Putting the Cycle Back into Business Cycle Analysis: Online Appendix

Paul Beaudry, Dana Galizia and Franck Portier

November 2016

`Version 7.7`

D Spectral Density Estimation

D.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 $[-\pi, \pi)$:

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

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

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

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.[2]

[1] 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.

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

D.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 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.[3]

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 $N = 1,024$.[4]

[3] 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 over all pattern of the sampled spectrum is in general unaffected by zero-padding.

[4] 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, $N = 1,000$.

D.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.[5] 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 over all 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 use a Hamming window of length $W = 13$ as the smoothing kernel.[6]

D.4 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, capacity utilization), 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

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

[5] 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).

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

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

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

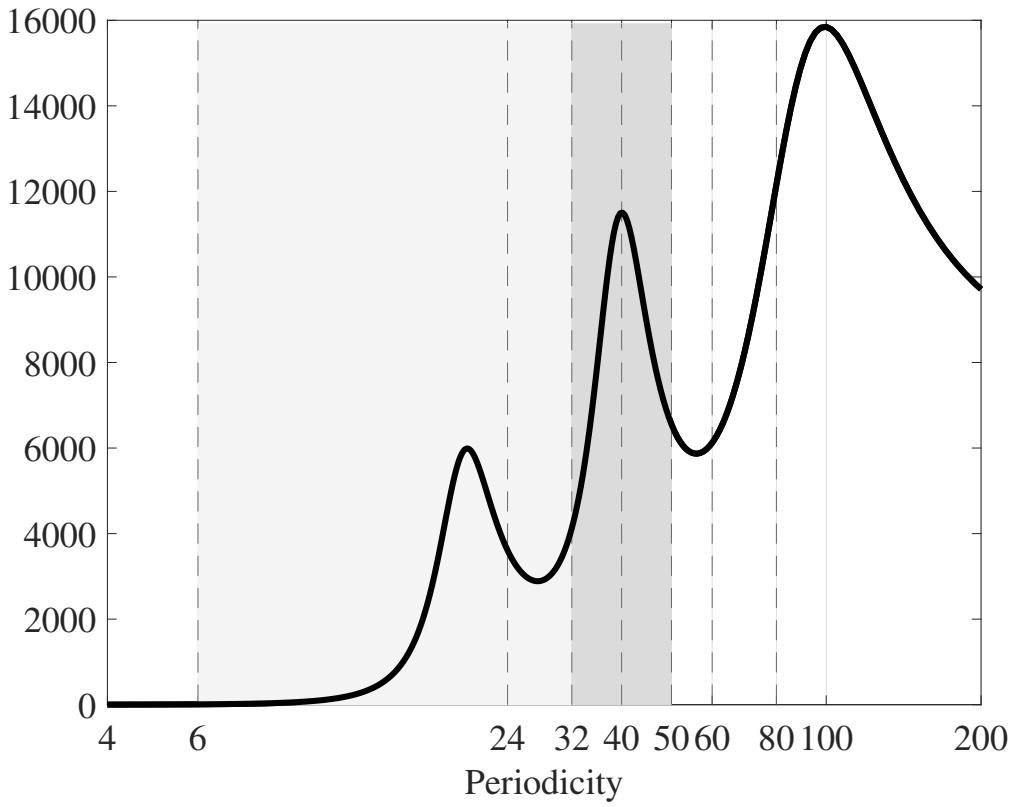
It can also be shown (see, e.g., Sargent [1987]) that for a given ρ_{i2}, the spectral density has a peak at frequency $\overline{\omega}_i$ if

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

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 1. As in the factor utilisation series we are using in the main text, the spectral density shows long-run fluctuations, but the bulk of the business cycle movements is explained by movements at the 40-quarter periodicity, although we observe another peak at periodicity 20 quarters.

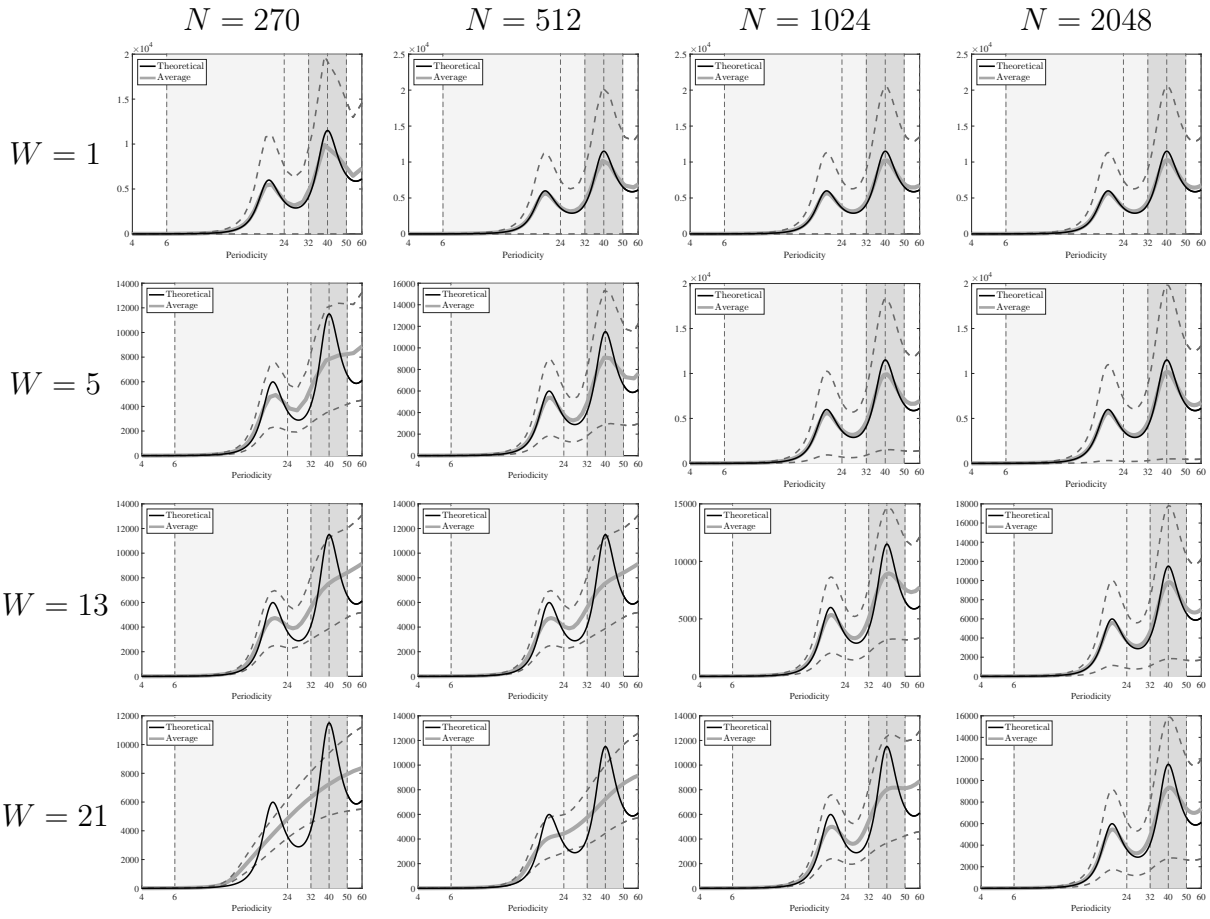
We simulate this process 1,000,000 times, with $T = 270$ for each simulation, which is the length of our observed macroeconomic series. We estimate the spectral density for various values of N (zero-padding) and W (length of the Hamming window). Higher N corresponds to higher resolution, and higher W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

