The Econometrics of the Hodrick-Prescott filter

Preliminary

Robert M. de Jong* Neslihan Sakarya†

May 19, 2013

Abstract

The Hodrick-Prescott (HP) filter is a commonly used tool in macroeconomics, and is used to extract a trend component from a time series. In this paper, we derive a new representation of the transformation of the data that is implied by the HP filter. This representation highlights that the HP filter is a weighted average plus a number of adjustments that are important near the begin and end of the sample. Using this representation, we characterize the large T behavior of the HP filter and find conditions under which it is asymptotically equivalent to a symmetric weighted average. We find that the cyclical component of the HP filter possesses weak dependence properties when the HP filter is applied to a stationary mixing process, a linear deterministic trend process and/or a process with a unit root. This justifies the use of the HP filter as a tool to achieve weak dependence in a time series and illustrates that the finding in empirical macro that data series tend to have deterministic trends and/or unit roots and the practice of using inference procedures based on the cyclical component of the HP filter are not contradictory. In addition, a large bandwidth approximation to the HP filter is derived, and using this approximation we find an alternative justification for the procedure given in Ravn and Uhlig (2001) for adjusting the bandwidth for the data frequency.

*Department of Economics, Ohio State University, 444 Arps Hall, Columbus, OH 43210, USA, Email de-jong.8@osu.edu.

†Department of Economics, Ohio State University, 410 Arps Hall, Columbus, OH 43210, USA, Email neslihan.sakarya@gmail.com.

1 Introduction

The Hodrick-Prescott (HP) filter is the standard technique in macroeconomics for separating the long run trend in a data series from short run fluctuations. While it seems intuitively clear that no smoothing technique should be equally well applicable to all types of trended macroeconomic data, the HP filter is universally used in macroeconomics, and while different types of criticism can be found in the literature on the HP filter, as Ravn and Uhlig (2002) state, "the HP-filter has withstood the test of time and the fire of discussion remarkably well." Well-cited papers in which the HP filter is applied are for example Kydland and Prescot (1990) and Backus and Kehoe (1992); however, the HP filter is at this time a commonplace tool in macroeconomics.

The HP filter smoothed series $\hat{\tau}_T = (\hat{\tau}_{T1}, \hat{\tau}_{T2}, \ldots, \hat{\tau}_{TT})'$ as introduced into Economics in Hodrick and Prescott (1980,1993) results from minimizing, over all $\tau \in \mathbb{R}^T$,

$$\sum_{t=1}^{T}(y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1}(\tau_{t+1} - 2\tau_t + \tau_{t-1})^2, \tag{1}$$

where T denotes sample size, λ is a (positive) smoothing parameter that for quarterly data is often chosen to equal 1600, and $y = (y_1, \ldots, y_T)'$ is the data series to be smoothed. It has been pointed out that in Whittaker (1923) a similar filtering technique was introduced. The $\hat{\tau}_{Tt}$ are typically referred to as the "trend component," while $y_t - \hat{\tau}_{Tt}$ is referred to as the "cyclical component." It can be shown that there exists a unique minimizer to the minimization problem of Equation (1), and that for a known positive definite $(T \times T)$ matrix F_T, letting I_T denote the $(T \times T)$ identity matrix,

$$y = (\lambda F_T + I_T)\hat{\tau}_T \text{ and } \hat{\tau}_T = (\lambda F_T + I_T)^{-1} y; \tag{2}$$

see for example Kim (2004). Therefore the trend component $\hat{\tau}_{Tt}$ and the cyclical component $y_t - \hat{\tau}_{Tt}$ are both weighted averages of the y_t, and in the sequel of this paper, we write $\hat{\tau}_{Tt} = \sum_{s=1}^{T} w_{Tts} y_s$; note that the dependence of w_{Tts} and $\hat{\tau}_{Tt}$ on λ is suppressed for notational convenience. However, the inability to find a simple expression for the elements of $(\lambda F_T + I_T)^{-1}$ has prevented researchers from finding a simple expression for the weights that are implicit in the HP filter. It also follows from the definition of the HP filter that $\hat{\tau}_{Tt}(y_1+1, y_2+1, \ldots, y_T+1) = \hat{\tau}_{Tt}(y_1, y_2, \ldots, y_T) + 1$, and therefore, $\sum_{s=1}^{T} w_{Tts} = 1$ for $t \in \{1, 2, \ldots, T\}$. Also from the definition of the HP filter it follows that $\hat{\tau}_{Tt}(1, 2, \ldots, T) = t$, implying that $\sum_{s=1}^{T} w_{Tts} s = t$ for $t \in \{1, 2, \ldots, T\}$.

The first order condition for $\hat{\tau}_{Tt}$, $t \in \{3, \ldots, T-2\}$ is

$$-2(y_t - \hat{\tau}_{Tt}) - 4\lambda(\hat{\tau}_{T,t+1} - 2\hat{\tau}_{Tt} + \hat{\tau}_{T,t-1}) + 2\lambda(\hat{\tau}_{Tt} - 2\hat{\tau}_{T,t-1} + \hat{\tau}_{T,t-2}) + 2\lambda(\hat{\tau}_{T,t+2} - 2\hat{\tau}_{T,t+1} + \hat{\tau}_{Tt}) = 0$$

$$\tag{3}$$

which, letting $\bar{B}$ denote the forward operator and B the backward operator, simplifies to

$$y_t = (\lambda\bar{B}^2 - 4\lambda\bar{B} + (1+6\lambda) - 4\lambda B + \lambda B^2)\hat{\tau}_{Tt}, \tag{4}$$

which can be written as

$$y_t = (\lambda|1-B|^4 + 1)\hat{\tau}_{Tt}. \tag{5}$$

Some papers in the literature that analyze the HP filter, such as King and Rebelo (1993), Cogley and Nason (1995), and McElroy (2009) take the above equation as the definition of the HP filter, thereby effectively studying an approximate procedure that intuitively makes sense in large samples for values of t away from the begin and end points of the sample. Therefore, such an approach gives an informal large T approximation to the HP filter without a formal large T justification; is not able to address end point issues; is unable to address the large T validity of the approximation; and cannot establish formal properties of the resulting cyclical or trend components, such as the weak dependence properties of the cyclical component. The latter is of particular interest, since weak dependence properties of the cyclical component guearantee that standard asymptotic inference procedures are valid when applied to the cyclical component.

2 The weights of the HP filter

In this section, we derive the exact weights w_{Tts} implied by the HP filter. Our analysis starts by noting that we can also equivalently minimize over $(\theta_1, \ldots, \theta_T)$, for a basis of functions $p_j(\cdot) : [0,1] \to \mathbb{R}$ for $T, j \in \mathbb{N}^+$,

$$\sum_{t=1}^{T}(y_t - \sum_{j=1}^{T}\theta_j p_j(t/T))^2 + \lambda\sum_{t=2}^{T-1}(\sum_{j=1}^{T}\theta_j p_j((t+1)/T) - 2\sum_{j=1}^{T}\theta_j p_j(t/T) + \sum_{j=1}^{T}\theta_j p_j((t-1)/T))^2$$

$$= \sum_{t=1}^{T}(y_t - \theta' p_{Tt})^2 + \lambda\sum_{t=2}^{T-1}(\theta'\Delta^2 p_{T,t+1})^2 \tag{6}$$

where $\theta' = (\theta_1, \ldots, \theta_T)$ and $p'_{Tt} = (p_1(t/T), p_2(t/T), \ldots, p_T(t/T))$. Differentiation with respect to θ now yields for the minimizer $\hat{\theta}_T$

$$0 = -2\sum_{t=1}^{T} y_t p_{Tt} + 2\sum_{t=1}^{T} p_{Tt}p'_{Tt}\hat{\theta}_T + 2\lambda\sum_{t=2}^{T-1}(\Delta^2 p_{T,t+1})(\Delta^2 p_{T,t+1})'\hat{\theta}_T, \tag{7}$$

which, if the inverse exists, is solved for

$$\hat{\theta}_T = (T^{-1}\sum_{t=1}^{T} p_{Tt}p_{Tt}' + \lambda T^{-1}\sum_{t=2}^{T-1}(\Delta^2 p_{T,t+1})(\Delta^2 p_{T,t+1})')^{-1}T^{-1}\sum_{t=1}^{T} y_t p_{Tt}. \tag{8}$$

Now we choose $p_1(t/T) = 1$ and $p_j(t/T) = \sqrt{2}\cos(\pi(j-1)(t-1/2)/T)$, $j = 2, \ldots, T$. Note that this is a minor abuse of notation, since formally, $p_j(\cdot)$ depends on T. We can then derive the following result. Below, I_T denotes the identity matrix of dimension $(T \times T)$.

Lemma 1. *Let $p_{Tt} = (p_1(t/T), \ldots, p_T(t/T))'$, where $p_1(t/T) = 1$ and $p_j(t/T) = \sqrt{2}\cos(\pi(j-1)(t-1/2)/T)$, $j = 2, \ldots, T$. Then*

$$T^{-1}\sum_{t=1}^{T} p_{Tt}p_{Tt}' = I_T \tag{9}$$

and

$$T^{-1}\sum_{t=2}^{T-1}(\Delta^2 p_{T,t+1})(\Delta^2 p_{T,t+1})'$$

$$= \mathrm{diag}(\{16\sin(\pi(j-1)/(2T))^4, j = 1, \ldots, T\}) - 32T^{-1}q_{T1}q_{T1}' - 32T^{-1}q_{T2}q_{T2}', \tag{10}$$

and for $j \in \{1, ..., T\}$, $q_{T1j} = \sin(\pi(j-1)/(2T))^2\cos(\pi(j-1)/(2T))$ and $q_{T2j} = \sin(\pi(j-1)/(2T))^2\cos(\pi(j-1)(T-1/2)/T)$.

The proof of this result and all proofs for this paper can be found in the Mathematical Appendix.

The importance of the above result is that the matrix to be inverted is now "close" to an easily invertible diagonal matrix (in the sense that two matrices of rank 1 have been added to the diagonal matrix), and therefore allows for a more tractable expression. When minimizing the criterion of Equation (1) over τ, no such structure occurs. It is well-known that explicit formulas can be obtained for the inverse of the sum of a matrix plus another matrix of rank 1, and such results can be adapted to deal with the inverse of a matrix plus a matrix of rank 2 as well. Our strategy will be to use such a result, as obtained in Miller (1981), in order to obtain a tractable expression for $\hat{\tau}_{Tt}$.

