Online Appendix to
Putting the Cycle Back into Business Cycle Analysis

Paul Beaudry*, Dana Galizia† and Franck Portier‡

June 2019

A Data

– U.S. Total Population: All Ages including Armed Forces Overseas, obtained from the FRED database (`POP`) from 1952Q1 to 2015Q2. Quarters from 1947Q1 to 1952Q1 are obtained from linear interpolation of the annual series of National Population obtained from U.S. Census, where the levels have been adjusted so that the two series match in 1952Q1.

– U.S. GDP and investment are obtained from the Bureau of Economic Analysis National Income and Product Accounts. The investment/GDP ratio is computed as the ratio of nominal gross private domestic investment to nominal GDP (Table 1.1.5). Real GDP (used in Appendix F) is computed as nominal GDP (Table 1.1.5) over prices (Table 1.1.4.). Sample is 1948Q1-2015Q2.

– U.S. Non-Farm Business Hours, Total Hours, Unemployment Rate (16 years and over), and the employment rate (employment-population ratio, 16 years and over) are obtained from the Bureau of Labor Statistics. Sample is 1948Q1-2015Q2.

– U.S. TFP: utilization-Adjusted quarterly-TFP series for the U.S. Business Sector, produced by John Fernald, series ID: `dtfp_util`, 1947Q1-2015Q2.

*Bank of Canada
†Department of Economics, Carleton University
‡Department of Economics, University College London and CEPR

- U.S. Spread: Moody's Seasoned Baa Corporate Bond Minus Federal Funds Rate, quarterly average, obtained from the FRED database, (`BAAFFM`). Sample is 1954Q3-2015Q2.

- U.S. AAA and BAA rates are from FRED (series names AAA and BAA). Samples are from 1920Q1-2015Q2. Spreads are computed by subtracting the T-bill rate, which is taken from Ramey and Zubairy [2018].

- U.S. National Financial Conditions Index: Chicago Fed National Financial Conditions Index, Index, Quarterly, Not Seasonally Adjusted, obtained from the FRED database, (`NFCI`). Sample is 1973Q1-2015Q2.

- U.S. National Financial Conditions Risk Subindex: Chicago Fed National Financial Conditions Risk Subindex, Index, Quarterly, Not Seasonally Adjusted, obtained from the FRED database, (`NFCIRISK`). Sample is 1973Q1-2015Q2.

- U.S. Policy rate : computed from the FRED database as Moody's Seasoned Baa Corporate Bond Yield (`BAAFFM`) minus the spread (`BAAFFM`). Sample is 1954Q3-2015Q2. The real policy rate is obtained by subtracting realized inflation, computed using the "Consumer Price Index for All Urban Consumers: All Items" (`CPIAUCSL` from the FRED database).

- Price Earning ratio : computed as Real S&P Composite Stock Price Index divided by Real S&P Composite Earnings. Data are described in Shiller [2015], and obtained from Robert Shiller's webpage (http://bit.ly/Shiller-data). Sample is 1971Q1-2015Q2.

- Business cycles reference dates: NBER for the U.S.A. (https://goo.gl/n7tS9s), Owyang, Ramey, and Zubairy [2013] for Canada, Economic Cycle Research Institute (https://goo.gl/RycWUt) for the other countries of the G7.

- Unemployment rates for the other countries of the G7: obtained from the FRED database: `LRUNTTTTDEQ156S` for Germany (sample: 1962Q1-2017Q3), `AURUKM` for the U.K. (sample: 1948Q1-2067Q4), `ITAURHARMQDSMEI` for Italy (sample: 1979Q4-2017Q3), `LRUN74TTFRQ156N` for France (sample: 1967Q4-2017Q3), `LRHUTTTTJPQ156S` for Japan (sample: 1960Q1-2017Q4), `LRHUTTTTCAQ156S` for Canada (sample: 1960Q1-2017Q4).

B Conditional Probability of a Recession

In this Appendix, we discuss how to compute confidence intervals for the conditional probabilities of recession (see panel (b) of Figure 1 in the main text), and present confidence intervals for the 3- and 4-quarter window cases. The intervals were constructed as follows. Fix a time horizon k and a window size x. Let r_t be an indicator for whether the economy is in recession at date t, and let $y_t^{k,x}$ be an indicator for whether the economy is in recession at some point in an x-quarter window around date $t+k$. We may then run the regression

$$y_t^{k,x} = \beta_0^{k,x} + \beta_1^{k,x} r_t + \epsilon_t^{k,x} .$$

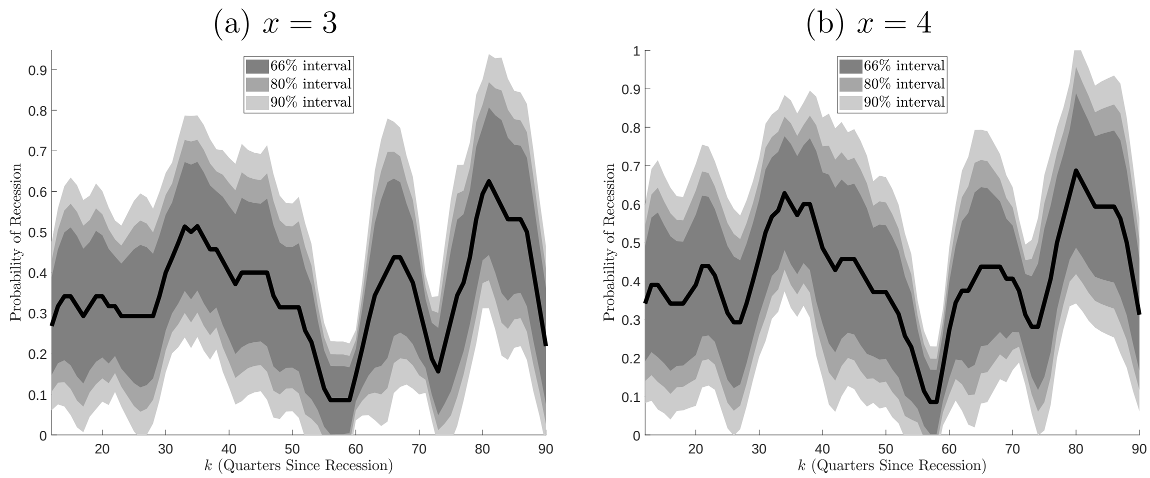
Letting $\widetilde{\beta}_0^{k,x}$ and $\widetilde{\beta}_1^{k,x}$ denote the true values of the regression parameters, the quantity $\widetilde{p}^{k,x} \equiv \widetilde{\beta}_0^{k,x} + \widetilde{\beta}_1^{k,x}$ is the true probability of being in recession some time in the x-quarter window around date $t+k$, conditional on being in recession at date t. Letting hats denote estimates of these parameters, $\widehat{p}^{k,x} \equiv \widehat{\beta}_0^{k,x} + \widehat{\beta}_1^{k,x}$ is a point estimate of that probability which, it can be verified, coincides asymptotically with the point estimates shown in Figure 1.[i] Letting $\widehat{\Omega}$ denote an estimate of the variance-covariance matrix of $(\widehat{\beta}_0^{k,x}, \widehat{\beta}_1^{k,x})'$, and $\eta \equiv (1,1)'$, then $\widehat{p}^{k,x}$ is approximately distributed as $N(\widetilde{p}^{k,x}, \eta'\widehat{\Omega}\eta)$. We can then construct approximate confidence intervals from this distribution in the usual way. In practice, we use the Newey and West [1994] HAC estimator for $\widehat{\Omega}$.

In Figure B.1, we present 66%, 80%, and 90% confidence intervals for the $x = 3$ and $x = 4$ quarter windows, point estimates for which are presented in panel (a) of Figure 1. In both cases, the confidence intervals indicate that the general pattern we highlighted in Figure 1—namely, a local peak in the probability for $k \approx$ 36-40 quarters, followed by a local trough around 56-60 quarters—is unlikely to have occurred simply by chance. Furthermore, in conjunction with panel (b) of Figure 1 we see that as the window size increases, the estimates become less uncertain and the aforementioned pattern becomes more starkly apparent.

We also perform a joint test of the null hypothesis that the conditional probability is constant in k, focusing on the case where $x = 5$ (size of the window) for the sake of brevity. The tests are structured as follows. Let $K = \{k_1, \ldots, k_n\}$ denote a set of "quarters since

[i] Indeed, even in our finite sample the two different ways of estimating this probability produce nearly identical results.

Figure B.1: Confidence Intervals for Recession Probability Plots



Notes: Panels (a) and (b) shows 66%, 80%, and 90% confidence intervals (gray shaded areas) for window widths $x = 3$ and $x = 4$, respectively. The solid black line in each panel gives the point estimates.

recession". For a window x, let ϕ_j denote the true probability of being in a recession in the window $t + k_j \pm x$, and p_j our estimate of that parameter. We test the null hypothesis

$$H_0 : \phi_j = \bar{p}, \quad \forall j = 1, \ldots, n$$

where $\bar{p} \equiv n^{-1} \sum_{i=1}^{n} p_j$. That is, H_0 is the null that all the probabilities are equal to the average estimated probability. Let Φ and P denote the vectors of probabilities and their estimates, $\Sigma = Var\,(P)$, and

$$\eta \equiv \Sigma^{-1/2} \left(\hat{P} - \bar{p} e_n \right)$$

where e_n is an n-vector of 1's. Then under the null, η is approximately $N\,(0, I_n)$, and thus $\eta' \eta$ is approximately χ_n^2.

We report results using two different estimates $\hat{\Sigma}$ of the unknown Σ. The first is a Newey-West HAC estimator (the same method used to generate the CIs reported in the paper). One might argue that the standard asymptotic HAC estimator might not be the right approach for this type of data (i.e., binary data with a strong degree of persistence in both the regressor and regressand). For this reason, we also provide results from using a bootstrap procedure to estimate Σ. We do so in the following way. We split the actual data series into full cycles from peak to peak (of which there are 11 total for the post-war US

Table B.1: Joint Test for Recession Probability Plots

	χ^2 stat.	r (d.o.f.)	p-value
Case 1			
Asymptotic	7.973	2	0.0186**
Bootstrap	5.878	2	0.0529*
Case 2			
Asymptotic	56.110	8	0.0000***
Bootstrap	24.109	8	0.0022***

Notes: This Table reports test values and p-values for joint test of the null hypothesis that the conditional probability is constant in k, focusing on the case where $x = 5$.

data). We then generate a simulated business cycle "history" by sequentially drawing full cycles until the history is at least as long as the true data history (i.e., 286 quarters). From this we can derive a series of zero-one recession indicators (truncating the histories at 286 quarters) and use it to estimate the recession probabilities just as in the data. We do this many ($N = 10,000$) times, and then use it to obtain a covariance matrix for the estimated probabilities.

Results using both the asymptotic HAC estimator and the bootstrap procedure are reported in Table B.1:

In Case 1, we compare whether the probability of a recession at 38 quarters (always $\pm x = 5$ quarters) is different form the probability of a recession at 56 quarters (i.e., $K = \{38, 56\}$). In Case 2, we compare whether the probability of a recession between 36 and 39 quarters is equal to the probability of a recession between 55 and 58 quarters (i.e., $K = \{36\text{-}39, 55\text{-}58\}$). In both cases these tests provide evidence against the null hypothesis that recessions are time-independent, in favor of some form of cyclicality. One could argue that such a testing procedure is cherry-picking particular horizons. We agree with this view, and that is why we mainly view these test as providing the basis for a hypothesis about the potential timing of cyclical forces. We thus maintain this range of timing when we go on to test for humps in the spectral densities across different series. We also performed the above χ^2 tests on a longer sample that goes back to 1850 for the US. Here again we found evidence – at the 5% level or less — against the hypothesis of a time-independent conditional probability.

