A property of the Hodrick-Prescott filter and its application

Neslihan Sakarya* Robert M. de Jong†

September 4, 2017

Abstract

This paper explores a remarkable property of the Hodrick-Prescott filter: the cyclical component of the Hodrick-Prescott filter is equal to the trend component of the Hodrick-Prescott filter when applied to the fourth differences, plus an additional term. Our result is remarkable due to its simplicity and its strength in explaining many aspects of the HP filter that have not previously been studied rigorously. We first use our result to analyze the consequences of a deterministic trend break. We find that the effect of a deterministic trend break on the cyclical component is asymptotically negligible for the points that are away from the break point, while for the points in the neighborhood of the break point, the effect is not negligible even asymptotically and a characterization is provided for it. Second, we apply our result to show that the cyclical component of the Hodrick-Prescott filter when applied to series that are integrated up to order 2 is weakly dependent, while the situation for the series that are integrated up to order 3 or more is more subtle. This result contrasts with the conjecture in the literature that the HP filter renders a cyclical component that is stationary when it is applied to series that are integrated up to order 4. Third, we characterize the behavior of the Hodrick-Prescott filter when applied to deterministic polynomial trends and show that the cyclical component reduces the order of the polynomial by order 4. Finally, we give a characterization of the Hodrick-Prescott filter when applied to an exponential deterministic trend, and this characterization shows that the filter is effectively incapable of dealing with a trend that increases this fast. This result suggests that the HP filter should not be used for series that are measured in nominal terms such as GDP, consumption, investment, and house prices.

Keywords: cyclical component, deterministic exponential trend, polynomial trend, structural break, weak dependence.

*Department of Economics, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK, Email `sakarya.2@osu.edu`.

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

1 Introduction

The Hodrick-Prescott (HP) filter is a long-standing standard technique in macroeconomics for separating the long run trend in a data series from short-run fluctuations. Introduced initially by Whittaker (1923) and popularized in economics by Hodrick and Prescott (1997), the HP filter is universally used in macroeconomics. The cited paper by Hodrick and Prescott has thousands of citations; yet, the impact of this work may go beyond that, since the HP filter has become an obliteration by incorporation. While the HP filter has a long and venerable history, it has recently being analyzed more formally in de Jong and Sakarya (2016), Cornea-Madeira (2016), and Hamilton (2016). These papers analyze the properties of the HP filter rigorously and reconsider its usefulness in the context of macroeconomics.

The HP filter calculates the trend of a series y_t, $t = 1, \ldots, T$ by minimizing

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

over $\tau = (\tau_1, \ldots, \tau_T)$. The parameter λ here is a smoothing parameter that for quarterly data is typically chosen to equal 1600. The minimizer, which we will label $\hat{\tau}_{Tt}$, is referred to in the literature as the "trend component," while $\hat{c}_{Tt} = y_t - \hat{\tau}_{Tt}$ is referred to as the "cyclical component." By writing the minimization problem as a vector differentiation problem, it follows that there exists a unique minimizer. The trend component $\hat{\tau}_{Tt}$ and the cyclical component $\hat{c}_{Tt}$ are both weighted averages of $\{y_t\}_{t=1}^{T}$, and in de Jong and Sakarya (2016), an exact formula for the weights is found. This paper also explores the statistical properties of the cyclical component when the HP filter is applied to a unit root process, and considers adjusting the smoothing parameter for the data frequency. Cornea-Madeira (2016) also provides an exact formula for the weights by using the Sherman-Morrison formula.

In this paper, we derive a remarkable property of the HP filter which allows us to derive more general results that are not present in the literature. We first use this elegant result to explore the effect of a deterministic trend break on the cyclical component. It is shown that the cyclical component consists of two terms when the HP filter is applied to a series that has a deterministic trend break at an unknown time point. The first part comprises the residue of the structural break, while the second part is equivalent to the cyclical component in the absence of a structural break. Our main result is also applied to the processes that are integrated up to order 4. We show that series that are integrated up to order 2 exhibit weak dependence properties, while for series that are integrated of order 3 or 4 the law of large numbers does not hold for a large class of unbounded functions of the cyclical component. This result shows that the widely known conjecture "... the HP filter will render stationary series that are integrated (up to fourth order) ..." by King and Rebelo (1993) is incorrect in some dimensions. The authors' conjecture depends on assuming that the first order condition of the minimization problem in Equation (1) that is valid for $t = 3, 4, \ldots, T-2$ holds for every $t \in \mathbb{Z}$. We call this approach the "heuristic approach" throughout the paper. On the other hand, our representation provides a rigorous analysis of the HP filter. In addition, we give a closed form formula for the cyclical component of a polynomial trend and an exponential deterministic trend by using our new representation.

In Section 2 of the paper, we provide an explanation of the "heuristic approach." In Section 3 of the paper, we establish our main result. Section 4 explores the consequences of a structural break; we show that a deterministic trend break has asymptotically no effect on the cyclical component $\hat{c}_{Tt}$ if t is away from the structural break point. In Section 5, we show that for series that are integrated up to order 2, the cyclical component possesses weak dependence properties; on the other hand, for the series that are integrated of order 3 or 4, the law of large numbers fails to hold for a large class of unbounded functions of the cyclical component. In Section 6, we characterize the behavior of the HP filter when applied to a polynomial trend and an exponential deterministic trend. Section 7 summarizes the findings of the paper.

2 Explanation of the heuristic approach

Letting $\bar{B}$ and B denote the forward and the backward operators, respectively, the first order conditions of the problem in Equation (1) can be written as

$$\left((1+\lambda) - 2\lambda\bar{B} + \lambda\bar{B}^2\right)\hat{\tau}_{T1} = y_1, \tag{2}$$

$$\left(-2\lambda B + (1+5\lambda) - 4\lambda\bar{B} + \lambda\bar{B}^2\right)\hat{\tau}_{T2} = y_2, \tag{3}$$

$$\left(-2\lambda\bar{B} + (1+5\lambda) - 4\lambda B + \lambda B^2\right)\hat{\tau}_{T,T-1} = y_{T-1}, \tag{4}$$

$$\left((1+\lambda) - 2\lambda B + \lambda B^2\right)\hat{\tau}_{T,T} = y_T, \tag{5}$$

while for $t = 3, 4, \ldots, T-2$,

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

By defining $|1-B|^2 = (1-B)(1-\bar{B})$, the first order condition in Equation (6) can be written as

$$y_t = (\lambda|1-B|^4 + 1)\hat{\tau}_{Tt}. \tag{7}$$

Analyses of the HP filter based on the first order condition of Equation (7) are for example King and Rebelo (1993), Cogley and Nason (1995), McElroy (2008), Phillips (2010), and Phillips and Jin (2015). Such an analysis cannot be more than a conjecture, since the first order conditions of Equations (2)-(5) are ignored. King and Rebelo (1993)'s conjecture (i.e., the HP filter will render stationary series that are integrated up to fourth order) is based on a simple manipulation of the first order condition of Equation (7) and the identity $y_t = \hat{\tau}_{Tt} + \hat{c}_{Tt}$, which give

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

Thus, one might conjecture that $\hat{c}_{Tt}$ should possess stationarity properties if y_t is integrated up to order 4, because $|1-B|^4 y_t = (1-B)^2(1-\bar{B})^2 B^2 \bar{B}^2 y_t = \Delta^4 y_{t+2}$. Conjecturing along these lines, we might also suspect that the HP filter is capable of removing a quadratic trend, since $|1-B|^4 y_t = |1-B|^4 t^2 = 0$. However, we will show that both conjectures are incorrect in general, since the first order condition of Equation (7) fails to hold for $t = 1, 2$ and $t = T-1, T$. After all, it is easy to see (for example, by calculating the cyclical component of a quadratic time trend in a software package) that the cyclical component of the HP filter when applied to a quadratic trend is not equal to zero. Similarly, we will show that the heuristic reasoning needs to be refined when considering the cyclical component of processes that are integrated of order 3 or more, since the cyclical component at the end of sample is integrated of order 1 or more rather than being integrated of order 0 as claimed in the heuristic approach.

The results of the aforementioned papers should be interpreted as the results derived from an approximate problem that will likely be valid for values of t away from the begin and end points of the sample and for large values of T. However, such findings cannot render exact results for the HP filter. This paper will seek to derive an exact result for the HP filter that allows us to address those issues formally. Note that the approach taken in this paper towards the analysis of the HP filter is completely different from the approach of de Jong and Sakarya (2016) and Cornea-Madeira (2016).

3 Main result

The main result of the paper introduces a property of the HP filter in the following theorem.

Theorem 1. *Let* $\tilde{y}_{T1} = \Delta^2 y_3$, $\tilde{y}_{T2} = \Delta^2 y_4 - 2\Delta^2 y_3$, $\tilde{y}_{T,T-1} = \Delta^2 y_{T-1} - 2\Delta^2 y_T$, $\tilde{y}_{T,T} = \Delta^2 y_T$, *and for* $t = 3, 4, \ldots, T-2$, $\tilde{y}_{Tt} = \Delta^4 y_{t+2}$. *Then for* $t = 1, 2, \ldots, T$

$$\hat{c}_{Tt}(y_1, y_2, \ldots, y_T) = \lambda\, \hat{\tau}_{Tt}(\tilde{y}_{T1}, \tilde{y}_{T2}, \ldots, \tilde{y}_{TT}).$$

This simple but elegant result provides insights into the structure of the cyclical term. To the best of our knowledge, this property of the cyclical term has not been established before. The result shows that the cyclical component is the trend in the fourth difference of the original series plus an additional term that only affects the first and last two observations, for which the fourth difference of y_{t+2} is undefined. This property can shed light on the behavior of the cyclical component that is obtained from various data generating processes.

The idea behind our result is the following. Using the first order condition in Equation (7) and the identity $y_t = \hat{\tau}_{Tt} + \hat{c}_{Tt}$, it follows that

$$\begin{aligned}\hat{c}_{Tt}(y_1, \ldots, y_T) &= \lambda|1 - B|^4 \hat{\tau}_{Tt}(y_1, \ldots, y_T)\\ &= \lambda(1 - B)^2(1 - \bar{B})^2 B^2 \bar{B}^2 \hat{\tau}_{Tt}(y_1, \ldots, y_T)\end{aligned}$$

$$= \lambda(1-B)^4\bar{B}^2\hat{\tau}_{Tt}(y_1,\dots,y_T).$$

Ignoring the fact that y_{-1}, y_0, y_{T+1} and y_{T+2} are undefined, we can now conjecture that the last expression is approximately equal to

$$\lambda(1-B)^4\hat{\tau}_{Tt}(y_3,\dots,y_{T+2}),$$

which can be conjectured to approximately equal to

$$\lambda\hat{\tau}_{Tt}(\Delta^4 y_3,\dots,\Delta^4 y_{T+2}).$$

Therefore, the conjecture presents itself that the cyclical component in a series y_t is approximately equal to the trend in the fourth difference. Theorem 1 corrects and formalizes the conjecture above by taking into account the first order conditions of Equations (2)-(5) as well.

Note that the last two observations $\tilde{y}_{T,T-1}$ and $\tilde{y}_{TT}$ are important especially if the original series is strongly trended; for example if y_t is integrated of order 3 or 4. In that case, those two observations possess the properties of a unit root process or an I(2) process depending on the integration order of the original series. In Section 5, we will elaborate this observation to show that King and Rebelo (1993)'s conjecture is incorrect in some dimensions.

de Jong and Sakarya (2016) gives an analysis of the weak dependence properties of the cyclical component of a unit root process. It is unclear how to extend this analysis to series that are integrated of order 2 or more. Also, the analysis of de Jong and Sakarya (2016) does not give a route for a characterization of structural breaks or deterministic trends (such as polynomial or exponential trends). However, the result that is given in Theorem 1 allows us to give formal results for processes integrated up to order 4 and for deterministic trends in a simple and elegant way by providing a full characterization for the cyclical component of any series.

4 The effect of a structural break

We analyze the effects of the structural breaks to the cyclical component. Specifically, we focus on the intercept break which is assumed to occur at an unknown date in the middle of the sample. The next result formalizes this.