Figure 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 ± one standard deviation. 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).

D.5 Smoothing and Zero-Padding with Non-Farm Business Hours per Capita

Figure 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.

D.6 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 4.

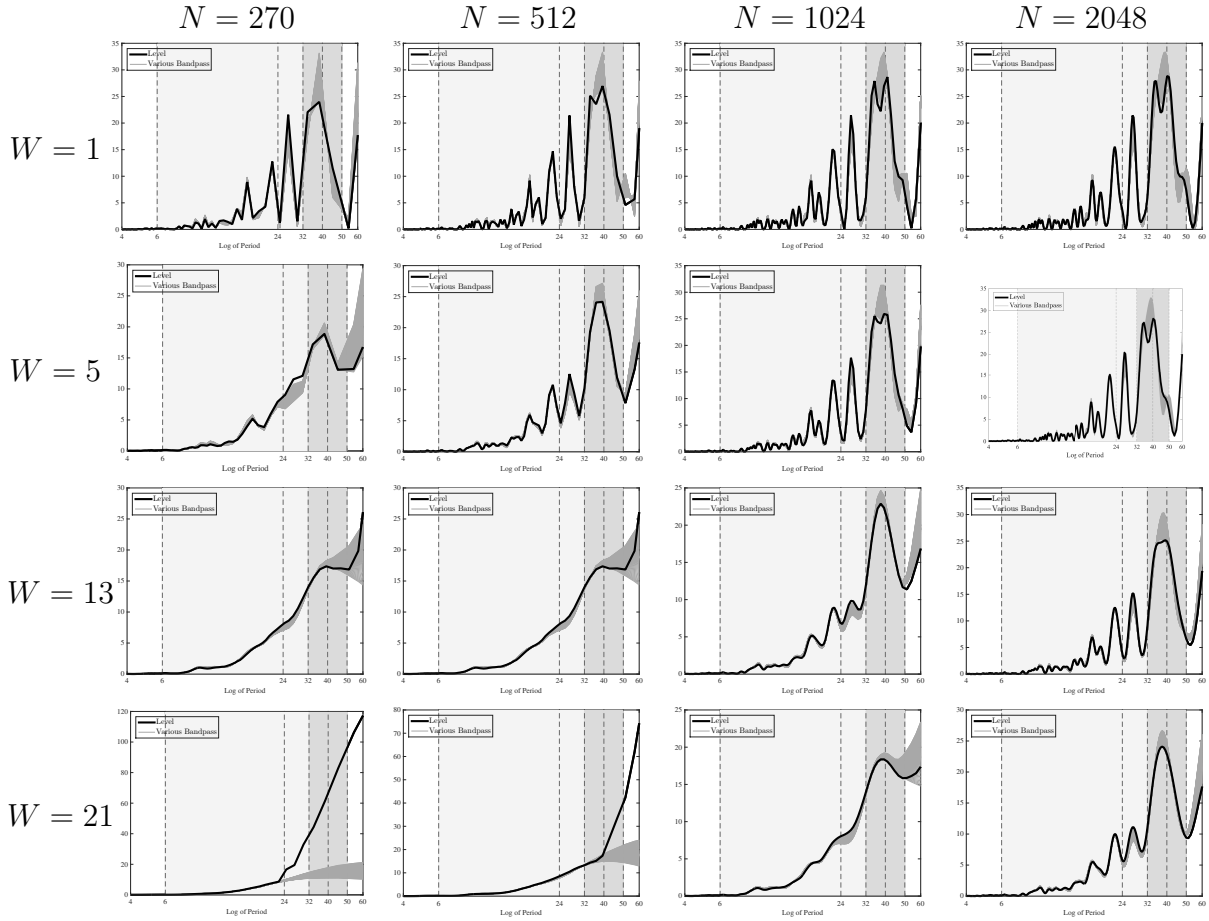
D.7 Alternative Estimators

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.[7] 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.[8] Figure 5 shows the estimated spectrum of Non Farm Business hours for the

[7] 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 $|k| \geq M$.