For the theorem below in which we derive the exact weights of the HP filter, we need to define, for $m \in \mathbb{Z}$,

$$f_{T\lambda}(m) = 1/(2T) + (-1)^m(2T)^{-1}(1+16\lambda)^{-1}$$

$$+T^{-1}\sum_{j=2}^{T}\cos(\pi(j-1)m/T)(1+16\lambda\sin(\pi(j-1)/(2T))^4)^{-1} \tag{11}$$

and

$$g_{T\lambda}(m) = T^{-1}\sum_{j=1}^{T}\sqrt{2}\cos(\pi(j-1)(m-1/2)/T)q_{T1j}(1+16\lambda\sin(\pi(j-1)/(2T))^4)^{-1}. \tag{12}$$

In addition, define the sequences

$$\delta_{T\lambda} = T^{-1}q'_{T1}(I_T+\lambda D_T)^{-1}q_{T1}, \tag{13}$$

$$\eta_{T\lambda} = T^{-1}q'_{T1}(I_T+\lambda D_T)^{-1}q_{T2}, \tag{14}$$

$$\xi_{T\lambda} = 32\lambda(1-64\lambda\delta_{T\lambda})(1-64\lambda\delta_{T\lambda}+32^2\lambda^2(\delta_{T\lambda}^2-\eta_{T\lambda}^2))^{-1}$$

$$+32^2\lambda^2(1-64\lambda\delta_{T\lambda}+32^2\lambda^2(\delta_{T\lambda}^2-\eta_{T\lambda}^2))^{-1}\delta_{T\lambda}, \tag{15}$$

and

$$\phi_{T\lambda} = 32^2\lambda^2(1-64\lambda\delta_{T\lambda}+32^2\lambda^2(\delta_{T\lambda}^2-\eta_{T\lambda}^2))^{-1}\eta_{T\lambda}. \tag{16}$$

Note that $f_{T\lambda}(m) = f_{T\lambda}(-m)$ for all $m \in \mathbb{Z}$. With these definitions in place, we can now obtain the following result.

Theorem 1. *$\hat{\tau}_{Tt} = \sum_{s=1}^{T} w_{Tts}y_s$, where*

$$\begin{aligned} w_{Tts} &= f_{T\lambda}(t-s)+f_{T\lambda}(T)I(t+s-1=T) \\ &+ f_{T\lambda}(s+t-1)I(t+s-1<T)+f_{T\lambda}(2T-t-s+1)I(t+s-1>T) \\ &+ \xi_{T\lambda}g_{T\lambda}(t)g_{T\lambda}(s)+\phi_{T\lambda}g_{T\lambda}(T-t+1)g_{T\lambda}(s) \\ &+ \phi_{T\lambda}g_{T\lambda}(t)g_{T\lambda}(T-s+1)+\xi_{T\lambda}g_{T\lambda}(T-t+1)g_{T\lambda}(T-s+1) \\ &= \sum_{j=1}^{8} w_{Tts}^{j} \end{aligned} \tag{17}$$

and for $m \in \{1,2,\ldots,T\}$, $|f_{T\lambda}(m)| \leq Cm^{-3}$ and $|g_{T\lambda}(m)| \leq Cm^{-3}$ for some constant C not depending on T.

A simple computer program to verify the correctness of the formula of Equation (17), as well as other material related to this paper, can be found at `http://dl.dropbox.com/u/2159931/hp.html`.

The above formula highlights that a key feature of the HP filter is the creation of a symmetric weighted average of the data with w^1_{Tts}. w^2_{Tts} is of little consequence since it is bounded by CT^{-3} by the bound on $f_{T\lambda}(\cdot)$. w^3_{Tts} is small as long as t is away from the begin point of the sample, since it is bounded by $C(t+s-1)^{-3}$. Similarly, w^4_{Tts} is small as long as t is away from T. By the bound on $g_{T\lambda}(\cdot)$, w^5_{Tts} and w^7_{Tts} are also small for t away from 1, and w^6_{Tts} and w^8_{Tts} are small for t away from T. This can be formalized by noting that for any $\gamma \in (0, 1/2)$ (writing "$t \in [\gamma T, (1-\gamma)T]$" to mean "$t \in [\gamma T, (1-\gamma)T] \cap \mathbb{Z}$" as t will take integer values throughout this paper)

$$\sup_{T\geq 1, s\in\{1,\ldots,T\}, t\in[\gamma T,(1-\gamma)T]} |T^3 \sum_{j=2}^{8} w^j_{Tts}| < \infty. \tag{18}$$

In conclusion, from the above theorem it follows that the HP filter behaves like a symmetric weighted average with several correction terms that are important near the begin and end of the sample.

Note that since $\cos(\pi(j-1)(2T-t-s+1)/T) = \cos(\pi(j-1)(t+s-1)/T)$, it follows that $f_{T\lambda}(s+t-1) = f_{T\lambda}(2T-t-s+1)$, and therefore,

$$f_{T\lambda}(s+t-1) = f_{T\lambda}(2T-t-s+1) = f_{T\lambda}(T)I(t+s-1=T)$$

$$+f_{T\lambda}(s+t-1)I(t+s-1<T) + f_{T\lambda}(2T-t-s+1)I(t+s-1>T); \tag{19}$$

however, the formula of Theorem 1 brings out the fact that the HP filter is invariant to reversing the time order, i.e., makes it clear that

$$\hat{\tau}_{Tt}(y_1, y_2, \ldots, y_{T-1}, y_T) = \hat{\tau}_{T,T-t+1}(y_T, y_{T-1}, \ldots, y_2, y_1),$$

and is more useful when considering large T and large λ approximations to the HP filter. By noting that $\lim_{\lambda\to\infty} f_{T\lambda}(m) = (2T)^{-1}$ and that $\lim_{\lambda\to\infty} g_{T\lambda}(m) = 0$, it follows that $\lim_{\lambda\to\infty} \hat{\tau}_{Tt} = T^{-1}\sum_{t=1}^{T} y_t$, which can also be observed directly from the definition of the HP filter.

In Tables 1 and 2 in Appendix 1, values for $\xi_{T\lambda}$ and $\phi_{T\lambda}$ for various values of λ and T are listed. For both sequences, it appears that as $T \to \infty$, limit values are reached, sometimes with values near the limit reached for sample sizes as small as $T = 50$, and this phenomenon will be discussed in Section 3. Appendix 1 also contains graphs (Figure 1 and Figure 2) of the $f_{T\lambda}(m)$ and $g_{T\lambda}(m)$ functions. It is notable that for $\lambda = 1600$, we found that

$$\max_{m=-50,\ldots,50} |f_{50,1600}(m) - f_{10000,1600}(m)| = 1.033 \cdot 10^{-4}$$

and

$$\max_{m=-50,\ldots,50} |g_{50,1600}(m) - g_{10000,1600}(m)| = 2.428 \cdot 10^{-7}$$

while $f_{50,1600}(0) = 0.05607$ and $g_{50,1600}(0) = 4.4041 \cdot 10^{-5}$, and therefore we only provided one graph for $\lambda = 1600$ and $T \geq 50$. The Matlab programs used for arriving at this conclusion and used for creating Tables 1 and 2 and Figure 1 and Figure 2 can be found at `http://dl.dropbox.com/u/2159931/hp.html`.

3 Large T results

In this section, we gather some large T results that will be used later on and are also of independent interest, as they show that it is possible to formulate weighted average procedures with weights that do not depend on T that are asymptotically equivalent to the HP filter. First, we prove that the constants present in Theorem 1 achieve limit values:

Theorem 2. *For all $\lambda > 0$, $\lim_{T\to\infty} \eta_{T\lambda} = 0$, $\lim_{T\to\infty} \phi_{T\lambda} = 0$,*

$$\lim_{T\to\infty} \delta_{T\lambda} = \int_0^1 \sin(\pi r/2)^4 \cos(\pi r/2)^2 (1 + 16\lambda \sin(\pi r/2)^4)^{-1} dr = \delta_\lambda, \tag{20}$$

and

$$\lim_{T\to\infty} \xi_{T\lambda} = \frac{32\lambda}{1 - 32\lambda\delta_\lambda} = \xi_\lambda. \tag{21}$$

Part of the assertion of Theorem 2 is that $1 - 32\lambda\delta_\lambda \neq 0$. Together with Theorem 1, the above result implies

$$\sup_{T\geq 1, 1\leq t\leq T} \sum_{s=1}^{T} |w_{Tts}| < \infty. \tag{22}$$

It can also be shown that the functions $f_{T\lambda}(\cdot)$ and $g_{T\lambda}(\cdot)$ converge pointwise to limit functions:

Theorem 3. *Pointwise in* (λ, m),

$$\lim_{T\to\infty} f_{T,\lambda}(m) = f_\lambda(m) = \int_0^1 \cos(\pi r m)(1 + 16\lambda \sin(\pi r/2)^4)^{-1} dr \tag{23}$$

and

$$\lim_{T\to\infty} g_{T,\lambda}(m) = g_\lambda(m) = 2^{1/2} \int_0^1 \cos(\pi r(m-1/2)) \sin(\pi r/2)^2 \cos(\pi r/2)(1+16\lambda \sin(\pi r/2)^4)^{-1} dr. \tag{24}$$

Given these results, we can now formulate a weighted average procedure that, away from the begin and end of the sample, is asymptotically equivalent to the HP filter procedure. In order to be able to formulate our result, we replaced t/T by r in the theorem below. Also, we used a scaling factor of $E|y_t|$ to account for the possibly increasing nature of y_t (and $\hat{\tau}_{Tt}$).

Theorem 4. *Assume that* γ *is a constant in* $(0, 1/2)$*, and assume* $r \in [\gamma, 1-\gamma]$*. In addition assume that* $E|y_t|$*,* $t \geq 1$*, is nondecreasing in* t *and that for any* $r \in [\gamma, 1-\gamma]$*,* $\limsup_{T\to\infty} E|y_T|(E|y_{[rT]}|)^{-1} < \infty$*. Then*

$$\limsup_{T\to\infty} (E|y_{[rT]}|)^{-1} E|\sum_{s=1}^{T} y_s(w_{T,[rT],s} - f_\lambda([rT] - s))| = 0. \tag{25}$$

Therefore, noting that $f_\lambda(m) = f_\lambda(-m)$, away from the end points, under the conditions of the theorem the HP filter is asymptotically equivalent to the symmetric weighted average

$$\sum_{s=1}^{T} y_s f_\lambda(t-s)$$

as long as the regularity conditions hold. These regularity conditions holds for a wide class of processes. For a series $y_t = \alpha + \beta t + \gamma z_t + u_t$ for which $\beta \neq 0$, $\sup_{t\geq 1} E|u_t| < \infty$, and $E|z_t| = o(t^{1/2})$ (such as in the case where z_t is a unit root process that satisfies mild regularity conditions), we have

$$\beta r T(E|y_{[rT]}|)^{-1} \to 1 \quad \text{and} \quad E|y_T|(E|y_{[rT]}|)^{-1} \to r^{-1},$$

implying that the conditions of the theorem hold. In the case where $\beta = 0$ and z_t is a unit root process such that $t^{-1/2}E|z_t| \to c$ for some constant $c > 0$, we have

$$\gamma c T^{1/2} r^{1/2} (E|y_{[rT]}|)^{-1} \to 1 \quad \text{and} \quad E|y_T|(E|y_{[rT]}|)^{-1} \to r^{-1/2},$$

again implying that the conditions of the theorem hold. Similar results can be established if the regressors are integrated of finite order and/or polynomials of the time trend.

