

# A property of the Hodrick-Prescott filter and its application

Neslihan Sakarya\*      Robert M. de Jong<sup>†</sup>

September 28, 2019

## Abstract

This paper explores a simple property of the Hodrick-Prescott filter: when the HP filter is applied to a series, the cyclical component is equal to the HP filtered trend of the fourth difference of the series, except for the first and last two observations, for which different formulas are needed. We use this result to derive small sample results and asymptotic results for a fixed smoothing parameter. We first apply this property to analyze the consequences of a deterministic break. We find that the effect of a deterministic 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. Second, we apply this property 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 series that are integrated up to order 3 or 4 is more subtle. Third, we characterize the behavior of the Hodrick-Prescott filter when applied to deterministic polynomial trends and show that in the middle of the sample, the cyclical component reduces the order of the polynomial by 4, while the end point behavior is different. 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. Our results are compared to those of Phillips and Jin (2015).

*Keywords:* Hodrick-Prescott filter, cyclical component, structural break, weak dependence, deterministic trend.

---

\*Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK, Email [ns17003@essex.ac.uk](mailto:ns17003@essex.ac.uk).

<sup>†</sup>Department of Economics, Ohio State University, 444 Arps Hall, Columbus, OH 43210, USA, Email [de-jong.8@osu.edu](mailto: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 Phillips and Jin (2002), Phillips (2010), Phillips and Jin (2015), de Jong and Sakarya (2016), Cornea-Madeira (2017), and Hamilton (2018). These papers analyze the properties of the HP filter rigorously and reconsider its usefulness in the context of macroeconomics. Note that Phillips and Jin (2002) and Phillips and Jin (2015) do not only study the HP filter, but the more general Whittaker filter.

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

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

over  $\tau = (\tau_1, \dots, \tau_T)$ . The parameter  $\lambda$  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  $y_t$ , 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 (2017) also provides an exact formula for the weights by using the Sherman-Morrison formula.

In this paper, we derive a property of the HP filter which allows us to derive more general results that are not present in the literature. A feature of our results is that we will assume  $\lambda$  to be fixed with  $T$ . If we interpret  $\lambda$  as a bandwidth type parameter, the assumption that  $\lambda$  grows with  $T$  is natural. This case is studied in detail in Phillips and Jin (2002) and Phillips and Jin (2015).<sup>1</sup> In our setting  $\lambda$  is assumed to be fixed, even while deriving our asymptotic results. One justification for this approach is that regardless of the sample size, in the macro literature values for  $\lambda$  are fixed; typically  $\lambda = 1600$  for quarterly data or  $\lambda = 6.25$  for annual data.

We first use this elegant result to explore the effect of an intercept break on the cyclical component. We show that the effect of a structural break on the cyclical component dies out as one moves a finite number of observations away from the break point, also asymptotically. If we specialize the general setting of Phillips and Jin (2015) to the case of an intercept break, their results consider the case where the size of the break grows at rate  $T^{1/2}$  and  $\lambda$  grows with  $T$ , while we assume that the break size and  $\lambda$  are fixed with  $T$ .

Our main result is also applied to processes that are integrated up to order 4. Again assuming a fixed value for  $\lambda$ , we show that the cyclical component of series that are integrated up to order 2 exhibits 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

---

<sup>1</sup>Note that Phillips and Jin (2002) and Phillips and Jin (2015) also contain fixed  $\lambda$  and  $T$  results for the more general case of the Whittaker filter.

of the cyclical component. This result contrasts with Phillips and Jin (2002) and Phillips and Jin (2015) for the case  $\lambda = O(T^4)$ , where it is shown that for this case, the HP filter “fails to eliminate” the unit root. Our result shows that for fixed  $\lambda$ , King and Rebelo (1993)’s widely known conjecture “... the HP filter will render stationary series that are integrated (up to fourth order) ...” is incorrect in the sense that the law of large numbers does not hold for a large class of functions of the cyclical component when the original series is integrated of order 3 or 4.<sup>2</sup> The reasoning used by King and Rebelo is based on assuming that the first order condition of the minimization problem in Equation (1) that is valid for  $t = 3, 4, \dots, T - 2$  holds for every  $t \in \mathbb{Z}$ . We call this approach the “heuristic approach” throughout the paper. Our representation provides a rigorous analysis of the HP filter that does not ignore begin and end point issues. In addition, we give a closed form formula for the cyclical component of a polynomial trend for fixed  $\lambda$ . Phillips and Jin (2015) again consider the large  $\lambda$  case. We also consider the case of 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. In Section 5, we discuss weak dependence properties of the cyclical component when the HP filter is applied to I(2), I(3), or I(4) series. 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

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

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

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

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

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

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

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

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

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). 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

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

---

<sup>2</sup>Phillips and Jin (2002) were the first to note that King and Rebelo (1993)’s conjecture is incorrect for the case  $\lambda = O(T^4)$ .

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 4.

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.

### 3 Main result

The following theorem presents a property of the HP filter.

**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  $t = 3, 4, \dots, T - 2$ ,  $\tilde{y}_{Tt} = \Delta^4 y_{t+2}$ . Then for  $t = 1, 2, \dots, T$*

$$\hat{c}_{Tt}(y_1, y_2, \dots, y_T) = \lambda \hat{\tau}_{Tt}(\tilde{y}_{T1}, \tilde{y}_{T2}, \dots, \tilde{y}_{TT}). \quad (8)$$

This simple but elegant result provides insights into the structure of the cyclical term. This result shows that when the HP filter is applied to a series, the cyclical component of the series equals the trend of the fourth difference of the series, where the first and last two observations need a different formula. This property can shed light on the behavior of the cyclical component that is obtained from various data generating processes. As pointed out by a perceptive referee, this property can also be established using matrix algebra. Given the relative simplicity of the minimization problem involved, it is perhaps not surprising that matrix based treatments of the HP filter (such as e.g. in Phillips and Jin (2015), in particular Remark 3) come close to establishing the result of Theorem 1.

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

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

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 may be relatively large, because in the formulas for these objects the second difference instead of the fourth difference shows up. In Section 5, we will elaborate on 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 that are 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 structural breaks to the cyclical component. Specifically, we focus on an intercept break that is assumed to occur at an unknown date in the middle of the sample. The next result formalizes the effect of an intercept break.

**Theorem 2.** *Let*

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

where  $4 \leq [rT] \leq T - 5$ ,  $r \in (0, 1)$  and  $[\cdot]$  is the floor function. Then for  $t = 1, 2, \dots, T$

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

where  $w_{Tts}$  is defined in Equation (17) of Appendix 1, and for  $k \in \mathbb{Z}$

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

where  $f_\lambda(\cdot)$  is defined in Equation (22) of Appendix 1.

The first result shows that the presence of the structural break alters the cyclical component  $\hat{c}_{Tt}$  by  $-\lambda\mu\Delta^3 w_{Tt, [rT]+2}$ . From results (5) and (6) of Appendix 1, we know that  $|f_\lambda(k)| \leq C|k|^{-3}$  for  $|k| \geq 1$ ; therefore, the cyclical component  $\hat{c}_{T, [rT]+k}$  for values of  $k$  that are away from zero is not affected much asymptotically by the structural break. However, for small values of  $k$ , 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 result (6) of Appendix 1. In Figure 1, the effect of an intercept break on the cyclical component is illustrated by plotting  $\lambda\mu\Delta^3 f_\lambda(k+1)$  for  $\lambda = 1600$  and  $\mu = 1$ . It shows that the structural break mainly impacts  $\hat{c}_{Tt}$ 's that are less than 10 time points away from the structural break point. Note that the problem of analyzing multiple breaks can be split into the analyses of single breaks by the additivity of the HP filter cyclical component.