C Spectral Density Estimation

C.1 Schuster's Periodogram

We estimate the spectral density of series $\{x_t\}_{t=0}^{T-1}$ of finite length T by first computing the Discrete Fourier Transform (DFT) X_k, which results from sampling the Discrete Time Fourier Transform (DTFT) at frequency intervals $\Delta\omega = \frac{2\pi}{T}$ in $[-\pi, \pi)$:

$$X_k = X\left(e^{i\frac{2\pi}{T}k}\right) = \sum_{t=0}^{T-1} x_t e^{-i\frac{2\pi}{T}kt}, \tag{C.1}$$

for $k = 1, ..., T-1$. We then compute samples of the Sample Spectral Density (SSD) S_k from samples of Schuster's periodogram I_k[ii] according to

$$S_k = I_k = \frac{1}{T}\left|X_k\right|^2 \tag{C.2}$$

Taking advantage of the fact that X is even, this amounts to evaluating the spectral density at T frequencies equally spaced between 0 and π.[iii]

C.2 Zero-Padding to Increase the Graphic Resolution of the Spectrum

As we have computed only T samples of the DTFT $X\left(e^{i\omega}\right)$, we might not have a detailed enough picture of the shape of the underlying function $X\left(e^{i\omega}\right)$, and therefore of the spectral density $\left|X\left(e^{i\omega}\right)\right|^2$. This problem is particularly acute if one is interested in the behavior of the spectrum at longer periodicities (i.e., lower frequencies). Specifically, since we uniformly sample frequencies, and since the periodicity p corresponding to frequency ω is given by $p = \frac{2\pi}{\omega}$, the spectrum is sparser at longer periodicities (and denser at shorter ones). While the degree of accuracy with which the samples of X_k describe the shape of $X\left(e^{i\omega}\right)$ is dictated and limited by the length T of the data set, we can nonetheless increase the number of points at which we sample the DTFT in order to increase the graphic resolution of the spectrum. One common (and numerically efficient) way to do this is to add a number of zeros to the end

[ii] Another approach for obtaining the spectral density is to take a Fourier transform of the sequence of autocovariances of x. We show below that this method gives essentially the same result when applied to our hours series.

[iii] See Priestley [1981] for a detailed exposition of spectral analysis, Alessio [2016] for practical implementation and Cochrane [2005] for a quick introduction.

of the sequence x_t before computing the DFT. This operation is referred to as zero-padding. As an example, suppose that we add exactly T zeros to the end of the length-T sequence $\{x_t\}$. One can easily check that this has no effect on the DFT computed at the original T sampled frequencies, instead simply adding another set of T sampled frequencies at the midpoints between each successive pair of original frequencies.[iv]

If one is interested in the behavior of the spectral density at long enough periodicities, zero-padding in this way is useful. We will denote by N the number of samples at which the DTFT (and thus the SSD) is sampled, meaning that $T' = N - T$ zeros will be added to the sequence $\{x_t\}$ before computing the DFT. In the main text, we have set $N = 1,024$.[v]

C.3 Smoothed Periodogram Estimates

We obtain the raw spectrum estimate of a series non-parametrically as the squared modulus of the DFT of the (zero-padded) data sequence, divided by the length of the data set.[vi] This estimate is called Schuster's periodogram, or simply the periodogram. It turns out that the periodogram is asymptotically unbiased, but is not a consistent estimate of the spectrum, and in particular the estimate of the spectrum at a given frequency ω_k is generally quite unstable (i.e., it has a high standard error). Notwithstanding this fact, the overall pattern of the spectrum is much more stable, in the sense that the average value of the estimated spectrum within a given frequency band *surrounding* ω_k is in fact consistent. In order to obtain a stable and consistent estimate of the spectrum, we exploit this fact by performing frequency-averaged smoothing. In particular, we obtain our estimate of the SSD $S(\omega)$ by kernel-smoothing the periodogram $I(\omega)$ over an interval of frequencies centered at ω. Since the errors in adjacent values of $I(\omega)$ are uncorrelated in large samples, this operation reduces the standard errors of the estimates without adding too much bias. In our estimations, we

[iv]This is true when the number of zeros added to the end of the sample is an integer multiple of T. When instead a non-integer multiple is added, the set of frequencies at which the padded DFT is computed no longer contains the original set of points, so that the two cannot be directly compared in this way. Nonetheless, the overall pattern of the sampled spectrum is in general unaffected by zero-padding.

[v]As is well known, standard numerical routines for computing the DFT (i.e., those based on the Fast Fourier Transform algorithm) are computationally more efficient when N is a power of 2, which is why we set $N = 1{,}024$ rather than, say, $N = 1{,}000$.

[vi]Note that we divide by the original length of the series (i.e., T), rather than by the length of the zero-padded series (i.e., N).

use a Hamming window of length $W = 13$ as the smoothing kernel.[vii]

C.4 Confidence Intervals

We compute bootstrap confidence intervals for our spectral density estimates as follows. Let $X = (x_1, \ldots, x_T)'$ be a sample from a (mean-zero) data series with (true) spectrum $S(\omega)$, and let $\widehat{S}(\omega_k; X)$, be our consistent estimator of this spectrum at the discrete frequencies $\omega_1, \ldots, \omega_K$ computed from the sample X. We would like to know the sampling distribution of the ratio of the estimated spectrum to the true smoothed spectrum, i.e., of

$$\frac{\widehat{S}\left(\omega_k; X\right)}{S\left(\omega_k\right)} . \tag{C.3}$$

To estimate this distribution, suppose that $\widehat{S}(\cdot; X)$ is the "true" spectrum, and then simulate N data series of length T from a process that has this "true" spectrum. Letting X_j denote the j-th simulated series, we can then use the empirical distribution of

$$\frac{\widehat{S}\left(\omega_k; X_j\right)}{\widehat{S}\left(\omega_k; X\right)}$$

as an approximation to the sampling distribution of (C.3).

To simulate data from a series with"true" spectrum $\widehat{S}(\cdot; X)$, we first take the inverse Fourier transform of $\widehat{S}(\cdot; X)$ to recover the autocovariance function (ACF), $\widehat{R}_k \equiv \mathbb{E}[x_t, x_{t+k}]$, that is associated with $\widehat{S}$.[viii] Note that a series drawn from this estimated ACF will necessarily have precisely the spectrum $\widehat{S}(\cdot; X)$. For a series of length T, this estimated ACF implies a covariance matrix given by

$$\widehat{\Sigma} = \begin{pmatrix} \widehat{R}_0 & \widehat{R}_1 & \cdots & \widehat{R}_{T-1} \\ \widehat{R}_1 & \widehat{R}_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \widehat{R}_1 \\ \widehat{R}_{T-1} & \cdots & \widehat{R}_1 & \widehat{R}_0 \end{pmatrix} .$$

Letting B denote the Cholesky decomposition of $\widehat{\Sigma}$, and ϵ_j a vector drawn from the $N\left(0, I_T\right)$ distribution, the vector $X_j \equiv B\epsilon_j$ is drawn from the above ACF function.

[vii] Using alternative kernel functions makes little difference to the results.

[viii] It can be verified that, since in our case $K > 2T - 1$, given the way we compute the estimated spectrum the inverse Fourier transform of the spectrum does indeed recover the associated ACF.

D Other G7 Countries

In this appendix we examine whether the motivating evidence reported in the main text for the U.S. also appears in other countries. In Figure D.1, we report report the probability of a recession at time $t + i$ conditional (plus or minus a window) on a recession at time t. The data on business cycles reference dates are taken from Owyang, Ramey, and Zubairy [2013] for Canada and from the Economic Cycle Research Institute for the other countries. As can be seen, for most of these countries the conditional probability of a recession initially increases, peaks around 40 quarters, and then starts to decrease, which is quite close to what we observed for the U.S..

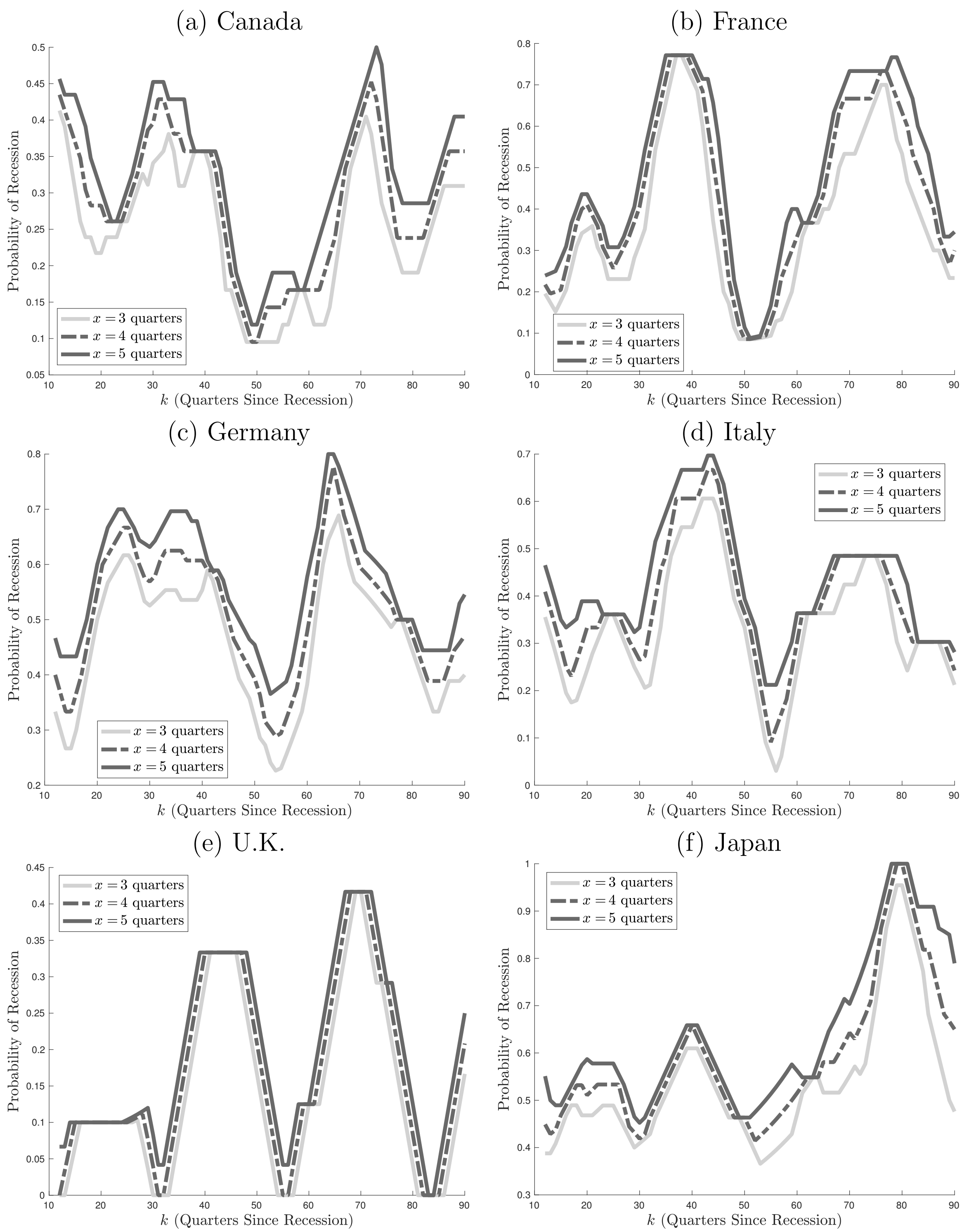
We also examine whether the spectral densities of cyclical sensitive variables in other G7 countries echo the observations for the U.S. labor market. Getting measures of hours worked per capita over a sufficiently long period is difficult in many of these countries. We therefore focus on the unemployment rate. Table D.1 shows p-values for the hump-shape tests with peak at 32-40 quarters. With the exception of Canada and Italy, the results are again overall supportive of the presence of important cyclical fluctuations in the 32-40 quarter range.

