A Unified Framework to Real-Time Signal Extraction and Data Revisions

Marc Wildi

Institute of Data Analysis and Process Design

September 30, 2011

Abstract

Real-Time Signal Extraction (RTSE) addresses a series of important prospective estimation problems ranging from 'simple' forecasting up to the timely and/or accurate extraction of components - trends or cycles -, towards the sample end of a time series. In practice, the optimal design of RTSE filters is challenged by two major issues: one-sided filters are subject to undesirable leakage (imperfect noise suppression) and they suffer from undesirable phase shifts (timeliness). Revisions add an extra complexity layer by generating particular non-stationarities along vintages. We here devise a unified framework to tackle RTSE and revisions based on the Multivariate Direct Filter Approach (MDFA). Strengths of the resulting approach are threefold when compared to the generic model-based paradigm: it simplifies calculations/computations by integrating out redundant model structure; it emphasizes a fundamental 'timeliness-reliability' dilemma which generalizes the traditional mean-square paradigm; it offers a flexible and powerful user-interface which allows to operationalize important research priorities. Our approach is generic in the sense that it could accommodate for parametric as well as non-parametric representation of the data.

Keywords. Real-time signal extraction, data revisions, Multivariate Direct Filter Approach (MDFA), timeliness-reliability dilemma, customized optimization criteria, cointegration, revision triangle.

Disclaimer : Preliminary/Draft.

1 Introduction

Real-time signal extraction (RTSE) concerns the determination and the estimation of important components, such as the trend or the cycle, towards the sample end of a time series. The optimal design of RTSE filters is challenged by two major issues in practice: one-sided filters are subject

to undesirable leakage (imperfect noise suppression) and they suffer from undesirable phase shifts (time delays). Wildi (2005, 2008.1, 2011) proposes a novel approach, the so-called Direct Filter Approach (DFA), which is based on a set of customized optimization criteria[1]. We here propose to extend the scope of this approach by allowing for general revision processes in the multivariate DFA (MDFA) as proposed in Wildi (2008.2).

Cunningham et al. (2009) state "Most macroeconomic data are uncertain - they are estimates rather than perfect measures of underlying economic variables -. One symptom of that uncertainty is the propensity of statistical agencies to revise their estimates in the light of new information or methodological advances". McKenzie (2006) notes eight reasons for revisions of official statistics. Statistical properties of revisions are analyzed in Aruoba (2008) where it is shown that the revision process contradicts simple rationality requirements, implying that future revisions can be forecasted, to some extent. Anderson and Gascon (2009) complement this view by emphasizing the determination of the 'true' value "Statistical agencies face a tradeoff between accuracy and timely reporting of macroeconomic data. As a result, agencies release their best estimates of the 'true' unobserved series in the proceeding month, quarter, or year with some measurement error. As agencies collect more information, they revise their estimates, and the data are said to be more 'mature'. As the reported data mature, the estimates, on average, are assumed to converge toward the 'true' unobserved values". Croushore (2009) expands the previous particular perspectives by distinguishing five general research areas entitled "Data Revisions", "Structural Modeling", "Forecasting", "Monetary and Fiscal Policy" as well as "Current (real-time) Analysis"[2]. The author argues "Optimal forecasts and indicators require a model of the data revision process (see Croushore 2006 for a survey). Some authors cast the data revision process in state-space form, which then allows the use of standard filtering techniques for forecasting, estimation, inference, smoothing, estimation of missing data, etc". Jacobs and van Norden (2011) distinguish three strands of the literature, namely "Data Description", "Optimal Forecasting and Inference" and "Cycle Trend Decompositions" and they propose a general state-space model which "further integrates all three strands of this literature". We position our approach by noting its contribution to the "Cycle Trend Decompositions" theme, evoked by the latter authors, as well as to the "Current Analysis" topic suggested by Croushore.

VAR, VECM and restricted VAR representations offer an alternative to the aforementioned state-space models, see for example Pain (1994), Patterson and Heravi (1991, 2004) and Garratt

---

[1]The potential of the approach has been documented in a series of recent applications, see for example http://www.idp.zhaw.ch/usri (real-time US-indicator) and http://www.neural-forecasting-competition.com/NN3/results.htm as well as http://www.neural-forecasting-competition.com/NN5/results.htm (forecasting).

[2]The author maintains and up-dates and comprehensive literature review on each of these themes, see https://facultystaff.richmond.edu/∼dcrousho/docs/realtime_lit.pdf.

et al. (2008, 2009). Hecq and Jacobs (2009) propose a unifying framework which reconciles two main streams: observation balanced systems and vintage balanced systems. When operating with non-stationary time series, we take inspiration from the latter authors who consider cointegration across *releases* (not vintages) of a non-stationary time series. More specifically, we propose a set of filter constraints which are able to link 'common trends' across releases.

