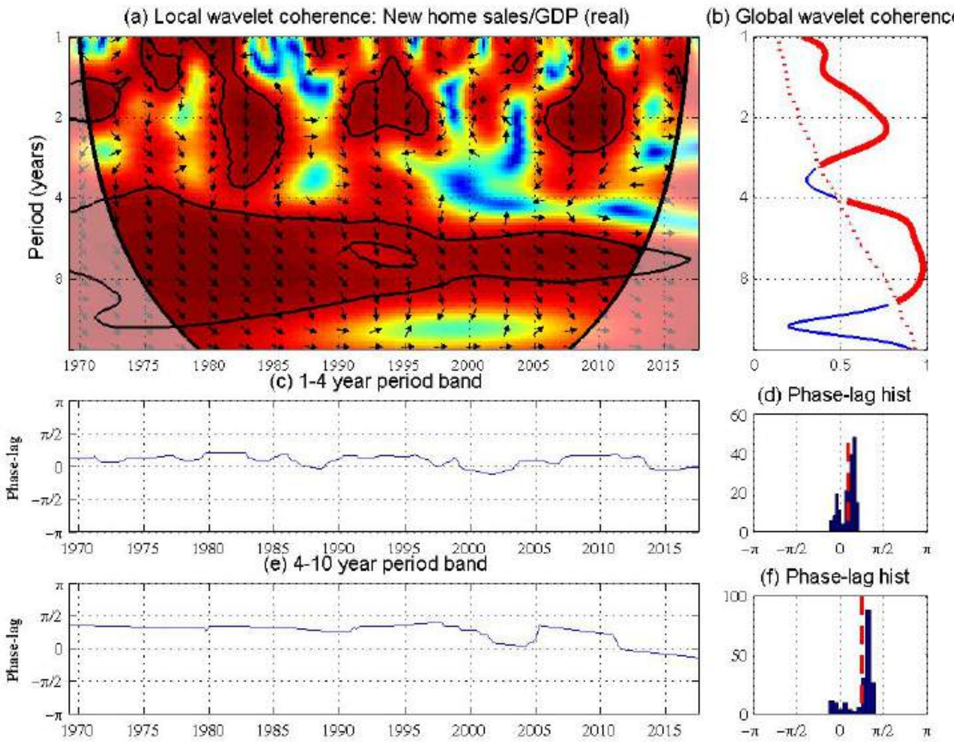



**Figure 21:** (a) Wavelet coherency between US new house sales (---) and GDP; (b) Global wavelet coherence; (c) and (e) Phase-differences at 1-4 and 4-10 year periodicity bands; (d) and (f) Histograms of the phase-difference statistics presented in (c) and (e). Data is quarterly for 1968-2017.

## 5.5 Network analysis

### Table 1: Network edge-list, 4 – 10 year periodicity range

| Source | Target | Type | Peak GWC | Phase-lead | Phase relation |
|---|---|---|---|---|---|
| EHS | HPI | Directed | 0.873205902 | 1.401565669 | in |
| EHS | PRFI | Directed | 0.898031239 | 0.441830162 | in |
| EHS | PNRFI | Directed | 0.879009154 | 1.28100398 | in |
| EHS | Durables | Directed | 0.959394669 | 1.085639144 | in |
| EHS | Non-durables | Directed | 0.784801614 | 1.362675803 | in |
| EHS | M | Directed | 0.977767164 | 1.204058502 | in |
| EHS | CC | Directed | 0.940480316 | 1.037676293 | in |
| EHS | Mi | Directed | 0.938586867 | 1.339486566 | in |
| PRFI | HPI | Directed | 0.922052921 | 0.989941168 | in |

| | | | | | |
|---|---|---|---|---|---|
| PNRFI | HPI | Directed | 0.881762747 | 0.269187647 | in |
| Durables | HPI | Directed | 0.802529079 | 0.243031644 | in |
| HPI | M | Directed | 0.809207365 | 0.406619204 | in |
| LTi | HPI | Directed | 0.849120169 | 0.031356293 | in |
| HPI | Mi | Directed | 0.869811429 | 0.032906525 | in |
| PRFI | PNRFI | Directed | 0.892822884 | 1.044605769 | in |
| PRFI | Durables | Directed | 0.936727032 | 0.473606137 | in |
| PRFI | M | Directed | 0.893836158 | 1.269327101 | in |
| PRFI | CC | Directed | 0.82545729 | 1.014315502 | in |
| Durables | PNRFI | Directed | 0.877448732 | 0.638334436 | in |
| PNRFI | Non-durables | Directed | 0.857667064 | 0.489254184 | in |
| PNRFI | CC | Directed | 0.751198973 | 0.347969466 | in |
| PNRFI | Mi | Directed | 0.856307984 | 0.315884131 | in |
| Durables | M | Directed | 0.960152905 | 0.844338328 | in |
| Durables | CC | Directed | 0.974649014 | 0.9838477 | in |
| CC | M | Directed | 0.929305634 | 0.562106848 | in |
| M | Mi | Directed | 0.850688775 | 0.758749918 | in |
| CC | Mi | Directed | 0.86308292 | 0.242247131 | in |

## 5.6 Periodicity band selection

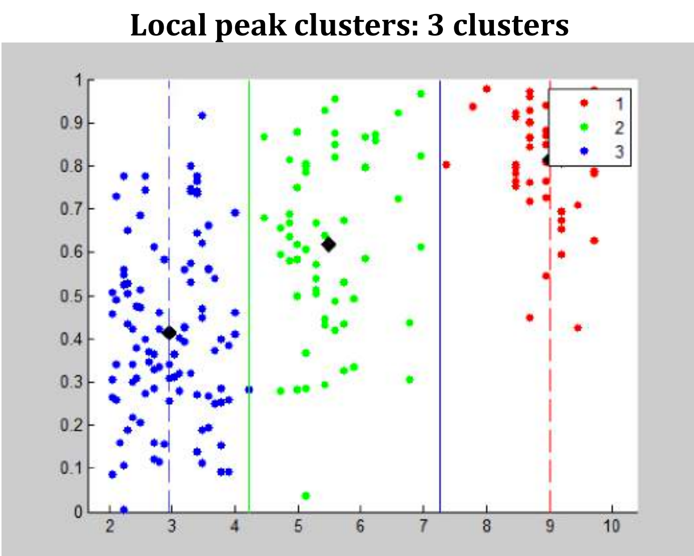



**Figure 22:** Clusters based on k-means. Three centroids have periodicity of 9, 5.5 and 3 years leading to periodicity band splits of 1-4.2, 4.2- 7.3 and 7.3 to 10. This k=3 clustering produces stable result, where k=2 unstable.