Table D.1: p-values for the Hump-Shaped Test with peak at 32-40 quarters in the Other G7 Countries

	Flat Null	$AR(1)$
U.K.	0.1%	5.2%
Germany	0.6%	4.1%
France	2.2%	11%
Japan	0.2%	3.8%
Canada	18%	31%
Italy	5.3%	19%

Notes: Table displays simulated p-values for the hump-shaped test with peak at 32-40 quarters, under the flat and $AR(1)$ null hypotheses.

Figure D.1: Conditional Probability of Being in a Recession: Other G7 countries

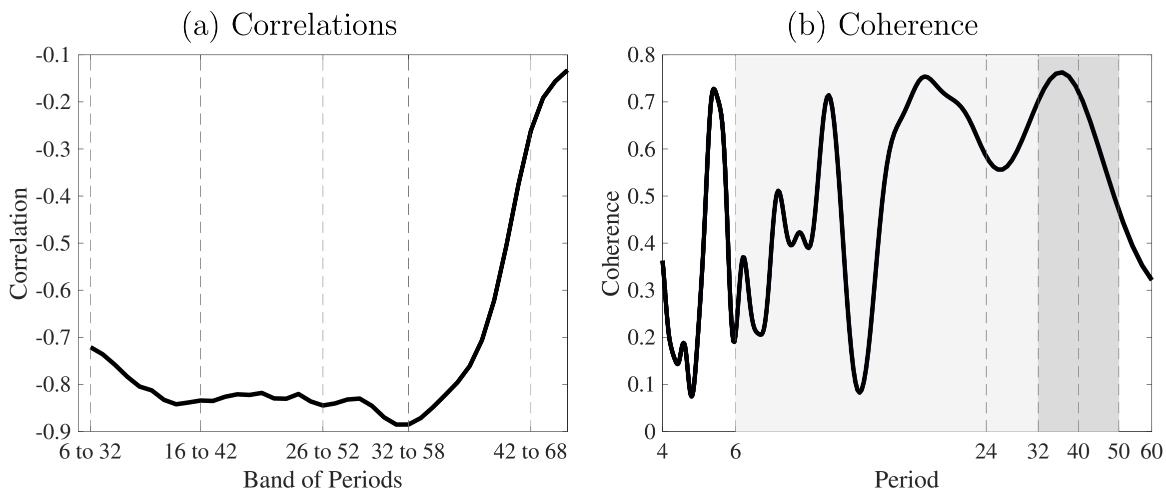


Notes: Each panel displays the fraction of time the economy was in a recession within an x-quarter window around time $t + k$, conditional on being in a recession at time t, where x is allowed to vary between 3 and 5 quarters.

E Counter-Cyclicality of the AAA/FFR Spread at Peak Frequencies

In this Appendix, we look at the co-movement between the BAA/FF spread and our hours worked measured at various frequencies. To do so, we first filter the two series with a Band-Pass filter $(P, P + 26)$, for P going from 6 quarters to 32, and compute the correlation between those two filtered series. The first correlation corresponds to the typical (6,32) "business cycle" filter, while the last one (32,58) corresponds to the peak in spectral density of hours and second peak in the spectral density of the spread. Panel (a) of Figure E.1 displays the correlation between the two filtered series when varying the band of the filter. The spread is always counter-cyclical, even more so for the (32,58) band of periods that correspond to the peak in spectral densities. Panel (b) shows the coherence between the two series. The spectral coherence is a measure of the degree of relationship, as a function of frequency, between the two time series. As we can see, over the periodicities 4 to 60 quarters, coherence is maximum at 38 quarters.

Figure E.1: Correlations and Coherence Between (Filtered for Correlations, Level for Coherence) Hours and BAA/FFR Spread

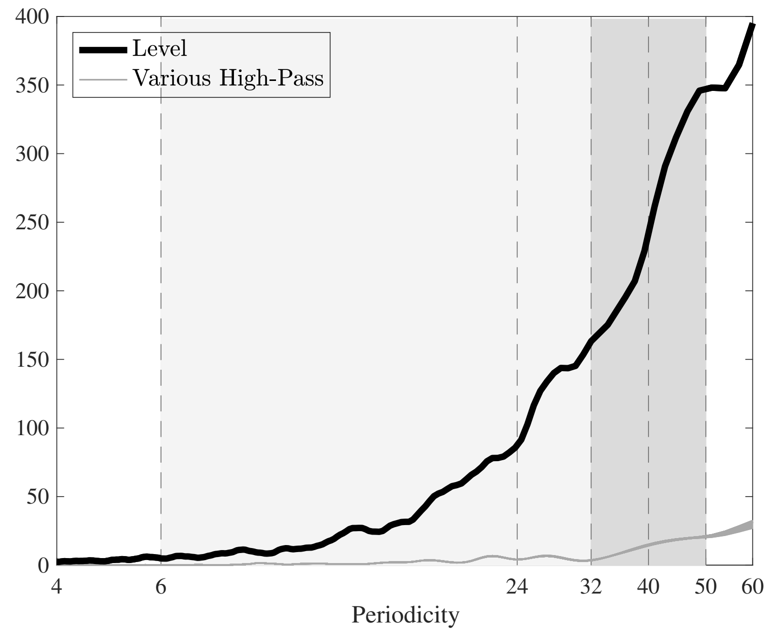


Notes: Panel (a) of this Figure displays the correlation between filtered Hours and BAA/FFR Spread, when the filter is a Band-Pass filter $(P, P+26)$, for P going from 6 quarters to 42. Panel (b) displays the coherence between the two series in levels. Sample is 1954Q3-2015Q2.

F Spectral Properties of Detrended GDP

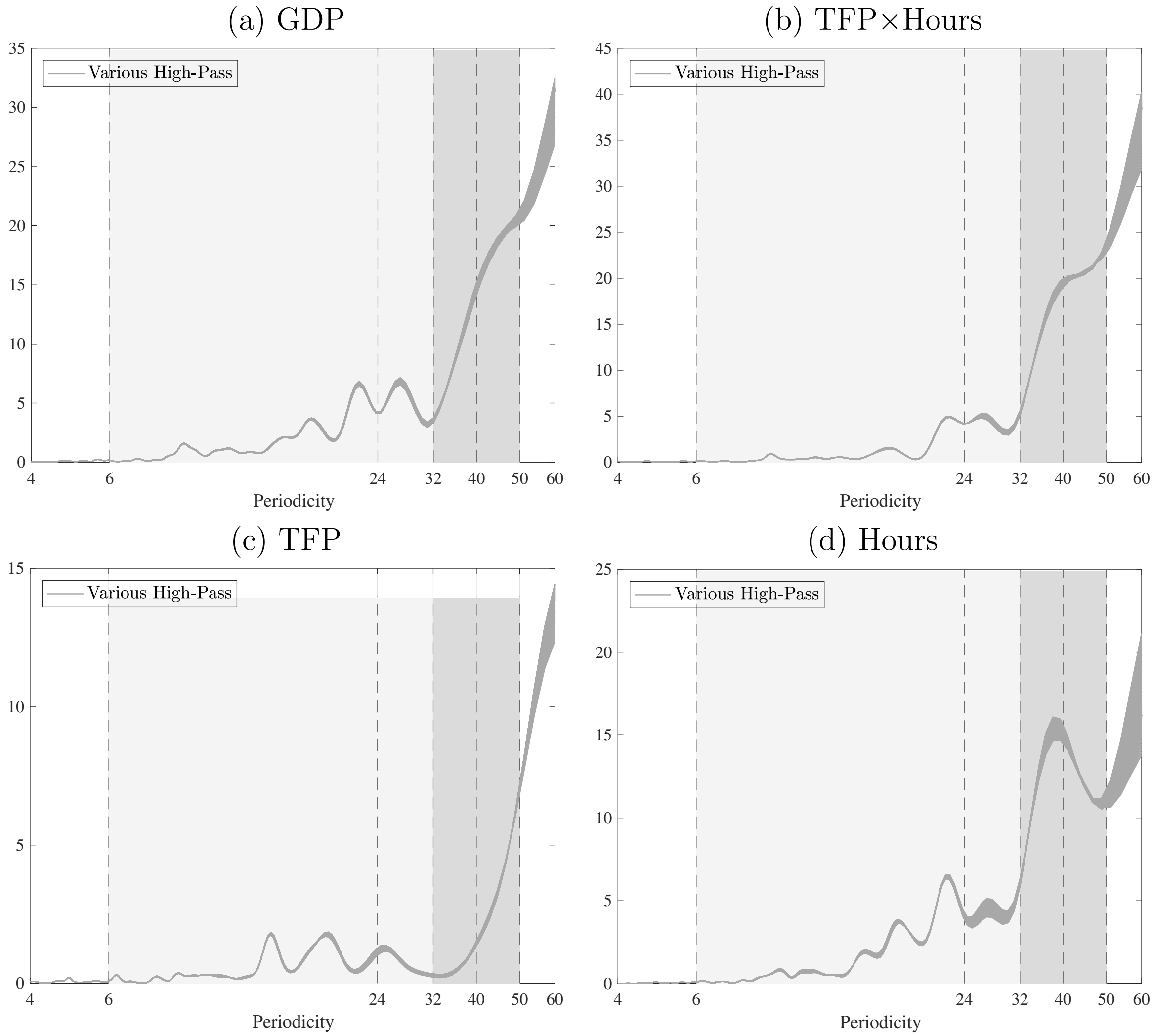
Our observations of a distinct peak in the spectral density of a set of macroeconomic variables may appear somewhat at odds with conventional wisdom. In particular, it is well known, at least since Granger [1966], that several macroeconomic variables do not exhibit such peaks, and for this reason the business cycle is often defined in terms of co-movement between variables instead of reflecting somewhat regular cyclical behavior. According to us, this perspective on business cycle dynamics may be biased by the fact that it often relies on examining the spectral properties of transformed non-stationary variables, such as detrended GDP. We instead have focused on variables—which we like to call cyclically sensitive variables—where business cycle fluctuations are large in relation to slower "trend" movements. For such variables, the breakdown between low-frequency trend and cycle is potentailly less problematic if the series can still be considered stationary. In contrast, if one focuses on quantity variables, for example GDP, one needs to believe that the detrending proceedure used to make it stationary is allowing one to isolate the relevant cyclical properties. This is certainly questionable as the detrending proceedure most often changes the spectral properties dramatically. To see this, in Figure F.1 we report the same information we reported before regarding the spectral density, but in this case the series is real per capita GDP. Here we see that the spectra of the non-detrended data and of the filtered data have very little in common with each other. Since it does not make much sense to report the spectral density of non-detrended GDP (it is clearly non-stationary), in Panel (a) of Figure F.2 we focus on the spectral density of GDP after removing low-frequency movements using the various high-pass filters. These spectral densities are in line with conventional wisdom: even when we have removed very low-frequency movements, we do not detect any substantial peak in the spectral density of GDP around 40 quarters. How can this be? What explains the different spectral properties of filter-output versus the level of hours worked? There are at least two possibilities. First, it could be that the filtering we implemented on GDP is simply not isolating cyclical properties. Alternatively, if one believes that such a filtering procedure is isolating cyclical properties, the answer to the puzzle may lie in the behavior of (detrended) TFP. Panel (b) of Figure F.2 plots the spectral density of the (log of the) product of TFP by hours, after having removed low-frequency movements in the same way we have done for GDP and other variables. Note that the spectral density of this variable is very similar to the GDP one, as if, for periodicity below 80 quarters, GDP could be approximatively seen as being produced linearly with hours only, whose productivity would be scaled by TFP. If Hours and TFP were uncorrelated, then the spectral density of GDP (in logs) would be the sum of the ones of (log) Hours and TFP. This is approximatively what we have, as shown in panels (c) and (d). The spectral density of TFP shows a quick pick-up it just above periodicities of 40 quarters. As with GDP, we do not see any marked peaks in the spectral density of TFP. An interesting aspect to note is that if we add the spectral density of hours worked to that of TFP, we get almost exactly that of GDP. This suggests that looking at the spectral density of GDP may be a much less informative way to understand business cycle phenomena than looking at the behavior of cyclically sensitive variables such as hours worked. Instead, GDP may be capturing two distinct processes: a business cycle process associated with factor usage and a lower-frequency process associated with movement in TFP. For this reason, we believe that business cycle analysis may gain by focusing more closely on explaining the behavior of cyclically sensitive variables at business cycle frequencies.