Phillips and Jin (2015) consider a much more general form of trend and analyze the effect of a structural break. They assume  $\lambda = \mu T^4$  where  $\mu \in (0, \infty)$  and consider the behavior of the trend at a distance  $k = O(T)$  away from the break. Our result assumes both  $\lambda$  and  $k$  constant.

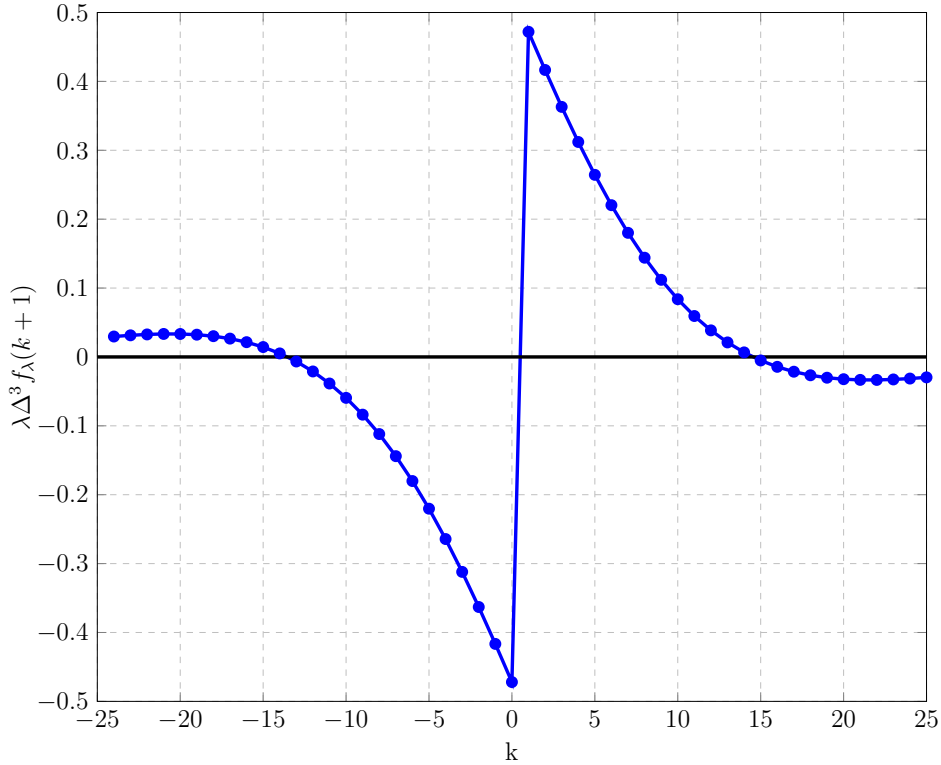


Figure 1: The effect of a deterministic trend break of size 1 when  $\lambda = 1600$

## 5 The HP filter when applied to integrated processes

We consider the weak dependence properties of the cyclical component 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 that are 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  $\hat{c}_{Tt}$  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  $\hat{c}_{Tt}^m$  for  $t = 1, 2, \dots, T$  as

$$\hat{c}_{Tt}^m = \lambda \sum_{s=3}^{T-2} w_{Tts} \tilde{y}_{Ts} I(|t-s| \leq m). \quad (9)$$

Note that  $\tilde{y}_{Ts} = \Delta^4 y_{s+2}$  for  $s = 3, \dots, T-2$ . Since  $\Delta^4 y_{s+2} = \Delta^{4-q} u_{s+2}$  for  $q = 1, 2, 3, 4$  is a function of  $u_{s-2+q}, \dots, u_{s+2}$ , the approximator  $\hat{c}_{Tt}^m$  depends only on  $u_{t-2-m+q}, \dots, u_{t+2+m}$ . 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. Everywhere in this paper,  $\|\cdot\|_p$  denotes  $(E|\cdot|^p)^{1/p}$ .

**Theorem 3.** *Let  $p \geq 1$ . 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  $m \geq 1$ ,*

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

and

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

The term

$$w_{Tt1} \tilde{y}_{T1} + w_{Tt2} \tilde{y}_{T2} + w_{Tt,T-1} \tilde{y}_{T,T-1} + w_{TtT} \tilde{y}_{TT} \quad (10)$$

will be small under standard conditions if  $y_t$  is integrated up to order 4. This is because I(2) processes are  $O_p(T^{3/2})$  under standard conditions (such as when a functional central limit theorem holds for the  $u_t$ ), while from results (5) and (8) of Appendix 1, it follows that  $|w_{Ttj}| \leq \tilde{C}T^{-3}$  for  $j = 1, 2, T-1, T$  and  $t \in [\gamma T, (1-\gamma)T]$ . Therefore,  $\hat{c}_{Tt}$  can be approximated well using  $\hat{c}_{Tt}^m$ , which is a function of the  $u_s$  for values of  $s$  close to  $t$ , and therefore has weak dependence properties. This reasoning can be used to establish 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  $E(|\tilde{y}_{1T}| + |\tilde{y}_{2T}| + |\tilde{y}_{T-1,T}| + |\tilde{y}_{TT}|) = O(T^{3/2})$ . Let  $g(\cdot)$  be a function that is bounded and Lipschitz continuous on  $\mathbb{R}$ . Then*

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

The result above can be viewed as a formalization of King and Rebelo's conjecture. Note that in the above result, the term of Equation (10) plays no role asymptotically as long as  $g(\cdot)$  is a bounded function. In the case that  $g(\cdot)$  is an unbounded function, the fact that  $\tilde{y}_{TT}$  is asymptotically large if  $y_t$  is integrated of order 3 or 4 can break the law of large numbers for functions of the cyclical component. The following theorem formalizes this.

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

$$T^{-1} \sum_{t=1}^T g(\hat{c}_{Tt}) \geq C\lambda^p T^{-1} \left| \sum_{s=1}^T w_{TTs} \tilde{y}_{Ts} \right|^p. \quad (11)$$

If  $y_t$  is an  $I(3)$  process, then

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

If  $y_t$  is an  $I(4)$  process, then

$$T^{-3p/2} \left| \sum_{s=1}^T w_{TTs} \tilde{y}_{Ts} \right|^p \xrightarrow{d} |((f_\lambda(0) - f_\lambda(2)) + \xi_\lambda g_\lambda(1)(g_\lambda(1) - g_\lambda(2))) \int_0^1 U(r) dr|^p, \quad (13)$$

where  $g_\lambda(\cdot)$  and  $\xi_\lambda$  are defined in results (7) and (8) of Appendix 1, respectively.

It is easy to verify that  $f_\lambda(0) - f_\lambda(2) + \xi_\lambda g_\lambda(1)(g_\lambda(1) - g_\lambda(2))$  takes the value of 0.254, 0.022, and 0.002 for  $\lambda = 6.25$ , 1600, and 129,600, respectively.<sup>3</sup> 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  $p \geq 2$ , the law of large numbers does not hold for  $g(\hat{c}_{Tt})$ . 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 again implies that the law of large numbers does not hold for  $g(\hat{c}_{Tt})$  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 the cyclical component when the HP filter is applied to  $I(3)$  or  $I(4)$  processes.