Theorem 2. *Let*

$$y_t = \begin{cases} u_t & \text{for } t = 1, 2, \ldots, [rT] \\ \mu + u_t & \text{for } t = [rT]+1, [rT]+2, \ldots, T, \end{cases}$$

where $4 \leq [rT] \leq T-5$ *and* $[\cdot]$ *is the floor function. Then for* $t = 1, 2, \ldots, T$

$$\hat{c}_{Tt}(y_1, \ldots, y_T) = -\lambda\mu\Delta^3 w_{Tt,[rT]+2} + \hat{c}_{Tt}(u_1, u_2, \ldots, u_T),$$

where $w_{Tt,[rT]+2}$ *is defined in Theorem 1 of de Jong and Sakarya (2016), and for* $k \in \mathbb{Z}$

$$\lim_{T\to\infty} |\hat{c}_{T,[rT]+k}(y_1, \ldots, y_T) - \hat{c}_{T,[rT]+k}(u_1, u_2, \ldots, u_T)| = \lambda|\mu\Delta^3 f_\lambda(k+1)| \ a.s.,$$

where $f_\lambda(\cdot)$ *is defined in Theorem 3 of de Jong and Sakarya (2016).*

The first result shows that the presence of the structural break alters the cyclical component $\hat{c}_{Tt}$ by $-\lambda\mu\Delta^3 w_{Tt,[rT]+2}$. Following Theorem 1 of de Jong and Sakarya (2016), $|w_{Tts}| \leq C|t-s|^{-3}$ for $t \neq s$; therefore, the cyclical component $\hat{c}_{Tt}$ for values of t that are away from $[rT]$ is not affected much by the structural break. For values of t close to $[rT]$, the second result shows that the cyclical component is altered by $\lambda\mu\Delta^3 f_\lambda(k+1)$ asymptotically. It is possible to calculate $\lambda\Delta^3 f_\lambda(k+1)$ by using the formula for $f_\lambda(k+1)$ in Theorem 3 of de Jong and Sakarya (2016). In Figure 1, the effect of an intercept break in the cyclical component is illustrated by plotting $\lambda\mu\Delta^3 f_\lambda(k+1)$ for $\lambda = 1600$ and $\mu = 1$. The figure shows that the structural break mainly impacts $\hat{c}_{Tt}$'s that are 10 time points away from the structural break point.The figure illustrates that the cyclical component $\hat{c}_{Tt}$ is affected by the structural break when t is in the neighborhood of $[rT]$, but this affect dies out as t gets away from the structural break point.

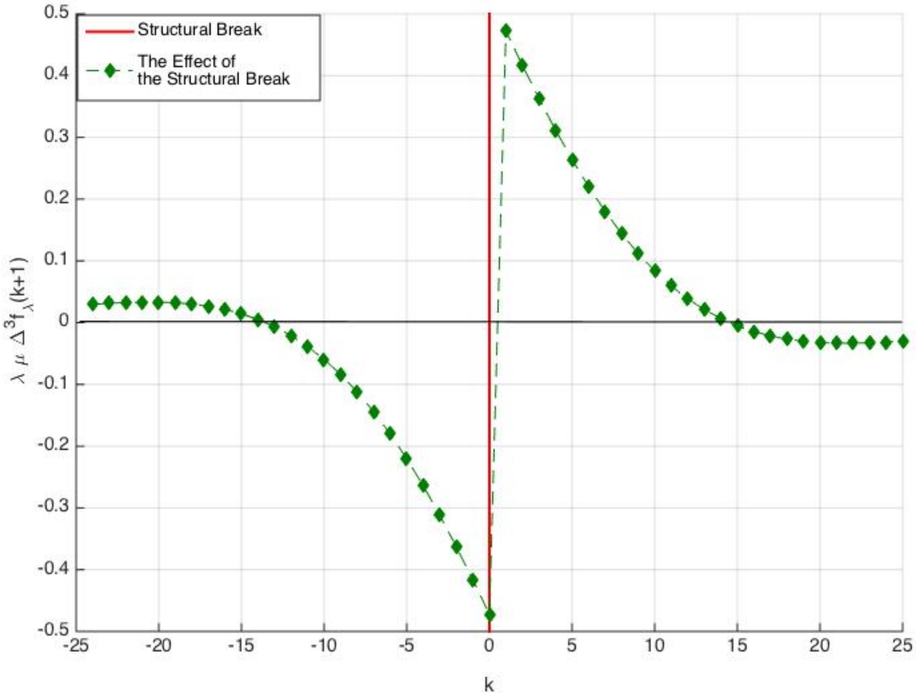

Figure 1: The effect of a deterministic trend break of size 1 when $\lambda = 1600$

5 The HP filter when applied to integrated processes

We consider the weak dependence properties of the cyclical term which is obtained from processes y_t that are integrated up to order 4; that is, $\Delta^q y_t = u_t$, for $q = 1, 2, 3$, or 4, where we assume that u_t has some stationarity or weak dependence properties. King and Rebelo (1993) have conjectured, based on considering the first order condition in Equation (7) only, that the cyclical component has weak dependence properties for processes integrated up to order 4. In this section, we show that this can be made precise for bounded functions of the cyclical component when the process is integrated up to order 4. On the other hand, for unbounded functions of the cyclical component, we show that the law of large numbers does not necessarily hold. Therefore, the picture is more subtle than suggested by King and Rebelo (1993).

The cyclical component $\{\hat{c}_{Tt}\}_{t=1}^{T}$ is a triangular array, and therefore it cannot be a strictly stationary sequence. On the other hand, it is possible to derive a near epoch dependence type result. The near-epoch dependence idea goes back to Ibragimov (1962) and is formalized by Billingsley (1968), McLeish (1975), Bierens (1983), Gallant and White (1988), Andrews (1988), and Pötscher and Prucha (1991) with different approaches. The idea behind the near epoch dependence concept is that the process can be approximated arbitrarily well by a function of a finite number of strong mixing variables. Such a function is called the approximator.

To formulate our result, for $m \geq 1$ define the approximator $\hat{c}_{Tt}^m$ as

$$\hat{c}_{Tt}^m = \lambda \sum_{s=3}^{T-2} w_{Tts} \tilde{y}_{Ts} I(|t-s| \leq m). \tag{8}$$

Note that $\tilde{y}_{Ts} = \Delta^4 y_{s+2}$ for $s = 3, \ldots, T-2$. Since $\Delta^4 y_{s+2} = \Delta^{4-q} u_{s+2}$ for $q = 1, 2, 3, 4$ is a function of $u_{s-2+q}, \ldots, u_{s+2}$, the approximator $\hat{c}_{Tt}^m$ depends only on $u_{t-2-m+q}, \ldots, u_{t+2+m}$. Also, note that the approximator cannot be defined as $\lambda \sum_{s=1}^{T} w_{Tts} \tilde{y}_{Ts} I(|t-s| \leq m)$ because then the approximator would include integrated terms (i.e., $\tilde{y}_{T,T-1}$ and $\tilde{y}_{TT}$) if y_t is an I(3) or I(4) process.

The next result shows that the cyclical component has an approximability property when the HP filter is applied to a process that is integrated of order 4 or less.

Theorem 3. *Assume that $\Delta^q y_t = u_t$, for $q = 1, 2, 3,$ or 4, where $\sup_{s \geq 1} \| u_s \|_p < \infty$. Then for any $\gamma \in (0, 1/2)$, there exists constant $C_1 > 0$ such that for any $m \geq 1$,*

$$\sup_{T \geq 1} \sup_{t \in [\gamma T, (1-\gamma)T]} \| \hat{c}_{Tt} - w_{Tt1} \tilde{y}_{T1} - w_{Tt2} \tilde{y}_{T2} - w_{Tt,T-1} \tilde{y}_{T,T-1} - w_{TtT} \tilde{y}_{TT} - \hat{c}_{Tt}^m \|_p \leq C_1 m^{-2}$$

and

$$\sup_{T \geq 1} \sup_{t \in [\gamma T, (1-\gamma)T]} \| \hat{c}_{Tt} - w_{Tt1} \tilde{y}_{T1} - w_{Tt2} \tilde{y}_{T2} - w_{Tt,T-1} \tilde{y}_{T,T-1} - w_{TtT} \tilde{y}_{TT} \|_p < \infty.$$

The assumption that y_t is integrated up to order 4 at most implies that, under standard moment assumptions, $\| \tilde{y}_{T,T-1} \|_p + \| \tilde{y}_{TT} \|_p = O(T^{3/2})$. This is because the definition of $\tilde{y}_{T,T-1}$ and $\tilde{y}_{TT}$ involves the second difference of the original process, making $\tilde{y}_{T,T-1}$ and $\tilde{y}_{TT}$ integrated of order 2 at most, and I(2) processes are $O_p(T^{3/2})$ under standard conditions.

Therefore, under such assumptions,

$$w_{Tt1}\tilde{y}_{T1} + w_{Tt2}\tilde{y}_{T2} + w_{Tt,T-1}\tilde{y}_{T,T-1} + w_{TtT}\tilde{y}_{TT} \tag{9}$$

will be of a small order for $t \in [\gamma T, (1-\gamma)T]$, since $|w_{Tts}| \le C|t-s|^{-3}$ for $t, s = 1, 2, \ldots, T$ and $t \neq s$ by Theorem 1 of de Jong and Sakarya (2016).

The result above is, to the best of our knowledge, the first formalization of King and Rebelo's conjecture. We can now prove the following weak law of large numbers for bounded and continuous functions of the cyclical component:

Theorem 4. *Assume that y_t satisfies $\Delta^q y_t = u_t$ for $q = 1, 2, 3,$ or 4, and assume that u_t is strong mixing. In addition, assume that $E(|\tilde{y}_{1T}| + |\tilde{y}_{2T}| + |\tilde{y}_{T-1,T}| + |\tilde{y}_{T,T}|) = O(T^{3/2})$. Let $g(\cdot)$ be a function that is bounded and Lipschitz continuous on $\mathbb{R}$. Then*

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

Note that in the above result, the term of Equation (9) plays no role asymptotically as long as $g(\cdot)$ is a bounded function. In the case that $g(\cdot)$ is an unbounded function, the convergence rate of $\tilde{y}_{TT}$ takes over if y_t is integrated of order 3 or more, and Theorem 4 does not hold anymore. The following theorem formalizes this.

Theorem 5. *Assume that y_t satisfies $\Delta^q y_t = u_t$ for $q = 3$ or 4, and $\sup_{k\ge1} E|u_k| < \infty$. Let $g(x) \ge C|x|^p$ where $x \in \mathbb{R}$ and $C > 0$ is a constant. Also, assume that $U_T(r) := T^{-1/2}\sum_{t=1}^{[rT]} u_t \Rightarrow U(r)$ on $r \in [0,1]$. Then,*

$$T^{-1}\sum_{t=1}^{T} g(\hat{c}_{Tt}) \ge C\lambda^p T^{-1}\left|\sum_{s=1}^{T} w_{TTs}\tilde{y}_{Ts}\right|^p. \tag{10}$$

If y_t is an I(3) process, then

$$T^{-p/2}\left|\sum_{s=1}^{T} w_{TTs}\tilde{y}_{Ts}\right|^p \xrightarrow{d} \left|\left((f_\lambda(0) - f_\lambda(1)) + \xi_\lambda g_\lambda(1)(g_\lambda(1) - g_\lambda(2))\right) U(1)\right|^p. \tag{11}$$

If y_t is an I(4) process, then

$$T^{-3p/2}\left|\sum_{s=1}^{T} w_{TTs}\tilde{y}_{Ts}\right|^p \xrightarrow{d} \left|((f_\lambda(0)-f_\lambda(1))+\xi_\lambda g_\lambda(1)(g_\lambda(1)-g_\lambda(2)))\int_0^1 U(r)dr\right|^p, \qquad (12)$$

where ξ_λ and $g_\lambda(\cdot)$ are defined in Theorem 2 and 3 of de Jong and Sakarya (2016), respectively.

It is easy to verify that $f_\lambda(0) - f_\lambda(1) + \xi_\lambda g_\lambda(1)(g_\lambda(1) - g_\lambda(2))$ takes the value of 0.254, 0.022, and 0.002 for $\lambda = 6.25$, 1600, and 129,600, respectively. The above result gives conditions under which $T^{-1}\sum_{t=1}^{T} g(\hat{c}_{Tt})$ is explosive. In the first result, we show that $T^{-1}\sum_{t=1}^{T} g(\hat{c}_{Tt})$ is bounded below by a process that is $O_p(T^{(p/2-1)})$ if y_t is an I(3) process. This implies that for $p \geq 2$, Theorem 4 does not hold. Similarly, $T^{-1}\sum_{t=1}^{T} g(\hat{c}_{Tt})$ is bounded below by a process that is $O_p(T^{(3p/2-1)})$ if y_t is an I(4) process, which in turn implies that Theorem 4 does not hold for $p \geq 2/3$.

Therefore, Theorem 5 provides a partial converse to King and Rebelo's conjecture, as it illustrates that the law of large numbers can fail for unbounded functions of cyclical components when the HP filter is applied to I(3) and I(4) processes.