Figure F.1: Spectral Density of U.S. GDP per Capita, Levels and Various High-Pass Filters



Notes: This figure shows an estimate of the spectral density of U.S. GDP per capita in levels (black line) and for 101 series that are high-pass (P) filtered version of the levels series, with P between 100 and 200 (grey lines). A high-pass (P) filter removes all fluctuations of period greater than P.

Figure F.2: Decomposing the Spectral Density of GDP



Notes: This figure shows an estimate of the spectral density of U.S. GDP per capita (panel (a) and Total Factor Productivity (panel (b)) for 101 series that are high-pass (P) filtered version of the levels series, with P between 100 and 200 (grey lines). A high-pass (P) filter removes all fluctuations of period greater than P. TFP is the corrected quarterly TFP series of Fernald [2014].

G More on Spectral Density Estimation

G.1 Smoothing and Zero-Padding with a Multi-Peaked Spectral Density

To illustrate the effects of smoothing and zero-padding, in this section we compare the estimated spectral density with the known theoretical one for a process that exhibits peaks in the spectral density at periods 20, 40 and 100 quarters. We think this is a good description of the factor variables we are studying (i.e., hours worked, unemployment), that display both business cycle movements and lower-frequency movements unrelated to the business cycle. We construct our theoretical series as the sum of three independent stationary AR(2) processes, denoted x_1, x_2 and x_3.

Each of the x_i follows an $AR(2)$ process

$$x_{it} = \rho_{i1} x_{it-1} + \rho_{i2} x_{it-2} + \varepsilon_{it},$$

where ε_i is *i.i.d.* $N(0, \sigma_i^2)$. The spectral density of this process can be shown to be given by

$$S(\omega) = \sigma_i^2 \left\{ 2\pi[1 + \rho_{i1}^2 + \rho_{i2}^2 + 2(\rho_{i1}\rho_{i2} - \rho_{i1})\cos(\omega) - 2\rho_{i2}\cos(2\omega)] \right\}^{-1}$$

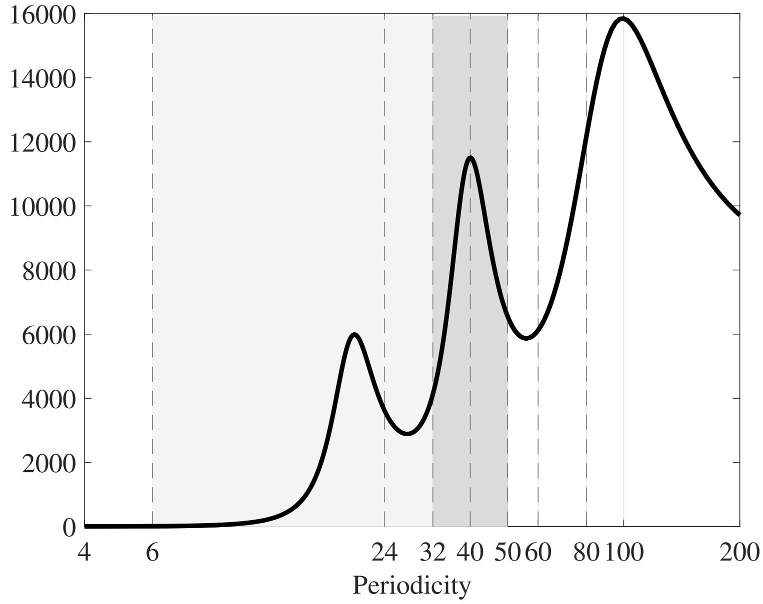
It can also be shown (see, e.g., Sargent [1987]) that for a given ρ_{i2}, the spectral density has a peak at frequency $\overline{\omega}_i$ if

$$\rho_{i1} = -\frac{4\rho_{i2}\cos(2\pi/\overline{\omega}_i)}{1 - \rho_{i2}}$$

We set $\overline{\omega}_i$ equal to 20, 40, and 100 quarters, respectively, for the three processes, and ρ_{i2} equal to -0.9, -0.95, and -0.95. The corresponding values for ρ_{i1} are 1.802, 1.9247, and 1.9449. We set σ_i equal to 6, 2, and 1. We then construct $x_t = x_{1t} + x_{2t} + x_{3t}$. The theoretical spectral density of x is shown in Figure G.1. While the spectral density shows some important long-run fluctuations, the bulk of the business cycle movements is explained by movements at the 40-quarter periodicity, although we do observe another peak at periodicity 20 quarters.

We simulate this process 1,000,000 times, with $T = 270$ for each simulation, which is the length of our observed macroeconomic series. We estimate the spectral density for various values of N (zero-padding) and W (length of the Hamming window). Higher N corresponds to higher resolution, and higher W to more smoothing. On each panel of Figure G.2, we

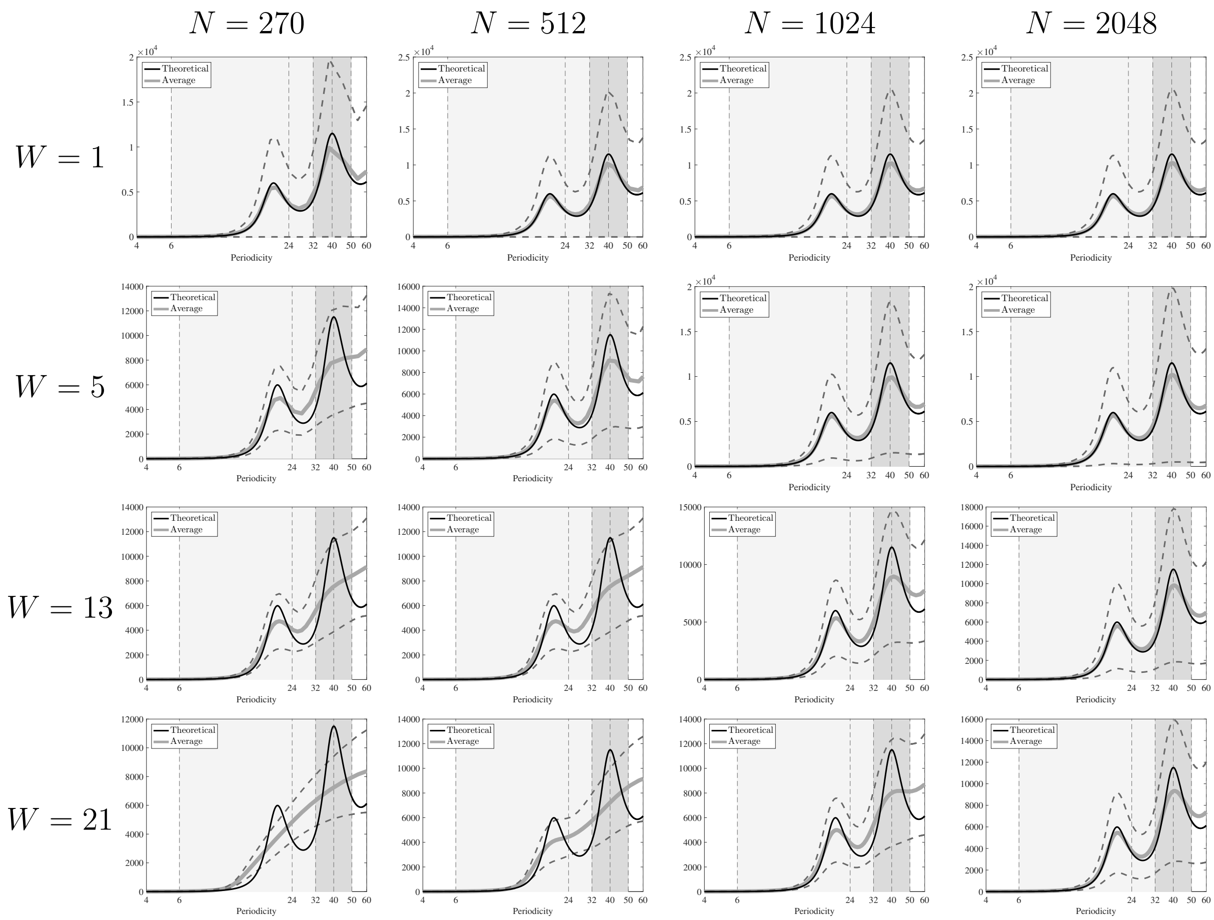
Figure G.1: Theoretical Spectral Density (Sum of Three AR(2))



Notes: Figure shows the theoretical spectral density of the sum of three independent AR(2) processes, which have peaks in their spectral densities at, respectively, 20, 40 and 100 quarters.

report the mean of the estimated spectrum over the 1,000,000 simulations (solid grey line), the mean $\pm$ one standard deviation (dashed lines), and the theoretical spectrum (solid black line). As we can see moving down the figure (i.e., for increasing W), more smoothing tends to reduce the error variance, but at the cost of increasing bias. Effectively, the additional smoothing "blurs out" the humps in the true spectrum. For example, with no zero-padding ($N = 270$), the peak in the spectral density at 40 quarters is (on average) hardly detected once we have any smoothing at all. Meanwhile, moving rightward across the figure (i.e., for increasing N), we see that more zero-padding tends to reduce the bias (and in particular, allows for the humps surrounding the peaks to be better picked up on average), but typically increases the error variance. As these properties suggest, by appropriately choosing the combination of zero-padding and smoothing, one can minimize the error variance while maintaining the ability to pick up the key features of the true spectrum (e.g., the peaks at 20 and 40 quarters).