Phillips and Jin (2015)'s Theorem 4 considers the  $I(1)$  case. For the case  $\lambda = \mu T^4$ , they conclude in Remark 9 that the scaled partial sum process of the cyclical component satisfies a functional central limit theorem. For the case of smaller  $\lambda$ , Phillips and Jin (2015) conclude that the HP filter "... eliminates the stochastic trend giving a zero 'cyclical' process."

## 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  $\tilde{y}_{Ts} = 0$  for  $t = 1, \dots, T$ , which by Theorem 1 implies that  $\hat{c}_{Tt} = 0$ . For higher order polynomials, the result is more complex. In Theorem 6 and below, we adopt the convention that the summations over empty index sets are zero.

<sup>3</sup>The Matlab program used to calculate these values can be found at [https://neslihansakarya.weebly.com/uploads/5/9/5/5/59554687/matlab\\_files\\_for\\_theorem\\_5.zip](https://neslihansakarya.weebly.com/uploads/5/9/5/5/59554687/matlab_files_for_theorem_5.zip)



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

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

where

$$c_{pk} = \binom{p}{k} (2^{p-k} - 2), \quad (14)$$

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

$$a_{pk} = \begin{cases} \binom{p}{k} (2^{p-k+1} - 8) & \text{if } p - k \text{ is even} \\ 0 & \text{if } p - k \text{ is odd,} \end{cases} \quad (15)$$

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

Note that for  $p = 2$ ,  $\hat{c}_{Tt} = \lambda C_{Tt2} = 2\lambda \hat{\tau}_{Tt}(1, -1, 0, \dots, 0, -1, 1) = 2\lambda(w_{Tt1} - w_{Tt2} - w_{Tt,T-1} + w_{TtT})$ . This result and results (5) and (8) of Appendix 1 together imply that  $\hat{c}_{Tt}$  takes a value close to zero if  $t$  is sufficiently away from the begin and end points. In the case of a cubic trend, Theorem 6 gives  $\hat{c}_{Tt} = \lambda C_{Tt3} = 6\lambda \hat{\tau}_{Tt}(2, -1, 0, \dots, 0, -T, (T-1)) = 6\lambda(2w_{Tt1} - w_{Tt2} - Tw_{Tt,T-1} + (T-1)w_{TtT})$ , 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, by results (5) and (8) of Appendix 1 and therefore,  $\hat{c}_{Tt}$  is roughly equal to  $24\lambda$  in that case. However, since by Lemma 2, for fixed  $k \geq 0$ ,  $\lim_{T \rightarrow \infty} T^{-2} C_{T,T-k,4}$  exists and does not equal 0,  $C_{T,T-k,4}$  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. The case  $p > 4$  follows a similar pattern.

Next, we provide an explicit formula for the cyclical component of a quadratic trend. In Theorem 7 and below,  $\bar{z}$  denotes the complex conjugate of  $z$ .

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

$$\begin{aligned} \hat{c}_{Tt} = & (C_{1T} + \bar{C}_{1T})|z_1|^t \cos(t\theta) + i(C_{1T} - \bar{C}_{1T})|z_1|^t \sin(t\theta) \\ & + (C_{1T} + \bar{C}_{1T})|z_1|^{T-t+1} \cos((T-t+1)\theta) \\ & + i(C_{1T} - \bar{C}_{1T})|z_1|^{T-t+1} \sin((T-t+1)\theta), \end{aligned} \quad (16)$$

where  $i^2 = -1$ ,

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

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

$$\begin{aligned}
C_{1T} &= 2\lambda (\bar{b}_T + \bar{a}_T) / (a_T \bar{b}_T - \bar{a}_T b_T), \\
z_1 &= 1 - \frac{\sqrt{\sqrt{1+16\lambda}-1}}{2\sqrt{2\lambda}} + i \left( \frac{\sqrt{2}}{\sqrt{\sqrt{1+16\lambda}-1}} - \frac{1}{2\sqrt{\lambda}} \right), \\
\theta &= \tan^{-1} \left( 2^{1/2} \left( \sqrt{1+16\lambda}-1 \right)^{-1/2} \right).
\end{aligned}$$

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 the sample, the last two terms in Equation (16) take over. When  $t$  is in the middle of a large sample,  $\hat{c}_{Tt}$  takes relatively smaller values since all terms in Equation (16) are small. This is because  $|z_1| < 1$  for  $0 < \lambda < \infty$ . Also, it is easy to see that  $\hat{c}_{Tt}(1, 2^2, \dots, T^2) = \hat{c}_{T, T-t+1}(1, 2^2, \dots, T^2)$  for  $t = 1, 2, \dots, T$ . 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:

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

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  $C_{2T}^*$ .

Phillips and Jin (2015)'s Theorem 4 and the subsequent discussion provide an asymptotic analysis of a general higher order polynomial case, and they conclude that "... the asymptotic forms of the filters project the higher order polynomials onto lower order polynomials ... and apply the smoother to the residual process." However, the results of this section considered fixed  $T$  results.

## 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  $t = 1, 2, \dots, T$ . Then*

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

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  $C_4 = C_1 \exp(-2)$ .

Since the above result implies that 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. As was noted by the Editor, given the fact that differencing an exponential trend gives a multiple of an exponential trend, and given the role of the fourth difference in the HP filter, it is not surprising that the cyclical component retains an exponential trend (the exact nature of which is given in Theorem 8). Results such as Phillips and Jin (2015)'s Theorem 1 and Remark 3 can also be shown to imply the result of Theorem 8.

The next result shows that the cyclical component of an exponentially deterministic trend is as explosive as the exponential trend itself when  $t$  is close to the end of the sample.

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

$$\begin{aligned} & \lim_{T \rightarrow \infty} \frac{\hat{c}_{T, T-k}(\exp(1), \dots, \exp(T))}{\exp(T-k)} \\ &= C\lambda \exp(k) \sum_{j=0}^{\infty} (f_{\lambda}(k-j) + f_{\lambda}(k+j+1) + \xi_{\lambda} g_{\lambda}(k+1) g_{\lambda}(j+1)) \exp(-j) \\ & \quad + C\lambda \exp(k-1)(C_3 - 1) (f_{\lambda}(k-1) + f_{\lambda}(k+2) + \xi_{\lambda} g_{\lambda}(k+1) g_{\lambda}(2)) \\ & \quad + C\lambda \exp(k)(C_4 - 1) (f_{\lambda}(k) + f_{\lambda}(k+1) + \xi_{\lambda} g_{\lambda}(k+1) g_{\lambda}(1)). \end{aligned}$$

This result shows that the HP filter is not capable of removing the trend in deterministic exponential trends, in the sense that  $\hat{c}_{Tt}$  is as explosive as  $y_t = \exp(t)$  when  $t$  is close to the end of the sample. For  $\lambda = 1600$  and  $k = 0, 1, \dots, 5$ , the limit in Theorem 9 takes the values 0.703, 0.277, -0.739, -3.130, -8.679 and -21.363, respectively.<sup>4</sup>

In the setting of an exponential time trend, a trivial remedy is to take the logarithm of  $y_t$ , 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  $y_t$  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. We use this result to derive small sample results and fixed lambda asymptotic results.