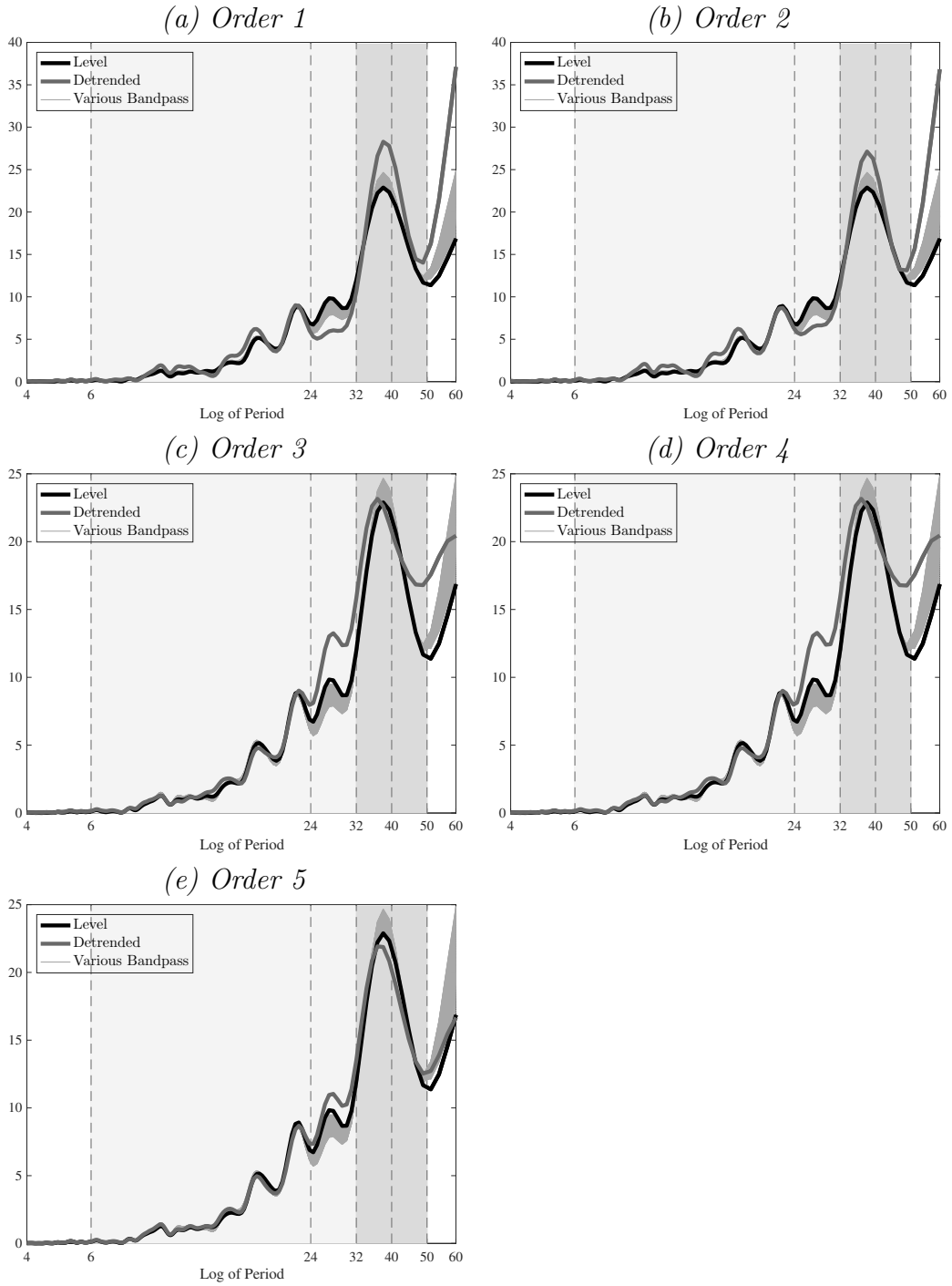
[8] Note that, in contrast to the kernel-smoothing case, in this case a wider window corresponds to less smoothing.

Figure 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).

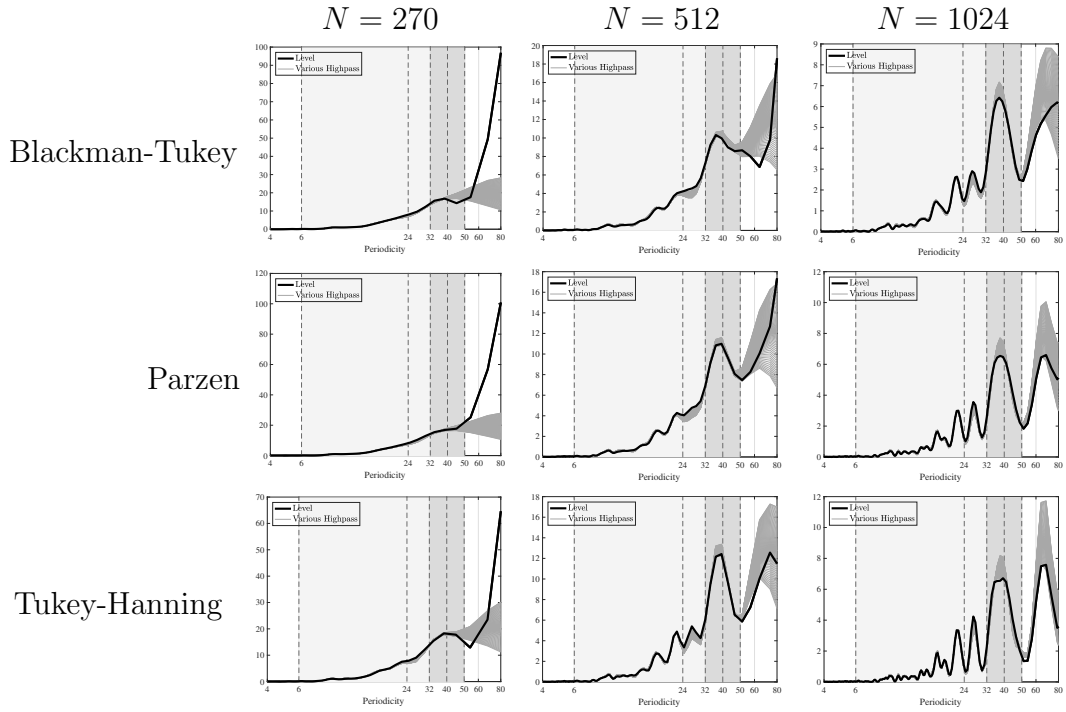
Figure 4: Using a Polynomial Trend of Various Orders for Benchmark Smoothing ($W = 13$) and Zero-Padding ($N = 1024$)



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).