Figure G.2: Effects of Smoothing and Zero-Padding (Sum of Three AR(2))

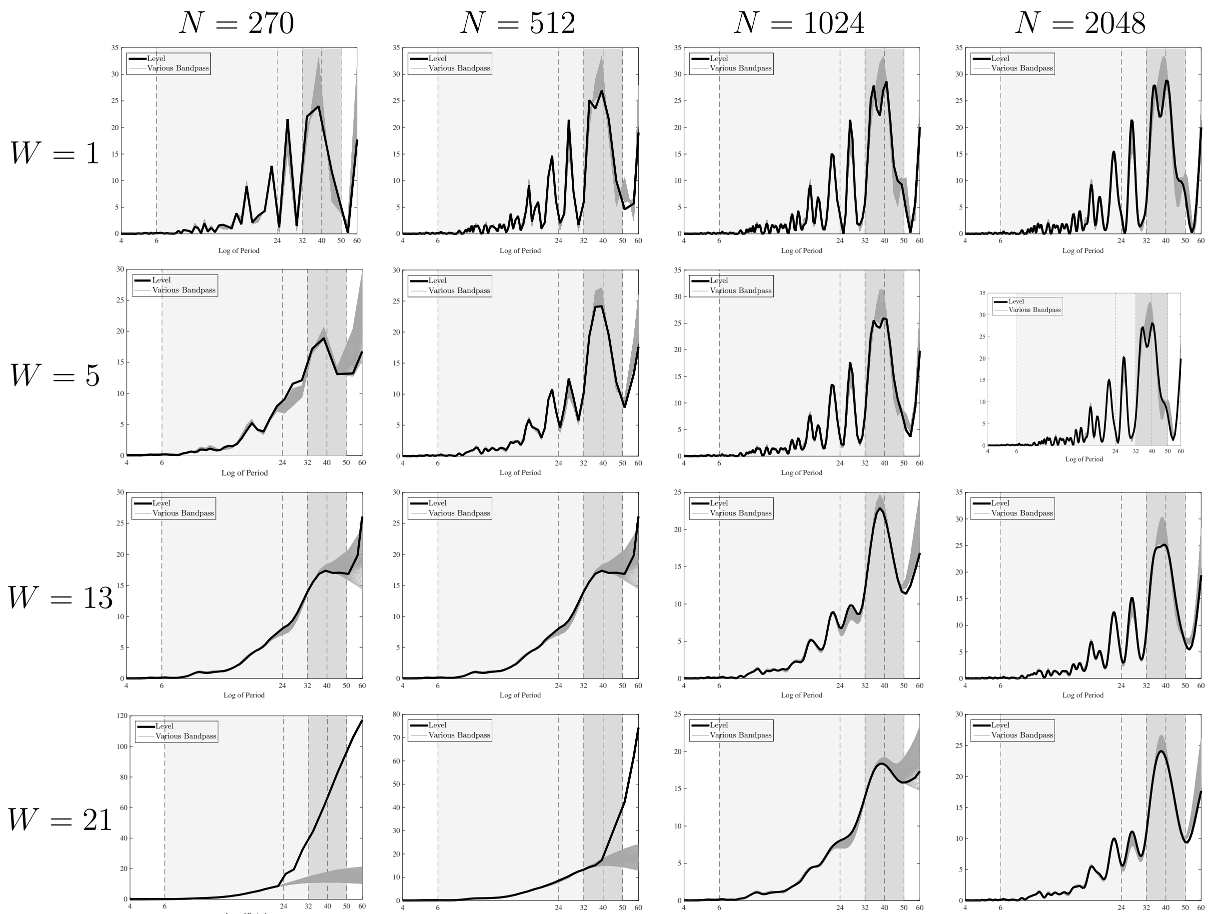


Notes: Figure shows estimates of the spectral density using simulations of the sum of three independent AR(2), which have peaks in their spectral densities at, respectively, 20, 40 and 100 quarters. The black line is the theoretical spectrum, the solid grey line is the average estimated spectrum over 1,000,000 simulations, and the dotted grey lines corresponds to that average $\pm$ *one standard deviation, and bounded below by zero.* W *is the length of the Hamming window (smoothing parameter) and* N *is the number of points at which the spectral density is evaluated (zero-padding parameter).*

G.2 Smoothing and Zero-Padding with Non-Farm Business Hours per Capita

Figure G.3 presents estimates of the spectral density of U.S. non-farm business hours per Capita for different choices of the zero-padding parameter (N) and different lengths of the Hamming window (W). The results indicate that, as long as the amount of zero-padding is not too small (i.e., N larger), we systematically observe the peak at around 40 quarters in the spectral density. In fact, it is only with minimal zero-padding (N low) and a wide smoothing window (W high) that the peak is entirely washed out. We take this as evidence of the robustness of that peak.

Figure G.3: Changing Smoothing and Zero-Padding

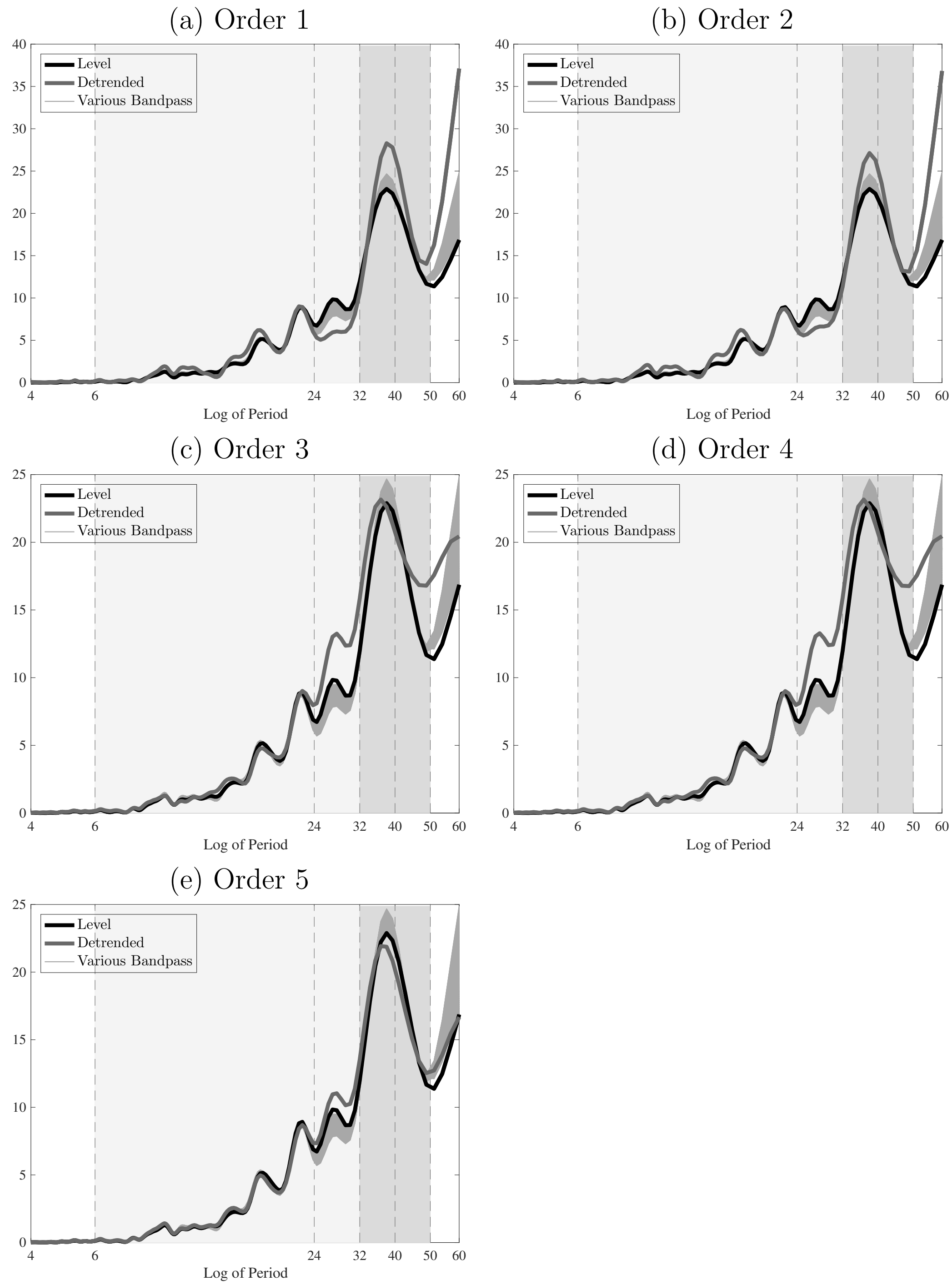


Notes: Figure shows estimates of the spectral density of U.S. Non-Farm Business Hours per Capita over the sample 1947Q1-2015Q2. The different lines correspond to estimates of the spectral density of hours in levels (black line) and of 101 series that are high-pass (P) filtered version of the levels series, with P between 100 and 200 (thin grey lines). W is the length of the Hamming window (smoothing parameter) and N is the number of points at which the spectral density is evaluated (zero-padding parameter).

G.3 Detrending with a Polynomial Trend

In this section, we check that detrending our hours series with a polynomial trend of degrees 1 to 5 does not affect our main finding; namely, the existence of a peak in the spectrum at a periodicity around 40 quarters. Plots confirming that our finding is robust to polynomial detrending are shown in Figure G.4.

Figure G.4: Using a Polynomial Trend of Various Orders for Benchmark Smoothing ($W = 13$) and Zero-Padding ($N = 1024$)



Notes: Figure shows estimates of the spectral density of U.S. Non-Farm Business Hours per Capita over the sample 1947Q1-2015Q2, when polynomial trends of order 1 to 5 have been removed from the data. The different lines correspond to estimates of the spectral density of hours in levels (black line), of hours detrended with a polynomial trend (thick grey line) and of 101 series that are high-pass (P) filtered version of the levels series, with P between 100 and 200 (thin grey lines).