Our main result is used to analyze the effect of an intercept break. We find that an intercept 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 integrated processes of order up to 4. We conclude that the cyclical component of series that are integrated of order 1 or 2 possesses weak dependence properties and that the law of large numbers holds for transformations of the cyclical component. On the other hand, the situation is more subtle when the HP filter is applied to processes that are integrated of order 3 or 4. We find that in this case, the law of large numbers fails to hold for unbounded transformations of the cyclical component. Our main result allows us to derive a closed form formula for the cyclical component of deterministic polynomial trends. It is shown that the HP filter reduces the order of the polynomial by 4 in the middle of the sample. We show that the HP filter is not capable of removing the trend in deterministic exponential trends.

Results such as those of Phillips and Jin (2002), Phillips and Jin (2015) and this paper should help economists gain a full understanding of the issues involved in the analysis of HP filtered macroeconomic data. In particular, these results can inform applied macroeconomists about the effect of applying the HP filter in the presence of structural breaks, unit roots, and/or deterministic polynomial or exponential trends.

## References

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

---

<sup>4</sup>The Matlab program used to calculate these values can be found at [https://neslihansakarya.weebly.com/uploads/5/9/5/5/59554687/matlab\\_files\\_for\\_theorem\\_9.zip](https://neslihansakarya.weebly.com/uploads/5/9/5/5/59554687/matlab_files_for_theorem_9.zip)

- 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.
- 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. (2017): “The Explicit Formula for the Hodrick-Prescott Filter in a Finite Sample,” *Review of Economics and Statistics*, 99, 314–318.
- 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. (2018): “Why you should never use the Hodrick-Prescott filter,” *Review of Economics and Statistics*, 100, 831–843.
- 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. B. (2010): “Two New Zealand pioneer econometricians,” *New Zealand Economic Papers*, 44, 1–26.
- PHILLIPS, P. C. B. AND S. JIN (2002): “Limit Behavior of the Whittaker (HP) Filter,” Working Paper, Yale University.
- (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.
- WHITTAKER, E. T. (1923): “On a new method of graduation,” *Proceedings of the Edinburgh Mathematical Society*, 41, 63–75.

# Appendix 1

In this section, we introduce the notation and the results from de Jong and Sakarya (2016) that are used throughout this paper.

1. For  $m \in \mathbb{Z}$ ,  $\lambda \in [0, \infty)$ , and  $T \geq 1$ ,

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

2. For  $m \geq 1$ ,  $\lambda \in [0, \infty)$ , and  $T \geq 1$ ,

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

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

3. The sequences  $\delta_{T\lambda}$ ,  $\eta_{T\lambda}$ ,  $\xi_{T\lambda}$  and  $\phi_{T\lambda}$  are defined as

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

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

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

4. When the HP filter is applied to a time series  $y_t$ , it calculates a trend that is the weighted average of the original series such that  $\hat{\tau}_{Tt} = \sum_{s=1}^T w_{Tts} y_s$  for  $t = 1, 2, \dots, T$  where

$$w_{Tts} = f_{T\lambda}(t-s) + f_{T\lambda}(T)I(t+s-1=T) \\ + f_{T\lambda}(t+s-1)I(t+s-1 < T) + f_{T\lambda}(2T-t-s+1)I(t+s-1 > T) \\ + \xi_{T\lambda} g_{T\lambda}(t) g_{T\lambda}(s) + \phi_{T\lambda} g_{T\lambda}(T-t+1) g_{T\lambda}(s) \\ + \phi_{T\lambda} g_{T\lambda}(t) g_{T\lambda}(T-s+1) + \xi_{T\lambda} g_{T\lambda}(T-t+1) g_{T\lambda}(T-s+1) \quad (17)$$

$$= f_{T\lambda}(t-s) + \sum_{j=2}^8 w_{Tts}^j. \quad (18)$$

5.  $|f_{T\lambda}(0)| \leq 1$ ; for  $m \in \{1, 2, \dots, T\}$ ,  $|f_{T\lambda}(m)| \leq Cm^{-3}$  and  $|g_{T\lambda}(m)| < Cm^{-3}$  for some constant  $C > 0$  independent  $T$  by Theorem 1 of de Jong and Sakarya (2016); from the above definitions and inequalities, it follows that

$$\sup_{T \geq 1, t, s \in \{1, 2, \dots, T\}} |w_{Tts}| < \infty, \quad (19)$$

$$\sup_{T \geq 1} \sum_{s=1}^{T-2} |w_{Tts}| < \infty, \quad (20)$$

and for  $\gamma \in (0, 1/2)$

$$\sup_{T \geq 1} \sup_{1 \leq s \leq T} \sup_{r \in [\gamma, (1-\gamma)]} |T^3 \sum_{j=2}^8 w_{T, [rT], s}^j| < \infty. \quad (21)$$

6. For all  $\lambda \geq 0$  and  $m \in \mathbb{Z}$ ,

$$\lim_{T \rightarrow \infty} f_{T\lambda}(m) = f_\lambda(m) = \int_0^1 \cos(\pi r m) (1 + 16\lambda \sin(\pi r/2)^4)^{-1} dr. \quad (22)$$

7. For all  $\lambda \geq 0$  and  $m \geq 1$ ,

$$\begin{aligned} \lim_{T \rightarrow \infty} g_{T\lambda}(m) &= g_\lambda(m) \\ &= \sqrt{2} \int_0^1 \cos(\pi r(m - 1/2)) \sin(\pi r/2)^2 \cos(\pi r/2) (1 + 16\lambda \sin(\pi r/2)^4)^{-1} dr. \end{aligned}$$

8. For all  $\lambda \geq 0$ ,  $\lim_{T \rightarrow \infty} \phi_{T\lambda} = 0$ ,

$$\lim_{T \rightarrow \infty} \xi_{T\lambda} = \xi_\lambda = \frac{32\lambda}{1 - 32\lambda\delta_\lambda},$$

where  $1 - 32\lambda\delta_\lambda \neq 0$ , and

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

9. Since  $\hat{\tau}_{Tt} = \sum_{s=1}^T w_{Tts} y_s$  for  $t = 1, 2, \dots, T$ , for sequences  $x_t$  and  $y_t$ , we have

$$\hat{\tau}_{Tt}(x_1 + y_1, x_2 + y_2, \dots, x_T + y_T) = \hat{\tau}_{Tt}(x_1, x_2, \dots, x_T) + \hat{\tau}_{Tt}(y_1, y_2, \dots, y_T), \quad (23)$$

and also since  $\hat{c}_{Tt} = y_t - \hat{\tau}_{Tt}$  we write

$$\hat{c}_{Tt}(x_1 + y_1, x_2 + y_2, \dots, x_T + y_T) = \hat{c}_{Tt}(x_1, x_2, \dots, x_T) + \hat{c}_{Tt}(y_1, y_2, \dots, y_T), \quad (24)$$

for  $t = 1, 2, \dots, T$ .