In conclusion, it appears that away from the end points of the sample for a wide range of data-generating processes, the HP filter is effectively a symmetric weighted average with weights $f_\lambda(t-s)$. It should be noted that $f_\lambda(\cdot)$ takes on negative values, and therefore, some observations are given negative weights, even asymptotically.

4 Weak dependence of the cyclical component

Given the fact that the HP filter is a weighted average of T observations, it follows that the cyclical component is a triangular array. Also, for every value of t, the HP filter necessarily weighs $t-1$ earlier observations and $T-t$ later observations; and therefore, it is not possible for the cyclical component of the HP filter to be strictly stationary or weakly stationary. However, it is possible to show that the cyclical component of the HP filter has weak dependence properties when the HP filter is applied to the sum of a stationary mixing process, a linear trend process, and a unit root process. Such a property then ensures that laws of large numbers and central limit theorems can hold for functions of the cyclical component. This provides a justification for the use of the HP filter, as it implies that inference based on the cyclical component can be asymptotically correct. Therefore, the results below illustrate that the finding in empirical macro that time series tend to possess time trends and/or unit roots, and the practice of using inference procedures based on the cyclical component of the HP filter, are not contradictory.

In addition, the weak dependence of the cyclical component indicates that the unit root process has essentially been absorbed into the trend component. This can be taken to imply that the HP filter is capable of removing unit roots from a data-generating process.

In this section, we consider weak dependence properties of the cyclical component when the HP filter has been applied to the sum of a stationary mixing process, a linear trend process, and a unit root process; that is, we assume that y_t satisfies the following assumption:

Assumption 1.

$$y_t = \alpha + \beta t + \gamma z_t + u_t,$$

where $z_t = \sum_{j=1}^{t} \varepsilon_j$, $\sup_{s \geq 1} \parallel \varepsilon_s \parallel_p < \infty$ *and* $\sup_{s \geq 1} \parallel u_s \parallel_p < \infty$.

We explicitly allow for the cases $u_s = 0$, $\alpha = 0$, $\beta = 0$ and/or $\varepsilon_s = 0$. Let the cyclical component of y_t after application of the HP filter to $y_t = \alpha + \beta t + u_t + z_t$ be denoted by $\hat{c}_{Tt}$, i.e. $\hat{c}_{Tt} = y_t - \hat{\tau}_{Tt} = y_t - \sum_{s=1}^{T} w_{Tts} y_s$, and define

$$\hat{c}_{Tt}^m = u_t - \sum_{s=1}^{T} w_{Tts} u_s I(|s-t| \leq m) - \sum_{j=1}^{T} \varepsilon_j (\sum_{s=j}^{T} w_{Tts} - I(j \leq t)) I(|j-t| \leq m).$$

Note that $\hat{c}_{Tt}^m$ is an approximation to $\hat{c}_{Tt}$ uses only $(v_{\max(1,t-m)}, \ldots, v_{\min(T,t+m)})$, where $v_t = (u_t, \varepsilon_t)'$. Therefore, $\hat{c}_{Tt}^m$ is an approximation to $\hat{c}_{Tt}$ that uses only information that is at most m time periods away from the current time period. Concepts of closeness to such an approximation have a long history in statistics and econometrics; see for example Ibragimov (1962), Billingsley (1968), McLeish (1975), Gallant and White (1988), Wooldridge and White (1988), Andrews (1988), and Pötscher and Prucha (1997), among others. Near epoch dependence is a simpler condition than mixing to verify, but the "base" (here, v_t) needs to satisfy a mixing condition in order for the condition to be useable for showing laws of large numbers and/or central limit theorems. Our result is the following:

Theorem 5. *Assume that y_t satisfies Assumption 1. Then for any $\gamma \in (0, 1/2)$,*

$$\sup_{T \geq 1, t \in [\gamma T, (1-\gamma)T]} \| \hat{c}_{Tt} - \hat{c}_{Tt}^m \|_p = O(m^{-1})$$

and

$$\sup_{T \geq 1, t \in [\gamma T, (1-\gamma)T]} \| \hat{c}_{Tt} \|_p < \infty. \tag{26}$$

Note that if $\varepsilon_t = 0$, we can improve the $O(m^{-1})$ to $O(m^{-2})$ in the above theorem.

In Gallant and White (1988) and Pötscher and Prucha (1997), the authors demonstrate how a complete theory of inference can be based on the near epoch dependence property as demonstrated above.

In the above theorem, we had to limit ourselves to sample values away from the begin and end of the sample. However, it is also possible to choose the γ arbitrarily small, and use an asymptotic negligibility argument for the begin and the end of the sample, and along those lines, a complete theory of inference based on the near epoch dependence property using the entire sample is feasible also. As an example, we prove the following weak law of large numbers for functions of the cyclical component:

Theorem 6. *Assume that y_t satisfies Assumption 1 for $p = 2$, and assume that $(u_t, \varepsilon_t)'$ is strong mixing. Let $f(\cdot)$ be a function that is bounded and continuous on $\mathbb{R}$. Then*

$$T^{-1} \sum_{t=1}^{T} (f(\hat{c}_{Tt}) - Ef(\hat{c}_{Tt})) \xrightarrow{p} 0.$$

5 Results for large T and large λ

While in Section 3 we found that the HP filter for large T is equivalent to a weighted average, the weight function depended on λ. When considering large values of λ, it turns out to also be possible to find an approximation to the HP filter using a weighted average that only uses λ as a bandwidth type parameter. This result is both of interest in its own right and provides a route towards a formal result on adjusting the HP filter for the frequency of the observations. Our result is based on the following theorem:

Theorem 7. *Pointwise in m for all $m \geq 0$,*

$$\lim_{\lambda\to\infty} \lambda^{1/4} f_\lambda(\lambda^{1/4} m) = f(m) = 2^{-3/2} \exp(-2^{-1/2} m)(\sin(2^{-1/2} m) + \cos(2^{-1/2} m))$$

and

$$\lim_{\lambda\to\infty} \lambda^{3/4} g_\lambda(\lambda^{1/4} m) = g(m) = 2^{-3} \exp(-2^{-1/2} m)(\cos(2^{-1/2} m) - \sin(2^{-1/2} m)).$$

Furthermore, for $m \geq 1$, $f_\lambda(m) \leq C\lambda^{1/4} m^{-2}$ for some constant $C > 0$ not depending on λ, and for all $K > 0$,

$$\lim_{\lambda\to\infty} \sup_{|m|\leq K} |\lambda^{1/4} f_\lambda(\lambda^{1/4} m) - f(m)| = 0.$$

Note that in the above result, $f(\cdot)$ and $g(\cdot)$ are not dependent on λ, and also note that $\int_{-\infty}^{\infty} f(m)dm = 1$ and $\int_{-\infty}^{\infty} g(m)dm = 0$. Using Theorem 7 as a basis, we can now establish the following result.

Theorem 8. *Assume that γ is a constant in $(0, 1/2)$, and assume $r \in [\gamma, 1 - \gamma]$. In addition assume that $E|y_t|$, $t \geq 1$, is nondecreasing in t and that for any $r \in [\gamma, 1 - \gamma]$, $\limsup_{T\to\infty} E|y_T|(E|y_{[rT]}|)^{-1} < \infty$. Then*

$$\limsup_{\lambda\to\infty} \limsup_{T\to\infty} (E|y_{[rT]}|)^{-1} E|\sum_{s=1}^{T} y_s (w_{T,[rT],s} - \lambda^{-1/4} f(\lambda^{-1/4}([rT] - s)))| = 0. \quad (27)$$

Theorem 8 suggests that for large T and λ, away from the beginning and end of our sample,

$$\hat{\tau}_{Tt} \approx \sum_{s=1}^{T} y_s \lambda^{-1/4} f(\lambda^{-1/4}(t-s)),$$

which for $\lambda = 1600$ simplifies to

$$\hat{\tau}_{Tt} \approx 0.1581 \sum_{s=1}^{T} y_s \ f(0.1581(t-s)).$$

The above theorem also implies that the negative weights remain a feature of the HP filter even when both T and λ are assumed large.

6 Adjusting the HP filter for the frequency of the observations

Hodrick and Prescott's suggestion to use a quarterly smoothing parameter of 1600 raises the question as to what smoothing parameter is appropriate for annual or monthly data. For annual data, Backus and Kehoe (1992) suggested a value of 100, and Correia, Neves, and Rebelo (1992) and Cooley and Ohanian (1991) suggested a value of 400, while Baxter and King (1999) suggested a value of 10. Ravn and Uhlig (2002) suggested that a value of 6.25 for the annual smoothing parameter corresponds to a quarterly smoothing parameter of 1600. They obtain this value as $1600 \times 4^{-4} = 6.25$. In this paragraph, we will give a different justification for rescaling the smoothing parameter using a fourth power when adjusting for the data frequency.

Assume for simplicity that we have $4T$ observations of quarterly data and apply the HP filter using a smoothing value λ_Q, implying that we have observations for T years. Then, letting $w_{Tts}(\lambda_Q)$ denote the weights corresponding to λ_Q, the HP filter trend in the tth year of data in quarter k, $t = 1, ..., T$, $k = 1, 2, 3, 4$ would equal

$$\sum_{s=1}^{4T} y_s w_{4T,4t-4+k,s}(\lambda_Q),$$

while using annual data $(1/4)\sum_{s=1}^{4} y_{4t-4+s}$ and smoothing parameter λ_A, we would use

$$\sum_{i=1}^{T}((1/4)\sum_{s=1}^{4} y_{4i-4+s})w_{Tti}(\lambda_A).$$

Note that this assumes that the quarterly data are measured on a per year basis. Alternatively of course, we can assume that our annual data are measured on a quarterly basis. The asymptotic equivalence of both procedures for

$$\lambda_A = 4^{-4}\lambda_Q$$

is then shown in the following theorem:

Theorem 9. *Assume that* γ *is a constant in* $(0, 1/2)$*, and assume* $r \in [\gamma, 1-\gamma]$*. In addition assume that* $E|y_t|$*,* $t \geq 1$*, is nondecreasing in* t *and that for any* $r \in [\gamma, 1-\gamma]$*,* $\limsup_{T\to\infty} E|y_T|(E|y_{[rT]}|)^{-1} < \infty$*. Then for* $\lambda_A = 4^{-4}\lambda_Q$ *and* $k = 1, 2, 3,$ *or* 4*,*

$$\lim_{\lambda_Q\to\infty} \limsup_{T\to\infty} (E|y_{[rT]}|)^{-1} E|\sum_{s=1}^{4T} y_s w_{4T,4t-4+k,s}(\lambda_Q) - \sum_{i=1}^{T}((1/4)\sum_{s=1}^{4} y_{4i-4+s}) w_{Tti}(\lambda_A)| = 0. \quad (28)$$

It should be noted that our asymptotic equivalence argument using $\lambda_A = 4^{-4}\lambda_Q$ only follows if we use quarterly data y_i and annual data $(1/4)\sum_{s=1}^{4} y_{4i-4+s}$. This implies that quarterly and annual data must be measured both on a quarterly or an annual basis for the equivalence to hold. Therefore, if we have quarterly data on domestic car sales, then the asymptotic equivalence between the HP filter for quarterly data using $\lambda = 1600$ and the HP filter for annual data using $\lambda = 6.25$ only holds if we measure domestic car sales on a per year or a per quarter basis for both applications of the HP filter. However, when using quarterly sales and employing the HP filter with $\lambda = 1600$ and using annual sales and $\lambda = 6.25$, no asymptotic equivalence is obtained.