6 The HP filter when applied to deterministic trends

6.1 Deterministic polynomial trends

Theorem 1 also allows us to establish the behavior of the HP filter when applied to deterministic polynomial trends. From Theorem 1, a result for the case of a linear trend $y_t = a + bt$ immediately follows. After all, for that case, $\Delta^2 y_t = 0$, implying that $\tilde{y}_{Tt} = 0$ for $t = 1, \ldots, T$, which by Theorem 1 implies that $\hat{c}_{Tt} = 0$. For higher order polynomials, the result is more complex:

Theorem 6. *Suppose that $y_t = t^p$ for $t = 1, 2, \ldots, T$ and $p \in \mathbb{N}$. Then,*

$$\hat{\tau}_{Tt}(1, 2^p, \ldots, T^p) = \begin{cases} t^p - \lambda C_{Ttp} & \text{if } p = 2, 3 \\ t^p - \lambda \sum_{k=0}^{p-4} a_{pk}\hat{\tau}_{Tt}(1, 2^k, \ldots, T^k) - \lambda C_{Ttp} + \lambda H_{Ttp} & \text{if } p \geq 4, \end{cases}$$

where

$$c_{pk} = \binom{p}{k}(2^{p-k} - 2), \tag{13}$$

$$C_{Ttp} = \sum_{k=0}^{p-2} c_{pk} \hat{\tau}_{Tt}\left(1, 2^k - 2, 0, \ldots, 0, (T-3)^k - 2(T-2)^k, (T-2)^k\right),$$

$$a_{pk} = \begin{cases} \binom{p}{k}\left(2^{p-k+1} - 8\right) & \textit{if } p-k \textit{ is even} \\ 0 & \textit{if } p-k \textit{ is odd,} \end{cases} \tag{14}$$

$$H_{Ttp} = \sum_{k=0}^{p-4} a_{pk} \hat{\tau}_{Tt}\left(1, 2^k, 0, \ldots, 0, (T-1)^k, T^k\right).$$

The above result shows that the cyclical component of a polynomial trend of order p is

$$\hat{c}_{Tt}(1, 2^p, \ldots, T^p) = \begin{cases} \lambda C_{Ttp} & \text{if } p = 2, 3 \\ \lambda \sum_{k=0}^{p-4} a_{pk} \hat{\tau}_{Tt}(1, 2^k, \ldots, T^k) + \lambda C_{Ttp} - \lambda H_{Ttp} & \text{if } p \geq 4. \end{cases}$$

It follows that for $p = 2$, $\hat{c}_{Tt} = \lambda C_{Tt2} = 2\lambda \hat{\tau}_{Tt}(1, -1, 0, \ldots, 0, -1, 1) = 2\lambda(w_{Tt1} - w_{Tt2} - w_{Tt,T-1} + w_{TtT})$. This result and Theorem 1 of de Jong and Sakarya (2016) together imply that $\hat{c}_{Tt}$ takes a value close to zero if t is sufficiently away from the begin and end points because $|w_{Tts}| \leq C|t-s|^{-3}$ for $t \neq s$. In the case of a cubic trend, Theorem 6 gives $\hat{c}_{Tt} = \lambda C_{Tt3} = 6\lambda \hat{\tau}_{Tt}(2, -1, 0, \ldots, 0, -T, (T-1)) = 6\lambda(2w_{Tt1} - w_{Tt2} - Tw_{Tt,T-1} + (T-1)w_{TtT})$, which suggests that the cyclical component approaches zero slower than the cyclical component of a quadratic trend in the middle of a large sample. The heuristic approach that we explained in Section 2 incorrectly suggests that the cyclical component of a polynomial trend of order 3 or less equals zero. Another implication of the above result is that for $p = 4$, $\hat{c}_{Tt} = 24\lambda + \lambda C_{Tt4} - \lambda H_{Tt4}$. Note that $C_{Tt4} = 50w_{Tt1} + 10w_{Tt2} + (-12(T-2)^2 + 70)w_{Tt,T-1} + (12(T-2)^2 + 24T - 34)w_{TtT}$ and $H_{Tt4} = 24(w_{Tt1} + w_{Tt2} + w_{Tt,T-1} + w_{TtT})$. In the middle of a large sample, C_{Tt4} and H_{Tt4} are $O(T^{-1})$ and $O(T^{-3})$, respectively, since $|w_{Tts}| \leq C|t-s|^{-3}$ for $t \neq s$ by Theorem 1 of de Jong and Sakarya (2016). Therefore, C_{Tt4} and H_{Tt4} are asymptotically negligible in the middle of a large sample. However, C_{Tt4} is

$O(T^2)$ when t is close to the end points, since for fixed $k, j \geq 0$ $\sup_{T\geq 1} |w_{T,T-k,T-j}| = O(1)$. This implies that the $\hat{c}_{Tt}$ is roughly equals 24λ in the middle of a large sample, while it is $O(T^2)$ near the end of the sample. The reduction of the polynomial order by 4 therefore only happens in the middle of the sample.

Next, we provide an explicit formula for the cyclical component of a quadratic trend.

Theorem 7. *Suppose that $y_t = t^2$ for $t = 1, 2, \ldots, T$, and $0 < \lambda < \infty$. Then*

$$\begin{aligned}\hat{c}_{Tt} =& (C_{1T} + \bar{C}_{1T})|z_1|^t \cos(t\theta) + i(C_{1T} - \bar{C}_{1T})|z_1|^t \sin(t\theta) \\ &+ (C_{1T} + \bar{C}_{1T})|z_1|^{T-t+1} \cos((T-t+1)\theta) + i(C_{1T} - \bar{C}_{1T})|z_1|^{T-t+1} \sin((T-t+1)\theta),\end{aligned} \tag{15}$$

where $i^2 = -1$,

$$a_T = (1+\lambda)(z_1 + z_1^T) - 2\lambda(z_1^2 + z_1^{T-1}) + \lambda(z_1^3 + z_1^{T-2}),$$

$$b_T = -2\lambda(z_1 + z_1^T) + (1+5\lambda)(z_1^2 + z_1^{T-1}) - 4\lambda(z_1^3 + z_1^{T-2}) + \lambda(z_1^4 + z_1^{T-3}),$$

$$C_{1T} = 2\lambda \left(\bar{b}_T + \bar{a}_T\right) / \left(a_T \bar{b}_T - \bar{a}_T b_T\right),$$

$$z_1 = 1 - \frac{\sqrt{\sqrt{1+16\lambda} - 1}}{2\sqrt{2\lambda}} + i\left(\frac{\sqrt{2}}{\sqrt{\sqrt{1+16\lambda}-1}} - \frac{1}{2\sqrt{\lambda}}\right),$$

$$\theta = \tan^{-1}\left(2^{1/2}\left(\sqrt{1+16\lambda} - 1\right)^{-1/2}\right),$$

and $\bar{C}_{1T}$ is the complex conjugate of C_{1T}.

Note that $C_{1T} + \bar{C}_{1T}$ and $i(C_{1T} - \bar{C}_{1T})$ that appear in Theorem 7 are both real-valued. The above result implies that when t is small and T is large, the cyclical component of a quadratic trend is approximately equal to $(C_{1T}+\bar{C}_{1T})|z_1|^t \cos(t\theta)+i(C_{1T}-\bar{C}_{1T})|z_1|^t \sin(t\theta)$, and when t is close to the end of sample, the last two terms in Equation (15) take over. When t is in the middle of a large sample, $\hat{c}_{Tt}$ takes relatively smaller values since all terms in Equation (15) are small. This is because $|z_1| < 1$ for $0 < \lambda < \infty$. Also, it is easy to see that $\hat{c}_{Tt}(1, 2^2, \ldots, T^2) = \hat{c}_{T,T-t+1}(1, 2^2, \ldots, T^2)$ for $t = 1, 2, \ldots, T$. This property of the cyclical component appears only in the quadratic trend case.

Note that it is also possible to derive an expression for the cyclical component of a cubic trend along the lines of Theorem 7:

$$\begin{aligned}\hat{c}_{Tt} =&(C^*_{1T} + \bar{C}^*_{1T})|z_1|^t \cos(t\theta) + i(C^*_{1T} + \bar{C}^*_{1T})|z_1|^t \sin(t\theta)\\ &+ (C^*_{2T} + \bar{C}^*_{2T})|z_1|^{T-t+1} \cos((T-t+1)\theta) + i(C^*_{2T} + \bar{C}^*_{2T})|z_1|^{T-t+1} \sin((T-t+1)\theta),\end{aligned}$$

where C^*_{1T} and C^*_{2T} are the complex-valued terms that depend only on z_1 and T, and $C^*_{1T} \neq C^*_{2T} \neq C_{1T}$. For the sake of brevity, we do not provide the explicit formulas for C^*_{1T} and C^*_{2T}.

6.2 Deterministic exponential trends

The following result characterizes an exponential deterministic trend by using the result of Theorem 1.

Theorem 8. *Let* $y_t = \exp(t)$ *for* $t = 1, 2, \ldots, T$. *Then*

$$\hat{c}_{Tt} = C\lambda\hat{\tau}_{Tt}(C_1 \exp(1), C_2 \exp(2), \exp(3), \ldots, \exp(T-2), C_3 \exp(T-1), C_4 \exp(T)) \tag{16}$$

where $C = \exp(2)(1-\exp(-1))^4$, $C_1 = (1-\exp(-1))^{-2}$, $C_2 = C_1(1-2\exp(-1))$, $C_3 = 1 - C_1$, *and* $C_4 = C_1 \exp(-2)$.

The result above implies that the cyclical component of an exponential deterministic trend can be formulated in terms of its HP filter trend. Since the cyclical component equals the trend in an exponentially increasing sequence, Theorem 8 suggests that the HP filter is not capable of removing an exponential deterministic trend from a series.

The next result shows that the cyclical component of an exponentially deterministic trend is as explosive as its HP filter trend when t is close to the end of sample.

Theorem 9. *Let* $y_t = \exp(t)$ *for* $t = 1, 2, \ldots, T$. *Then for* $k \geq 0$

$$\begin{aligned}&\lim_{T\to\infty} \frac{\hat{c}_{T,T-k}\left(\exp(1), \exp(2), \ldots, \exp(T)\right)}{\hat{\tau}_{T,T-k}\left(\exp(1), \exp(2), \ldots, \exp(T)\right)}\\ =&C\lambda + C(C_3-1)\exp(-1)\lambda \frac{f_\lambda(k-1) + f_\lambda(k+2) + \xi_\lambda g_\lambda(k+1)g_\lambda(2)}{\sum_{j=0}^{\infty}(f_\lambda(k-j) + f_\lambda(k+j+1) + \xi_\lambda g_\lambda(k+1)g_\lambda(j+1))\exp(-j)}\end{aligned}$$

$$+ C(C_4 - 1)\lambda \frac{f_\lambda(k) + f_\lambda(k+1) + \xi_\lambda g_\lambda(k+1) g_\lambda(1)}{\sum_{j=0}^{\infty} (f_\lambda(k-j) + f_\lambda(k+j+1) + \xi_\lambda g_\lambda(k+1) g_\lambda(j+1)) \exp(-j)}.$$

This result shows that the HP filter is not capable of removing the trend in deterministic exponential trends, since $\hat{c}_{Tt}$ is as explosive as $\hat{\tau}_{Tt}$ when t is close to the end of sample. For example, for $\lambda = 1600$ and $k = 0, 1, \ldots, 5$, the limit in Theorem 9 takes the values $2.37, 0.38, -0.42, -0.76, -0.9$ and -0.95, respectively. A trivial remedy to this problem is to take the logarithm of $\{y_t\}_{t=1}^T$ which will become a linear trend after the logarithmic transformation. It was argued in the beginning of Section 6 that the HP filter is capable of removing a linear trend; therefore, the cyclical component of the transformed $\{y_t\}_{t=1}^T$ would be zero.

7 Conclusion

This paper derives a simple but elegant property of the HP filter, which highlights the behavior of the cyclical component when the HP filter is applied to various processes. Our result is remarkable due to its simplicity and its strength in explaining many aspects of the HP filter which have not been studied rigorously. Our main result is used to analyze the effect of a deterministic trend break. We find that a deterministic trend break affects the cyclical component, and this effect is not negligible even when the sample size is large. Next, our main result is applied to the integrated processes of order up to 4. We conclude that the cyclical component of the series that are integrated of order up to 2 possesses weak dependence properties and the Law of Large Number and the Central Limit Theorem type results would hold. On the other hand, the situation is more subtle when the HP filter is applied to the processes that are integrated of order 3 or more. We find that the Law of Large Number and the Central Limit Theorem type results might fail to hold for the unbounded transformation of the cyclical component of the processes that are integrated of order 3 more. Lastly, our main result allows us to derive a closed form formula for the cyclical component of the deterministic polynomial trends such as quadratic or cubic trend, and deterministic exponential trends. It is shown that the HP filter reduces the order of polynomial by 4 in the middle of the sample when it is applied to the polynomial trends. We show that the HP filter is not capable of removing the trend in the deterministic exponential trends.