10. For  $t = 1, 2, \dots, T$

$$\hat{\tau}_{Tt}(y_1, y_2, \dots, y_{T-1}, y_T) = \hat{\tau}_{T, T-t+1}(y_T, y_{T-1}, \dots, y_2, y_1). \quad (25)$$

## Appendix 2: 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

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

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)

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

and therefore, it suffices to minimize

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

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

$$((1 + \lambda) - 2\lambda\bar{B} + \lambda\bar{B}^2) \hat{c}_{T1} = \lambda\Delta^2 y_3 = \lambda\tilde{y}_{T1}, \quad (26)$$

$$(-2\lambda B + (1 + 5\lambda) - 4\lambda\bar{B} + \lambda\bar{B}^2) \hat{c}_{T2} = \lambda(\Delta^2 y_4 - 2\Delta^2 y_3) = \lambda\tilde{y}_{T2}, \quad (27)$$

$$(-2\lambda\bar{B} + (1 + 5\lambda) - 4\lambda B + \lambda B^2) \hat{c}_{T,T-1} = \lambda(\Delta^2 y_{T-1} - 2\Delta^2 y_T) = \lambda\tilde{y}_{T,T-1}, \quad (28)$$

$$((1 + \lambda) - 2\lambda B + \lambda B^2) \hat{c}_{TT} = \lambda\Delta^2 y_T = \lambda\tilde{y}_{TT}, \quad (29)$$

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

$$(\lambda\bar{B}^2 - 4\lambda\bar{B} + (1 + 6\lambda) - 4\lambda B + \lambda B^2) \hat{c}_{Tt} = \lambda\Delta^4 y_{t+2} = \lambda\tilde{y}_{Tt}. \quad (30)$$

The analogy between the first order conditions of Equations (26)-(30) and the first conditions of Equations (2)-(6) now reveals that  $\hat{c}_{Tt}(y_1, y_2, \dots, y_T) = \hat{\tau}_{Tt}(\lambda\tilde{y}_{T1}, \lambda\tilde{y}_{T2}, \dots, \lambda\tilde{y}_{TT}) = \lambda\hat{\tau}_{Tt}(\tilde{y}_{T1}, \tilde{y}_{T2}, \dots, \tilde{y}_{TT})$ .  $\square$

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

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

*Proof of Lemma 1.* Using Equation (17), we write

$$\begin{aligned}
& \lim_{T \rightarrow \infty} w_{T, [rT] + k, [rT] + l} \\
&= \lim_{T \rightarrow \infty} (f_{T\lambda}(k - l) + f_{T\lambda}(T)I(2[rT] + k + l - 1 = T)) \\
&\quad + \lim_{T \rightarrow \infty} f_{T\lambda}(2[rT] + k + l - 1)I(2[rT] + k + l - 1 < T) \\
&\quad + \lim_{T \rightarrow \infty} f_{T\lambda}(2(T - [rT]) - k - l + 1)I(2[rT] + k + l - 1 > T) \\
&\quad + \lim_{T \rightarrow \infty} (\xi_{T\lambda}g_{T\lambda}([rT] + k)g_{T\lambda}([rT] + l) + \phi_{T\lambda}g_{T\lambda}(T - [rT] - k + 1)g_{T\lambda}([rT] + l)) \\
&\quad + \lim_{T \rightarrow \infty} \phi_{T\lambda}g_{T\lambda}([rT] + k)g_{T\lambda}(T - [rT] - l + 1) \\
&\quad + \lim_{T \rightarrow \infty} \xi_{T\lambda}g_{T\lambda}(T - [rT] - k + 1)g_{T\lambda}(T - [rT] - l + 1) \\
&= f_{\lambda}(k - l),
\end{aligned}$$

by results (5)-(8) in Appendix 1. □

*Proof of Theorem 2.* Because  $y_t = \mu I(t \geq [rT] + 1) + u_t$  and Equation (24)

$$\hat{c}_{Tt}(y_1, y_2, \dots, y_T) = \hat{c}_{Tt}(u_1, u_2, \dots, u_T) + \hat{c}_{Tt}(\mu I(1 \geq [rT] + 1), \mu I(2 \geq [rT] + 1), \dots, \mu I(T \geq [rT] + 1)).$$

By Theorem 1, and because  $4 \leq [rT] \leq T - 5$  was assumed,

$$\begin{aligned}
& \hat{c}_{Tt}(\mu I(1 \geq [rT] + 1), \mu I(2 \geq [rT] + 1), \dots, \mu I(T \geq [rT] + 1)) \\
&= \lambda \mu \hat{\tau}_{Tt}(0, 0, \Delta^4 I(5 \geq [rT] + 1), \dots, \Delta^4 I(T \geq [rT] + 1), 0, 0).
\end{aligned}$$

Noting that  $\Delta^4 I(t + 2 \geq [rT] + 1) = \Delta^3 I(t + 2 = [rT] + 1)$ , it follows that the last expression equals

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

where the first equality is obtained by the weighted average representation of the trend provided in result (9) of Appendix 1. This establishes the first result of theorem.

Next, we need to show the second result of the theorem. Using the first result of the theorem, we find that

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

where the third equality follows from Lemma 1. □



*Proof of Theorem 3:* By Theorem 1 and the definition of  $\hat{c}_{Tt}^m$  in Equation (9),

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

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

$$\begin{aligned}
& \sup_{t \in [\gamma T, (1-\gamma)T]} \sum_{s=3}^{T-2} |w_{Tts}| \|\tilde{y}_{Ts}\|_p I(|t-s| > m) I(T \geq m) \\
& \leq C \sup_{t \in [\gamma T, (1-\gamma)T]} \sum_{s=3}^{T-2} |f_{T\lambda}(t-s)| I(|t-s| > m) \sup_{s \geq 1} \|u_s\|_p \\
& + C \sup_{t \in [\gamma T, (1-\gamma)T]} \sum_{s=3}^{T-2} \left| \sum_{j=2}^8 w_{Tts}^j \right| I(|t-s| > m) \sup_{s \geq 1} \|u_s\|_p I(T \geq m) \\
& \leq C \sum_{j=m}^{T-2} |f_{T\lambda}(j)| \sup_{s \geq 1} \|u_s\|_p + C \sum_{s=3}^{T-2} \sup_{t \in [\gamma T, (1-\gamma)T]} \left| \sum_{j=2}^8 w_{Tts}^j \right| \sup_{s \geq 1} \|u_s\|_p \\
& \leq C_1 \sup_{s \geq 1} \|u_s\|_p \sum_{j=m}^{\infty} j^{-3} \\
& + CT^{-2} \sup_{s \geq 1} \|u_s\|_p \sup_{T \geq 1} \sup_{1 \leq s \leq T} \sup_{t \in [\gamma T, (1-\gamma)T]} |T^3 \sum_{j=2}^8 w_{Tts}^j| I(T \geq m) \\
& = O(m^{-2}),
\end{aligned}$$

by the results of Equations (18)-(21). This shows the first assertion of the theorem. To show the second assertion,