Using a similar argument, going from quarterly to monthly data, we find that the smoothing parameter λ_M asymptotically corresponds to $(1/3)^{-4}\lambda_Q = 81\lambda_Q$, and for $\lambda_Q = 1600$, this then suggests $\lambda_M = 129600$. This also corresponds to the suggestion of Ravn and Uhlig (2002). We omit a formal theorem to this effect, as it will be analogous to Theorem 9. Again, for the asymptotic equivalence to hold, we need to measure both monthly and annual data on either a monthly or an annual basis. Similar statements can of course be made for adjustments to the smoothing parameter when considering different frequency adjustments.

Appendix 1: the constants of Section 2

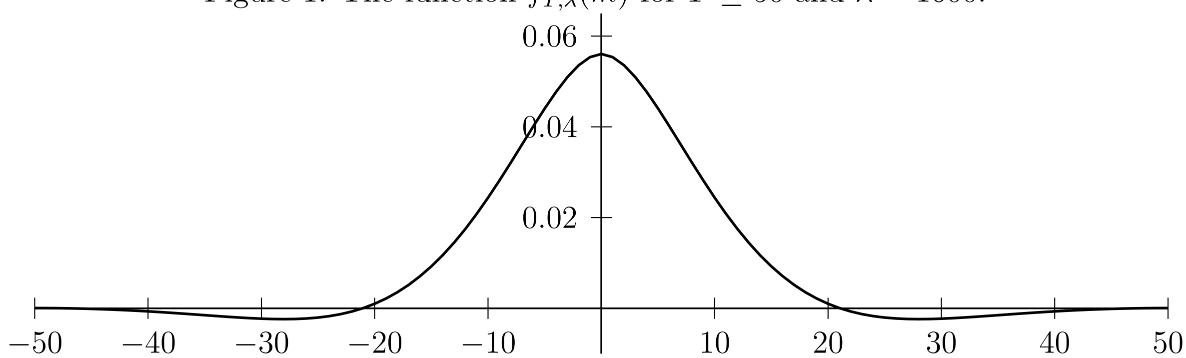
Table 1: Values of $\xi_{T\lambda}$ for various values of λ and T. The Matlab program used to calculate these values (and values for the $\delta_{T\lambda}$ and $\eta_{T\lambda}$ sequences) can be found at `https://dl.dropbox.com/u/2159931/hp.html`.

	$T = 25$	$T = 50$	$T = 100$	$T = 1000$
$\lambda = 100$	14491.2	14490.8	14490.8	14490.8
$\lambda = 400$	81643.7	81461.9	81461.8	81461.8
$\lambda = 1600$	461398	459391	459380	459380

Table 2: Values of $\phi_{T\lambda}$ for various values of λ and T. The Matlab program used to calculate these values (and values for the $\delta_{T\lambda}$ and $\eta_{T\lambda}$ sequences) can be found at `https://dl.dropbox.com/u/2159931/hp.html`.

	$T = 25$	$T = 50$	$T = 100$	$T = 1000$
$\lambda = 100$	-8.52750	0.33904	6.37970×10^{-6}	-5.86296×10^{-11}
$\lambda = 400$	4421.81	-57.7142	0.02290	-4.69732×10^{-10}
$\lambda = 1600$	33352.1	-403.167	10.3726	-3.73839×10^{-9}

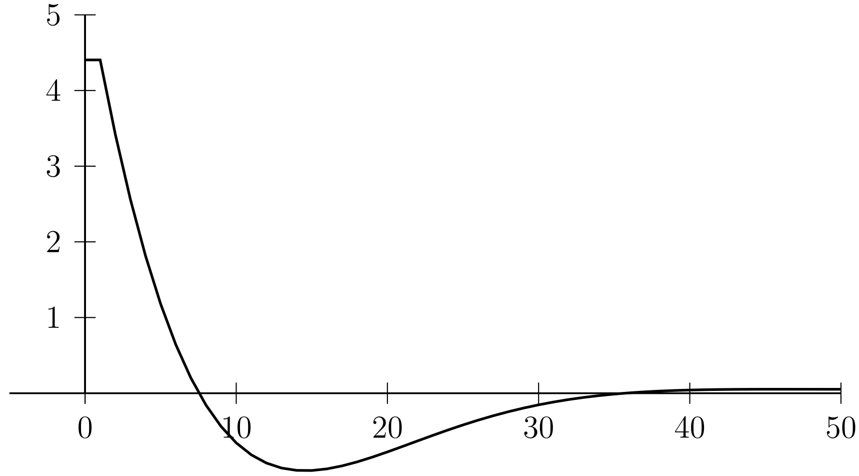
Figure 1: The function $f_{T,\lambda}(m)$ for $T \geq 50$ and $\lambda = 1600$.



Appendix 2: Mathematical proofs

Lemma 2. *Let* $p_1(t/T) = 1$ *and* $p_j(t/T) = \sqrt{2}\cos(\pi(j-1)(t-1/2)/T)$, $j = 2, \ldots, T$. *Then* $T^{-1}\sum_{t=1}^{T} p_{Tt}p_{Tt}' = I_T$.

Figure 2: The function $g_{T,\lambda}(m)$ for $T \geq 50$ and $\lambda = 1600$; y axis needs to be multiplied by 10^{-5}.



Proof. This follows from Lemma 3.4 of Eubank (1999, p. 143). □

Lemma 3. *Assume that* $h : [0,1] \to \mathbb{R}$ *is continuous on* $[0,1]$. *Then*

$$\lim_{T\to\infty} T^{-1} \sum_{j=1}^{T} h((j-1)/T) = \int_0^1 h(r)dr.$$

Proof. This follows because by continuity of $h(\cdot)$ and setting $j = rT$,

$$T^{-1} \sum_{j=1}^{T} h((j-1)/T) = T^{-1} \int_0^T h([j]/T)dj = \int_0^1 h([rT]/T)dr,$$

and because $\sup_{r\in[0,1]} |h(r)| < \infty$ by continuity and because $r - 1/T \leq [rT]/T \leq r$, the result now follows by the dominated convergence theorem. □

Lemma 4. *For* $h(r) = 1/(1 + 16\lambda \sin(\pi r/2)^4)$,

$$T^{-1} \sum_{j=2}^{T} \cos(\pi(j-1)m/T)h((j-1)/T)$$

$$= -(2T)^{-1}h(1/T)$$
$$-(2T)^{-1}(-1)^m h((T-1)/T)$$
$$-T^{-1}(2\sin(\pi m/(2T))^{-2}(h(2/T)-h(1/T))\cos(\pi m/T)$$
$$+T^{-1}(2\sin(\pi m/(2T))^{-2}(h(1-1/T)-h(1-2/T))\cos(\pi m(1-1/T))$$
$$+T^{-1}(2\sin(\pi m/(2T)))^{-3}(h(3/T)-2h(2/T)+h(1/T))\sin(\pi(3/2)m/T)$$
$$-T^{-1}(2\sin(\pi m/(2T)))^{-3}(h(1-1/T)-2h(1-2/T)+h(1-3/T))\sin(\pi(T-3/2)m/T)$$
$$+T^{-1}(2\sin(\pi m/(2T)))^{-3}\sum_{j=2}^{T-3}\Delta^3 h((j+2)/T)\sin(\pi(j+1/2)m/T).$$

Proof. The proof of this result is somewhat tedious and can be found at `https://dl.dropbox.com/u/2159931/hp.html`. □

Proof of Lemma 1: By Lemma 2, $T^{-1}\sum_{t=1}^{T} p_{Tt}p'_{Tt} = I_T$, and therefore it only remains to prove the second part of Lemma 1. To show this, note that for $j \in \{1, 2, ..., T\}$, remembering that $p_1(t/T) = 1$ for $t \in \mathbb{Z}$,

$$p_j(t/T) - p_j((t-1)/T) = \sqrt{2}\cos(\pi(j-1)(t-1/2)/T) - \sqrt{2}\cos(\pi(j-1)(t-3/2)/T),$$

and because

$$\cos(\alpha+\beta) - \cos(\alpha-\beta) = -2\sin(\alpha)\sin(\beta)$$

and setting $\alpha+\beta = \pi(j-1)(t-1/2)/T$ and $\alpha-\beta = \pi(j-1)(t-3/2)/T$, we get

$$p_j(t/T) - p_j((t-1)/T) = -2\sqrt{2}\sin(\pi(j-1)(t-1)/T))\sin(\pi(j-1)/(2T)).$$

Therefore,

$$p_j((t+1)/T) - 2p_j(t/T) + p_j((t-1)/T)$$
$$= (p_j((t+1)/T) - p_j(t/T)) - (p_j(t/T) - p_j((t-1)/T))$$
$$= 2\sqrt{2}\sin(\pi(j-1)/(2T))(\sin(\pi(j-1)(t-1)/T) - \sin(\pi(j-1)t/T)).$$

Now because

$$\sin(\alpha+\beta)-\sin(\alpha-\beta)=2\sin(\beta)\cos(\alpha)$$

and setting $\alpha+\beta=\pi(j-1)(t-1)/T$ and $\alpha-\beta=\pi(j-1)t/T)$, we have

$$\sin(\pi(j-1)(t-1)/T)-\sin(\pi(j-1)t/T)$$
$$=-2\sin(\pi(j-1)/(2T))\cos(\pi(j-1)(t-1/2)/T).$$

Therefore, for $j\in\{1,...,T\}$ and $t\in\mathbb{Z}$,

$$p_j((t+1)/T)-2p_j(t/T)+p_j((t-1)/T)=-4\sqrt{2}\sin(\pi(j-1)/(2T))^2\cos(\pi(j-1)(t-1/2)/T),$$

implying that, for $j,k\in\{1,2,...,T\}$,

$$[\sum_{t=2}^{T-1}(p_{T,t+1}-2p_{Tt}+p_{T,t-1})(p_{T,t+1}-2p_{Tt}+p_{T,t-1})']_{jk}$$

$$=32\sin(\pi(j-1)/(2T))^2\sin(\pi(k-1)/(2T))^2\sum_{t=2}^{T-1}\cos(\pi(j-1)(t-1/2)/T)\cos(\pi(k-1)(t-1/2)/T).$$