We here propose a novel generic approach which combines 'revision' filtering (determination of the true data) and signal extraction filtering (determination of the signal) in a common Multivariate Direct Filter Approach (MDFA): the proposed optimization criteria address *filter* parameters - not *model* parameters - 'directly'. Potential strengths of the resulting shift of perspective are threefold: the particular arrangement of parameters in the reduced-form simplifies computations by 'integrating out' redundant (model) structure; control on critical RTSE performances (timeliness, noise suppression) can be exerted explicitly; finally, the method offers a natural and flexible user-interface which redefines and revaluates the role of the user. The approach is generic in the sense that we do not emphasize a particular model philosophy or data-representation or target signal. The user could supply his preferred model-based representation (for example TRAMO, X-12-ARIMA, STAMP, VECM, multivariate state space), he could specify his preferred signal target (model-based or any of the classical filters: Hodrick-Prescott, Christiano Fitzgerald, Baxter-King,...) and he could replicate corresponding real-time performances perfectly with or without revisions involved; he could 'customize' any preferred design(s) by emphasizing timeliness and/or reliability issues of early (real-time) trend or cycle estimates; alternatively, the user could supply a non-parametric representation (a sufficient statistic) of the data and customize the resulting optimization criterion to his individual research priorities. Although the approach remains completely generic, we take the liberty to mark our personal preference - in notational terms - in the text.

The paper is organized as follows: section 2 discusses alternative data-arrangements which are useful in 'organizational' terms; section 3 reviews the generic filter-background and refers to the literature on the topic; section 4 combines vintage filtering and signal extraction filtering in a common framework and section 5 concludes by a short wrap-up.

2 Data Organization

We first introduce simple notational conventions for working with vintages. The data can be organized in the so-called revision triangle, as proposed in Jacobs and van Norden (2011)

$$
\begin{pmatrix}
x_1^0 & x_1^1 & ... & x_1^h & ... & x_1^{T-1} \\
 & x_2^0 & ... & x_2^{h-1} & ... & ... \\
 & & ... & ... & ... & ... \\
 & & & x_{T-h}^0 & ... & x_{T-h}^h \\
 & & & & ... & ... \\
 & & & & & x_T^0
\end{pmatrix}
\tag{1}
$$

where columns are indexing release time - collect vintages - and rows are indexing run-time from $t = 1$ to $t = T$: x_t^l designates historical data t released in $t + l$. The most recent vintage is denoted by x_{T-j}^j, where $j = 0, ..., T - 1$. Final 'true' data is denoted by x_t^∞. Our notation slightly departs from the custom notation x_t^{t+l}, found in the literature, because the time-dimension t does not have a direct meaning in the frequency-domain - where our approach resides - whereas the lag l, alone, will have. Moreover, our notation emphasizes the diagonals of the revision triangle which are interesting because the resulting series are (supposed to be) congruent in statistical terms - stationary or difference stationary -, see the appendix (section 6.1) for reference on this topic. Publication lags can be accounted for by corresponding shifts of the subscripts, $x_{t-\text{lag}}^l$ (for notational ease we here ignore lags). We also frequently rely on the following alternative data arrangement in our developments

$$
\begin{pmatrix}
x_1^0 & x_1^1 & ... & ... & x_1^{T-2} & x_1^{T-1} \\
x_2^0 & x_2^1 & ... & ... & x_2^{T-2} & \\
... & ... & ... & ... & & \\
x_{T-h}^0 & x_{T-h}^1 & ... & x_{T-h}^h & & \\
... & ... & & & & \\
x_T^0 & & & & &
\end{pmatrix}
\tag{2}
$$

where the h-th column corresponds to the h-th upper diagonal in 1 or, equivalently, to the time series of historical h-th releases (accounting for the first h revisions). Obviously, the latter representation is also a 'revision triangle'. In order to distinguish 1 from 2 we call the former '*vintage* triangle' and the latter '*release* triangle' thereby emphasizing the importance of the respective columns. Our preference for the latter data-organization is based on the aforementioned statistical congruency argument.

3 (M)DFA Background

When attempting to combine signal extraction and data-revisions, the literature suggests that state space models have become something like a 'golden standard'. A recent approach proposed by Jacobs and van Norden (2011) is sketched in the appendix. Their sophisticated model allows to isolate unobserved revision components (news, noise, spillovers) as well as unobserved time series components (trend, cycle, noise). The sheer elegance of the model is very appealing and the compelling idea of identifying 'components' with (structural) 'dynamics' of the latent state-vector suggests that the analyst can rely on a powerful statistical tool in order to analyze 'components'. In contrast, our reduced-form approach shifts the research priority: we are not primarily interested in the "Data Description" theme, as evoked in the introduction, but in RTSE instead[3]. The following method tackles RTSE 'directly', it allows for sophisticated customization of the optimization criterion and it offers a flexible user-interface for operationalizing important research priorities.