$$\begin{aligned}
& \sup_{T \geq 1} \sup_{t \in [\gamma T, (1-\gamma)T]} \left\| \hat{c}_{Tt} - \lambda w_{Tt1} \tilde{y}_{T1} - \lambda w_{Tt2} \tilde{y}_{T2} - \lambda w_{Tt, T-1} \tilde{y}_{T, T-1} - \lambda w_{TtT} \tilde{y}_{TT} \right\|_p \\
& \leq \lambda \sup_{T \geq 1} \sup_{t \in [\gamma T, (1-\gamma)T]} \left\| \sum_{s=3}^{T-2} w_{Tts} \tilde{y}_{Ts} \right\|_p \\
& \leq 32\lambda \sum_{j=0}^{T-2} |f_{T\lambda}(j)| \sup_{s \geq 1} \|u_s\|_p + 16\lambda \sum_{s=3}^{T-2} \sup_{t \in [\gamma T, (1-\gamma)T]} \left| \sum_{j=2}^8 w_{Tts}^j \right| \sup_{s \geq 1} \|u_s\|_p = O(1),
\end{aligned}$$

by a reasoning similar to that of the proof of the first assertion.  $\square$

*Proof of Theorem 4:* We write that

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

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

$$\begin{aligned} & T^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} (g(\hat{c}_{Tt}) - Eg(\hat{c}_{Tt})) \\ = & T^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} (g(\hat{c}_{Tt}) - g(\hat{c}_{Tt} - a_{Tt})) \end{aligned} \quad (31)$$

$$+ T^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} (g(\hat{c}_{Tt} - a_{Tt}) - Eg(\hat{c}_{Tt} - a_{Tt})) \quad (32)$$

$$+ T^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} (Eg(\hat{c}_{Tt} - a_{Tt}) - Eg(\hat{c}_{Tt})). \quad (33)$$

by the triangle inequality. For the expression in Equation (32), the weak law of large numbers follows analogously to the proof of Theorem 6 of de Jong and Sakarya (2016). Therefore, we only need to show that the expression in Equation (31) is  $o_p(1)$  and the expression in Equation (33) converges to 0. Both results follow if

$$T^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} E|g(\hat{c}_{Tt}) - g(\hat{c}_{Tt} - a_{Tt})| \longrightarrow 0. \quad (34)$$

In order to establish this, note that since  $g(\cdot)$  is Lipschitz continuous, we have

$$T^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} E|g(\hat{c}_{Tt}) - g(\hat{c}_{Tt} - a_{Tt})| \leq LT^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} E|a_{Tt}|, \quad (35)$$

and by the definition of  $a_{Tt}$  and Equation (18), we have

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

$$\leq C_1 T^{3/2} \sup_{s \in \{1, 2, T-1, T\}} |f_{T\lambda}(t-s)| + C_2 T^{3/2} \sup_{s \in \{1, 2, T-1, T\}} \left| \sum_{k=2}^8 w_{Tts}^k \right|, \quad (36)$$

where the last line follows from the assumption  $E(|\tilde{y}_{T1}| + |\tilde{y}_{T2}| + |\tilde{y}_{T,T-1}| + |\tilde{y}_{TT}|) = O(T^{3/2})$ . Therefore, we use the upper bound for  $E|a_{Tt}|$  in Equation (36) to bound the expression in Equation (35), as follows:

$$\begin{aligned} & \sum_{t \in [\gamma T, (1-\gamma)T]} E|g(\hat{c}_{Tt}) - g(\hat{c}_{Tt} - a_{Tt})| \\ & \leq C_1 L T^{3/2} T^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} \sup_{s \in \{1, 2, T-1, T\}} |f_{T\lambda}(t-s)| \\ & + C_2 L T^{3/2} T^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} \sup_{s \in \{1, 2, T-1, T\}} \left| \sum_{k=2}^8 w_{Tts}^k \right|. \end{aligned}$$

Note that

$$\begin{aligned} & \lim_{T \rightarrow \infty} T^{3/2} T^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} \sup_{s \in \{1, 2, T-1, T\}} |f_{T\lambda}(t-s)| \\ & \leq C_3 \lim_{T \rightarrow \infty} T^{-3/2} T^{-1} ((1-2\gamma)T+1) \sup_{r \in [\gamma, (1-\gamma)]} \sup_{s \in \{1, 2, T-1, T\}} T^3 |[rT] - s|^{-3} \\ & = 0, \end{aligned}$$

where the last inequality follows from  $|f_{T\lambda}(m)| \leq C|m|^{-3}$  for  $m \in \{1, 2, \dots, T\}$  as given in result (5) of Appendix 1. Similarly, we write that

$$\begin{aligned} & \lim_{T \rightarrow \infty} T^{3/2} T^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} \sup_{s \in \{1, 2, T-1, T\}} \left| \sum_{k=2}^8 w_{Tts}^k \right| \\ & \leq \lim_{T \rightarrow \infty} T^{3/2} T^{-1} \sum_{t \in [\gamma T, (1-\gamma)T]} \sup_{r \in [\gamma, (1-\gamma)]} \sup_{s \in \{1, 2, T-1, T\}} \left| \sum_{k=2}^8 w_{T, [rT], s}^k \right| \\ & = \lim_{T \rightarrow \infty} T^{-3/2} T^{-1} ((1-2\gamma)T+1) \sup_{r \in [\gamma, (1-\gamma)]} \sup_{s \in \{1, 2, T-1, T\}} \left| T^3 \sum_{k=2}^8 w_{T, [rT], s}^k \right| \\ & = 0, \end{aligned}$$

since  $\sup_{T \geq 1} \sup_{1 \leq s \leq T} \sup_{r \in [\gamma, (1-\gamma)]} |T^3 \sum_{j=2}^8 w_{T, [rT], s}^j| < \infty$  as given in result (5) of Appendix 1. Therefore, we have now shown the result of Equation (34), and the proof is complete.  $\square$

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

$$\begin{aligned} & \lim_{T \rightarrow \infty} w_{T, T-k, j+1} = 0, \\ & \lim_{T \rightarrow \infty} w_{T, T-k, T-j} = f_\lambda(k-j) + f_\lambda(k+j+1) + \xi_\lambda g_\lambda(k+1) g_\lambda(j+1). \end{aligned}$$

*Proof of Lemma 2.* Using Equation (17), we write

$$\begin{aligned}
& \lim_{T \rightarrow \infty} w_{T, T-k, j+1} \\
&= \lim_{T \rightarrow \infty} (f_{T\lambda}(T-k-j-1) + f_{T\lambda}(T)I(T-k+j=T)) \\
&\quad + \lim_{T \rightarrow \infty} (f_{T\lambda}(T-k+j)I(T-k+j < T) + f_{T\lambda}(T+k-j)I(T-k+j > T)) \\
&\quad + \lim_{T \rightarrow \infty} (\xi_{T\lambda}g_{T\lambda}(T-k)g_{T\lambda}(j+1) + \phi_{T\lambda}g_{T\lambda}(k+1)g_{T\lambda}(j+1)) \\
&\quad + \lim_{T \rightarrow \infty} (\phi_{T\lambda}g_{T\lambda}(T-k)g_{T\lambda}(T-j) + \xi_{T\lambda}g_{T\lambda}(k+1)g_{T\lambda}(T-j)) = 0,
\end{aligned}$$

by results (5)-(8) of Appendix 1. Similarly, we write that

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

since  $\lim_{T \rightarrow \infty} f_{T\lambda}(m) = f_\lambda(m)$  and  $\lim_{T \rightarrow \infty} g_{T\lambda}(m) = g_\lambda(m)$  for all  $\lambda \geq 0$  and  $m \in \mathbb{Z}$ ,  $\lim_{T \rightarrow \infty} \xi_{T\lambda} = \xi_\lambda$ , and  $\lim_{T \rightarrow \infty} \phi_{T\lambda} = 0$  as given in results (6)-(8) of Appendix 1.  $\square$