Also, for $j,k\in\{1,2,\ldots,T\}$, by Lemma 2,

$$\sum_{t=2}^{T-1}\cos(\pi(j-1)(t-1/2)/T)\cos(\pi(k-1)(t-1/2)/T)$$
$$=(1/2)TI(j=k)-\cos(\pi(j-1)/(2T))\cos(\pi(k-1)/(2T))$$
$$-\cos(\pi(j-1)(T-1/2)/T)\cos(\pi(k-1)(T-1/2)/T),$$

implying that

$$T^{-1}\sum_{t=2}^{T-1}(p_{T,t+1}-2p_{Tt}+p_{T,t-1})(p_{T,t+1}-2p_{Tt}+p_{T,t-1})'$$
$$=\text{diag}(\{16\sin(\pi(j-1)/(2T))^4,j=1,\ldots,T\})-32T^{-1}q_{T1}q_{T1}'-32T^{-1}q_{T2}q_{T2}'$$
$$=D_T-32T^{-1}q_{T1}q_{T1}'-32T^{-1}q_{T2}q_{T2}',$$

where $q_{T1j}=\sin(\pi(j-1)/(2T))^2\cos(\pi(j-1)(1/2)/T)$ and $q_{T2j}=\sin(\pi(j-1)/(2T))^2\cos(\pi(j-1)(T-1/2)/T)$. □

Proof of Theorem 1: It follows from Lemma 1 that

$$\hat{\theta}_T = [T^{-1}\sum_{t=1}^{T} p_{Tt}p'_{Tt} + \lambda T^{-1}\sum_{t=2}^{T-1}(p_{T,t+1} - 2p_{Tt} + p_{T,t-1})(p_{T,t+1} - 2p_{Tt} + p_{T,t-1})']^{-1}T^{-1}\sum_{t=1}^{T} y_t p_{Tt}$$

$$= (I_T + \lambda D_T + H_T(I_T + \lambda D_T)]^{-1}\ T^{-1}\sum_{t=1}^{T} y_t p_{Tt}$$

$$= (I_T + \lambda D_T)^{-1}(I_T + H_T)^{-1}T^{-1}\sum_{t=1}^{T} y_t p_{Tt}$$

where I_T is the identity matrix of dimension $(T \times T)$,

$$D_T = \text{diag}(\{16\sin(\pi(j-1)/(2T))^4, j = 1, \ldots, T\})$$

and

$$H_T = (-32\lambda T^{-1}q_{T1}q'_{T1} - 32\lambda T^{-1}q_{T2}q'_{T2})(I_T + \lambda D_T)^{-1}.$$

Now according to Miller (1981), for a matrix H_T of rank 0, 1 or 2, we have

$$(I_T + H_T)^{-1} = I_T - (a_T + b_T)^{-1}(a_T H_T - H_T^2)$$

for $a_T = 1 + \text{tr}(H_T)$ and $2b_T = (\text{tr}(H_T))^2 - \text{tr}(H_T^2)$. Therefore,

$$(I_T + \lambda D_T + H_T(I_T + \lambda D_T))^{-1} = (I_T + \lambda D_T)^{-1}(I_T - (a_T + b_T)^{-1}(a_T H_T - H_T^2))$$

$$= (I_T + \lambda D_T)^{-1}$$
$$+32\lambda a_T(a_T + b_T)^{-1}(I_T + \lambda D_T)^{-1}(T^{-1}q_{T1}q'_{T1})(I_T + \lambda D_T)^{-1}$$
$$+32\lambda a_T(a_T + b_T)^{-1}(I_T + \lambda D_T)^{-1}(T^{-1}q_{T2}q'_{T2})(I_T + \lambda D_T)^{-1}$$
$$+32^2\lambda^2(a_T + b_T)^{-1}(I_T + \lambda D_T)^{-1}(T^{-1}q_{T1}q'_{T1})(I_T + \lambda D_T)^{-1}(T^{-1}q_{T1}q'_{T1})(I_T + \lambda D_T)^{-1}$$
$$+32^2\lambda^2(a_T + b_T)^{-1}(I_T + \lambda D_T)^{-1}(T^{-1}q_{T1}q'_{T1})(I_T + \lambda D_T)^{-1}(T^{-1}q_{T2}q'_{T2})(I_T + \lambda D_T)^{-1}$$
$$+32^2\lambda^2(a_T + b_T)^{-1}(I_T + \lambda D_T)^{-1}(T^{-1}q_{T2}q'_{T2})(I_T + \lambda D_T)^{-1}(T^{-1}q_{T1}q'_{T1})(I_T + \lambda D_T)^{-1}$$
$$+32^2\lambda^2(a_T + b_T)^{-1}(I_T + \lambda D_T)^{-1}(T^{-1}q_{T2}q'_{T2})(I_T + \lambda D_T)^{-1}(T^{-1}q_{T2}q'_{T2})(I_T + \lambda D_T)^{-1}$$
$$= \sum_{i=1}^{7} M_{Ti},$$

say, and we can write

$$\hat{\tau}_{Tt} = p'_{Tt}\hat{\theta}_T = p'_{Tt}\sum_{i=1}^{7} M_{Ti}T^{-1}\sum_{s=1}^{T} p_{Ts}y_s$$

$$= \sum_{s=1}^{T} y_s \sum_{i=1}^{7} T^{-1}p'_{Tt}M_{Ti}p_{Ts} = \sum_{s=1}^{T} y_s \sum_{i=1}^{7} w^i_{Tts}$$

where

$$w^1_{Tts} = T^{-1}p'_{Tt}M_{T1}p_{Ts} = T^{-1}p'_{Tt}(I_T + \lambda D_T)^{-1}p_{Ts},$$

$$w^2_{Tts} = T^{-1}p'_{Tt}M_{T2}p_{Ts}$$
$$= 32\lambda a_T(a_T + b_T)^{-1}(T^{-1}p'_{Tt}(I_T + \lambda D_T)^{-1}q_{T1})(T^{-1}q'_{T1}(I_T + \lambda D_T)^{-1}p_{Ts}),$$

$$w^3_{Tts} = T^{-1}p'_{Tt}M_{T3}p_{Ts}$$
$$= 32\lambda a_T(a_T + b_T)^{-1}(T^{-1}p'_{Tt}(I_T + \lambda D_T)^{-1}q_{T2})(T^{-1}q'_{T2}(I_T + \lambda D_T)^{-1}p_{Ts}),$$

$$w^4_{Tts} = T^{-1}p'_{Tt}M_{T4}p_{Ts}$$
$$= 32^2\lambda^2(a_T+b_T)^{-1}(T^{-1}p_{Tt}(I_T+\lambda D_T)^{-1}q_{T1})(T^{-1}q'_{T1}(I_T+\lambda D_T)^{-1}q_{T1})(T^{-1}q'_{T1}(I_T+\lambda D_T)^{-1}p_{Ts}),$$

$$w^5_{Tts} = T^{-1}p'_{Tt}M_{T5}p_{Ts}$$
$$= 32^2\lambda^2(a_T+b_T)^{-1}(T^{-1}p_{Tt}(I_T+\lambda D_T)^{-1}q_{T1})(T^{-1}q'_{T1}(I_T+\lambda D_T)^{-1}q_{T2})(T^{-1}q'_{T2}(I_T+\lambda D_T)^{-1}p_{Ts}),$$

$$w^6_{Tts} = T^{-1}p'_{Tt}M_{T6}p_{Ts}$$
$$= 32^2\lambda^2(a_T+b_T)^{-1}(T^{-1}p_{Tt}(I_T+\lambda D_T)^{-1}q_{T2})(T^{-1}q'_{T2}(I_T+\lambda D_T)^{-1}q_{T1})(T^{-1}q'_{T1}(I_T+\lambda D_T)^{-1}p_{Ts}),$$

and

$$w^7_{Tts} = T^{-1}p'_{Tt}M_{T7}p_{Ts}$$

$$= 32^2\lambda^2(a_T+b_T)^{-1}(T^{-1}p_{Tt}(I_T+\lambda D_T)^{-1}q_{T2})(T^{-1}q'_{T2}(I_T+\lambda D_T)^{-1}q_{T2})(T^{-1}q'_{T2}(I_T+\lambda D_T)^{-1}p_{Ts}).$$

Simplification of w^1_{Tts} can be achieved by noting that

$$\cos(\alpha)\cos(\beta) = (1/2)\cos(\alpha + \beta) + (1/2)\cos(\alpha - \beta),$$

which implies that for $j \geq 2$,

$$p_j(t/T)p_j(s/T) = 2\cos(\pi(j-1)(t-1/2)/T)\cos(\pi(j-1)(s-1/2)/T)$$

$$= \cos(\pi(j-1)(t+s-1)/T) + \cos(\pi(j-1)(t-s)/T),$$

and therefore, remembering that $p_1(t/T) = 1$ for $t \in \{1, 2, \ldots, T\}$,

$$w^1_{Tts} = T^{-1} p'_{Tt}(I_T + \lambda D_T)^{-1} p_{Ts}$$

$$= T^{-1} \sum_{j=2}^{T} \cos(\pi(j-1)(t+s-1)/T)(1 + 16\lambda \sin(\pi(j-1)/(2T))^4)^{-1}$$

$$+ T^{-1} \sum_{j=2}^{T} \cos(\pi(j-1)(t-s)/T)(1 + 16\lambda \sin(\pi(j-1)/(2T))^4)^{-1} + T^{-1}$$

$$= f_{T\lambda}(t+s-1) - (2T)^{-1} - (2T)^{-1}(-1)^{t+s-1}(1+16\lambda)^{-1}$$

$$+ f_{T\lambda}(t-s) - (2T)^{-1} - (2T)^{-1}(-1)^{t-s}(1+16\lambda)^{-1} + T^{-1}$$

$$= f_{T\lambda}(t+s-1) + f_{T\lambda}(t-s)$$

because $(-1)^{t+s-1} + (-1)^{t-s} = (-1)^t((-1)^{s-1} + (-1)^s) = 0$. Note that by symmetry of D_T and I_T, $T^{-1} q'_{T1}(I_T + \lambda D_T)^{-1} q_{T2} = T^{-1} q'_{T2}(I_T + \lambda D_T)^{-1} q_{T1} = \eta_T$. Also, since for $j \in \{1, \ldots, T\}$

$$q_{T2j} = \cos(\pi(j-1)) q_{T1j}$$

which is trivial for $j = 1$ and for $j \geq 2$ follows because

$$q_{T2j} = \sin(\pi(j-1)/(2T))^2 \cos(\pi(j-1)) \cos(\pi(j-1)/(2T)) = \cos(\pi(j-1)) q_{T1j},$$

and

$$p_j((T-t)/T) = \sqrt{2} \cos(\pi(j-1)((T-t-1/2)/T))$$

$$= \sqrt{2} \cos(\pi(j-1)) \cos(\pi(j-1)(t+1/2)/(2T)) = \cos(\pi(j-1)) p_j((t+1)/T),$$

it follows that

$$T^{-1} p'_{Tt}(I_T + \lambda D_T)^{-1} q_{T2} = \sum_{j=1}^{T} T^{-1} p_j(t/T) q_{T2j} (1 + \lambda d_{jj})^{-1}$$

$$= T^{-1} p'_{T-t+1}(I_T + \lambda D_T)^{-1} q_{T1} = g_T(T-t+1),$$

and

$$T^{-1} q'_{T2}(I_T + \lambda D_T)^{-1} q_{T2} = T^{-1} q'_{T1}(I_T + \lambda D_T)^{-1} q_{T1} = \delta_T.$$