References

ANDREWS, D. W. (1988): "Laws of large numbers for dependent non-identically distributed random variables," *Econometric Theory*, 4, 458–467.

BIERENS, H. J. (1983): "Uniform consistency of kernel estimators of a regression function under generalized conditions," *Journal of the American Statistical Association*, 78, 699–707.

BILLINGSLEY, P. (1968): *Convergence of probability measures*, John Wiley & Sons New York.

COGLEY, T. AND J. M. NASON (1995): "Effects of the Hodrick-Prescott filter on trend and difference stationary time series: implications for business cycle research," *Journal of Economic Dynamics and Control*, 19, 253–278.

CORNEA-MADEIRA, A. (2016): "The explicit formula for the Hodrick-Prescott filter in finite sample," *Review of Economics and Statistics*, forthcoming.

DE JONG, R. M. AND N. SAKARYA (2016): "The econometrics of the Hodrick-Prescott filter," *Review of Economics and Statistics*, 98, 310–317.

GALLANT, A. R. AND H. WHITE (1988): *A unified theory of estimation and inference for nonlinear dynamic models*, Blackwell.

HAMILTON, J. D. (2016): "Why you should never use the Hodrick-Prescott filter," Working paper.

HODRICK, R. J. AND E. C. PRESCOTT (1997): "Postwar US business cycles: an empirical investigation," *Journal of Money, Credit and Banking*, 29, 1–16.

IBRAGIMOV, I. A. (1962): "Some limit theorems for stationary processes," *Theory of Probability and its Applications*, 7, 349–382.

King, R. G. and S. T. Rebelo (1993): "Low frequency filtering and real business cycles," *Journal of Economic Dynamics and Control*, 17, 207–231.

McElroy, T. (2008): "Exact formulas for the Hodrick-Prescott filter," *Econometrics Journal*, 11, 209–217.

McLeish, D. (1975): "A maximal inequality and dependent strong laws," *Annals of Probability*, 3, 829–839.

Phillips, P. C. (2010): "Two New Zealand pioneer econometricians," *New Zealand Economic Papers*, 44, 1–26.

Phillips, P. C. and S. Jin (2015): "Business cycles, trend elimination, and the HP Filter," Cowles Foundation Discussion Paper.

Pötscher, B. M. and I. R. Prucha (1991): "Basic structure of the asymptotic theory in dynamic nonlinear econometric models, part I: consistency and approximation concepts," *Econometric Reviews*, 10, 125–216.

Rudin, W. (1976): *Principles of mathematical analysis*, vol. 3, McGraw-Hill New York.

Whittaker, E. T. (1923): "On a new method of graduation," *Proceedings of the Edinburgh Mathematical Society*, 41, 63–75.

Appendix 1: Mathematical proofs

Throughout this Appendix, we define summations with empty index sets to equal 0.

Proof of Theorem 1. First, we rewrite the minimization problem of Equation (1) as a minimization problem over $c_t = y_t - \tau_t$. We then obtain

$$\sum_{t=1}^{T} c_t^2 + \lambda \sum_{t=2}^{T-1} (c_{t+1} - 2c_t + c_{t-1})^2 - 2\lambda \sum_{t=2}^{T-1} (y_{t+1} - 2y_t + y_{t-1})(c_{t+1} - 2c_t + c_{t-1})$$
$$+\lambda \sum_{t=2}^{T-1} (y_{t+1} - 2y_t + y_{t-1})^2.$$

The last term of the expression above is irrelevant to the minimization problem. Applying summation by parts twice gives (Rudin, 1976, Theorem 3.41 on p. 70)

$$
\begin{aligned}
&\sum_{t=2}^{T-1}(y_{t+1}-2y_t+y_{t-1})(c_{t+1}-2c_t+c_{t-1}) \\
=&\sum_{t=2}^{T-1}\Delta^2 y_{t+1}\Delta^2 c_{t+1} \\
=&\sum_{t=3}^{T-1}(\Delta^2 y_t-\Delta^2 y_{t+1})\Delta c_t+\Delta^2 y_T\Delta^2 c_T-\Delta^2 y_3\Delta c_2 \\
=&\sum_{t=3}^{T-2}\Delta^4 y_{t+2}c_t+\Delta^2 y_3 c_1+\left(\Delta^2 y_4-2\Delta^2 y_3\right)c_2+\left(\Delta^2 y_{T-1}-2\Delta^2 y_T\right)c_{T-1}+\Delta^2 y_T c_T,
\end{aligned}
$$

and therefore, it suffices to minimize

$$
\begin{aligned}
&\sum_{t=1}^{T}c_t^2+\lambda\sum_{t=2}^{T-1}(c_{t+1}-2c_t+c_{t-1})^2-2\lambda\sum_{t=3}^{T-2}\Delta^4 y_{t+2}c_t \\
&-2\lambda\Delta^2 y_3 c_1-2\lambda\left(\Delta^2 y_4-2\Delta^2 y_3\right)c_2-2\lambda\left(\Delta^2 y_{T-1}-2\Delta^2 y_T\right)c_{T-1}-2\lambda\Delta^2 y_T c_T \\
=&\sum_{t=1}^{T}c_t^2+\lambda\sum_{t=2}^{T-1}(c_{t+1}-2c_t+c_{t-1})^2-2\lambda\sum_{t=3}^{T-2}\tilde{y}_{Tt}c_t-2\lambda\tilde{y}_{T1}c_1-2\lambda\tilde{y}_{T2}c_2-2\lambda\tilde{y}_{T,T-1}c_{T-1}-2\lambda\tilde{y}_{TT}c_T,
\end{aligned}
$$

over $(c_1,\dots,c_T)$ where $\{\tilde{y}_{Tt}\}_{t=1}^{T}$ is defined in Theorem 1. The first order conditions for c_t for $t=1,2,T-1,T$ are

$$\left((1+\lambda)-2\lambda\bar{B}+\lambda\bar{B}^2\right)\hat{c}_{T1}=\lambda\Delta^2 y_3=\lambda\tilde{y}_{T1}, \tag{17}$$

$$\left(-2\lambda B+(1+5\lambda)-4\lambda\bar{B}+\lambda\bar{B}^2\right)\hat{c}_{T2}=\lambda\left(\Delta^2 y_4-2\Delta^2 y_3\right)=\lambda\tilde{y}_{T2}, \tag{18}$$

$$\left(-2\lambda\bar{B}+(1+5\lambda)-4\lambda B+\lambda B^2\right)\hat{c}_{T,T-1}=\lambda\left(\Delta^2 y_{T-1}-2\Delta^2 y_T\right)=\lambda\tilde{y}_{T,T-1}, \tag{19}$$

$$\left((1+\lambda)-2\lambda B+\lambda B^2\right)\hat{c}_{TT}=\lambda\Delta^2 y_T=\lambda\tilde{y}_{TT}, \tag{20}$$

respectively. Also, the first order condition for c_t for $t=3,\dots,T-2$ is

$$\left(\lambda\bar{B}^2-4\lambda\bar{B}+(1+6\lambda)-4\lambda B+\lambda B^2\right)\hat{c}_{Tt}=\lambda\Delta^4 y_{t+2}=\lambda\tilde{y}_{Tt}. \tag{21}$$

The analogy between the first order conditions of Equations (17)-(21) and the first conditions of Equations (2)-(6) now reveals that $\hat{c}_{Tt}(y_1, y_2, \ldots, y_T) = \hat{\tau}_{Tt}(\lambda\tilde{y}_{T1}, \lambda\tilde{y}_{T2}, \ldots, \lambda\tilde{y}_{TT}) = \lambda\hat{\tau}_{Tt}(\tilde{y}_{T1}, \tilde{y}_{T2}, \ldots, \tilde{y}_{TT})$. □

Lemma 1. *For $r \in (0, 1)$ and $k, l \in \mathbb{Z}$,*

$$\lim_{T\to\infty} w_{T,[rT]+k,[rT]+l} = f_\lambda(k-l).$$

Proof of Lemma 1. We use the definition of weights in Theorem 1 of de Jong and Sakarya (2016) and write that

$$\lim_{T\to\infty} w_{T,[rT]+k,[rT]+l}$$

$$= \lim_{T\to\infty} (f_{T\lambda}(k-l) + f_{T\lambda}(T)I(2[rT]+k+l-1 = T))$$

$$+ \lim_{T\to\infty} f_{T\lambda}(2[rT]+k+l-1)I(2[rT]+k+l-1 < T)$$

$$+ \lim_{T\to\infty} f_{T\lambda}(2(T-[rT])-k-l+1)I(2[rT]+k+l-1 > T)$$

$$+ \lim_{T\to\infty} (\xi_{T\lambda}g_{T\lambda}([rT]+k)g_{T\lambda}([rT]+l) + \phi_{T\lambda}g_{T\lambda}(T-[rT]-k+1)g_{T\lambda}([rT]+l))$$

$$+ \lim_{T\to\infty} \phi_{T\lambda}g_{T\lambda}([rT]+k)g_{T\lambda}(T-[rT]-l+1)$$

$$+ \lim_{T\to\infty} \xi_{T\lambda}g_{T\lambda}(T-[rT]-k+1)g_{T\lambda}(T-[rT]-l+1)$$

$$= f_\lambda(k-l),$$

since $\lim_{T\to\infty} f_{T\lambda}(k-l) = f_\lambda(k-l)$, $\lim_{T\to\infty} f_{T\lambda}(T) = 0$, $\lim_{T\to\infty} g_{T\lambda}(T) = 0$, $\lim_{T\to\infty} \xi_{T\lambda} = \xi_\lambda$, and $\lim_{T\to\infty} \phi_{T\lambda} = 0$ by Theorem 1-3 of de Jong and Sakarya (2016). □

Proof of Theorem 2. First, we write that

$$y_t = \mu I(t \geq [rT]+1) + u_t,$$

where we assume that $4 \leq [rT] \leq T-5$. By using the result of Theorem 1, it follows that

$$\hat{c}_{Tt}(y_1, y_2, \ldots, y_T)$$

$$
\begin{aligned}
=&\lambda\hat{\tau}_{Tt}(\Delta^2 u_3, \Delta^2 u_4 - 2\Delta^2 u_3, \mu\Delta^4 I(5 \geq [rT]+1) + \Delta^4 u_5, \\
&\ldots, \mu\Delta^4 I(T \geq [rT]+1) + \Delta^4 u_T, \Delta^2 u_{T-1} - 2\Delta^2 u_T, \Delta^2 u_T).
\end{aligned}
$$

Note that

$$
\hat{\tau}_{Tt}(x_1+y_1, x_2+y_2, \ldots, x_T+y_T) = \hat{\tau}_{Tt}(x_1, x_2, \ldots, x_T) + \hat{\tau}_{Tt}(y_1, y_2, \ldots, y_T),
$$

therefore,

$$
\begin{aligned}
&\hat{c}_{Tt}(y_1, y_2, \ldots, y_T) \\
=&\lambda\hat{\tau}_{Tt}(0, 0, \mu\Delta^4 I(5 \geq [rT]+1), \ldots, \mu\Delta^4 I(T \geq [rT]+1), 0, 0) \qquad (22) \\
&+ \lambda\hat{\tau}_{Tt}(\Delta^2 u_3, \Delta^2 u_4 - 2\Delta^2 u_3, \Delta^4 u_5, \ldots, \Delta^4 u_T, \Delta^2 u_{T-1} - 2\Delta^2 u_T, \Delta^2 u_T).
\end{aligned}
$$

Also, note that $\Delta^4 I(t+2 \geq [rT]+1) = \Delta^3 I(t+2 = [rT]+1)$, thus the first term after the equality in Equation (22) is equivalent to

$$
\begin{aligned}
&\lambda\hat{\tau}_{Tt}(0, 0, \mu\Delta^3 I(5 = [rT]+1), \ldots, \mu\Delta^3 I(T = [rT]+1), 0, 0) \\
&= \lambda \sum_{s=3}^{T-2} \mu\Delta^3 I(s+2 = [rT]+1) w_{Tts} \\
&= \lambda\mu \sum_{s=3}^{T-2} \left(I(s+2 = [rT]+1) - 3I(s+1 = [rT]+1)\right) w_{Tts} \\
&+\lambda\mu \sum_{s=3}^{T-2} \left(3I(s = [rT]+1) - I(s-1 = [rT]+1)\right) w_{Tts} \\
&= -\lambda\mu\Delta^3 w_{Tt,[rT]+2},
\end{aligned}
$$

where the first equality is obtained by the weighted average representation of the trend given in Theorem 1 of de Jong and Sakarya (2016). The above expression gives that

$$
\begin{aligned}
&\hat{c}_{Tt}(y_1, y_2, \ldots, y_T) \\
=&-\lambda\mu\Delta^3 w_{Tt,[rT]+2} + \lambda\hat{\tau}_{Tt}(\Delta^2 u_3, \Delta^2 u_4 - 2\Delta^2 u_3, \Delta^4 u_5, \ldots, \Delta^4 u_T, \Delta^2 u_{T-1} - 2\Delta^2 u_T, \Delta^2 u_T)
\end{aligned}
$$

$$= - \lambda\mu\Delta^3 w_{Tt,[rT]+2} + \hat{c}_{Tt}(u_1, u_2, \ldots, u_T),$$

where the last equality follows from Theorem 1.