*Proof of Theorem 5.* First, we write that

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

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 (11). Next, we will show that the results in Equations (12) and (13) hold. If  $y_t$  is an I(3) or an I(4) process, we have

$$E \left| \sum_{s=1}^{T-2} w_{Tts} \tilde{y}_{Ts} \right| \leq \sum_{s=1}^{T-2} |w_{Tts}| \sup_{1 \leq s \leq T-2} E |\tilde{y}_{Ts}|.$$

Furthermore,  $\sup_{1 \leq s \leq T-2} E |\tilde{y}_{Ts}| \leq C_2 \sup_{s \geq 1} E |u_s| < \infty$  by the definition of  $\tilde{y}_{Ts}$  given in Theorem 1, which together with the result of Equation (20) implies that

$$\left| \sum_{s=1}^{T-2} w_{Tts} \tilde{y}_{Ts} \right| = O_p(1). \tag{37}$$

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 (37), 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

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

by Lemma 2 and by the continuous mapping theorem. This shows the result in Equation (12).

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 (37). The second equality follows from Theorem 1 and the assumption that  $\Delta^4 y_t = u_t$  which implies 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

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

by Lemma 2 and by the continuous mapping theorem. This completes the proof.  $\square$

**Lemma 3.**  $\Delta^2(t+2)^p = \sum_{k=0}^{p-2} c_{pk} t^k$  for  $t = 1, 2, \dots, T$ , where  $c_{pk}$  is defined in Equation (14).

*Proof of Lemma 3.* The binomial theorem gives the following equality

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

Therefore,

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

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

**Lemma 4.**  $\Delta^4(t+2)^p = \sum_{k=0}^{p-4} a_{pk} t^k$  for  $t = 1, 2, \dots, T$ , where  $a_{pk}$  is defined in Equation (15).

*Proof of Lemma 4.* By the binomial theorem, we have

$$\begin{aligned}
\Delta^4(t+2)^p &= (t+2)^p - 4(t+1)^p + 6t^p - 4(t-1)^p + (t-2)^p \\
&= \sum_{k=0}^p \binom{p}{k} t^k 2^{p-k} - 4 \sum_{k=0}^p \binom{p}{k} t^k + 6t^p \\
&\quad - 4 \sum_{k=0}^p \binom{p}{k} t^k (-1)^{p-k} + \sum_{k=0}^p \binom{p}{k} t^k (-2)^{p-k} \\
&= \sum_{k=0}^{p-1} a_{p,k} t^k + (t^p - 4t^p + 6t^p - 4t^p + t^p) \\
&= \sum_{k=0}^{p-1} a_{pk} t^k,
\end{aligned}$$

where

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

This implies that

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

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

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

$\square$

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

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

where by Lemma 3

$$\tilde{y}_{T1} = 3^p - 2^{p+1} + 1 = \sum_{k=0}^{p-2} c_{pk}, \quad (38)$$

$$\tilde{y}_{T2} = (4^p - 2 \times 3^p + 2^p) - 2(3^p - 2^{p+1} + 1) = \sum_{k=0}^{p-2} c_{pk}(2^k - 2), \quad (39)$$

$$\begin{aligned} \tilde{y}_{T,T-1} &= ((T-1)^p - 2(T-2)^p + (T-3)^p) - 2(T^p - 2(T-1)^p + (T-2)^p) \\ &= \sum_{k=0}^{p-2} c_{pk}((T-3)^k - 2(T-2)^k), \end{aligned} \quad (40)$$

$$\tilde{y}_{TT} = T^p - 2(T-1)^p + (T-2)^p = \sum_{k=0}^{p-2} c_{pk}(T-2)^k, \quad (41)$$

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

$$\tilde{y}_{Tt} = \Delta^4(t+2)^p = \sum_{k=0}^{p-4} a_{pk}t^k. \quad (42)$$

Also, by Equation (23) it follows that

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

By replacing  $\tilde{y}_{Tt}$  with the expression in Equation (42) for  $t = 3, 4, \dots, T-2$ , we obtain

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

where the last equality follows from Equation (23). Similarly, by replacing  $\tilde{y}_{T1}, \tilde{y}_{T2}, \tilde{y}_{T,T-1}$  and  $\tilde{y}_{TT}$  with the expressions in Equations (38)-(41) and Equation (23) we obtain

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

Thus, we have

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

By using the identity in Equation (23), we write that  $\hat{\tau}_{Tt}(0, 0, 3^k, \dots, (T-2)^k, 0, 0) = \hat{\tau}_{Tt}(1, 2^k, \dots, T^k) - \hat{\tau}_{Tt}(1, 2^k, 0, \dots, 0, (T-1)^k, T^k)$ , which gives

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

Therefore, the result is shown.  $\square$

**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

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

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

$$\lambda(z-1)^4 + z^2 = \left( \sqrt{\lambda}(z-1)^2 + iz \right) \left( \sqrt{\lambda}(z-1)^2 - iz \right).$$

Two tedious calculation show that

$$\left( \sqrt{\lambda}(z-1)^2 + iz \right) = \sqrt{\lambda}(z-z_1)(z-z_2)$$

and

$$\left( \sqrt{\lambda}(z-1)^2 - iz \right) = \sqrt{\lambda}(z-z_3)(z-z_4)$$

where  $z_1, z_2, z_3$  and  $z_4$  are as defined in the lemma.  $\square$

*Proof of Theorem 7.* For  $t = 3, 4, \dots, T-2$ , the first order condition for  $\hat{c}_{Tt}$  given in Equation (30) is

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

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. Since zero is a particular solution to the difference equation, and using the fact that  $z_2 = z_1^{-1}$ ,  $z_3 = \bar{z}_1$ , and  $z_4 = \bar{z}_1^{-1}$ , it follows that

$$\hat{c}_{Tt} = C_{1T} z_1^t + C_{2T} z_1^{-t} + C_{3T} \bar{z}_1^t + C_{4T} \bar{z}_1^{-t}.$$

Note that the HP filter always produces trends and cyclical components that are real valued as long as  $y_t$  is real for  $t = 1, 2, \dots, T$ . Thus,  $\hat{c}_{Tt} = \bar{\hat{c}}_{Tt}$ , which implies that

$$C_{1T} z_1^t + C_{2T} z_1^{-t} + C_{3T} \bar{z}_1^t + C_{4T} \bar{z}_1^{-t} = \bar{C}_{1T} \bar{z}_1^t + \bar{C}_{2T} \bar{z}_1^{-t} + \bar{C}_{3T} z_1^t + \bar{C}_{4T} z_1^{-t},$$



and therefore  $C_{3T} = \bar{C}_{1T}$  and  $C_{4T} = \bar{C}_{2T}$ .