Therefore,

$$w_{Tts}^2 = 32\lambda a_T(a_T+b_T)^{-1}g_T(t)g_T(s),$$

$$w_{Tts}^3 = 32\lambda a_T(a_T+b_T)^{-1}g_T(T-t+1)g_T(T-s+1),$$

$$w_{Tts}^4 = 32^2\lambda^2(a_T+b_T)^{-1}\delta_T g_T(t)g_T(s),$$

$$w_{Tts}^5 = 32^2\lambda^2(a_T+b_T)^{-1}\eta_T g_T(t)g_T(T-s+1),$$

$$w_{Tts}^6 = 32^2\lambda^2(a_T+b_T)^{-1}\eta_T g_T(T-t+1)g_T(s),$$

$$w_{Tts}^7 = 32^2\lambda^2(a_T+b_T)^{-1}\delta_T g_T(T-t+1)g_T(T-s+1).$$

Therefore, the assertion of Equation (17) now follows and

$$\xi_{T\lambda} = 32\lambda a_T(a_T+b_T)^{-1} + 32^2\lambda^2(a_T+b_T)^{-1}\delta_T,$$

$$\phi_{T\lambda} = \alpha_{T3} = 32^2\lambda^2(a_T+b_T)^{-1}\eta_T,$$

and the formulas for $\xi_{T\lambda}$ and $\phi_{T\lambda}$ given prior to Theorem 1 now follow by noting that

$$a_T = 1 + \mathrm{tr}(H_T) = 1 + \mathrm{tr}((-32\lambda T^{-1}q_{T1}q_{T1}' - 32\lambda T^{-1}q_{T2}q_{T2}')(I_T+\lambda D_T)^{-1})$$

$$= 1 - 32\lambda T^{-1}q_{T1}'(I_T+\lambda D_T)^{-1}q_{T1} - 32\lambda T^{-1}q_{T2}'(I_T+\lambda D_T)^{-1}q_{T2}$$

$$= 1 - 64\lambda\delta_T$$

and

$$b_T = (1/2)(\mathrm{tr}(H_T))^2 - (1/2)\mathrm{tr}(H_T^2)$$

$$= (1/2)(-64\lambda\delta_T)^2 - (1/2)32^2\lambda^2(2\delta_T^2 + 2\eta_T^2)$$

$$= 32^2\lambda^2(\delta_T^2 - \eta_T^2),$$

which completes the proof of the representation for w_{Tts}. By Lemma 4, it follows that for $h(r) = 1/(1+16\lambda\sin(\pi r/2)^4)$,

$$f_{T\lambda}(m) = -(2T)^{-1}(h(1/T) - h(0))$$

$$-(2T)^{-1}(-1)^m(h((T-1)/T) - h(1))$$

$$-T^{-1}(2\sin(\pi m/(2T))^{-2}(h(2/T) - h(1/T))\cos(\pi m/T)$$

$$+T^{-1}(2\sin(\pi m/(2T))^{-2}(h(1-1/T) - h(1-2/T))\cos(\pi m(1-1/T))$$

$$+T^{-1}(2\sin(\pi m/(2T)))^{-3}(h(3/T) - 2h(2/T) + h(1/T))\sin(\pi(3/2)m/T)$$

$$-T^{-1}(2\sin(\pi m/(2T)))^{-3}(h(1-1/T)-2h(1-2/T)+h(1-3/T))\sin(\pi(T-3/2)m/T)$$

$$+T^{-1}(2\sin(\pi m/(2T)))^{-3}\sum_{j=2}^{T-3}\Delta^3 h((j+2)/T)\sin(\pi(j+1/2)m/T).$$

Noting that $h'(0)=h'(1)=0$ and that $\sup_{x\in[0,1]}|h'''(x)|<\infty$, it follows that the first and second terms are bounded in absolute value by

$$(1/2)T^{-3}\sup_{x\in[0,1]}|h''(x)|.$$

Similarly, the third and fourth terms are bounded by a multiple of

$$T^{-3}(2\sin(\pi m/(2T))^{-2},$$

and the fifth, sixth and seventh terms by a multiple of

$$T^{-3}(2\sin(\pi m/(2T))^{-3}.$$

By noting that

$$\sup_{r\in[0,1]}|r/\sin(\pi r/2)|=1$$

and that $m\le T$, it now follows that $|f_{T\lambda}(m)|\le Cm^{-3}$. The proof of the bound for $g_{T\lambda}(m)$ is analogous. □

Proof of Theorem 2: By Lemma 3,

$$\delta_{T\lambda}=T^{-1}q_{T1}'(I_T+\lambda D_T)^{-1}q_{T1}$$

$$=T^{-1}\sum_{j=1}^{T}\sin(\pi(j-1)/(2T))^4\cos(\pi(j-1)/(2T))^2(1+16\lambda\sin(\pi(j-1)/(2T))^4)^{-1}$$

$$\to\int_0^1\sin(\pi r/2)^4\cos(\pi r/2)^2(1+16\lambda\sin(\pi r/2)^4)^{-1}dr=\delta_\lambda$$

and

$$\eta_{T\lambda}=T^{-1}q_{T1}'(I_T+\lambda D_T)^{-1}q_{T2}$$

$$= T^{-1}\sum_{j=1}^{T} q_{T1j}q_{T2j}(1+16\lambda\sin(\pi(j-1)/(2T))^4)^{-1},$$

and because $q_{T2j} = \cos(\pi(j-1))q_{T1j}$,

$$\eta_{T\lambda} = T^{-1}\sum_{j=1}^{T}\cos(\pi(j-1))q_{T1j}^2(1+16\lambda\sin(\pi(j-1)/(2T))^4)^{-1}$$

$$= T^{-1}\sum_{j=1,j\text{ odd}}^{T} q_{T1j}^2(1+16\lambda\sin(\pi(j-1)/(2T))^4)^{-1}$$

$$-T^{-1}\sum_{j=1,j\text{ even}}^{T} q_{T1j}^2(1+16\lambda\sin(\pi(j-1)/(2T))^4)^{-1}$$

$$= T^{-1}\sum_{j=0}^{[(T-1)/2]}\sin(\pi j/T)^4\cos(\pi j/T)^2(1+16\lambda\sin(\pi j/T)^4)^{-1}$$

$$-T^{-1}\sum_{j=1}^{[T/2]}\sin(\pi(2j-1)/(2T))^4\cos(\pi(2j-1)/(2T))^2(1+16\lambda\sin(\pi(2j-1)/(2T))^4)^{-1},$$

and by an argument similar to that of Lemma 3, it now follows that both terms converge to the same number, implying that $\lim_{T\to\infty}\eta_{T\lambda} = 0$. Note that $\sin(\pi r/2)^4\lambda(1+16\lambda\sin(\pi r/2)^4)^{-1}$ is increasing in λ for $\lambda \geq 0$ for any $r \neq 0$, and therefore bounded by 1/16, it follows that

$$\lambda\delta_\lambda \leq (1/16)\int_0^1\cos(\pi r/2)^2 dr = 1/32.$$

Therefore,

$$\xi_{T\lambda} = 32\lambda(1-64\lambda\delta_{T\lambda})(1-64\lambda\delta_{T\lambda}+32^2\lambda^2(\delta_{T\lambda}^2-\eta_{T\lambda}^2))^{-1}+32^2\lambda^2(1-64\lambda\delta_{T\lambda}+32^2\lambda^2(\delta_{T\lambda}^2-\eta_{T\lambda}^2))^{-1}\delta_{T\lambda}$$

$$\to 32\lambda(1-64\lambda\delta_\lambda)(1-64\lambda\delta_\lambda+32^2\lambda^2\delta_\lambda^2)^{-1}+32^2\lambda^2(1-64\lambda\delta_\lambda+32^2\lambda^2\delta_\lambda^2)^{-1}\delta_\lambda$$

$$= \frac{32\lambda(1-64\lambda\delta_\lambda)+32^2\lambda^2\delta_\lambda}{(1-64\lambda\delta_\lambda+32^2\lambda^2\delta_\lambda^2)} = \frac{32\lambda}{1-32\lambda\delta_\lambda},$$

and

$$\phi_{T\lambda} = 32^2\lambda^2(1-64\lambda\delta_{T\lambda}+32^2\lambda^2(\delta_{T\lambda}^2-\eta_{T\lambda}^2))^{-1}\eta_{T\lambda}$$

$$\to 32^2\lambda^2(1-32\lambda\delta_\lambda)^{-2}\times 0 = 0.$$

□

Proof of Theorem 3: This is a direct consequence of Lemma 3. □

Proof of Theorem 4: First note that, because $E|y_t|$ is nondecreasing in t,

$$E|(E|y_t|)^{-1}\sum_{s=1}^{T} y_s(w_{Tts}-f_\lambda(t-s))|$$

$$\leq (E|y_t|)^{-1}\sum_{s=1}^{T} E|y_s||w_{Tts}-f_\lambda(t-s)| \leq \sum_{s=1}^{T}|w_{Tts}-f_\lambda(t-s)|$$

$$\leq \sum_{s=1}^{T}|f_{T\lambda}(t-s)-f_\lambda(t-s))| + \sum_{s=1}^{T}\sum_{j=2}^{8}|w_{Tts}^{j}|.$$

$$\leq 2\sum_{m=0}^{\infty}|f_{T\lambda}(m)-f_\lambda(m)| + \sum_{s=1}^{T}\sum_{j=2}^{8}|w_{Tts}^{j}|.$$

The second term, by the result of Equation (18), is $O(T^{-1})$. By Theorem 1, $|f_{T\lambda}(m)| \leq C(|m|+1)^{-3}$ for a constant C not depending on T, and therefore the result now follows by the dominated convergence theorem and because $|f_\lambda(m)| \leq C(|m|+1)^{-2}$ by ... and because for all $m \geq 0$, $f_{T\lambda}(m) \to f_\lambda(m)$ by ...