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.

Figure 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).*

E Steps in Deriving the Equilibrium of the New Keynesian Model with Financial Frictions and Durables

E.1 The Model

Households

There is a continuum of mass one of households indexed by i who purchase consumption services from the market. The preferences of agent i are given by $\mathbb{E}_0 \sum_t \beta^t \xi_{t-1}[U\left(C_{it} - \gamma C_{t-1}\right) + \nu(1 - L_{it})]$, with $U'(\cdot) > 0, U''(\cdot) < 0$, $\nu'(\cdot)$, $\nu''(\cdot) < 0$ and $0 \leq \gamma < 1 - \delta$. C_{it} represents the consumption services purchased by household i in period t, C_t 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. Household i's problem takes the form

$$\max \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \xi_{t-1} \left[U\left(C_{it} - \gamma C_{t-1}\right) + \nu(1 - L_{it})\right]$$

subject to

$$P_t C_{it} + \left(1 + r_{t-1}\right) B_{it} + P_t^x D_{it} = B_{it+1} + r_t^x X_{it} + w_t L_{it} + \Gamma_{it},$$
$$X_{it+1} = (1 - \delta) X_{it} + D_{it}.$$

Here, B_{it+1} represents the borrowing by the household at time t, to be repaid with interest at $t + 1$, r_t represents the nominal interest rate faced by the household on such borrowing, $w_t L_{it}$ is labor income, X_{it} represents the stock of durable goods (or houses) held by the household at the beginning of period t, Γ_{it} are firm profits that are returned to households, and D_{it} the quantity of new durable goods purchased by the household at t. P_t and P_t^x are the (nominal) prices of consumption services and new durable goods, respectively, at date t. Households are assumed to buy all of their consumption services—including those derived from durable goods—from the market. Specifically, households do not consume the services of their durable goods directly. Instead, they rent X_{it} out to firms each period at nominal rental rate r_t^x. Firms then combine the rented stock of durables with labor in order to produce consumption services, C, as well as new durable goods, D, in a manner to be described shortly. Under these conditions, dropping i subscripts, the household's Euler

equation with respect to their optimal choice of C takes the familiar form:

$$U'\left(C_t - \gamma C_{t-1}\right) = \beta_t \mathbb{E}_t \left[U\left(C_{t+1} - \gamma C_t\right)\left(1 + r_t - \pi_{t+1}\right)\right], \tag{E.3}$$

where π_{t+1} is the inflation rate from t to $t+1$, and $\beta_t \equiv \beta \frac{\xi_t}{\xi_{t-1}}$. Note that we have here used the approximation $\frac{1+r}{1+\pi} = 1 + r - \pi$.

The households problem also leads a labor supply decisions represented by

$$\frac{\nu'(L_t)}{U'\left(C_t - \gamma C_{t-1}\right)} = \frac{w_t}{P_t}$$

and an arbitrage condition between the holding of bonds and capital given by

$$\mathbb{E}_t \left[\frac{U'\left(C_{t+1} - \gamma C_t\right)}{P_{t+1}} \left\{(1 + r_t) - \left(\frac{r^x_{t+1} + (1-\delta)P^x_{t+1}}{P^x_t}\right)\right\}\right] = 0.$$

Final Good producers

The final good sector is competitive. This sector provides consumption services to households by buying a set of differentiated intermediate services, denotes C_{jt}, from intermediate good firms. We assume a measure one of intermediate good firms, indexed by j. The technology of the final good firms is

$$C_t = \left(\int_0^1 C_{jt}^{\eta} dj\right)^{\frac{1}{\eta}}, \tag{E.4}$$

with $\eta \in (0, 1)$. The objective of the final good firm is thus to solve

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

subject to (E.4), where p_{jt} is the price of intermediate good j. This gives rise to demand for intermediate good j given by

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

Intermediate good firms

Intermediate good producers are monopolistically competitive and take the demand from final good firms as given. We assume that intermediate firms produce an intermediate factor M according to the technology

$$M_{jt} = BF\left(\Theta_t L_{jt}\right).$$

A fixed fraction $(1-\varphi)$ of that factor is transformed one to one into consumption services, whereas the use of the rented stock of durables also produces consumption services with a one to one technology, so that the amount of consumption services produced by intermediate firm j is

$$C_{jt} = X_{jt} + (1-\varphi)M_{jt} = X_{jt} + (1-\varphi)BF\left(\Theta_t L_{jt}\right). \tag{E.5}$$

The remaining fraction φ of the intermediate factor M is transformed one to one into durable goods D, so that

$$D_{jt} = \varphi M_{jt} = \varphi BF\left(\Theta_t L_{jt}\right).$$

Period-t profits of an intermediate good producer are given by

$$\Gamma_{jt} = P_{jt}C_{jt} + P_t^X D_{jt} - r_t^x X_{jt} - w_t L_{jt}.$$

Cost minimisation implies that

$$w_t = \left((1-\varphi)r_t^x + \varphi P_t^x\right)\Theta_t BF'\left(\Theta_t L_{jt}\right).$$