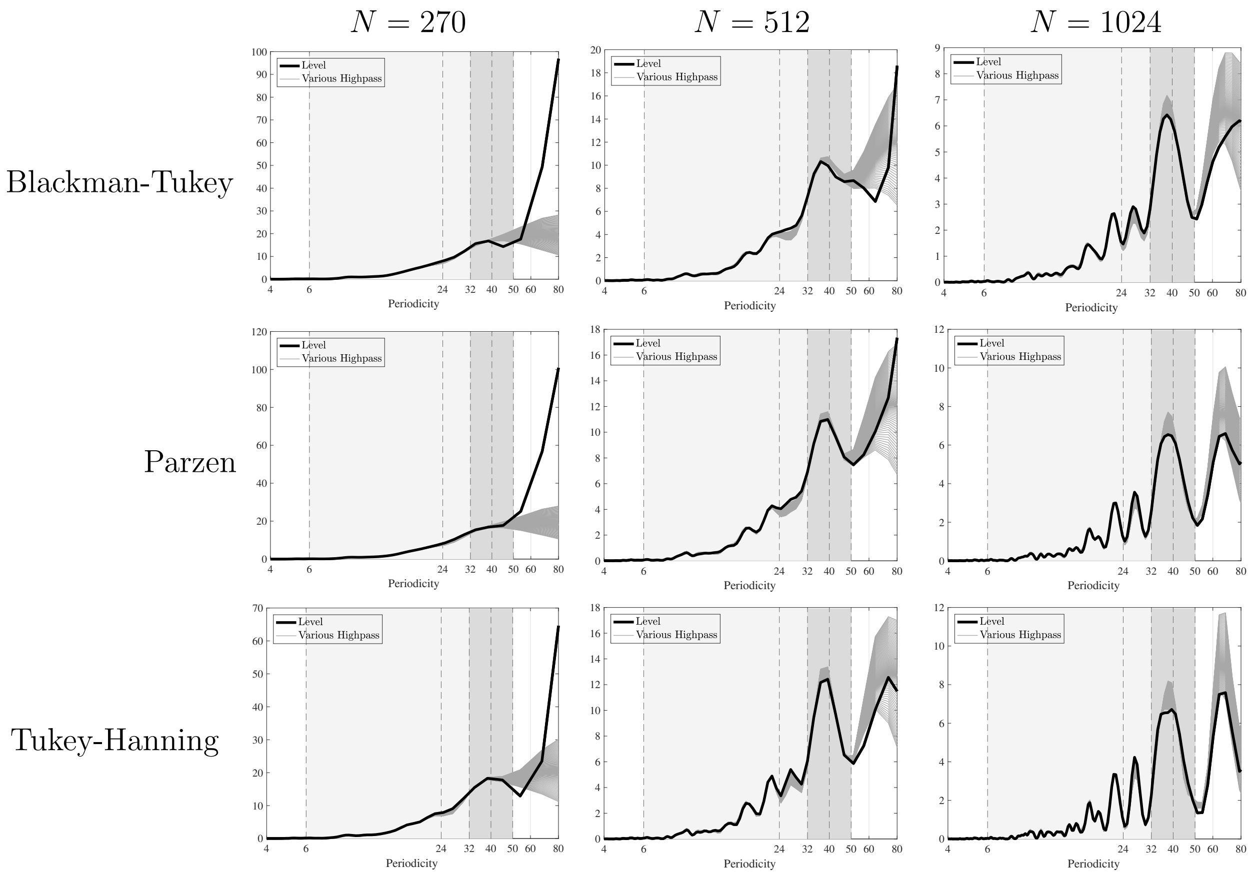
G.4 Alternative Estimators

As another robustness test, we estimate the spectrum using the `SPECTRAN` package (for Matlab), which is described in Marczak and Gómez [2012]. The spectrum is computed in this case as the Fourier transform of the covariogram (rather the periodogram as we have done thus far). Smoothing is achieved by applying a window function of length M to the covariogram before taking its Fourier transform.[ix] Three different window shapes are proposed: the Blackman-Tukey window, Parzen window, and Tukey-Hanning window. The width of the window used in estimation is set as a function of the number of samples of the spectrum. In the case where no zero-padding is done ($N = 270$), these "optimal" widths correspond to lengths of, respectively, $M = 68$, 89, and 89 quarters for the three methods.[x] Figure G.5 shows the estimated spectrum of Non Farm Business hours for the three windows and with or without zero-padding ($N = 270$, 512, or 1024). Results again confirm the existence of a peak at a periodicity around 40 quarters, as long as there is enough zero-padding.

[ix] Specifically, the k-th-order sample autocovariance is first multiplied by $w(|k|)$, where the window function w is an even function with the property that $max_k w(k) = w(0) = 1$, and the window length $M > 0$ is such that $w(|k|) \neq 0$ for $|k| = M - 1$ and $w(|k|) = 0$ for $|k| \geq M$.

[x] Note that, in contrast to the kernel-smoothing case, in this case a wider window corresponds to less smoothing.

Figure G.5: Non-Farm Business Hours with Various Windows and Estimation Using the Covariogram



Notes: Notes: This Figure shows estimates of the spectral density of U.S. Non Farm Business Hours per Capita over the sample 1947Q1-2015Q2, as computed from the covariogram using the `SPECTRAN` *package. The different lines correspond to estimate of the spectral density of hours in levels (black line) and of 101 series that are high-pass (P) filtered version of the levels series, with P between 100 and 200 (thin grey lines). N is the number of points at which is evaluated the spectral density (zero padding).*

H Non Normality

Here we also examined whether the business cycle fluctuations which we have been focusing upon are approximately normally distributed. To this end, we perform Bai and Ng [2005] test, which combines skewness and kurtosis into a single statistic for serially dependant data. This is done on hours and spread after using a high-pass linear filter to remove periodicities greater than 60 quarters, which allows us to retain all the variation that we have argued is relevant for the business cycle, while removing more medium- and long-run fluctuations. The null hypothesis for the test is normality. For non-farm business hours and spread, the p-values we find are, respectively, 1.2% and .02%. This indicates that linear-Gaussian models might not be an appropriate way to describe these data, and that one may need to allow for some type of non-linearity in order to explain these movements. Accordingly, we explore a class of explanation that allows for such non-linearities. Moreover, when we estimate our model, we use the skewness and kurtosis properties of the data to help identify parameters.

I Steps in Deriving the Model of Section III

The model can be seen as an extension of the canonical three-equation New Keynesian model. The extra elements we add are habit persistence, the accumulation of durable goods (or houses), and a credit market imperfection that generates a counter-cyclical risk premium.

I.1 Model Summary

The economy is populated with a continuum of households of mass one, final good and intermediate services firms, commercial banks and a central bank. Each household (or *family*) will be composed of a continuum of mass one of worker members and an atomless family head who coordinates family activities. Time is discrete, and a period of time is divided into four sub-periods.

Final good firms

Final good firms buy consumption services from the set of intermediate firms, and bundle them into a final good according to a Dixit-Stiglitz aggregator.

Intermediate firms

Intermediate firms produce a set of differentiated consumption services using durable goods as input. These durable goods will be composed of existing durable goods rented from the households and new ones produced by these intermediate firms using labor. After the production of consumption services occurs, the remaining un-depreciated stock of new durable goods will be sold to households. We assume that firms are price takers on the market for durable goods, but that they are monopolistic competitors when selling their variety of consumption services. As standard in a New Keynesian framework, they will only be able to adjust their prices upon the arrival of price change opportunity, which occurs according to a Poisson distribution.

Commercial Banks

Orders for final goods and new durable goods are made by each individual household member before going to the labor market, and these orders are financed by credit granted by banks at interest rate r. To finance their lending, banks take deposits on which they pay the risk-free nominal rate i. Because some workers will not find a job and may as a result default on their debt, the interest rate r charged by banks to borrowers will include a risk premium r^p over and above the risk-free rate. We assume free entry and perfect competition in the banking market, so that banks make no profits.

Central Bank

The central bank sets the risk free nominal interest rate i according to a type of Taylor rule.

Households

A family purchases consumption services and durable goods. At the beginning of each period, after the resolution of aggregate uncertainty and the clearing of the last period's debts and deposits, each worker in the family places orders for consumption services and newly produced durable goods, financed by borrowing from banks against their uncertain labor income, with the imperfect backing of their family. Because of a financial imperfection that we describe below, the cost of borrowing will be the risky rate $r = (1 + r^p)(1 + i) - 1$.

The uncertainty at this stage is only idiosyncratic, and comes from the fact that jobs are indivisible. All the workers inelastically supply one unit of labor, but firms will demand only $e_t \leq 1$ jobs, so that a fraction $1 - e_t$ of workers will be unemployed, and will not be able to repay their debt the next period. We assume that each employed worker works a fixed number of hours normalized to one, so that e_t is also hours.

The financial friction

We assume that for workers that were employed the last period, loan contracts are fully enforceable, so that their debt is always repaid. The financial friction takes the form of a costly enforcement of debt repayments by the family for the household members that do not find a job. In particular, not all non-performing loans can be recovered by going after the family. When a household member cannot pay back a loan—which will be the case when they were unable to find a job—then with exogenous probability ϕ the bank can pay a cost $\Phi < 1$ per unit of the loan to recover the funds from the household (which the bank will always choose to do), while with probability $1 - \phi$ it is too costly to pursue the household, in which case the bank is forced to accept a default. To fix ideas, we assume that, for unemployed workers who have their debt covered by the household, the household transfers to the unemployed worker the funds necessary to cover the debt.[xi]

I.2 Timing

Borrowing is in one-period bonds. A family enters the period with outstanding debt and bank deposits for each of its workers, and a stock of durable goods that is managed by the family head.

In the first sub-period, interest on last-period deposits are paid, and past debts with banks are settled. Workers that were employed in the previous period pay back their loans, and return any remaining cash balances to the household.

In the second sub-period, workers first borrow from banks, and then use the proceeds to order consumption services and durable goods. Final good services firms receive demand (paid orders) and make orders to intermediate good firms, which also receive the orders for

[xi] An equivalent formulation would have the household simply absorbing that workers debt, to be repaid in a future period.

new durable goods from the households. A worker who wants to borrow can try to do it on his own, or can get backing from his family. If there is no backing by the family, a worker who borrows but does not find a job will not be able to repay loans. Backing by the family will therefore reduce the risk premium on loans, and in equilibrium will always be chosen by the workers. As we have explained above, backing by the family will be imperfect: with probability ϕ the bank will pursue the household, and with probability $1-\phi$ it is too costly to pursue the household.

In the third sub-period, intermediate good firms rent durable goods from the household head and hire workers to produce any new durable goods. As noted above, there will always be a measure one of workers looking for a job, but because of job indivisibility, only a fraction e_t of workers will find a job. Within a match, the firm makes a take-it-or-leave-it wage offer to the worker, the equilibrium level of which will be at a reservation wage, which will be decided by the family head.

Finally, in the fourth sub-period, production takes place, wages are paid to workers, rental payments for durable goods are made to the household head, orders are fulfilled, and all consumption services and durable goods are brought back to the household and shared equally among family members. Firms' and banks' profits, if any, are returned to the household.

I.3 Households

There is a continuum of mass one of households (families) indexed by h who purchase consumption services from the market and accumulate durable goods. The preferences of family h are given by $U(h) = \mathbb{E}_0 \sum_t \beta^t \xi_{t-1}[U(C_{ht} - \gamma C_{t-1}) + \nu(1 - e_{ht})]$, with $U'(\cdot) > 0, U''(\cdot) < 0$, $\nu'(\cdot)$, $\nu''(\cdot) < 0$ and $0 \leq \gamma < 1 - \delta$, where δ is the depreciation rate of the durable good. C_{ht} represents the consumption services purchased by household h in period t, $C_t \equiv \int_0^1 C_{ht} dh$ denotes the average level of consumption in the economy, β is the discount factor, and ξ_t denotes an exogenous shock to the discount factor at date t. Note that this preference structure assumes the presence of external habit. Each family owns a stock of durable goods X_{ht} that is rented to firms in order to produce intermediate services. This stock of durable goods is accumulated according to the equation

$$X_{ht+1} = (1 - \delta) X_{ht} + I_{ht}\,. \tag{I.1}$$

A family is composed of a continuum of measure one of workers indexed by j. All the members of the family share the utility $U(h)$, where $C_{ht} = \int_0^1 C_{jht} dj$ is the total consumption services bought by the workers (C_{jht} for each worker) in household h, and total purchases of new durable goods is similarly given by $I_{ht} = \int_0^1 I_{jht} dj$.

We go into details of a family behavior starting with the second sub-period of period t, when family i has inherited from the past a net debt position given by D_{ht}; that is, the family owes a bank the amount D_{ht}. It is useful to note at this stage that, because all families are identical, D_{ht} will be zero in equilibrium.

The outstanding debt of the family D_{ht} is equally shared among the unit measure of workers. Worker j is told by the family head to go to bank with its share of existing debt D_{jht} $(= D_{ht})$ to apply for a new loan $L_{jht} \geq D_{jht}$, where the first D_{jht} units of that loan are used to settle the past debt and $L_{jht} - D_{jht}$ is left and available for spending. When granting the loan, the bank opens the worker a checking account (which cannot have a negative balance), where the initial amount in the deposit would be $L_{jht} - D_{jht}$. As noted above, a worker can apply for their loan either with or without family backing. We only consider the case where all workers apply for loans with family backing. This is not restrictive, as the risk premia on backed and non-backed loans will be such that workers will never choose a non-backed loan.