Proof of Theorem 4: First note that

$$E|(E|y_{[rT]}|)^{-1}\sum_{s=1}^{T} y_s(w_{Tts}-f_\lambda(t-s))|$$

$$\leq (\limsup_{T\to\infty}(E|y_T|E|(E|y_{[rT]}|)^{-1}))\sum_{s=1}^{T}|w_{Tts}-f_\lambda(t-s)|$$

$$\leq \sum_{s=1}^{T} |f_{T\lambda}(t-s) - f_\lambda(t-s))| + \sum_{s=1}^{T}\sum_{j=2}^{8} |w_{Tts}^j|.$$

The first term now vanishes by the dominated convergence theorem and because for $s \neq t$,

$$|f_{T\lambda}(t-s) - f_\lambda(t-s)| \leq C|t-s|^{-3}.$$

Noting that the second term, by the result of Equation (18), is $O(T^{-1})$ now completes the argument. □

Proof of Theorem 5: Note that, by the properties $\sum_{s=1}^{T} w_{Tts} = 1$ and $\sum_{s=1}^{T} w_{Tts}s = t$ that were noted in Section 2,

$$\hat{c}_{Tt} = y_t - \sum_{s=1}^{T} w_{Tts}(\alpha + \beta s + u_s + z_s)$$

$$= u_t - \sum_{s=1}^{T} w_{Tts}u_s + z_t - \sum_{s=1}^{T} w_{Tts}z_s$$

$$= u_t - \sum_{s=1}^{T} w_{Tts}u_s + \sum_{j=1}^{t} \varepsilon_j - \sum_{s=1}^{T}\sum_{j=1}^{s} \varepsilon_j w_{Tts}$$

$$= u_t - \sum_{s=1}^{T} w_{Tts}u_s - \sum_{j=1}^{T} \varepsilon_j (\sum_{s=j}^{T} w_{Tts} - I(j \leq t)).$$

Since $\hat{c}_{Tt} = \hat{c}_{Tt}^m$ for $m > T$, it therefore suffices to show that for every $m \geq 0$,

$$\sup_{T:T\geq m} \| \sum_{s=1}^{T} w_{Tts}u_s I(|t-s| \geq m) \|_p \leq C(m+1)^{-2}$$

and

$$\sup_{T:T\geq m} \| \sum_{j=1}^{T} \varepsilon_j (\sum_{s=j}^{T} w_{Tts} - I(j \leq t)) I(|j-t| \geq m) \|_p \leq C(m+1)^{-1}.$$

To see the first result, note that by the result of Equation (18) and because $m \leq T$,

$$\sup_{T:T\geq m} \| \sum_{s=1}^{T} w_{Tts}u_s I(|t-s| \geq m) \|_p$$

$$\leq \sup_{T:T\geq m} \sum_{j=2}^{p} \sum_{s=1}^{T} |w_{Tts}^{j}| \sup_{s\geq 1} \parallel u_s \parallel_p + \sup_{T:T\geq m} \sum_{s=1}^{T} |f_{T\lambda}(t-s)| I(|t-s| \geq m) \sup_{s\geq 1} \parallel u_s \parallel_p$$

$$\leq C_1 m^{-2} + C_2 \sum_{k=m}^{\infty} k^{-3} \leq C m^{-2}.$$

For the second result, note that

$$z_t - \sum_{s=1}^{T} z_s w_{Tts}$$

$$= \sum_{j=1}^{t} \varepsilon_j - \sum_{s=1}^{T} \sum_{j=1}^{s} \varepsilon_j w_{Tts} = -\sum_{j=1}^{T} \varepsilon_j (\sum_{s=j}^{T} w_{Tts} - I(j \leq t)),$$

and therefore the second result follows because

$$\parallel \sum_{j=1}^{T} \varepsilon_j (\sum_{s=j}^{T} w_{Tts} - I(j \leq t)) I(|j-t| \geq m) \parallel_p$$

$$\leq \sum_{j=1}^{t} I(|j-t| \geq m) |\sum_{s=j}^{T} w_{Tts} - 1| \sup_{j\geq 1} \parallel \varepsilon_j \parallel_p$$

$$+ \sum_{j=t+1}^{T} I(|j-t| \geq m) |\sum_{s=j}^{T} w_{Tts}| \sup_{j\geq 1} \parallel \varepsilon_j \parallel_p$$

$$\leq \sum_{j=1}^{t-m} I(|j-t| \geq m) |\sum_{j=2}^{8} \sum_{s=1}^{j-1} w_{Tts}^{j}| \sup_{j\geq 1} \parallel \varepsilon_j \parallel_p + \sum_{j=t+m}^{T} |\sum_{j=2}^{8} \sum_{s=j}^{T} w_{Tts}^{j}| \sup_{j\geq 1} \parallel \varepsilon_j \parallel_p$$

$$+ \sum_{j=1}^{t-m} |\sum_{s=1}^{j-1} f_{T\lambda}(t-s)| \sup_{j\geq 1} \parallel \varepsilon_j \parallel_p + \sum_{j=t+m}^{T} |\sum_{s=j}^{T} f_{T\lambda}(t-s)| \sup_{j\geq 1} \parallel \varepsilon_j \parallel_p .$$

By the result of Equation (18, the first two expressions are bounded by a multiple of T^{-1}, and since $T \geq m$, are therefore also bounded by a multiple of $(m+1)^{-1}$. The third summation, divided by $\sup_{j\geq 1} \parallel \varepsilon_j \parallel_p$, is bounded by

$$\sum_{j=1}^{t-m} \sum_{s=1}^{j-1} C(t-s)^{-3} = \sum_{s=1}^{t-m-1} \sum_{j=s+1}^{t-m} C(t-s)^{-3}$$

$$\leq \sum_{s=1}^{t-m-1} C(t-s)^{-2} \leq C \sum_{k=m+1}^{\infty} k^{-2} = O((m+1)^{-1}).$$

For the fourth term, a similar argument holds. Therefore, the conclusion of the theorem now follows. □ □

Proof of Theorem 6: Write

$$T^{-1}\sum_{t=1}^{T}(f(\hat{c}_{Tt}) - Ef(\hat{c}_{Tt}))$$

$$= T^{-1} \sum_{t\in\{1,\ldots,T\},t\notin[\gamma T,(1-\gamma)T]} (f(\hat{c}_{Tt}) - Ef(\hat{c}_{Tt})),$$

$$+T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} (f(\hat{c}_{Tt}) - Ef(\hat{c}_{Tt}))$$

and note that the first term is bounded in absolute value by

$$4\gamma \sup_{x\in\mathbb{R}} |f(x)|,$$

and because γ can be chosen arbitrarily small, it therefore suffices to show a weak law of large numbers for the second term. To show this, first note that for all $K > 0$ and $\eta > 0$,

$$\sup_{T\geq 1,t\in[\gamma T,(1-\gamma)T]} \| f(\hat{c}_{Tt}) - E(f(\hat{c}_{Tt})|v_t,\ldots,v_{t-m}) \|_2$$

$$\leq \sup_{T\geq 1,t\in[\gamma T,(1-\gamma)T]} \| f(\hat{c}_{Tt}) - E(f(\hat{c}_{Tt})|v_t,\ldots,v_{t-m})I(|\hat{c}_{Tt}| > K)| \|_2$$

$$+ \sup_{T\geq 1,t\in[\gamma T,(1-\gamma)T]} \| f(\hat{c}_{Tt}) - E(f(\hat{c}_{Tt})|v_t,\ldots,v_{t-m})I(|\hat{c}_{Tt}| \leq K)I(|\hat{c}_{Tt} - \hat{c}^m_{Tt}| > \eta)| \|_2$$

$$+ \sup_{T\geq 1,t\in[\gamma T,(1-\gamma)T]} \| f(\hat{c}_{Tt}) - E(f(\hat{c}_{Tt})|v_t,\ldots,v_{t-m})I(|\hat{c}_{Tt}| \leq K)I(|\hat{c}_{Tt} - \hat{c}^m_{Tt}| \leq \eta)| \|_2 .$$

Because $f(\cdot)$ is bounded in absolute value,

$$\sup_{T\geq 1,t\in[\gamma T,(1-\gamma)T]} \| f(\hat{c}_{Tt}) - E(f(\hat{c}_{Tt})|v_t,\ldots,v_{t-m})I(|\hat{c}_{Tt}| > K)| \|_2^2$$

$$\leq \sup_{x\in\mathbb{R}} |f(x)| \sup_{T\geq 1,t\in[\gamma T,(1-\gamma)T]} P(|\hat{c}_{Tt}| > K)$$

$$\leq \sup_{x\in\mathbb{R}} |f(x)| K^{-2} \sup_{T\geq 1, t\in[\gamma T,(1-\gamma)T]} E|\hat{c}_{Tt}|^2,$$

while similarly,

$$\sup_{T\geq 1, t\in[\gamma T,(1-\gamma)T]} \parallel f(\hat{c}_{Tt}) - E(f(\hat{c}_{Tt})|v_t,\dots,v_{t-m}) I(|\hat{c}_{Tt}| \leq K) I(|\hat{c}_{Tt} - \hat{c}^m_{Tt}| > \eta)| \parallel_2^2$$

$$\leq \sup_{x\in\mathbb{R}} |f(x)| \eta^{-2} \sup_{T\geq 1, t\in[\gamma T,(1-\gamma)T]} E|\hat{c}_{Tt} - \hat{c}^m_{Tt}|^2,$$

and

$$\sup_{T\geq 1, t\in[\gamma T,(1-\gamma)T]} \parallel f(\hat{c}_{Tt}) - E(f(\hat{c}_{Tt})|v_t,\dots,v_{t-m}) I(|\hat{c}_{Tt}| \leq K) I(|\hat{c}_{Tt} - \hat{c}^m_{Tt}| \leq \eta)| \parallel_2$$

$$\leq \sup_{|x|\leq K} \sup_{x'\in\mathbb{R}:|x-x'|\leq\eta} |f(x) - f(x')|.$$

Therefore, by making first m approach infinity, the making η approach 0, and then K approach infinity, it now follows that

$$\lim_{m\to\infty} \sup_{T\geq 1, t\in[\gamma T,(1-\gamma)T]} \parallel f(\hat{c}_{Tt}) - E(f(\hat{c}_{Tt})|v_t,\dots,v_{t-m}) \parallel_2 = 0.$$

Therefore, $f(\hat{c}_{Tt})$ is near epoch dependent on v_t for $t \in [\gamma T, (1-\gamma)T]$, and it therefore is a bounded L_1-mixingale as defined in Andrews (1988). Therefore, Andrews' weak law of large numbers applies, and the proof of the theorem is complete. □

Proof of Theorem 7:] Since $f_\lambda(m) = \int_0^1 \cos(\pi r m) h_\lambda(r) dr$ where $h_\lambda(r) = l(2\lambda^{1/4}\sin(\pi r/2))$ for $l(x) = (1+x^4)^{-1}$, it follows by partial integration that for any integer m

$$f_\lambda(m) = (\pi m)^{-1} \int_0^1 h_\lambda(r) d\sin(\pi r m) = -(\pi m)^{-1} \int_0^1 \sin(\pi r m) h'_\lambda(r) dr$$

$$= (\pi m)^{-2} \int_0^1 h'_\lambda(r) d\cos(\pi r m)$$

$$= (\pi m)^{-2}\ h'_\lambda(r)\cos(\pi r m)\ |_0^1 - (\pi m)^{-2} \int_0^1 \cos(\pi r m) dh'_\lambda(r).$$

Now

$$h'_\lambda(r) = l'(2\lambda^{1/4}\sin(\pi r/2)) 2\lambda^{1/4}\cos(\pi r/2))(\pi/2),$$

and therefore

$$|(\pi m)^{-2}\ h'_\lambda(r)\cos(\pi rm)\ |_0^1\ | \leq 2(\pi m)^{-2}\sup_{r\in[0,1]}|l'(r)|\pi\lambda^{1/4}.$$