Hence, assuming $F'' < 0$, all intermediate good firms will hire the same amount of labor (i.e., $L_{jt} = L_t$), which implies that the aggregate capital stock will satisfy

$$X_{t+1} = (1-\delta)X_t + \varphi BF\left(\Theta_t L_t\right). \tag{E.6}$$

With the normalisation $B = \frac{1}{1-\varphi}$ and denoting $\psi = \frac{\varphi}{1-\varphi}$, we obtain

$$C_t = X_t + F\left(\Theta_t L_t\right) \tag{E.7}$$

and

$$X_{t+1} = (1-\delta)X_t + \psi F\left(\Theta_t L_t\right). \tag{E.8}$$

Setting of interest rates

We assume that the interest rate faced by households is equal to the policy rate set by the central bank (which we denote i_t) plus a risk premium r_t^p; that is, we assume that $r_t = i_t + r_t^p$. We directly allow for counter-cyclical risk premium by positing r_t^p to be a non-increasing function of the level of employment in the economy, L_t, via

$$r_t^p = R(L_t), \qquad R'(L_t) \leq 0.$$

In setting the policy rate i_t, we assume that the central bank follows a slightly modified Taylor-type rule where the interest rate reacts to expected inflation and to expected labor market conditions via

$$i_t = \phi_\pi \mathbb{E}_t \left[\pi_{t+1}\right] + \phi_\ell \mathbb{E}_t \left[G\left(L_{t+1}\right)\right], \qquad \phi_\pi, \phi_\ell, G' > 0. \tag{E.9}$$

Equilibrium in the absence of sticky prices

In the absence of sticky prices, all intermediate good producers will act in the same manner and set P_{jt} as a mark-up over the marginal cost, implying that

$$P_{jt} = \frac{r_t^x}{1-\eta}.$$

The equilibrium values for $\{P_t^x, w_t, r_t^x, P_t, r_t, \pi_t, i_t, L_t, C_t, X_t\}$ are given as the solution to the set of equations:

$$U'\left(C_t - \gamma C_{t-1}\right) = \beta_t \mathbb{E}_t \left[U\left(C_{t+1} - \gamma C_t\right)\left(1 + r_t - \pi_{t+1}\right)\right],$$

$$\frac{\nu'(L_t)}{U'\left(C_t - \gamma C_{t-1}\right)} = \frac{w_t}{P_t},$$

$$\mathbb{E}_t \left[\frac{U'\left(C_{t+1} - \gamma C_t\right)}{P_{t+1}} \left\{(1 + r_t) - \left(\frac{r_{t+1}^x + (1-\delta)P_{t+1}^x}{P_t^x}\right)\right\}\right] = 0,$$

$$w_t = (r_t^x + \psi P_t^x)\Theta_t F'\left(\Theta_t L_t\right),$$

$$P_t = \frac{r_t^x}{1-\eta},$$

$$C_t = X_t + F\left(\Theta_t L_t\right),$$

$$X_{t+1} = (1-\delta)\, X_t + \psi F\left(\Theta_t L_t\right),$$

$$i_t = \phi_\pi \mathbb{E}_t \left[\pi_{t+1}\right] + \phi_\ell \mathbb{E}_t \left[G\left(L_{t+1}\right)\right],$$

$$\pi_t = \frac{P_t}{P_{t-1}},$$

$$r_t = i_t + R(L_t).$$

In fact, this system does not determine the level of P_t, but instead gives solutions for $\{\frac{P_t^x}{P_t}, \frac{w_t}{P_t}, \frac{r_t^x}{P_t}, \pi_t, i_t, r_t, L_t, C_t, X_t\}$.

With sticky prices

Assume now that intermediate firms are confronted to price stickiness, in that the arrival of the option to change their price follows a Poisson process. In that case, the previous equilibrium conditions are not affected except for the two equations

$$P_t = \frac{r_t^x}{1-\eta} \tag{E.10}$$

$$C_t = X_t + F\left(\Theta_t L_t\right); \tag{E.11}$$

With sticky prices, the pricing equation (E.10) is replaced by a pricing equation for a firm that adjusts its price, with the result that the price is set to essentially be a markup on expected discounted marginal cost. From this we can derive a Phillips curve following standard steps.

Aggregate output in the presence of sticky prices is now be given by

$$C_t = \left(\int_0^1 \left(X_{jt} + F(\Theta_t L_{jt})\right)^{\rho} dj\right)^{\frac{1}{\rho}},$$

where this differs from (E.11) because of the distribution in X_{jt}. The X_{jt} will satisfy the equation

$$X_{jt} - F(\Theta_t L_{jt}) = \left(\frac{P_{jt}}{P_t}\right)^{\frac{-1}{1-\rho}} C_t$$

One important approximation step we are taking in the paper is that we disregard the impact of production dispersion on aggregate output, and are indeed assuming that (E.11) holds even in the case with sticky prices, while this is only approximatively true.

Deriving the linear reduced form

With the assumption $\phi_\pi = 1$, we can bypass the need to be explicit about the intermediate firm's pricing problem and its implications for inflation, as the realisations of inflation do not feed back into the determination of the quantity variables that are of interest to us. Specifically, substituting (E.9) (with $\phi_\pi = 1$) and (E.11) into the Euler Equation (E.3), and denoting $\varepsilon_{t+1}^{\pi} = \mathbb{E}_t\left[\pi_{t+1}\right] - \pi_{t+1}$, we obtain:

$$\begin{aligned} U'\big(X_t + F\left(\theta_t L_t\right) - \gamma\left(X_{t-1} + F\left(\theta_{t-1} L_{t-1}\right)\right)\big) = \\ \beta_t \mathbb{E}_t \Big\{ U'\big(X_{t+1} + F\left(\theta_{t+1} L_{t+1}\right) - \gamma\left(X_t + F\left(\theta_t L_t\right)\right)\big) \\ \times \big(1 + \varepsilon_{t+1}^{\pi} + \phi_\ell G\left(L_{t+1}\right) + R\left(L_t\right)\big) \Big\}. \end{aligned} \tag{E.12}$$