Banks can transfer the balances in the checking account to other agents, and these deposits earn the safe (central bank-determined) interest rate i at the beginning of next period. The workers can then use this bank money to order consumption goods C_{jht} at nominal price P_t and new durable goods I_{jht} at nominal price P_t^X. When a firm receives an order, the bank money is transferred to that firm, which then uses it to similarly place orders with the intermediate good firms (in the case of a final good firm) or to pay workers and rent durables (in the case of an intermediate good firm). With these latter payments, much of the bank money will be transferred back to (employed) workers and the household head (who maintains a family bank account) at the end of the period. The remaining bank money (reflecting intermediate good firm profits) is also transferred to the household head. After paying off their own loans, employed workers are asked by the family head to transfer their remaining bank money back to the family account.

The head of the family coordinates the family activities by telling workers how much to borrow, how much to purchase and how much of the bank money to transfer to the family. On top of that, the family head manages the stock of durable goods (or houses) X_{ht} of the family. She does so by renting it on the market at rate R_t^X.

In the third sub-period, firms wish to hire e_t workers from of the unit measure pool of workers, so that $1 - e_t$ workers will be left unemployed.[xii] Firms make take-it-or-leave-it offers to the workers in the form of a nominal wage W_t, who accept it as long as it is not below a reservation level set by the head of the family. In equilibrium, all offered jobs will be accepted. Because there is a unit measure of families with a unit measure of workers each, we will have $e_{ht} = e_t$, so that e_t is the probability of a worker in family h being employed.

Finally, as noted above, in the fourth sub-period, production takes place, all wage and rental payments are made, profits are transferred, an consumption and new durable goods are delivered to those who ordered them, which in turn are returned to the family to be split equally among its members.

In the first sub-period of the next period, debts need to be settled after interest on bank money is paid. Workers who were employed the last period ended it with a bank account balance of $W_t + L_{jht} - D_{jht} - P_t C_{jht} - P_t^X I_{jht}$. This balance receives an interest payment $1 + i_t$ at the beginning of period $t + 1$. The worker, meanwhile, owes the bank $(1 + r_t)L_{jht}$ at this point. The bank will limit its lending such that $(1 + r_t)L_{jht} \leq (1 + i_t)W_t$ so as to ensure that employed workers can always pay back their loans. Further, since $r_t \geq i_t$ (reflecting the risk premium), it will never be optimal for workers to borrow more than they intend to spend, and thus we will have $L_{jht} = D_{jht} + P_t C_{jht} + P_t^X I_{jht}$. Thus, an amount $T_{jht}^e = (1+i_t)W_t - (1+r_t)(D_{jht} + P_t C_{jht} + P_t^X I_{jht}) > 0$ will be transferred back to the family account by each employed worker.

The bank account balance of an unemployed worker at the end of period t, meanwhile,

[xii] As the supply of workers is always inelastically one, equilibrium e_t must be less than or equal to one. To simplify notation, we will also let e_t denote the probability that a worker is offered a job, whereas it technically should be written $\min\{e_t, 1\}$ to allow for off-equilibrium labor demand.

is given by $L_{jht} - D_{jht} - P_t C_{jht} - P_t^X I_{jht} = 0$, which as noted above is optimally set to zero. Further, as with employed workers, unemployed workers owe an amount $(1 + r_t)L_{jht}$ to the bank at the beginning of period $t+1$. For these workers, the bank decides whether or not to pursue the family for repayment. If it is too costly (which which happens with probability ϕ), the loan is not repaid. With probability $1 - \phi$, pursuing the household is worthwhile for the bank, in which case the family transfers $T^u_{jht} = (1 + r_t)(D_{jht} + P_t C_{jht} + P_t^X I_{jht})$ to the worker so they can repay the loan.

Therefore, family h's net debt at the beginning of period $t + 1$ will be given by

$$D_{ht+1} = (1 - e_t)\,\phi T^u_{jht} - e_t T^e_{jht} - (1 + i_t)\left(R_t^X X_{ht} + \Pi_t\right)$$

which rewrites

$$\begin{aligned} D_{ht+1} &= [e_t + (1 - e_t)\,\phi]\,(1 + r_t)\left(D_{ht} + P_t C_{ht} + P_t^X I_{ht}^X\right) \\ &- (1 + i_t)\left(e_t W_t + R_t^X X_{ht} + \Pi_t\right), \end{aligned} \tag{I.2}$$

where Π_t is total profits of all firms and banks.

I.4 Banks

Banks remunerate deposits at rate i_t, receive interest r_t on the fraction $e_t + (1-e_t)\phi$ of loans that are repaid, and incur processing costs of Φ per unit of loans to bailed-out workers (i.e., those made to the fraction $(1 - e_t)\phi$ of workers who end up unemployed and are bailed out by the household). Thus, bank profits are

$$\Pi_t^{\text{banks}} = \{[e_t + (1 - e_t)\,\phi]\,(1 + r_t) - (1 + i_t)\,[1 + (1 - e_t)\,\phi\Phi]\}\,L_t\,,$$

where L_t is the total volume of loans. Free entry into banking implies zero profits, i.e.,

$$1 + r_t = (1 + i_t)\,\frac{1 + (1 - e_t)\,\phi\Phi}{e_t + (1 - e_t)\,\phi}\,. \tag{I.3}$$

Defining the risk premium as $(1 + r_t) = (1 + i_t)(1 + r_t^p)$, the above equation implies

$$1 + r_t^p = \frac{1 + (1 - e_t)\,\phi\Phi}{e_t + (1 - e_t)\,\phi} \tag{I.4}$$

If there is no unemployment risk ($e_t = 1$), or no default or recovery costs ($\phi = 1$ and $\Phi = 0$), one can check that there is no risk premium ($r_t^p = 0$).

I.5 Firms

The final good sector is competitive. This sector provides consumption services to households by buying a set of differentiated intermediate services, denoted C_{kt}, from intermediate service firms and combining them using a Dixit-Stiglitz aggregator. We assume a measure one of intermediate service firms, indexed by k. The objective of the final good firm is thus to solve

$$\max P_t C_t - \int_0^1 P_{kt} C_{kt} dj$$

subject to

$$C_t = \left(\int_0^1 C_{kt}^{\eta} dk \right)^{\frac{1}{\eta}}$$

with $\eta \in (0, 1)$ and where P_{kt} is the nominal price of intermediate service k. This gives rise to demand for intermediate service k given by

$$C_{kt} = \left(\frac{P_{kt}}{P_t} \right)^{-\frac{1}{1-\eta}} C_t.$$

Details of intermediate firms are presented in the main text.

I.6 Equilibrium Outcomes

As noted in the text, the equilibrium outcome for this model is determined from a set of nine equations in the endogenous variables. The first two equilibrium conditions are the bank zero-profit condition (I.3) and the intermediate goods optimality condition (I.5).

$$R_t^X = \frac{W_t}{F_e(e_{kt}, \theta_t)} - \psi P_t^X \tag{I.5}$$

We derive the three next equilibrium conditions from the household's behavior. The head of family h maximizes

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \xi_{t-1} \left[U \left(C_{ht} - \gamma C_{t-1} \right) + v \left(1 - e_{ht} \right) \right]$$

subject to the intertemporal budget constraint (I.2) and the durables accumulation equation (I.1). We can use the first-order conditions of this problem and the equilibrium conditions (the bank zero profit condition (I.3), $D_{ht} = 0$ and $C_{ht} = C_t$) to obtain

$$U'\left(C_t - \gamma C_{t-1}\right) = \beta \frac{\xi_t}{\xi_{t-1}}\left(1 + i_t\right)\left[1 + \left(1 - e_t\right)\phi\Phi\right]\mathbb{E}_t\left[\frac{U'\left(C_{t+1} - \gamma C_t\right)}{1 + \pi_{t+1}}\right], \tag{I.6}$$

$$U'\left(C_t - \gamma C_{t-1}\right) = \beta \frac{\xi_t}{\xi_{t-1}}\mathbb{E}_t\left[\frac{U'\left(C_{t+1} - \gamma C_t\right)}{\left(1 + \pi_{t+1}\right)P_t^X}\left\{\frac{R_{t+1}^X}{1 + \left(1 - e_{t+1}\right)\phi\Phi} + \left(1 - \delta\right)P_{t+1}^X\right\}\right] \tag{I.7}$$

with an additional wage condition given by (I.8)

$$\frac{W_t}{P_t} \geq \frac{v'\left(1 - e_t\right)}{U'\left(C_t - \gamma C_{t-1}\right)}\frac{\left[e_t + \left(1 - e_t\right)\phi\right]\left(1 + r_t\right)}{\left(1 + i_t\right)} + \frac{1 + r_t}{1 + i_t}\left(1 - \phi\right)\left(C_t + \frac{P_t^X}{P_t}I_t^X\right). \tag{I.8}$$

Equation (I.6) is derived from the Euler equation for debt accumulation, (I.7) is derived from the Euler equation for the durable good accumulation, and the wage condition is the reservation wage of the household, which is in equilibrium the actual wage offered by firms. We also have the aggregate law of motion for the stock of durable goods (I.9)

$$X_{t+1} = (1 - \delta)X_t + \psi F(e_t, \theta_t), \tag{I.9}$$

and the consumption services market-clearing condition

$$C_t = s\left[X_t + F(e_t, \theta_t)\right]. \tag{I.10}$$

The last two equations will be given by the nominal interest rate policy rule of the central bank (I.11), and the optimal pricing decision of firms under Calvo adjustment costs which, as noted in the text, we do not need to explicitly obtain.

Block Recursive Structure Using the monetary policy rule (I.11),[xiii]

$$1 + i_t \approx \Theta\mathbb{E}_t\left[e_{t+1}^{\varphi_e}\left(1 + \pi_{t+1}\right)\right] \tag{I.11}$$

, the equilibrium equations have a block recursive structure whereby the variables C_t, e_t, X_{t+1} and r_t^p can be solved for first using equations (I.3), (I.4), (I.10) and the combination of

[xiii]The precise form of the Taylor rule we use to obtain the block recursive property is

$$1 + i_t = \Theta\mathbb{E}_t\left[e_{t+1}^{\varphi_e}\frac{U'\left(C_{t+1} - \gamma C_t\right)}{\mathbb{E}_t\left[\frac{U'\left(C_{t+1} - \gamma C_t\right)}{1 + \pi_{t+1}}\right]}\right].$$

This deviates slightly from the Taylor rule we write in the main text: $1 + i_t \approx \Theta\mathbb{E}_t\left[e_{t+1}^{\varphi_e}\left(1 + \pi_{t+1}\right)\right]$ due to Jensen's inequality, which is why that is expressed with an $\approx$ symbol.

(I.6) and (I.11), which is given by

$$U'\left(C_t - \gamma C_{t-1}\right) = \beta \frac{\xi_t}{\xi_{t-1}} \Theta \left[e_t + \left(1 - e_t\right)\phi\right]\left(1 + r_t^p\right) \mathbb{E}_t \left[U'\left(C_{t+1} - \gamma C_t\right) e_{t+1}^{\varphi_e}\right] \tag{I.12}$$

Given that the above equations determine the quantity variables, the rest of the system then simultaneously determines the remaining variables $\left\{\frac{R_t^X}{P_t}, \frac{P_t^X}{P_t}, \frac{W_t}{P_t}, \pi_t\right\}$. In particular, as we do not consider the implications of the model for inflation, we do not need to explicitly derive the optimal pricing behavior of firms. Finally, using (I.10) to eliminate C from the system, we end up with the three equations (15), (16) and (17) in the main text.