Furthermore,

$$h''_\lambda(r) = l''(2\lambda^{1/4}\sin(\pi r/2))(2\lambda^{1/4}\cos(\pi r/2))(\pi/2))^2 + l'(2\lambda^{1/4}\sin(\pi r/2))2\lambda^{1/4}\sin(\pi r/2))(\pi/2)^2,$$

and because $l'(x) \leq 0$ for $0 \leq x \leq 1$, $l''(x) = 4x^2(5x^4-3)(x^4+1)^{-3}$ and $l''(x) \geq 0$ for $0 \leq x \leq (3/5)^{1/4}$ and $l''(x) < 0$ for $(3/5)^{1/4} < x \leq 1$, it follows that

$$|(\pi m)^{-2}\int_0^1 \cos(\pi rm)dh'_\lambda(r)| \leq (\pi m)^{-2}\int_0^1 |h''_\lambda(r)|dr$$

$$= (\pi m)^{-2}\int_0^1 I(2\lambda^{1/4}\sin(\pi r/2) \leq (3/5)^{1/4})l''(2\lambda^{1/4}\sin(\pi r/2))(2\lambda^{1/4}\cos(\pi r/2))(\pi/2))^2dr$$

$$-(\pi m)^{-2}\int_0^1 I(2\lambda^{1/4}\sin(\pi r/2) > (3/5)^{1/4})l''(2\lambda^{1/4}\sin(\pi r/2))(2\lambda^{1/4}\cos(\pi r/2))(\pi/2))^2dr$$

$$-(\pi m)^{-2}\int_0^1 l'(2\lambda^{1/4}\sin(\pi r/2))2\lambda^{1/4}\sin(\pi r/2))(\pi/2)^2dr$$

$$\leq (\pi m)^{-2}\int_0^1 I(2\lambda^{1/4}\sin(\pi r/2) \leq (3/5)^{1/4})h''_\lambda(r)dr$$

$$-(\pi m)^{-2}\int_0^1 I(2\lambda^{1/4}\sin(\pi r/2) > (3/5)^{1/4})h''_\lambda(r)dr$$

$$+3(\pi m)^{-2}\int_0^1 |l'(2\lambda^{1/4}\sin(\pi r/2))|2\lambda^{1/4}\sin(\pi r/2))(\pi/2)^2dr$$

$$\leq 4(\pi m)^{-2}\sup_{r\in[0,1]}|h'_\lambda(r)| + ...,$$

and therefore,

$$|f_\lambda(m)| \leq Cm^{-2}\lambda^{1/4}.$$

From this it follows that

$$\lambda^{1/4}|f_\lambda(\lambda^{1/4}m) \leq C\lambda^{1/4}\lambda^{1/4}(\lambda^{1/4}m)^{-2} = Cm^{-2}.$$

Also,

$$\lambda^{1/4} f_\lambda(\lambda^{1/4} m) = \lambda^{1/4} \int_0^1 \cos(\pi r \lambda^{1/4} m) h_\lambda(r) dr = \int_0^{\lambda^{1/4}} \cos(\pi y m) h_\lambda(\lambda^{-1/4} y) dy,$$

so, noting that $f(m) = \int_0^\infty \cos(\pi y m) l(\pi y) dy$,

$$|\lambda^{1/4} f_\lambda(\lambda^{1/4} m) - f(m)|$$

$$\leq |\int_0^{\lambda^{1/14}} \cos(\pi y m)(l(2\lambda^{1/4} \sin(\pi \lambda^{-1/4} y/2)) - l(\pi y)) dy|$$

$$+|\int_{\lambda^{1/14}}^{\lambda^{1/4}} \cos(\pi y m) l(2\lambda^{1/4} \sin(\pi \lambda^{-1/4} y/2)) dy| + |\int_{\lambda^{1/14}}^{\infty} \cos(\pi y m) l(\pi y) dy|,$$

and for $y \in [0, \lambda^{1/14}]$,

$$|l(2\lambda^{1/4} \sin(\pi \lambda^{-1/4} y/2)) - l(\pi y)| \leq 2\lambda^{1/4} |\sin(\pi \lambda^{-1/4} y/2) - \pi y \lambda^{-1/4}/2| \sup_{x \geq 0} |l'(x)|$$

$$\leq 2\lambda^{1/4} C(\pi y \lambda^{-1/4}/2)^2 \sup_{x \geq 0} |l'(x)| = (1/2)\lambda^{-1/4+1/7} C\pi^2 \sup_{x \geq 0} |l'(x)|,$$

so

$$|\int_0^{\lambda^{1/14}} \cos(\pi y m)(l(2\lambda^{1/4} \sin(\pi \lambda^{-1/4} y/2)) - l(\pi y)) dy| \leq C\lambda^{-1/28}.$$

Furthermore,

$$|\int_{\lambda^{1/14}}^{\lambda^{1/4}} \cos(\pi y m) l(2\lambda^{1/4} \sin(\pi \lambda^{-1/4} y/2)) dy| \leq \lambda^{1/4} l(2\lambda^{1/4} \sin(\pi \lambda^{-1/4} \lambda^{1/14}/2)) \leq \lambda^{1/4} l(C\lambda^{1/14}),$$

and because $l(x) = O(x^{-4})$ as $x \to \infty$, the last term is $O(\lambda^{1/4-4/14}) = O(\lambda^{-1/28})$. For showing uniform convergence, note that

$$\lambda^{1/4} f_\lambda(\lambda^{1/4} m) = \lambda^{1/4} \int_0^1 \cos(m r \pi \lambda^{1/4}) h_\lambda(r) dr = \int_0^{\lambda^{1/4}} \cos(m y \pi) l(2\lambda^{1/4} \sin(\pi \lambda^{-1/4} y/2)) dy,$$

so, because $l(x)$ is nonincreasing on $[0, \infty)$ and $\sin(\pi x/2) \geq x$ for $x \in [0, 1]$,

$$\sup_{|m| \leq K, |m-m'| \leq \eta} |\lambda^{1/4} f_\lambda(\lambda^{1/4} m) - \lambda^{1/4} f_\lambda(\lambda^{1/4} m')|$$

$$\leq \int_0^{\lambda^{1/4}} \sup_{m,|m-m'|\leq\eta} |\cos(my\pi) - \cos(m'y\pi)| l(2\lambda^{1/4}\sin(\pi\lambda^{-1/4}y/2))dy$$

$$\leq \int_0^{\infty} \sup_{m,|m-m'|\leq\eta} |\cos(my\pi) - \cos(m'y\pi)| l(2y)dy,$$

and by the dominated convergence theorem, it now follows that

$$\limsup_{\eta\downarrow 0} \sup_{|m|\leq K,|m-m'|\leq\eta} |\lambda^{1/4}f_\lambda(\lambda^{1/4}m) - \lambda^{1/4}f_\lambda(\lambda^{1/4}m')| = 0.$$

□

Proof of Theorem 8: Because of the result of Theorem 4, it suffices to show that

$$E|(E|y_{[rT]}|)^{-1}\sum_{s=1}^{T} y_s(f_\lambda(t-s) - \lambda^{-1/4}f(\lambda^{-1/4}(t-s)))| \to 0.$$

Similarly to the proof of Theorem 4,

$$(E|y_{[rT]}|)^{-1}\sum_{s=1}^{T} y_s(f_\lambda(t-s) - \lambda^{-1/4}f(\lambda^{-1/4}(t-s)))$$

$$\leq C\sum_{s=1}^{T} |f_\lambda(t-s) - \lambda^{-1/4}f(\lambda^{-1/4}(t-s))|$$

$$\leq C\int_{s=1}^{T+1} |f_\lambda(t-[s]) - \lambda^{-1/4}f(\lambda^{-1/4}(t-[s]))|ds$$

and setting $y = \lambda^{-1/4}(t-s)$, which implies that $s = t - y\lambda^{1/4}$, it follows that the last expression is bounded by

$$C\int_{-\infty}^{\infty} |\lambda^{1/4}f_\lambda(t-[t-y\lambda^{1/4}]) - f(\lambda^{-1/4}(t-[t-y\lambda^{1/4}]))|dy.$$

Because $\lambda^{1/4}|f_\lambda(\lambda^{1/4}m)| \leq Cm^{-2}$, it follows that

$$\limsup_{\lambda\to\infty} \int_{|y|>K} |\lambda^{1/4}f_\lambda(t-[t-y\lambda^{1/4}]) - f(\lambda^{-1/4}(t-[t-y\lambda^{1/4}]))|dy$$

$$\leq \limsup_{\lambda\to\infty} 2C\int_{|y|>K} (\lambda^{-1/4}(t-[t-y\lambda^{1/4}]))^{-2}dy$$

$$\leq \limsup_{\lambda\to\infty} 2C \int_{|y|>K} (|y|-1)^{-2} dy.$$

Also, because $y-1 \leq \lambda^{-1/4}(t-[t-y\lambda^{1/4}]) \leq y+1$ if $\lambda \geq 1$,

$$\int_{|y|\leq K} |\lambda^{1/4} f_\lambda(t-[t-y\lambda^{1/4}]) - f(\lambda^{-1/4}(t-[t-y\lambda^{1/4}]))| dy$$

$$\leq \sup_{|x|\leq K+1} |\lambda^{1/4} f_\lambda(\lambda^{1/4}x) - f(x)| \to 0$$

by the uniform convergence of $\lambda^{1/4} f_\lambda(\lambda^{1/4}x)$ to $f(x)$ on compacta as asserted (where?)

Proof of Theorem 9: Note that

$$\sum_{s=1}^{4T} y_s w_{4T,4t-4+k,s}(\lambda_Q) = \sum_{i=1}^{T} \sum_{s=1}^{4} y_{4i-4+s} w_{4T,4t-4+k,4i-4+s}(\lambda_Q),$$

and analogously to the proof of Theorem ..., it therefore suffices to show that

$$\lim_{\lambda\to\infty} \limsup_{T\to\infty} \sum_{i=1}^{T} \sum_{s=1}^{4} |w_{T,4t-4+k,4i-4+s}(\lambda) - (1/4) w_{Tti}(4^{-4}\lambda)| = 0.$$

Now by Theorem 8,

$$\lim_{\lambda\to\infty} \limsup_{T\to\infty} \sum_{i=1}^{T} \sum_{s=1}^{4} |w_{T,4t-4+k,4i-4+s}(\lambda) - \lambda^{-1/4} f(\lambda^{-1/4}(4t-4+k-(4i-4+s))| = 0$$

and

$$\lim_{\lambda\to\infty} \limsup_{T\to\infty} \sum_{i=1}^{T} \sum_{s=1}^{4} |(1/4) w_{Tti}(4^{-4}\lambda) - (1/4)(4^{-4}\lambda)^{-1/4} f((4^{-4}\lambda)^{-1/4}(t-s))| = 0.$$

Therefore, it remains to show that, for $k \in \{1,2,3,4\}$,

$$\lim_{\lambda\to\infty} \limsup_{T\to\infty} \sum_{i=1}^{T} \sum_{s=1}^{4} |\lambda^{-1/4} f(\lambda^{-1/4}(4t+k-(4i+s)) - \lambda^{-1/4} f(\lambda^{-1/4}(4t-4s))| = 0,$$

and this follows from elementary calculus. □ □