We assume the following functional forms. The utility function is $U(z) = -\exp\left\{-\frac{z}{\omega}\right\}$, $\omega > 0$, the production of new goods is given by $F(\Theta L) = \ln(\Theta L)$, the interest rate policy function is $G(L) = \ln(L)$, and Θ_t is constant. With the approximation $\ln(1+z) \approx z$, we have

$$\ln\left(1 + \varepsilon_{t+1}^{\pi} + \phi_\ell G\left(L_{t+1}\right) + R\left(L_t\right)\right) \approx \varepsilon_{t+1}^{\pi} + \phi_\ell \ln L_{t+1} + R\left(L_t\right).$$

Therefore, (E.12) can be rewritten, with the approximation $\mathbb{E}[\ln(z)] \approx \ln(\mathbb{E}[z])$:

$$\begin{aligned}\frac{1}{\omega}\left(X_t + \ln\Theta + \ln L_t - \gamma\left(X_{t-1} + \ln\Theta + \ln L_{t-1}\right)\right) = \\ \ln\beta_t + \frac{1}{\omega}\left(\mathbb{E}_t\left[X_{t+1}\right] + \ln\Theta + \mathbb{E}_t\left[\ln L_{t+1}\right] - \gamma\left(X_t + \ln\Theta + \ln L_t\right)\right) \\ + \mathbb{E}_t\left[\varepsilon_{t+1}^{\pi}\right] + \phi_\ell \mathbb{E}_t\left[\ln L_{t+1}\right] + R\left(L_t\right). \quad \text{(E.13)}\end{aligned}$$

Given that $\mathbb{E}_t\left[\varepsilon_{t+1}^{\pi}\right] = 0$, this equations rewrites, using $\ell_t = \ln L_t$,

$$\ell_t = \mu_t - \widehat{\alpha}_1 X_t + \widehat{\alpha}_2 \ell_{t-1} + \widehat{\alpha}_3 \mathbb{E}_t\left[\ell_{t+1}\right] + \widehat{F}(\ell_t), \quad \text{(E.14)}$$

where $\mu_t = \frac{1}{\kappa}\left[\psi\left(1 - \frac{\gamma}{1-\delta}\right)\ln\Theta - \omega\ln\beta_t\right]$, $\widehat{\alpha}_1 = \frac{\delta}{\kappa}\left(1 - \frac{\gamma}{1-\delta}\right)$, $\widehat{\alpha}_2 = \frac{\gamma}{\kappa}\left(1 - \frac{\psi}{1-\delta}\right)$, $\widehat{\alpha}_3 = \frac{1-\omega\phi_\ell}{\kappa}$ and $\widehat{F}(\ell_t) = -\frac{\omega}{\kappa}\widetilde{R}\left(\ell_t\right)$, with $\kappa = 1 + \gamma - \psi > 0$.

Equation (E.14) and the accumulation Equation (E.8) together form a two-variable linear system in X and ℓ. It is worth noticing that inflation π does not appear anywhere. Since C and D can be obtained from X and L, it then follows that all quantity variables are determined independently from the inflation rate.

F Solving and Estimating the Model

F.1 Solution Method

F.1.1 Solving a Deterministic Version of the New Keynesian Model

Consider the deterministic version of our model, with all variables expressed as deviations from the steady state. We can write our dynamic economic system as

$$\begin{pmatrix} y_{t+1} \\ \ell_{t+1} \end{pmatrix} = \begin{pmatrix} f\left(y_t, \ell_t\right) \\ g\left(y_t, \ell_t\right) \end{pmatrix} \equiv h\left(y_t, \ell_t\right), \quad \text{(F.15)}$$

where $y_t = \left(X_t, \ell_{t-1}\right)'$ is the vector of predetermined variables, $h\left(0\right) = 0$, and y_0 is given. Letting $x_t \equiv \left(y_t', \ell_t\right)'$, a solution is a function $\phi : \mathbb{R}^2 \to \mathbb{R}$ such that, after setting $\ell_0 = \phi\left(y_0\right)$,

the resulting sequence $\{x_t\}$ obtained from (F.15) satisfies the transversality condition (TVC) $\limsup_{t\to\infty} \|x_t\| < \infty$ (i.e., the system remains bounded). Suppose this ϕ exists and is unique, and let $x(t;y) \equiv h^t(y, \phi(y))$[9] denote the state at date t when $y_0 = y$ and $\ell_0 = \phi(y)$ (i.e., we have $x_t = x(t;y_0)$). Letting

$$\mathcal{M} \equiv \left\{x(0;y) : y \in \mathbb{R}^2\right\} \subset \mathbb{R}^3,$$

it must be the case that $x(t;y) \in \mathcal{M}$ for all t, y.[10] That is, $\mathcal{M}$ is the unique 2-dimensional invariant manifold[11] of h such that system (F.15), when restricted to $\mathcal{M}$, produces non-explosive dynamics.[12] In this case, ϕ is the function that projects x onto $\mathcal{M}$ by choice of l. We henceforth make the following assumption about ϕ (if it exists), which is necessary if we are to solve for it using perturbation methods (as we do below): ϕ is analytic on a neighborhood of the steady state.

To find $\mathcal{M}$, let $A \equiv D_x h(0)$,[13] and write the linearized version of (F.15) as.

$$x_{t+1} = Ax_t. \tag{F.16}$$

Let $\mathcal{E} \subset \mathbb{R}^3$ be a 2-dimensional invariant subspace of A; that is, (a) $\mathcal{E} = \{\alpha z_1 + \beta z_2 : \alpha, \beta \in \mathbb{R}\}$ for some linearly independent basis vectors $z_1, z_2 \in \mathbb{R}^3$, and (b) if $x \in \mathcal{E}$ then $Ax \in \mathcal{E}$. Note that there are at most three possible such subspaces.[14] Given such an $\mathcal{E}$, we look for a candidate $\mathcal{M}$ as the 2-dimensional invariant manifold of (F.15) with the following properties:

[9] We use h^t to denote the t-th iterate of h; i.e., h^0 is the identity function and $h^t \equiv h \circ h^{t-1}$ for $t \geq 1$.