J Solving and Estimating the Model

J.1 Deriving the Estimated Equations

Substituting the functional forms $U(c) = (c^{1-\omega} - 1)/(\omega - 1)$ and $F(e, \theta) = \theta e^{\alpha}$, as well as the normalizations $s = \theta = 1$ and the definition $X_t^* \equiv X_t^{\psi/(\alpha\delta)}$ into equations (I.9) and (J.1) from the main text

$$\begin{aligned} U'\left(s\left(X_t + F(e_t, \theta_t)\right) - \gamma s\left(X_{t-1} + F(e_{t-1}, \theta_{t-1})\right)\right) &= \beta\Theta \frac{\xi_t}{\xi_{t-1}} \left[e_t + \left(1 - e_t\right)\phi\right]\left(1 + r_t^p\right) \\ &\times \mathbb{E}_t \left[U'\left(s\left(X_{t+1} + F(e_{t+1}, \theta_{t+1})\right) - \gamma s\left(X_t + F(e_t, \theta_t)\right)\right) e_{t+1}^{\varphi_e}\right], \end{aligned} \tag{J.1}$$

equations (J.2) and (J.3)

$$\widehat{X}_{t+1}^{\star} = \left(1 - \delta\right)\widehat{X}_t^{\star} + \psi \widehat{e}_t\,, \tag{J.2}$$

$$\widehat{e}_t = \alpha_1 \widehat{X}_t^{\star} + \alpha_2 \widehat{e}_{t-1} + \alpha_3 \widehat{e}_{t+1} - \alpha_4 \widehat{r}_t^p + \alpha_4 \mu_t\,, \tag{J.3}$$

can then be obtained by linearizing these two equations around the non-stochastic steady state with respect to $\log(e_t)$, $\log(X_t^*)$, r_t^p, and $\mu_t \equiv -\Delta\log(\xi_t)$. In particular, the coefficients in (J.3) are given by

$$\alpha_1 \equiv -\frac{\alpha\delta^2\left(1 - \delta - \gamma\right)}{\left(1 - \delta\right)\kappa}\,,$$

$$\alpha_2 \equiv \frac{\gamma\alpha\delta\left(1 - \delta - \psi\right)}{\left(1 - \delta\right)\kappa}\,,$$

$$\alpha_3 \equiv \frac{\alpha\delta - \tau\varphi_e}{\kappa}\,,$$

$$\alpha_4 \equiv \frac{\tau}{\kappa}\,,$$

where $\tau \equiv (1-\gamma)(\psi+\delta)/\omega$, $\kappa \equiv (1+\gamma-\psi)\alpha\delta + \tau\Xi$, $\Xi \equiv (1-\phi)e_s/[e_s+(1-e_s)\phi]$, and e_s is the steady state employment rate.

For the linear RP model, we may obtain the (log-)linearized version of equation (J.4)

$$1 + r_t^p = \frac{1+(1-e_t)\,\phi\Phi}{e_t+(1-e_t)\,\phi} \tag{J.4}$$

as $\widehat{r}_t^p = \varrho_1 \widehat{e}_t$, where $\varrho_1 = -(1-\phi+\phi\Phi)e_s/[e_s+(1-e_s)\,\phi]^2$. For the non-linear RP model, we have $\phi_t = \phi(e_t)$ and $r_t^p = f(e_t)$, where

$$f(e) = \frac{1+(1-e)\,\phi\,(e)\,\Phi}{e+(1-e)\,\phi\,(e)} - 1\,.$$

Letting $\tilde{e} \equiv \log e$ and $g(\tilde{e}) \equiv f(\exp\{\tilde{e}\})$, and noting that $\phi'(e_s) = 0$ by assumption, the third-order approximation to (J.4) is given by

$$\widehat{r}_t^p = \varrho_1 \widehat{e}_t + \frac{1}{2} g'' \widehat{e}_t^2 + \frac{1}{6} g''' \widehat{e}_t^3\,,$$

where ϱ_1 is as in the linear RP model, and the derivatives of g are evaluated at the steady state and are implicitly dependent on the corresponding derivatives of ϕ of the same order or lower. In practice, we directly estimate ϱ_2 and ϱ_3 in the expression $\widehat{r}_t^p = \varrho_1\widehat{e}_t + \varrho_2\widehat{e}_t^2 + \varrho_3\widehat{e}_t^3$, and then solve for the values of ϕ'' and ϕ''' such that $g'' = 2\varrho_2$ and $g''' = 6\varrho_3$.

J.2 Solution Method

For the three linear versions of the model (linear RP, no friction, and canonical), we use standard (linear) perturbation methods to solve the model (see, e.g., Fernández-Villaverde, Rubio-Ramírez, and Schorfheide [2016]). For the non-linear RP model, since we would like to allow for the possibility of local instability and limit cycles, standard non-linear perturbation methods are not appropriate. In particular, as a first step, standard methods require obtaining a rational-expectations solution to the linearized (i.e., linear RP, in this case) model. If the model is locally unstable (as is necessarily the case if it features limit cycles), such a solution cannot exist. We therefore instead use the perturbation method discussed in Galizia [2018]. The method differs from standard methods in that it does not require that the linear approximation to the solution of the non-linear model also be a solution to the linearized model, an assumption implicit in standard methods. See Galizia [2018] for further details.

In practice, we solve the non-linear RP model in this manner to a third-order approximation. Given a parameterization and an associated solution in state-space form, as part of the Galizia [2018] method we need to verify that the system indeed remains bounded (in expectation). It is not possible to confirm this analytically, so we employ numerical methods instead. In particular, to minimize computational burden, we simply check whether, for a given initial (non-zero) state of the system,[xiv] the deterministic path for hours (i.e., the one obtained by feeding in a constant stream of zeros for the shock) explodes, where we define an explosion as a situation where, within the first 270 simulated quarters (the length of our data set), the absolute value of hours exceeds $10 \times \bar{L}$, where $\bar{L}$ is the maximum absolute value of hours (in log-deviations from the mean) observed in our data set ($\bar{L} = 0.146$).

J.3 Estimation Procedure

To estimate the model, we use an indirect inference method as follows. Let $x_t \in \mathbb{R}^n$ denote a vector of date-t observations in our data set, $t = 1, \ldots, T$, and let $\mathbf{x}_T \equiv (x_1', \ldots, x_T')'$ denote the full data set in matrix form. Let $F : \mathbb{R}^{T \times n} \to \mathbb{R}^q$ be the function that generates the q-vector of features of the data we wish to match (i.e., $F(\mathbf{x}_T)$ is a vector containing all relevant spectrum values, plus, for the non-linear RP model, the correlation, skewness and kurtosis for hours and the risk premium).

Suppose we simulate M data sets of length T from the model using the parameterization θ. Collect the m-th simulated simulated data set in the matrix $\tilde{\mathbf{x}}_T^m(\theta) \in \mathbb{R}^{T \times n}$, $m = 1, \ldots, M$. The estimation strategy is to choose the parameter vector θ to minimize the Euclidean distance between $F(\mathbf{x}_T)$ and the average value of $F(\tilde{\mathbf{x}}_T^m(\theta))$, i.e., we seek the parameter vector

$$\widehat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \left[F(\mathbf{x}_T) - \frac{1}{M} \sum_{m=1}^{M} F(\tilde{\mathbf{x}}_T^m(\theta)) \right]' \left[F(\mathbf{x}_T) - \frac{1}{M} \sum_{m=1}^{M} F(\tilde{\mathbf{x}}_T^m(\theta)) \right],$$

where Θ is the parameter space. In practice, we simulate $M = 3{,}000$ data sets for each parameter draw, and estimate $\widehat{\theta}$ using Matlab 's `fminsearch` optimization function. We explored several different weighting matrices in estimation which all gave similar results. In the main text, we report results based on using an identity matrix as the weighting matrix.

[xiv] We use the same initial point for every parameterization.

The Parameter Space

We estimate the parameters of the model imposing several restrictions on the parameter space Θ. First, we require that $0 \leq \gamma, \psi < 1$. Second, we require that the policy rate reacts positively to expected hours, but not so strongly as to cause current hours to fall in response to an increase in expected hours, i.e., $0 < \varphi_e < \alpha\delta/\tau$, where τ is defined above. Third, we impose that the degree of complementarity is never so strong as to generate temporary multiple equilibria.[xv] This latter property is ensured if the function $\widehat{e}_t + \frac{\tau}{\kappa} R^p(\widehat{e}_t)$ (see equation (22) in the main text) is strictly increasing in $\widehat{\ell}$ (so that it is invertible). This in turn requires $1 + \frac{\tau}{\kappa} R_1^p > 0$ and $R_3^p \geq R_2^{p2}/[3(\kappa/\tau + R_1^p)]$. Fourth, we impose that the shock process is stationary, i.e., $|\rho| < 1$. Finally, we require that the parameters be such that a solution to the model exists and is unique (see Appendix J.2). Note that none of the estimated parameters in either the linear or non-linear RP models is on the boundary of the set of constraints we have imposed.

[xv] By temporary multiple equilibria, we mean a situation where, for a given X_t, e_{t-1} and expectation about e_{t+1}, there are multiple values of e_t consistent with the dynamic equilibrium condition.

References

Alessio, S. M. (2016): *Digital Signal Processing and Spectral Analysis for Scientists*. Springer, Cham.

Bai, J., and S. Ng (2005): "Tests for Skewness, Kurtosis, and Normality for Time Series Data," *Journal of Business & Economic Statistics*, 23(1), 49–60.

Cochrane, J. (2005): "Time Series for Macroeconomics and Finance," Unpublished manuscript.

Fernald, J. (2014): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," Federal Reserve Bank of San Francisco Working Paper 2012-19.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, and F. Schorfheide (2016): "Solution and estimation methods for DSGE models," *Handbook of Macroeconomics*, 2, 527–724.

Galizia, D. (2018): "Saddle Cycles: Solving Rational Expectations Models Featuring Limit Cycles (or Chaos) Using Perturbation Methods," Working Paper.

Granger, C. W. J. (1966): "The Typical Spectral Shape of an Economic Variable," *Econometrica*, 34(1), 150–161.

Marczak, M., and V. Gómez (2012): "SPECTRAN, a Set of Matlab Programs for Spectral Analysis," Discussion paper 60, Universität Hohenheim.

Newey, W. K., and K. D. West (1994): "Automatic Lag Selection in Covariance Matrix Estimation," *The Review of Economic Studies*, 61(4), 631–653.

Owyang, M., V. Ramey, and S. Zubairy (2013): "Are Government Spending Multipliers Greater During Periods of Slack? Evidence from 20th Century Historical Data," *American Economic Review*, 103(3), 129–34.

Priestley, M. (1981): *Spectral Analysis of Time Series*, vol. 1 Univariate Series. Academic Press, London.

Ramey, V. A., and S. Zubairy (2018): "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data," *Journal of Political Economy*, forthcoming.

Sargent, T. J. (1987): *Macroeconomic Theory*. Emerald Group Publishing Limited, second edn.

Shiller, R. J. (2015): *Irrational Exuberance*. Princeton University Press, 3rd edn.