Furthermore, Theorem 1 implies that  $\hat{c}_{Tt}(1, 2^2, \dots, T^2) = \lambda \hat{\tau}_{Tt}(1, -1, 0, \dots, 0, -1, 1)$  and  $\hat{c}_{T, T-t+1}(1, 2^2, \dots, T^2) = \lambda \hat{\tau}_{T, T-t+1}(1, -1, 0, \dots, 0, -1, 1)$ . Equation (25) implies that

$$\hat{\tau}_{Tt}(1, -1, 0, \dots, 0, -1, 1) = \hat{\tau}_{T, T-t+1}(1, -1, 0, \dots, 0, -1, 1).$$

Therefore, it follows that  $\hat{c}_{Tt}(1, 2^2, \dots, T^2) = \hat{c}_{T, T-t+1}(1, 2^2, \dots, T^2)$ , which means that

$$C_{1T}z_1^t + C_{2T}z_1^{-t} + \bar{C}_{1T}\bar{z}_1^t + \bar{C}_{2T}\bar{z}_1^{-t} = C_{1T}z_1^{T+1}z_1^{-t} + C_{2T}z_1^{-(T+1)}z_1^t + \bar{C}_{1T}\bar{z}_1^{T+1}\bar{z}_1^{-t} + \bar{C}_{2T}\bar{z}_1^{-(T+1)}\bar{z}_1^t.$$

Therefore,  $C_{2T} = C_{1T}z_1^{T+1}$ , which allows us to conclude

$$\hat{c}_{Tt}(1, 2^2, \dots, T^2) = C_{1T}(z_1^t + z_1^{T-t+1}) + \bar{C}_{1T}(\bar{z}_1^t + \bar{z}_1^{T-t+1}). \quad (43)$$

Next, we use the first order conditions for  $t = 1$  and  $2$ , given in Equations (26) and (27), to solve for  $C_{1T}$ . Defining

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

Equations (26) and (27) are equivalent to  $C_{1T}a_T + \bar{C}_{1T}\bar{a}_T = 2\lambda$  and  $C_{1T}b_T + \bar{C}_{1T}\bar{b}_T = -2\lambda$ , which implies that  $C_{1T} = 2\lambda(\bar{b}_T + \bar{a}_T)/(a_T\bar{b}_T - \bar{a}_Tb_T)$ . 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 (43), we obtain

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

□

**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.* The result is easy to verify and we omit its proof for the sake of brevity. □

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

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

where by Lemma 6,  $\tilde{y}_{T1} = CC_1 \exp(1)$ ,  $\tilde{y}_{T2} = CC_1(1 - 2\exp(-1))\exp(2)$ ,  $\tilde{y}_{T, T-1} = C(1 - C_1)\exp(T-1)$ ,  $\tilde{y}_{TT} = CC_1 \exp(-2)\exp(T)$ , and  $\tilde{y}_{Tt} = C \exp(t)$  for  $t = 3, 4, \dots, T-2$ . Using these expressions and using the definitions of  $C_2$ ,  $C_3$  and  $C_4$  in the theorem, we find

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

Therefore, we have

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

□

*Proof of Theorem 9.* Evaluating the result of Theorem 8 at  $t = T - k$  gives that

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

where the second equality follows from Equation (23) and the third equality is obtained by the weighted average formula of the trend given in result (9) of Appendix 1. Therefore it now follows that

$$\begin{aligned}
& \lim_{T \rightarrow \infty} \frac{\hat{c}_{T,T-k}(\exp(1), \dots, \exp(T))}{\exp(T-k)} \\
&= C\lambda \lim_{T \rightarrow \infty} \frac{\hat{\tau}_{T,T-k}(\exp(1), \exp(2), \dots, \exp(T))}{\exp(T-k)} \tag{44}
\end{aligned}$$

$$+ C\lambda \lim_{T \rightarrow \infty} \frac{((C_1 - 1)w_{T,T-k,1} \exp(1) + (C_2 - 1)w_{T,T-k,2} \exp(2))}{\exp(T-k)} \tag{45}$$

$$+ C\lambda \lim_{T \rightarrow \infty} \frac{((C_3 - 1)w_{T,T-k,T-1} \exp(T-1) + (C_4 - 1)w_{T,T-k,T} \exp(T))}{\exp(T-k)}. \tag{46}$$

The limit in Equation (44) satisfies

$$\begin{aligned}
& \lim_{T \rightarrow \infty} \frac{\sum_{s=1}^T w_{T,T-k,s} \exp(s)}{\exp(T-k)} = \lim_{T \rightarrow \infty} \sum_{s=1}^T w_{T,T-k,s} \exp(s+k-T) \\
&= \lim_{T \rightarrow \infty} \sum_{j=0}^{T-1} w_{T,T-k,T-j} \exp(k-j) = \lim_{T \rightarrow \infty} \sum_{j=0}^{\infty} w_{T,T-k,T-j} \exp(k-j) I(j \leq T-1) \\
&= \sum_{j=0}^{\infty} \lim_{T \rightarrow \infty} (w_{T,T-k,T-j} \exp(k-j) I(j \leq T-1)),
\end{aligned}$$

where the last equality follows by the dominated convergence theorem and Equation (19). Therefore, Lemma 2 implies that the limit in Equation (44) is equal to

$$\begin{aligned}
& C\lambda \sum_{j=0}^{\infty} (f_{\lambda}(k-j) + f_{\lambda}(k+j+1) + \xi_{\lambda} g_{\lambda}(k+1) g_{\lambda}(j+1)) \exp(k-j) \\
&= C\lambda \exp(k) \sum_{j=0}^{\infty} (f_{\lambda}(k-j) + f_{\lambda}(k+j+1) + \xi_{\lambda} g_{\lambda}(k+1) g_{\lambda}(j+1)) \exp(-j).
\end{aligned}$$

It is easy to see that the expression in Equation (45) vanishes since  $\lim_{T \rightarrow \infty} w_{T,T-k,j} = 0$  for  $j = 1, 2$  by Lemma 2. Lastly, we consider the expression in Equation (46). Dividing the numerator and the denominator of the expression by  $\exp(T)$ , we obtain

$$\begin{aligned}
& C\lambda \exp(k) \lim_{T \rightarrow \infty} ((C_3 - 1)w_{T,T-k,T-1} \exp(-1) + (C_4 - 1)w_{T,T-k,T}) \\
&= C\lambda \exp(k-1) (C_3 - 1) (f_{\lambda}(k-1) + f_{\lambda}(k+2) + \xi_{\lambda} g_{\lambda}(k+1) g_{\lambda}(2)) \\
&\quad + C\lambda \exp(k) (C_4 - 1) (f_{\lambda}(k) + f_{\lambda}(k+1) + \xi_{\lambda} g_{\lambda}(k+1) g_{\lambda}(1))
\end{aligned}$$

by Lemma 2. So we conclude that

$$\begin{aligned}
& \lim_{T \rightarrow \infty} \frac{\hat{c}_{T, T-k}(\exp(1), \dots, \exp(T))}{\exp(T-k)} \\
&= C\lambda \exp(k) \sum_{j=0}^{\infty} (f_{\lambda}(k-j) + f_{\lambda}(k+j+1) + \xi_{\lambda} g_{\lambda}(k+1) g_{\lambda}(j+1)) \exp(-j) \\
&\quad + C\lambda \exp(k-1)(C_3 - 1) (f_{\lambda}(k-1) + f_{\lambda}(k+2) + \xi_{\lambda} g_{\lambda}(k+1) g_{\lambda}(2)) \\
&\quad + C\lambda \exp(k)(C_4 - 1) (f_{\lambda}(k) + f_{\lambda}(k+1) + \xi_{\lambda} g_{\lambda}(k+1) g_{\lambda}(1)).
\end{aligned}$$

□