[10] To see this, let $y(t;y)$ denote the first two elements of $x(t;y)$ and $l(t;y)$ the third, and note that the trajectory passing through $x(t;y) \equiv \left(y(t;y)', l(t;y)\right)'$ remains bounded by construction. Furthermore, also by construction, the trajectory passing through $x(0; y(t;y)) \equiv \left(y(t;y)', \phi(y(t;y))\right)'$ remains bounded as well. If $l(t;y) \neq \phi(y(t;y))$, then if we had $y_0 = y(t;y)$ there would be two different solutions for ℓ_0 that produce bounded trajectories, which contradicts our supposition that ϕ is unique.

[11] A manifold $\tilde{\mathcal{M}}$ is said to be h-invariant if $x \in \tilde{\mathcal{M}}$ implies $h(x) \in \tilde{\mathcal{M}}$.

[12] That is, $\mathcal{M}$ is precisely the "non-explosive manifold" referred to in the text.

[13] Here, the notation D_z indicates the partial derivative operator with respect to the vector z.

[14] To ensure this, we must actually impose the following additional technical requirement in constructing $\mathcal{E}$: if the algebraic multiplicity of an eigenvalue λ of A is greater than one, and $\mathcal{E}$ contains a generalized eigenvector corresponding to λ, then $\mathcal{E}$ must contain *all* of the generalized eigenvectors corresponding to λ. We henceforth impose this requirement. As a consequence, if the three eigenvalues of A are all real and distinct, then there are three $\mathcal{E}$'s; namely, those associated with the three possible ways of choosing two linearly independent eigenvectors of A as the basis vectors z_1 and z_2. If instead A has an eigenvalue λ with an algebraic multiplicity of two, then there is only one $\mathcal{E}$; i.e., the one given by the generalized eigenspace of λ. If instead A has a complex conjugate pair of eigenvalues, then there is also only one $\mathcal{E}$; i.e., the one obtained by taking $\mathrm{Re}(v)$ and $\mathrm{Im}(v)$ as the basis vectors, where v is any eigenvector associated with one of the complex eigenvalues. Finally, if A has an eigenvalue with an algebraic multiplicity of three, then such an $\mathcal{E}$ does not exist.

(a) $\mathcal{M}$ is tangent to $\mathcal{E}$ at the steady state; and (b) the associated projection function ϕ is analytic on a neighborhood of the steady state.[15] Given a candidate $\mathcal{M}$ and the associated ϕ, we can then obtain an expression for the evolution of y_t on $\mathcal{M}$ as

$$y_{t+1} = \theta(y_t) \equiv f(y_t, \phi(y_t)).$$

Using this expression, we can check numerically whether indeed trajectories on the candidate $\mathcal{M}$ remain bounded. If they do, then $\mathcal{M}$ continues to be a candidate non-explosive manifold; otherwise, we rule it out. We can check all possible $\mathcal{E}$'s and their associated candidate $\mathcal{M}$'s in this way. If, after doing so, we have found more than one candidate $\mathcal{M}$, then we conclude that the solution is indeterminate (i.e., that ϕ is not unique). If, on the other hand, all candidate $\mathcal{M}$'s contain unbounded trajectories, then we conclude that no solution exists. Finally, if exactly one candidate $\mathcal{M}$ is found whose trajectories are bounded, then we conclude that $\mathcal{M}$ is indeed the non-explosive manifold we sought.[16]

F.1.2 Finding ϕ Using Perturbation Methods

Given a subspace $\mathcal{E}$ as defined above, the associated candidate ϕ (which fully determines $\mathcal{M}$) will in general not possess a closed-form solution. We thus solve for ϕ up to a k-th-order Taylor approximation around the steady state as follows. First, note that, by invariance of $\mathcal{M}$, ϕ must implicitly solve[17]

$$\phi(f(y, \phi(y))) = g(y, \phi(y)). \tag{F.17}$$

Differentiating (F.17) once with respect to y and evaluating at $y = 0$ yields

$$D_y\phi \cdot [D_y f + D_l f \cdot D_y\phi] = D_y g + D_l g \cdot D_y\phi, \tag{F.18}$$

which is a quadratic system in the two elements of $D_y\phi$. In general, this equation has a multiplicity of solutions. However, $D_y\phi$ is, by definition,[18] the function that projects x onto

[15] In general, there are an uncountably infinite number of manifolds satisfying requirement (a), but only one of these manifolds also satisfies requirement (b).

[16] If A has either an eigenvalue with an algebraic multiplicity of two or a complex conjugate pair of eigenvalues, so that there is only one $\mathcal{E}$, then we also need to verify that trajectories passing through points *not* on $\mathcal{M}$ *do* become unbounded. If they do not, then we also conclude in this case that the solution is indeterminate.

[17] This expression can be obtained from (F.15) by using the "invariance" relationships $\ell_t = \phi(y_t)$ and $\ell_{t+1} = \phi(y_{t+1})$, and the fact that $y_{t+1} = f(y_t, \ell_t)$.

[18] In particular, we are using here the fact that $\mathcal{M}$ is tangent to $\mathcal{E}$ at the steady state.

$\mathcal{E}$ by choice of l. It can be verified that this projection function—which can be obtained using standard eigenvalue-eigenvector methods—is indeed a solution to (F.18). This gives us a candidate ϕ up to a first-order approximation.

To obtain the desired k-th-order approximation, we proceed iteratively as follows. Suppose we have in hand a $(j-1)$-th order approximation. To obtain the j-th-order approximation, obtain all j-th-order derivatives of the expression (F.17) and evaluate them at $y = 0$. We can express the result as a system of linear equations in the j-th-order derivatives of ϕ, with coefficients that are functions of the known derivatives of ϕ up to $(j-1)$-th order. It is thus straightforward to solve this system for the j-th-order derivatives.