Next, we need to show the second result of the theorem. The first result of the theorem implies that

$$\begin{aligned}
&\lim_{T\to\infty} |\hat{c}_{T,[rT]+k}(y_1, y_2, \ldots, y_T) - \hat{c}_{T,[rT]+k}(u_1, u_2, \ldots, u_T)| \\
=&\lim_{T\to\infty} |\lambda\mu\Delta^3 w_{T,[rT]+k,[rT]+2}| \\
=&\lambda|\mu| \lim_{T\to\infty} |w_{T,[rT]+k,[rT]+2} - 3w_{T,[rT]+k,[rT]+1} + 3w_{T,[rT]+k,[rT]} - w_{T,[rT]+k,[rT]-1}| \\
=&\lambda|\mu||f_\lambda(k-2) - 3f_\lambda(k-1) + 3f_\lambda(k) - f_\lambda(k+1)| \\
=&\lambda|\mu\Delta^3 f_\lambda(k+1)| \quad a.s.,
\end{aligned}$$

where the third equality follows from Lemma 1. □

Proof of Theorem 3: By Theorem 1 and the definition of $\hat{c}^m_{Tt}$ in Equation (8),

$$\begin{aligned}
&\sup_{t\in[\gamma T,(1-\gamma)T]} \| \hat{c}_{Tt}(y_1, y_2, \ldots, y_T) - w_{Tt1}\tilde{y}_{T1} - w_{Tt2}\tilde{y}_{T2} - w_{Tt,T-1}\tilde{y}_{T,T-1} - w_{TtT}\tilde{y}_{TT} - \hat{c}^m_{Tt} \|_p \\
=& \sup_{t\in[\gamma T,(1-\gamma)T]} \| \lambda\hat{\tau}_{Tt}(\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_T) - w_{Tt1}\tilde{y}_{T1} - w_{Tt2}\tilde{y}_{T2} - w_{Tt,T-1}\tilde{y}_{T,T-1} - w_{TtT}\tilde{y}_{TT} - \hat{c}^m_{Tt} \|_p \\
=& \lambda \sup_{t\in[\gamma T,(1-\gamma)T]} \| \sum_{s=3}^{T-2} w_{Tts}\tilde{y}_{Ts} - \sum_{s=3}^{T-2} w_{Tts}\tilde{y}_{Ts} I(|t-s| \le m) \|_p \\
\le& \lambda \sup_{t\in[\gamma T,(1-\gamma)T]} \sum_{s=3}^{T-2} |w_{Tts}| \, \| \tilde{y}_{Ts} \|_p \, I(|t-s| > m) I(T \ge m).
\end{aligned}$$

From the discussion in de Jong and Sakarya (2016) following their Theorem 1, it follows that w_{Tts} can be split into eight parts, as

$$w_{Tts} = f_{T\lambda}(t-s) + \sum_{j=2}^{8} w^j_{Tts}, \tag{23}$$

where $|f_{T\lambda}(0)| \leq 1$; for $m \in \{1, 2, \ldots, T\}$,

$$|f_{T\lambda}(m)| \leq Cm^{-3} \tag{24}$$

for some constant $C > 0$ independent of T; and

$$\sup_{T\geq 1} \sup_{1\leq s\leq T} \sup_{r\in[\gamma T,(1-\gamma)T]} |T^3 \sum_{j=2}^{8} w^j_{T,[rT],s}| < \infty. \tag{25}$$

For $m \geq 1$ and $T \geq m$, noting that $\sup_{3\leq s\leq T-2} \parallel \tilde{y}_{Ts} \parallel_p \leq 16 \sup_{s\geq 1} \parallel u_s \parallel_p$ if $\Delta^p y_t = u_t$ for $p = 1, 2, 3$, or 4,

$$\sup_{t\in[\gamma T,(1-\gamma)T]} \sum_{s=3}^{T-2} |w_{Tts}| \parallel \tilde{y}_{Ts} \parallel_p I(|t-s| > m) I(T \geq m)$$

$$\leq 16 \sup_{t\in[\gamma T,(1-\gamma)T]} \sum_{s=3}^{T-2} |f_{T\lambda}(t-s)| I(|t-s| > m) \sup_{s\geq 1} \parallel u_s \parallel_p$$

$$+16 \sup_{t\in[\gamma T,(1-\gamma)T]} \sum_{s=3}^{T-2} |\sum_{j=2}^{8} w^j_{Tts}| I(|t-s| > m) \sup_{s\geq 1} \parallel u_s \parallel_p I(T \geq m)$$

$$\leq 32 \sum_{j=m}^{T-2} |f_{T\lambda}(j)| \sup_{s\geq 1} \parallel u_s \parallel_p + 16 \sum_{s=3}^{T-2} \sup_{t\in[\gamma T,(1-\gamma)T]} |\sum_{j=2}^{8} w^j_{Tts}| \sup_{s\geq 1} \parallel u_s \parallel_p$$

$$\leq 32C \sup_{s\geq 1} \parallel u_s \parallel_p \sum_{j=m}^{\infty} j^{-3}$$

$$+16T^{-2} \sup_{s\geq 1} \parallel u_s \parallel_p \sup_{T\geq 1} \sup_{1\leq s\leq T} \sup_{t\in[\gamma T,(1-\gamma)T]} |T^3 \sum_{j=2}^{8} w^j_{Tts}| I(T \geq m)$$

$$= O(m^{-2}),$$

by the results of Equations (23)-(25). This shows the first assertion of the theorem. To show the second assertion,

$$\sup_{T\geq 1} \sup_{t\in[\gamma T,(1-\gamma)T]} \parallel \hat{c}_{Tt} - w_{Tt1}\tilde{y}_{T1} - w_{Tt2}\tilde{y}_{T2} - w_{Tt,T-1}\tilde{y}_{T,T-1} - w_{TtT}\tilde{y}_{TT} \parallel_p$$

$$\leq \lambda \sup_{T\geq 1} \sup_{t\in[\gamma T,(1-\gamma)T]} \| \sum_{s=3}^{T-2} w_{Tts}\tilde{y}_{Ts} \|_p$$
$$\leq 32\lambda \sum_{j=0}^{T-2} |f_{T\lambda}(j)| \sup_{s\geq 1} \| u_s \|_p + 16\lambda \sum_{s=3}^{T-2} \sup_{t\in[\gamma T,(1-\gamma)T]} |\sum_{j=2}^{8} w_{Tts}^j| \sup_{s\geq 1} \| u_s \|_p = O(1),$$

by a reasoning similar to that of the proof of the first assertion. □

Proof of Theorem 4: We write that

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

The first term is bounded in absolute value by $4\gamma \sup_{x\in\mathbb{R}} |g(x)|$ where γ can be chosen arbitrarily small. Therefore, it is sufficient to show that the second term vanishes as $T \to \infty$. Let $a_{Tt} = w_{Tt1}\tilde{y}_{T1} + w_{Tt2}\tilde{y}_{T2} + w_{Tt,T-1}\tilde{y}_{T,T-1} + w_{TtT}\tilde{y}_{TT}$ and note that

$$T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} |g(\hat{c}_{Tt}) - Eg(\hat{c}_{Tt})|$$
$$\leq T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} |g(\hat{c}_{Tt}) - g(\hat{c}_{Tt} - a_{Tt})| \tag{26}$$
$$+ T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} |g(\hat{c}_{Tt} - a_{Tt}) - Eg(\hat{c}_{Tt} - a_{Tt})| \tag{27}$$
$$+ T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} |Eg(\hat{c}_{Tt} - a_{Tt}) - Eg(\hat{c}_{Tt})|, \tag{28}$$

by the triangle inequality. For the expression in Equation (27), it is sufficient to show a weak law of large numbers. This result follows analogously to the proof of Theorem 6 of de Jong and Sakarya (2016). Therefore, we only need to show that the expressions in Equation (26) and (28) vanish in the limit. In order to do that, we first establish the result for the expression in Equation (28) then the result for the expression in Equation (26) will follow.

We write that

$$\begin{aligned}
&T^{-1}\sum_{t\in[\gamma T,(1-\gamma)T]}|Eg(\hat{c}_{Tt}-a_{Tt})-Eg(\hat{c}_{Tt})|\\
\leq&T^{-1}\sum_{t\in[\gamma T,(1-\gamma)T]}E|g(\hat{c}_{Tt}-a_{Tt})-g(\hat{c}_{Tt})|.
\end{aligned}$$

Since $g(\cdot)$ is Lipschitz continuous, we have

$$T^{-1}\sum_{t\in[\gamma T,(1-\gamma)T]}E|g(\hat{c}_{Tt})-g(\hat{c}_{Tt}-a_{Tt})|\leq LT^{-1}\sum_{t\in[\gamma T,(1-\gamma)T]}E|a_{Tt}|, \tag{29}$$

and by the definition of a_{Tt} and Theorem 1 of de Jong and Sakarya (2016), we have

$$\begin{aligned}
&E|a_{Tt}|\\
&\leq|w_{Tt1}|E|\tilde{y}_{T1}|+|w_{Tt2}|E|\tilde{y}_{T2}|+|w_{Tt,T-1}|E|\tilde{y}_{T,T-1}|+|w_{TtT}|E|\tilde{y}_{TT}|\\
&\leq(|f_{T\lambda}(t-1)|+|\sum_{k=2}^{8}w_{Tt1}^{k}|)E|\tilde{y}_{T1}|+(|f_{T\lambda}(t-2)|+|\sum_{k=2}^{8}w_{Tt2}^{k}|)E|\tilde{y}_{T2}|\\
&+(|f_{T\lambda}(t-T+1)|+|\sum_{k=2}^{8}w_{Tt,T-1}^{k}|)E|\tilde{y}_{T,T-1}|+(|f_{T\lambda}(t-T)|+|\sum_{k=2}^{8}w_{TtT}^{k}|)E|\tilde{y}_{TT}|\\
&\leq\sup_{s\in\{1,2,T-1,T\}}|f_{T\lambda}(t-s)|E(|\tilde{y}_{T1}|+|\tilde{y}_{T2}|+|\tilde{y}_{T,T-1}|+|\tilde{y}_{TT}|)\\
&+\sup_{s\in\{1,2,T-1,T\}}|\sum_{k=2}^{8}w_{Tts}^{k}|E(|\tilde{y}_{T1}|+|\tilde{y}_{T2}|+|\tilde{y}_{T,T-1}|+|\tilde{y}_{TT}|)\\
&\leq C_1T^{3/2}\sup_{s\in\{1,2,T-1,T\}}|f_{T\lambda}(t-s)|+C_2T^{3/2}\sup_{s\in\{1,2,T-1,T\}}|\sum_{k=2}^{8}w_{Tts}^{k}|, \qquad (30)
\end{aligned}$$

where the last line follows from the fact that $E(|\tilde{y}_{T1}|+|\tilde{y}_{T2}|+|\tilde{y}_{T,T-1}|+|\tilde{y}_{TT}|)=O(T^{3/2})$ under standard moment assumptions if y_t is integrated up to order 4. Therefore, we use the upper bound for $E|a_{Tt}|$ in Equation (30) to rewrite for the expression in Equation (29) as follows

$$\sum_{t\in[\gamma T,(1-\gamma)T]} E|g(\hat{c}_{Tt}) - g(\hat{c}_{Tt} - a_{Tt})|$$

$$\le C_1 L T^{3/2} T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} \sup_{s\in\{1,2,T-1,T\}} |f_{T\lambda}(t-s)|$$

$$+ C_2 L T^{3/2} T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} \sup_{s\in\{1,2,T-1,T\}} |\sum_{k=2}^{8} w_{Tts}^k|.$$