F.1.3 Extension to the Stochastic Case

Write the stochastic version of the model as

$$\begin{pmatrix} y_{t+1} \\ \mathbb{E}_t\left[\ell_{t+1}\right] \end{pmatrix} = \begin{pmatrix} f\left(y_t, \ell_t, \mu_t; \sigma_t\right) \\ g\left(y_t, \ell_t, \mu_t; \sigma_t\right) \end{pmatrix} \equiv h\left(y_t, \ell_t, \mu_t; \sigma_t\right), \tag{F.19}$$

with $\mu_t = \rho\mu_{t-1} + \sigma_t\eta_t$, $\eta_t \sim N(0,1)$, and $\sigma_t = \sigma$.[19] Write the augmented state vector of this system as $z_t = (x_t', \mu_t, \sigma_t)'$. Whereas before the solution ϕ was the function that projected the system onto a manifold $\mathcal{M}$ in x-space, in the stochastic environment the corresponding manifold is in z-space, and in particular, $\phi : \mathbb{R}^4 \to \mathbb{R}$ is the function such that (a) $\ell_t = \phi\left(y_t, \mu_t; \sigma\right)$ for all t satisfies (F.19), and (b) the sequence generated by

$$y_{t+1} = \theta\left(y_t, \mu_t; \sigma\right) \equiv f\left(y_t, \phi\left(y_t, \mu_t; \sigma\right), \mu_t; \sigma\right)$$

satisfies the stochastic TVC, $\limsup_{t\to\infty} \mathbb{E}_0 \left\|y_t\right\| < \infty$.

Noting that $D_\sigma h\left(0\right) = 0$ in our setup, and taking date-t expectations of both sides of the linearized version of (F.19), we may obtain

$$\mathbb{E}_t\left[z_{t+1}\right] = \begin{pmatrix} 0 \\ 0 \\ \sigma \end{pmatrix} + \underbrace{\begin{pmatrix} A_x & A_\mu & 0 \\ 0 & \rho & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{A} z_t, \tag{F.20}$$

where $A_x \equiv D_x h\left(0\right)$ and $A_\mu \equiv D_\mu h\left(0\right)$. Note that ρ and 0 are eigenvalues of A, and let v_ρ and v_0 denote corresponding eigenvectors. Let $\mathcal{E}$ be a 4-dimensional invariant subspace of

[19] The reason we introduce σ_t as a (degenerate) state variable should become apparent shortly.

A such that $v_\rho, v_0 \in \mathcal{E}$; that is, (a) $\mathcal{E} = \left\{\sum_{j=1}^{4} \alpha_j z_j : \alpha_j \in \mathbb{R},\ j = 1, \ldots, 4\right\}$ for some linearly independent basis vectors $z_j \in \mathbb{R}^5$, (b) $v_\rho, v_0 \in \mathcal{E}$, and (c) if $x \in \mathcal{E}$ then $Ax \in \mathcal{E}$. Note that, as in the non-stochastic case, there are at most three possible such subspaces.

Given such an $\mathcal{E}$, we look for a candidate $\mathcal{M}$ (or, equivalently, the associated projection function ϕ) in a similar way to the non-stochastic case. That is, we seek the 4-dimensional manifold $\mathcal{M}$ with the following properties: (a) $\ell_t = \phi(y_t, \mu_t; \sigma)$ for all t satisfies the expression (F.19); (b) $\mathcal{M}$ is tangent to $\mathcal{E}$ at the non-stochastic steady state $z_t = 0$; and (c) the function ϕ is analytic on a neighborhood of the non-stochastic steady state. After obtaining the candidate $\mathcal{M}$ for each possible $\mathcal{E}$, we can then check numerically whether the stochastic TVC is satisfied for exactly one of these candidate $\mathcal{M}$'s, in which case we have found the desired solution.

As in the non-stochastic case, we cannot in general find ϕ analytically. However, from (F.19), we may obtain that ϕ implicitly solves

$$\mathbb{E}\left[\phi\left(f\left(y, \phi\left(y, \mu; \sigma\right), \mu; \sigma\right), \rho\mu + \sigma\eta; \sigma\right)\right] = g\left(y, \phi\left(y, \mu; \sigma\right), \mu; \sigma\right), \tag{F.21}$$

where the expectation on the left-hand side of (F.21) is taken over realizations of the i.i.d. $N(0,1)$ random variable η. Note also that the other time-varying variables in this expression (i.e., y and μ) are determined independently of η. We can thus easily solve for the k-th-order Taylor approximation to ϕ around the non-stochastic steady state $z = 0$ in a manner similar to the non-stochastic case by sequentially differentiating the expression (F.21) with respect to the vector $(y, \mu; \sigma)$.

F.2 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 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

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

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.

F.2.1 The Parameter Space

We estimate the nine parameters of the model imposing several restrictions on the parameter space Θ. First, we require that the habit parameter γ and durables-share parameter ψ be non-negative and less than one, i.e., $0 \leq \gamma, \psi < 1$. 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., $0 < \phi_\ell < 1/\omega$. Third, we impose that $\widetilde{R}'(0) \leq 0$ (i.e., we have complementarity near the steady state), but that the degree of complementarity is never so strong as to generate static multiple equilibria.[20] This latter property is ensured if the function $\widehat{\ell} + \frac{\omega}{\kappa}\widetilde{R}\left(\widehat{\ell}\right)$ is strictly increasing in $\widehat{\ell}$ (so that it is invertible), which requires $0 \geq \widetilde{R}_1 > -\frac{1+\gamma-\psi}{\omega}$ and $\widetilde{R}_3 > \frac{\omega\widetilde{R}_2^2}{3\left(1+\gamma-\psi+\omega\widetilde{R}_1\right)}$. 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 F.1). None of the estimated parameters is on the boundary of the set of constraints we have imposed.

[20] By static multiple equilibria, we mean a situation where, for a given $\widehat{X}_t$, $\widehat{\ell}_{t-1}$ and expectation about $\widehat{\ell}_{t+1}$, there are multiple values of $\widehat{\ell}_t$ consistent with the dynamic equilibrium condition.

References

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

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

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

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

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