Note that

$$\lim_{T\to\infty} T^{3/2} T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} \sup_{s\in\{1,2,T-1,T\}} |f_{T\lambda}(t-s)|$$

$$\le \lim_{T\to\infty} T^{3/2} T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} \sup_{r\in[\gamma,(1-\gamma)]} \sup_{s\in\{1,2,T-1,T\}} |f_{T\lambda}([rT]-s)|$$

$$= \lim_{T\to\infty} T^{3/2} T^{-1} ((1-2\gamma)T+1) \sup_{r\in[\gamma,(1-\gamma)]} \sup_{s\in\{1,2,T-1,T\}} |f_{T\lambda}([rT]-s)|$$

$$\le C_3 \lim_{T\to\infty} T^{-3/2} T^{-1} ((1-2\gamma)T+1) \sup_{r\in[\gamma,(1-\gamma)]} \sup_{s\in\{1,2,T-1,T\}} T^3 |[rT]-s|^{-3}$$

$$= 0,$$

where the last inequality follows from the fact that $|f_{T\lambda}(m)| \le C|m|^{-3}$ for $m \in \{1, 2, \ldots, T\}$ (de Jong and Sakarya, 2016, Theorem 1). Similarly, we write that

$$\lim_{T\to\infty} T^{3/2} T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} \sup_{s\in\{1,2,T-1,T\}} |\sum_{k=2}^{8} w_{Tts}^k|$$

$$\le \lim_{T\to\infty} T^{3/2} T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} \sup_{r\in[\gamma,(1-\gamma)]} \sup_{s\in\{1,2,T-1,T\}} |\sum_{k=2}^{8} w_{T,[rT],s}^k|$$

$$= \lim_{T\to\infty} T^{-3/2} T^{-1} ((1-2\gamma)T+1) \sup_{r\in[\gamma,(1-\gamma)]} \sup_{s\in\{1,2,T-1,T\}} |T^3 \sum_{k=2}^{8} w_{T,[rT],s}^k|$$

$$= 0,$$

since $\sup_{T\geq 1}\sup_{1\leq s\leq T}|T^3\sum_{k=2}^{8} w_{T,[rT],s}| < \infty$ for any $r \in (0,1)$ by Equation (18) of de Jong and Sakarya (2016). Therefore, we have shown that the expression in Equation (28) vanishes as $T \to \infty$.

Lastly, we need to show that the expression in Equation (26) vanishes in the limit, which is implied by

$$\lim_{T\to\infty} T^{-1} \sum_{t\in[\gamma T,(1-\gamma)T]} E|g(\hat{c}_{Tt} - a_{Tt}) - g(c_{Tt})| = 0.$$

□

Lemma 2. *For $k, j \geq 0$,*

$$\lim_{T\to\infty} w_{T,T-k,j+1} = 0,$$

$$\lim_{T\to\infty} w_{T,T-k,T-j} = f_\lambda(k-j) + f_\lambda(k+j+1) + \xi_\lambda g_\lambda(k+1)g_\lambda(j+1).$$

Proof of Lemma 2. We use the definition of weights in Theorem 1 of de Jong and Sakarya (2016) and write that

$$\lim_{T\to\infty} w_{T,T-k,j+1}$$

$$= \lim_{T\to\infty} (f_{T\lambda}(T-k-j-1) + f_{T\lambda}(T)I(T-k+j=T))$$

$$+ \lim_{T\to\infty} (f_{T\lambda}(T-k+j)I(T-k+j<T) + f_{T\lambda}(T+k-j)I(T-k+j>T))$$

$$+ \lim_{T\to\infty} (\xi_{T\lambda}g_{T\lambda}(T-k)g_{T\lambda}(j+1) + \phi_{T\lambda}g_{T\lambda}(k+1)g_{T\lambda}(j+1))$$

$$+ \lim_{T\to\infty} (\phi_{T\lambda}g_{T\lambda}(T-k)g_{T\lambda}(T-j) + \xi_{T\lambda}g_{T\lambda}(k+1)g_{T\lambda}(T-j)) = 0,$$

since $\lim_{T\to\infty} f_{T\lambda}(T) = 0$, $\lim_{T\to\infty} g_{T\lambda}(T) = 0$, $\lim_{T\to\infty} g_{T\lambda}(j+1) = g_\lambda(j+1)$, $\lim_{T\to\infty}\xi_{T\lambda} = \xi_\lambda$, and $\lim_{T\to\infty}\phi_{T\lambda} = 0$ by Theorems 1 and 2 of de Jong and Sakarya (2016).

Similarly, we write that

$$\lim_{T\to\infty} w_{T,T-k,T-j}$$

$$
\begin{aligned}
=&\lim_{T\to\infty}\left(f_{T\lambda}(k-j)+f_{T\lambda}(T)I(2T-k-j-1=T)\right)\\
&+\lim_{T\to\infty} f_{T\lambda}(2T-k-j-1)I(2T-k-j-1<T)\\
&+\lim_{T\to\infty} f_{T\lambda}(k+j+1)I(2T-k-j-1>T)\\
&+\lim_{T\to\infty}\left(\xi_{T\lambda}g_{T\lambda}(T-k)g_{T\lambda}(T-j)+\phi_{T\lambda}g_{T\lambda}(k+1)g_{T\lambda}(T-j)\right)\\
&+\lim_{T\to\infty}\left(\phi_{T\lambda}g_{T\lambda}(T-k)g_{T\lambda}(j+1)+\xi_{T\lambda}g_{T\lambda}(k+1)g_{T\lambda}(j+1)\right)\\
=&f_\lambda(k-j)+f_\lambda(k+j+1)+\xi_\lambda g_\lambda(k+1)g_\lambda(j+1),
\end{aligned}
$$

since $\lim_{T\to\infty} f_{T\lambda}(m)=f_\lambda(m)$ and $\lim_{T\to\infty} g_{T\lambda}(m)=g_\lambda(m)$ for all $\lambda\geq 0$ and $m\in\mathbb{Z}$, and $\lim_{T\to\infty}\xi_{T\lambda}=\xi_\lambda$ by Theorems 2 and 3 of de Jong and Sakarya (2016). □

Proof of Theorem 5. First, we write that

$$
\begin{aligned}
&T^{-1}\sum_{t=1}^{T} g(\hat{c}_{Tt})\geq CT^{-1}\sum_{t=1}^{T}|\hat{c}_{Tt}|^p\\
&=CT^{-1}\sum_{t=1}^{T}\left|\lambda\sum_{s=1}^{T} w_{Tts}\tilde{y}_{Ts}\right|^p\\
&\geq C\lambda^p T^{-1}\left|\sum_{s=1}^{T} w_{TTs}\tilde{y}_{Ts}\right|^p,
\end{aligned}
$$

where the first inequality follows from the fact that $g(x)\geq C|x|^p$, and the first equality is due to Theorem 1. This gives the result in Equation (10).

Next, we will show that the results in Equations (11) and (12) hold. If y_t is an I(3) or an I(4) process, we have

$$
\begin{aligned}
\left|\sum_{s=1}^{T-2} w_{TTs}\tilde{y}_{Ts}\right| \leq& \sum_{s=1}^{T-2}|w_{TTs}||\tilde{y}_{Ts}|\\
\leq& C_1|T-1|^{-3}|\tilde{y}_{T1}|+C_1|T-2|^{-3}|\tilde{y}_{T2}|+\sum_{s=3}^{T-2}|T-s|^{-3}|\tilde{y}_{Ts}|,
\end{aligned}
$$

since $|w_{TTs}|\leq C_1|T-s|^{-3}$ for $s=1,2,\ldots,T-2$ by Theorem 1 of de Jong and Sakarya (2016). Furthermore, $\sup_{1\leq s\leq T-2} E|\tilde{y}_{Ts}|\leq C_2\sup_k E|u_k|<\infty$ by the definition of $\tilde{y}_{Ts}$ given in Theorem 1. Since $\sum_{s=3}^{T-2}|T-s|^{-3} \leq \sum_{m=2}^{\infty} m^{-3} < \infty$, $\sum_{s=3}^{T-2}|w_{TTs}||\tilde{y}_{Ts}| = O_p(1)$ a.s. This implies that

$$\left|\sum_{s=1}^{T-2} w_{TTs}\tilde{y}_{Ts}\right| = O_p(1) \quad \text{a.s.} \tag{31}$$

If y_t is an I(3) process, then

$$\begin{aligned}
&\left|\sum_{s=1}^{T} w_{TTs}T^{-1/2}\tilde{y}_{Ts}\right|^p \\
&= \left|\sum_{s=1}^{T-2} w_{TTs}T^{-1/2}\tilde{y}_{Ts} + w_{TT,T-1}T^{-1/2}\tilde{y}_{T,T-1} + w_{TTT}T^{-1/2}\tilde{y}_{TT}\right|^p \\
&= \left|w_{TT,T-1}T^{-1/2}\tilde{y}_{T,T-1} + w_{TTT}T^{-1/2}\tilde{y}_{TT} + O_p(T^{-1/2})\right|^p \\
&= \left|(w_{TTT} - w_{TT,T-1})T^{-1/2}\sum_{k=1}^{T} u_k + O_p(T^{-1/2})\right|^p,
\end{aligned}$$

where the second equality is implied by the expression in Equation (31), and the third equality follows from Theorem 1 and the assumption that $\Delta^3 y_t = u_t$, since $\tilde{y}_{T,T-1} = \Delta^2 y_{T-1} - 2\Delta^2 y_T = -\sum_{k=1}^{T} u_k - u_T$ and $\tilde{y}_{TT} = \Delta^2 y_T = \sum_{k=1}^{T} u_k$. Since $U_T(1) \xrightarrow{d} U(1)$, we obtain

$$\begin{aligned}
&\left|(w_{TTT} - w_{TT,T-1})T^{-1/2}\sum_{k=1}^{T} u_k + o_p(1)\right|^p \\
&\xrightarrow{d} \left|\left((f_\lambda(0) - f_\lambda(1)) + \xi_\lambda g_\lambda(1)(g_\lambda(1) - g_\lambda(2))\right) U(1)\right|^p,
\end{aligned}$$

by Lemma 2 and by the continuous mapping theorem. This shows the result in Equation (11).

If y_t is an I(4) process, then

$$\begin{aligned}
\left|\sum_{s=1}^{T} w_{TTs}T^{-3/2}\tilde{y}_{Ts}\right|^p &= \left|w_{TT,T-1}T^{-3/2}\tilde{y}_{T,T-1} + w_{TTT}T^{-3/2}\tilde{y}_{TT} + O_p(T^{-3/2})\right|^p \\
&= \left|(w_{TTT} - w_{TT,T-1})T^{-3/2}\sum_{k=1}^{T}\sum_{l=1}^{k} u_l + O_p(T^{-3/2})\right|^p,
\end{aligned}$$

where the first equality is implied by the expression in Equation (31). The second equality follows from Theorem 1 and the assumption that $\Delta^4 y_t = u_t$ which imply that $\tilde{y}_{T,T-1} = \Delta^2 y_{T-1} - 2\Delta^2 y_T = -\sum_{k=1}^{T}\sum_{l=1}^{k} u_l - \sum_{l=1}^{T} u_l$ and $\tilde{y}_{TT} = \Delta^2 y_T = \sum_{k=1}^{T}\sum_{l=1}^{k} u_l$. Since $U_T(r) \Rightarrow U(r)$ on $r \in [0,1]$, we obtain

$$\begin{aligned}
&\left|(w_{TTT} - w_{TT,T-1})T^{-3/2}\sum_{k=1}^{T}\sum_{l=1}^{k} u_l + o_p(1)\right|^p \\
\xrightarrow{d}&\left|((f_\lambda(0) - f_\lambda(1)) + \xi_\lambda g_\lambda(1)(g_\lambda(1) - g_\lambda(2)))\int_0^1 U(r)dr\right|^p.
\end{aligned}$$

by Lemma 2 and by the continuous mapping theorem. This completes the proof. □

Lemma 3. $\Delta^2(t+2)^p = \sum_{k=0}^{p-2} c_{pk}t^k$ *for* $t = 1, 2, \ldots, T$, *where* c_{pk} *is defined in Equation (13).*

Proof of Lemma 3. The Binomial Theorem gives the following equality

$$(t+m)^p = \sum_{k=0}^{p}\binom{p}{k}t^k m^{p-k}.$$

By the Binomial Theorem,

$$\begin{aligned}
\Delta^2(t+2)^p &= (t+2)^p - 2(t+1)^p + t^p \\
&= \sum_{k=0}^{p}\binom{p}{k}t^k 2^{p-k} - 2\sum_{k=0}^{p}\binom{p}{k}t^k + t^p \\
&= \sum_{k=0}^{p-1} c_{pk}t^k + (t^p - 2t^p + t^p), \\
&= \sum_{k=0}^{p-2} c_{pk}t^k,
\end{aligned}$$

where $c_{pk} = \binom{p}{k}(2^{p-k} - 2)$. The last equality follows from the fact that $c_{p,p-1} = 0$. □

Lemma 4. $\Delta^4(t+2)^p = \sum_{k=0}^{p-4} a_{pk}t^k$ *for* $t = 1, 2, \ldots, T$, *where* a_{pk} *is defined in Equation (14).*

Proof of Lemma 4. By the Binomial Theorem, we have

$$\Delta^4(t+2)^p$$

$$= (t+2)^p - 4(t+1)^p + 6t^p - 4(t-1)^p + (t-2)^p$$

$$= \sum_{k=0}^{p} \binom{p}{k} t^k 2^{p-k} - 4\sum_{k=0}^{p} \binom{p}{k} t^k + 6t^p$$

$$-4\sum_{k=0}^{p} \binom{p}{k} t^k (-1)^{p-k} + \sum_{k=0}^{p} \binom{p}{k} t^k (-2)^{p-k}$$

$$= \sum_{k=0}^{p-1} a_{p,k} t^k + (t^p - 4t^p + 6t^p - 4t^p + t^p)$$

$$= \sum_{k=0}^{p-1} a_{pk} t^k,$$

where

$$a_{pk} = \binom{p}{k} \left(2^{p-k} - 4 - 4(-1)^{p-k} + (-2)^{p-k}\right).$$

This implies that

$$a_{pk} = \begin{cases} \binom{p}{k} \left(2^{p-k+1} - 8\right) & \text{if } p-k \text{ is even} \\ 0 & \text{if } p-k \text{ is odd.} \end{cases}$$

Note that $a_{pk} = 0$ for $k = p-1$ and $k = p-3$, since $p-k$ is odd in both cases. It is also easy to see that $a_{p,p-2} = 0$. Thus, we conclude that

$$\Delta^4(t+2)^p = \sum_{k=0}^{p-4} a_{p,k} t^k.$$

□

Proof of Theorem 6. Let $y_t = t^p$ for $t = 1, 2, \ldots, T$. Theorem 1 implies that

$$\hat{c}_{Tt}\,(1, 2^p, \ldots, T^p) = \lambda \hat{\tau}_{Tt}(\tilde{y}_{T1}, \tilde{y}_{T2}, \ldots, \tilde{y}_{TT}),$$

where

$$\tilde{y}_{T1} = 3^p - 2^{p+1} + 1 = \sum_{k=0}^{p-2} c_{pk}, \tag{32}$$

$$\tilde{y}_{T2} = (4^p - 2 \times 3^p + 2^p) - 2(3^p - 2^{p+1} + 1) = \sum_{k=0}^{p-2} c_{pk}(2^k - 2), \tag{33}$$

$$\tilde{y}_{T,T-1} = ((T-1)^p - 2(T-2)^p + (T-3)^p) - 2(T^p - 2(T-1)^p + (T-2)^p)$$

$$= \sum_{k=0}^{p-2} c_{pk}((T-3)^k - 2(T-2)^k), \tag{34}$$

$$\tilde{y}_{TT} = T^p - 2(T-1)^p + (T-2)^p = \sum_{k=0}^{p-2} c_{pk}(T-2)^k, \tag{35}$$

by Lemma 3, and for $t = 3, 4, \ldots, T-2$,

$$\tilde{y}_{Tt} = \Delta^4 (t+2)^p = \sum_{k=0}^{p-4} a_{pk} t^k, \tag{36}$$

by Lemma 4.

Also, note that

$$\hat{\tau}_{Tt}(x_1 + y_1, x_2 + y_2, \ldots, x_T + y_T) = \hat{\tau}_{Tt}(x_1, x_2, \ldots, x_T) + \hat{\tau}_{Tt}(y_1, y_2, \ldots, y_T), \tag{37}$$

therefore,

$$\hat{c}_{Tt}(1, 2^p, \ldots, T^p) = \lambda\hat{\tau}_{Tt}(0, 0, \tilde{y}_{T3}, \ldots, \tilde{y}_{T,T-2}, 0, 0) + \lambda\hat{\tau}_{Tt}(\tilde{y}_{T1}, \tilde{y}_{T2}, 0, \ldots, 0, \tilde{y}_{T,T-1}, \tilde{y}_{TT}).$$

By replacing $\tilde{y}_{Tt}$ with the expression in Equation (36) for $t = 3, 4, \ldots, T-2$, we obtain

$$\begin{aligned}
&\hat{\tau}_{Tt}(0, 0, \tilde{y}_{T3}, \ldots, \tilde{y}_{T,T-2}, 0, 0) \\
&= \hat{\tau}_{Tt}(0, 0, \sum_{k=0}^{p-4} a_{pk}3^k, \ldots, \sum_{k=0}^{p-4} a_{pk}(T-2)^k, 0, 0) \\
&= \sum_{k=0}^{p-4} a_{pk}\hat{\tau}_{Tt}(0, 0, 3^k, \ldots, (T-2)^k, 0, 0),
\end{aligned}$$

where the last equality follows from the linearity of $\hat{\tau}_{Tt}$ (de Jong and Sakarya, 2016, Theorem 1). Similarly, by replacing $\tilde{y}_{T1}, \tilde{y}_{T2}, \tilde{y}_{T,T-1}$ and $\tilde{y}_{TT}$ with the expressions in Equations (32)-(35) and by the linearity of $\hat{\tau}_{Tt}$, we obtain

$$\begin{aligned}
&\hat{\tau}_{Tt}(\tilde{y}_{T1}, \tilde{y}_{T2}, 0, \ldots, 0, \tilde{y}_{T,T-1}, \tilde{y}_{TT}) \\
=&\sum_{k=0}^{p-2} c_{pk}\hat{\tau}_{Tt}(1, 2^k - 2, 0 \ldots, 0, (T-3)^k - 2(T-2)^k, (T-2)^k).
\end{aligned}$$

Thus, we have

$$\begin{aligned}
&\hat{c}_{Tt}(1, 2^p, \ldots, T^p) \\
&= \lambda \sum_{k=0}^{p-4} a_{pk}\hat{\tau}_{Tt}(0, 0, 3^k, \ldots, (T-2)^k, 0, 0) \\
&+ \lambda \sum_{k=0}^{p-2} c_{pk}\hat{\tau}_{Tt}(1, 2^k - 2, 0 \ldots, 0, (T-3)^k - 2(T-2)^k, (T-2)^k).
\end{aligned}$$

By using the identity in equation (37), we write that $\hat{\tau}_{Tt}(0, 0, 3^k, \ldots, (T-2)^k, 0, 0) = \hat{\tau}_{Tt}(1, 2^k, \ldots, T^k) - \hat{\tau}_{Tt}(1, 2^k, 0, \ldots, 0, (T-1)^k, T^k)$, which gives

$$\hat{c}_{Tt}(1, 2^p, \ldots, T^p)$$

$$= \lambda \sum_{k=0}^{p-4} a_{pk} \hat{\tau}_{Tt}(1, 2^k, \dots, T^k)$$

$$+ \lambda \sum_{k=0}^{p-2} c_{pk} \hat{\tau}_{Tt}(1, 2^k - 2, 0 \dots, 0, (T-3)^k - 2(T-2)^k, (T-2)^k)$$

$$- \lambda \sum_{k=0}^{p-4} a_{pk} \hat{\tau}_{Tt}(1, 2^k, 0, \dots, 0, (T-1)^k, T^k).$$

In order to conclude the proof, we write that

$$\hat{\tau}_{Tt}(1, 2^p, \dots, T^p)$$

$$= t^p - \hat{c}_{Tt}(1, 2^p, \dots, T^p)$$

$$= t^p - \lambda \sum_{k=0}^{p-4} a_{p,k} \hat{\tau}_{Tt}(1, 2^k, \dots, T^k)$$

$$- \lambda \sum_{k=0}^{p-2} c_{pk} \hat{\tau}_{Tt}(1, 2^k - 2, 0 \dots, 0, (T-3)^k - 2(T-2)^k, (T-2)^k)$$

$$+ \lambda \sum_{k=0}^{p-4} a_{pk} \hat{\tau}_{Tt}(1, 2^k, 0, \dots, 0, (T-1)^k, T^k).$$

□

Lemma 5. $\lambda z^4 - 4\lambda z^3 + (1 + 6\lambda)z^2 - 4\lambda z + \lambda = 0$ *has four roots* z_1, z_2, z_3 *and* z_4 *where*

$$z_1 = 1 - \frac{\sqrt{\sqrt{1+16\lambda} - 1}}{2\sqrt{2\lambda}} + i \left(\frac{\sqrt{2}}{\sqrt{\sqrt{1+16\lambda} - 1}} - \frac{1}{2\sqrt{\lambda}} \right), \tag{38}$$

and $z_2 = z_1^{-1}$, $z_3 = \bar{z}_1$, $z_4 = \bar{z}_1^{-1}$.

Proof of Lemma 5. Note that $\lambda z^4 - 4\lambda z^3 + (1+6\lambda)z^2 - 4\lambda z + \lambda = \lambda(z-1)^4 + z^2$. Therefore, we write that

$$\begin{aligned} \lambda(z-1)^4 + z^2 &= \left(\sqrt{\lambda}(z-1)^2 + iz\right)\left(\sqrt{\lambda}(z-1)^2 - iz\right) \\ &= 0. \end{aligned} \tag{39}$$

First, we consider $\sqrt{\lambda}(z-1)^2+iz=0$, which is equivalent to $\sqrt{\lambda}z^2+(i-2\sqrt{\lambda})z+\sqrt{\lambda}=0$. A tedious calculation shows that this quadratic equation is equivalent to

$$\sqrt{\lambda}(z-z_1)(z-z_2)=0,$$

where z_1 is defined in Equation (38) and $z_2=z_1^{-1}$.

By the fundamental theorem of algebra, the polynomial equation in Equation (39) has four complex roots where two of the roots are the complex conjugate of the other two. Therefore,

$$\lambda(z-1)^4+z^2=(z-z_1)(z-z_1^{-1})(z-z_3)(z-z_4),$$

where $z_3=\bar{z}_1$ and $z_4=\bar{z}_1^{-1}$. □

Proof of Theorem 7. For $t=3,4,\ldots,T-2$, the first order condition for $\hat{c}_{Tt}$ given in Equation (21) as

$$\lambda\hat{c}_{T,t+2}-4\lambda\hat{c}_{T,t+1}+(1+6\lambda)\hat{c}_{Tt}-4\lambda\hat{c}_{T,t-1}+\lambda\hat{c}_{T,t-2}=0.$$

The above equation is a fourth order difference equation, which has four roots z_1,z_2,z_3 and z_4 given in Lemma 5. It is possible to write that

$$\hat{c}_{Tt}=C_{1T}z_1^t+C_2z_1^{-t}+C_3\bar{z}_1^t+C_4\bar{z}_1^{-t},$$

by using the fact that $z_2=z_1^{-1}$, $z_3=\bar{z}_1$, and $z_4=\bar{z}_1^{-1}$, which are implied by Lemma 5.

Note that the HP filter always produces trends and cyclical components that are real as long as the original series is real. Thus,

$$\begin{aligned}\hat{c}_{Tt}&=\bar{\hat{c}}_{Tt},\\ C_{1T}z_1^t+C_2z_1^{-t}+C_3\bar{z}_1^t+C_4\bar{z}_1^{-t}&=\bar{C}_{1T}\bar{z}_1^t+\bar{C}_2\bar{z}_1^{-t}+\bar{C}_3z_1^t+\bar{C}_4z_1^{-t},\end{aligned}$$

which implies that $C_3=\bar{C}_{1T}$ and $C_4=\bar{C}_2$.

Furthermore, Theorem 1 implies that $\hat{c}_{Tt}(1,2^2,\ldots,T^2)=\lambda\hat{\tau}_{Tt}(1,-1,0,\ldots,0,-1,1)$

and $\hat{c}_{T,T-t+1}(1,2^2,\ldots,T^2) = \lambda\hat{\tau}_{T,T-t+1}(1,-1,0,\ldots,0,-1,1)$. Equation (20) of de Jong and Sakarya (2016) implies that $\hat{c}_{Tt}(1,2^2,\ldots,T^2) = \hat{c}_{T,T-t+1}(1,2^2,\ldots,T^2)$.[1] Therefore, we write that

$$\hat{c}_{Tt}(1,2^2,\ldots,T^2) = \hat{c}_{T,T-t+1}(1,2^2,\ldots,T^2),$$

$$C_{1T}z_1^t + C_2 z_1^{-t} + \bar{C}_{1T}\bar{z}_1^t + \bar{C}_2\bar{z}_1^{-t} = C_{1T}z_1^{T+1}z_1^{-t} + C_2 z_1^{-(T+1)}z_1^t + \bar{C}_{1T}\bar{z}_1^{T+1}\bar{z}_1^{-t} + \bar{C}_2\bar{z}_1^{-(T+1)}\bar{z}_1^t,$$

which implies that $C_2 = C_{1T}z_1^{T+1}$. This identity allows us to write that

$$\hat{c}_{Tt}(1,2^2,\ldots,T^2) = C_{1T}(z_1^t + z_1^{T-t+1}) + \bar{C}_{1T}(\bar{z}_1^t + \bar{z}_1^{T-t+1}). \tag{40}$$

Next, we use the first order conditions for $t = 1$ and 2, given in Equations (17) and (18), to solve for constant C_{1T}. Let

$$\begin{aligned} a_T &= (1+\lambda)(z_1 + z_1^T) - 2\lambda(z_1^2 + z_1^{T-1}) + \lambda(z_1^3 + z_1^{T-2}) \\ b_T &= -2\lambda(z_1 + z_1^T) + (1+5\lambda)(z_1^2 + z_1^{T-1}) - 4\lambda(z_1^3 + z_1^{T-2}) + \lambda(z_1^4 + z_1^{T-3}), \end{aligned}$$

then Equations (17) and (18) are equivalent to $C_{1T}a_T + \bar{C}_{1T}\bar{a}_T = 2\lambda$ and $C_{1T}b_T + \bar{C}_{1T}\bar{b}_T = -2\lambda$, respectively, when $y_t = t^2$ for $t = 1,2,\ldots,T$. and $\hat{c}_{Tt} = C_{1T}(z_1^t + z_1^{T-t+1}) + \bar{C}_{1T}(\bar{z}_1^t + \bar{z}_1^{T-t+1})$. These imply that $C_{1T} = 2\lambda(\bar{b}_T + \bar{a}_T)/(a_T\bar{b}_T - \bar{a}_T b_T)$.

By using the polar coordinate form, we write that $z_1 = |z_1|(\cos(\theta) + i\sin(\theta))$ where

$$\theta = \tan^{-1}\left(2^{1/2}(\sqrt{1+16\lambda} - 1)^{-1/2}\right).$$

Then, when we replace z_1 with its polar coordinate form in Equation (40), we obtain

$$\begin{aligned} \hat{c}_{Tt} =& (C_{1T} + \bar{C}_{1T})|z_1|^t\cos(t\theta) + i(C_{1T} - \bar{C}_{1T})|z_1|^t\sin(t\theta) \\ &+ (C_{1T} + \bar{C}_{1T})|z_1|^{T-t+1}\cos((T-t+1)\theta) + i(C_{1T} - \bar{C}_{1T})|z_1|^{T-t+1}\sin((T-t+1)\theta). \end{aligned}$$

□

[1] Equation (20) of de Jong and Sakarya (2016) states that $\hat{\tau}_{Tt}(y_1,y_2,\ldots,y_T) = \hat{\tau}_{T,T-t+1}(y_T,y_{T-1},\ldots,y_2,y_1)$.

Lemma 6. *$\Delta^2 \exp(t+2) = CC_1 \exp(t)$ and $\Delta^4 \exp(t+2) = C\exp(t)$ where $C = \exp(2)\,(1-\exp(-1))^4$ and $C_1 = (1-\exp(-1))^{-2}$.*

Proof of Lemma 6. Note that

$$\Delta^2 \exp(t+2)$$

$$= \exp(t+2) - 2\exp(t+1) + \exp(t)$$

$$= \exp(t+2)\,(1 - 2\exp(-1) + \exp(-2))$$

$$= \exp(t+2)\,(1-\exp(-1))^2$$

$$= CC_1 \exp(t).$$

Similarly, we can write that

$$\Delta^4 \exp(t+2)$$

$$= \exp(t+2) - 4\exp(t+1) + 6\exp(t) - 4\exp(t-1) + \exp(t-2)$$

$$= \exp(t+2)\,(1 - 4\exp(-1) + 6\exp(-2) - 4\exp(-3) + \exp(-4))$$

$$= \exp(t+2)\,(1-\exp(-1))^4$$

$$= C\exp(t).$$

□

Proof of Theorem 8. Let $y_t = \exp(t)$ for $t = 1, 2, \ldots, T$. By Theorem 1, we write that

$$\hat{c}_{Tt}\,(\exp(1), \exp(2), \ldots, \exp(T)) = \lambda \hat{\tau}_{Tt}\,(\tilde{y}_{T1}, \tilde{y}_{T2}, \ldots, \tilde{y}_{TT})\,,$$

where by Lemma 6

$$\tilde{y}_{T1} = \exp(3) - 2\exp(2) + \exp(1) = CC_1 \exp(1), \tag{41}$$

$$
\begin{aligned}
\tilde{y}_{T2} &= (\exp(4) - 2\exp(3) + \exp(2)) - 2(\exp(3) - 2\exp(2) + \exp(1)) \\
&= CC_1 \exp(2) - 2CC_1 \exp(1) \\
&= CC_1(1 - 2\exp(-1))\exp(2),
\end{aligned} \tag{42}
$$

$$
\begin{aligned}
\tilde{y}_{T,T-1} =& (\exp(T-1) - 2\exp(T-2) + \exp(T-3)) \\
& - 2(\exp(T) - 2\exp(T-1) + \exp(T-2)) \\
=& CC_1 \exp(T-3) - 2CC_1 \exp(T-2) \\
=& CC_1(\exp(-2) - 2\exp(-1))\exp(T-1) \\
=& CC_1((1-\exp(-1))^2 - 1)\exp(T-1) \\
=& C(1 - C_1)\exp(T-1),
\end{aligned} \tag{43}
$$

$$
\tilde{y}_{TT} = \exp(T) - 2\exp(T-1) + \exp(T-2) = CC_1 \exp(-2)\exp(T), \tag{44}
$$

and for $t = 3, 4, \ldots, T-2$,

$$
\begin{aligned}
\tilde{y}_{Tt} &= \exp(t) - 4\exp(t-1) + 6\exp(t-2) - 4\exp(t-3) + \exp(t-4) \\
&= C\exp(t).
\end{aligned} \tag{45}
$$

By using the expressions in Equations (41)-(45), we write that

$$
\begin{aligned}
& \hat{\tau}_{Tt}(\tilde{y}_{T1}, \tilde{y}_{T2}, \ldots, \tilde{y}_{TT}) \\
=& \hat{\tau}_{Tt}(CC_1 \exp(1), CC_2 \exp(2), C\exp(3), \ldots, C\exp(T-2), CC_3 \exp(T-1), CC_4 \exp(T)) \\
=& C\hat{\tau}_{Tt}(C_1 \exp(1), C_2 \exp(2), \exp(3), \ldots, \exp(T-2), C_3 \exp(T-1), C_4 \exp(T)),
\end{aligned}
$$

where the last line follows by the linearity of $\hat{\tau}_{Tt}$, and $C_2 = (1-2\exp(-1))C_1$, $C_3 = (1-C_1)$ and $C_4 = C_1 \exp(-2)$.

Therefore, by Theorem 1 we have

$$
\hat{c}_{Tt}(\exp(1), \ldots, \exp(T))
$$

$$=C\lambda\hat{\tau}_{Tt}(C_1\exp(1), C_2\exp(2), \exp(3), \ldots, \exp(T-2), C_3\exp(T-1), C_4\exp(T)).$$

□

Proof of Theorem 9. By Theorem 8

$$\begin{aligned}
&\hat{c}_{Tt}(\exp(1), \ldots, \exp(T)) \\
=&C\lambda\hat{\tau}_{Tt}(C_1\exp(1), C_2\exp(2), \exp(3), \ldots, \exp(T-2), C_3\exp(T-1), C_4\exp(T)) \\
=&C\lambda\hat{\tau}_{Tt}(\exp(1), \exp(2), \ldots, \exp(T)) \\
+&C\lambda\hat{\tau}_{Tt}((C_1-1)\exp(1), (C_2-1)\exp(2), 0, \ldots, 0, (C_3-1)\exp(T-1), (C_4-1)\exp(T)),
\end{aligned}$$

because for any sequences $\{x_t\}_{t=1}^T$ and $\{y_t\}_{t=1}^T$, $\hat{\tau}_{Tt}(\{x_t+y_t\}_{t=1}^T) = \hat{\tau}_{Tt}(\{x_t\}_{t=1}^T)+\hat{\tau}_{Tt}(\{y_t\}_{t=1}^T)$ for all $t = 1, 2, \ldots, T$.

Evaluating the above expression at $t = T-k$, dividing it by $\hat{\tau}_{T,T-k}$, and taking the limit as $T \to \infty$ gives

$$\begin{aligned}
&\lim_{T\to\infty}\frac{\hat{c}_{T,T-k}(\exp(1), \exp(2), \ldots, \exp(T))}{\hat{\tau}_{T,T-k}(\exp(1), \exp(2), \ldots, \exp(T))} = C\lambda \\
&+ C\lambda\lim_{T\to\infty}\frac{\hat{\tau}_{T,T-k}((C_1-1)\exp(1), (C_2-1)\exp(2), 0, \ldots, 0, (C_3-1)\exp(T-1), (C_4-1)\exp(T))}{\hat{\tau}_{T,T-k}(\exp(1), \exp(2), \ldots, \exp(T))}.
\end{aligned}$$

By using the weighted average representation of the HP filter trend given in Theorem 1 of de Jong and Sakarya (2016), we have

$$\begin{aligned}
&\lim_{T\to\infty}\frac{\hat{\tau}_{T,T-k}((C_1-1)\exp(1), (C_2-1)\exp(2), 0, \ldots, 0, (C_3-1)\exp(T-1), (C_4-1)\exp(T))}{\hat{\tau}_{T,T-k}(\exp(1), \exp(2), \ldots, \exp(T))} \\
=&(C_1-1)\exp(1)\lim_{T\to\infty}\frac{w_{T,T-k,1}}{\sum_{s=1}^T w_{T,T-k,s}\exp(s)} &(46)\\
&+(C_2-1)\exp(2)\lim_{T\to\infty}\frac{w_{T,T-k,2}}{\sum_{s=1}^T w_{T,T-k,s}\exp(s)} &(47)\\
&+(C_3-1)\lim_{T\to\infty}\frac{\exp(T-1)w_{T,T-k,T-1}}{\sum_{s=1}^T w_{T,T-k,s}\exp(s)}+(C_4-1)\lim_{T\to\infty}\frac{\exp(T)w_{T,T-k,T}}{\sum_{s=1}^T w_{T,T-k,s}\exp(s)}. &(48)
\end{aligned}$$

The expression in Equation (46) can be written as

$$(C_1 - 1)\exp(1)\lim_{T\to\infty}\frac{\exp(-T)w_{T,T-k,1}}{\sum_{s=1}^{T} w_{T,T-k,s}\exp(s-T)}$$

$$= (C_1 - 1)\exp(1)\lim_{T\to\infty}\frac{\exp(-T)w_{T,T-k,1}}{\sum_{j=0}^{T-1} w_{T,T-k,T-j}\exp(-j)} = 0,$$

where the first equality is due to $j = T - s$. The result follows from the fact that $\lim_{T\to\infty} w_{T,T-k,1} = 0$, and $\lim_{T\to\infty}\sum_{j=0}^{T-1} w_{T,T-k,T-j}\exp(-j) = \sum_{j=0}^{\infty}(f_\lambda(k-j) + f_\lambda(k+j+1) + \xi_\lambda g_\lambda(k+1)g_\lambda(j+1))\exp(-j)$ by Lemma 2.

The expression in Equation (47) is zero by an argument similar to the one above.

Setting $j = T - s$, then the first term in Equation (48) equals

$$(C_3 - 1)\exp(-1)\lim_{T\to\infty}\frac{w_{T,T-k,T-1}}{\sum_{j=0}^{T-1} w_{T,T-k,T-j}\exp(-j)}$$

$$= (C_3 - 1)\exp(-1)\frac{f_\lambda(k-1) + f_\lambda(k+2) + \xi_\lambda g_\lambda(k+1)g_\lambda(2)}{\sum_{j=0}^{\infty}(f_\lambda(k-j) + f_\lambda(k+j+1) + \xi_\lambda g_\lambda(k+1)g_\lambda(j+1))\exp(-j)},$$

and the last term in Equation (48) equals

$$(C_4 - 1)\lim_{T\to\infty}\frac{w_{T,T-k,T}}{\sum_{j=0}^{T-1} w_{T,T-k,T-j}\exp(-j)}$$

$$= (C_4 - 1)\frac{f_\lambda(k) + f_\lambda(k+1) + \xi_\lambda g_\lambda(k+1)g_\lambda(1)}{\sum_{j=0}^{\infty}(f_\lambda(k-j) + f_\lambda(k+j+1) + \xi_\lambda g_\lambda(k+1)g_\lambda(j+1))\exp(-j)},$$

by Lemma 2. This completes the proof. □