3.1 Direct Filter Approach (DFA)

We assume that the target series is the output of a possibly bi-infinite (and generally symmetric) filter[4] applied to 'final' data x_t^∞:

$$y_T = \sum_{j=-\infty}^{\infty} \gamma_j x_{T-j}^\infty \tag{3}$$

Our task consists in estimating y_T by relying on past and present vintages x_t^h, $t = 1, ..., T$, $h = 0, ..., T-1$. The index T of y_T emphasizes the real-time aspect (*concurrent* filter): extensions to 'smoothing' - estimation of $y_{T-1}, y_{T-2}, ...$ - are straightforward and are not discussed explicitly.
If data is not revised, then $x_t^0 = x_t^\infty$ and the expression for the (real-time) finite-sample estimate is

$$\hat{y}_T = \sum_{j=0}^{L} b_j x_{T-j}^0 \tag{4}$$

where $L \leq T-1$ denotes the length of the moving average (MA-)filter. Optimal parameter estimates could be derived according to the mean-square norm

$$E[(y_t - \hat{y}_t)^2] = \int_{-\pi}^{\pi} |\Gamma(\omega) - \hat{\Gamma}(\omega)|^2 dS(\omega) \to \min_{b_0, ..., b_L}$$

---

[3]Obviously, a state space model of the data is able to tackle RTSE too and the corresponding real-time estimates would be optimal, in a mean-square sense, conditionally on the model being 'true'. But this crucial assumption can be challenged by the sheer complexity of the model.

[4]We avoid intentionally to specify the bi-infinite filter in order to keep the approach generic. Note that classical one-step forecasting could be replicated by setting $\gamma_{-1} = 1, \gamma_j = 0, j \neq -1$ (which defines a finite asymmetric target filter).

where $\Gamma(\omega) = \sum_{|j|<\infty} \gamma_j \exp(-ij\omega)$ is the transfer function of the symmetric filter (target), $\hat{\Gamma}(\omega) = \sum_{j=0}^{L} b_j \exp(-ij\omega)$ is the transfer function of the one-sided (real-time) filter and $S(\omega)$ is the spectral distribution of x_t[5]. In practice, the integral is approximated by a discrete (finite) sum[6]

$$\frac{2\pi}{T} \sum_{k=-T/2}^{T/2} |\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)|^2 \hat{dS}(\omega_k) \to \min_{b_0,\dots,b_L} \tag{5}$$

where $\omega_k = k2\pi/T$ is the standard discrete frequency-grid and $\hat{dS}(\omega)$ is an estimate of the (unknown) spectral distribution function. Possible candidates are explicit model-based spectral densities (as obtained by standard packages such as TRAMO or X-12-ARIMA or STAMP, for example) or implicit model-based densities (such as obtained from HP-, CF- or BK-filter designs, for example) or non-parametric estimates such as the periodogram $\hat{dS}(\omega_k) := \left|\Xi_{TX}^0(\omega_k)\right|^2$, see Wildi (2008.1):

$$\frac{2\pi}{T} \sum_{k=-T/2}^{T/2} |\Gamma(\omega_k)\Xi_{TX}^0(\omega_k) - \hat{\Gamma}(\omega_k)\Xi_{TX}^0(\omega_k)|^2 \tag{6}$$

$$= \frac{2\pi}{T} \sum_{k=-T/2}^{T/2} |\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)|^2 \left|\Xi_{TX}^0(\omega_k)\right|^2 \to \min_{b_0,\dots,b_L} \tag{7}$$

where $\Xi_{TX}^0(\omega_k)$ denotes the discrete Fourier Transform of the data x_t^0:

$$\Xi_{TX}^0(\omega_k) = \frac{1}{\sqrt{T2\pi}} \sum_{t=1}^{T} x_t^0 \exp(-it\omega_k)$$

Tackling revisions addresses a generalization of the frequency-domain expression

$$\hat{\Gamma}(\omega_k)\Xi_{TX}^0(\omega_k) = \left(\sum_{j=0}^{L} b_j \exp(-ij\omega_k) \right) \Xi_{TX}^0(\omega_k) \tag{8}$$

see section 4. In contrast to classical maximum likelihood approaches, criterion 5 emphasizes the filter error - not a model residual - and it addresses filter parameters - not model parameters -. By avoiding to specify the precise type of spectral estimate (model-based or not) we could keep the approach generic and in a sense agnostic. However, we shall mark our personal long-term preference for the DFT here.

Wildi (1998, 2005, 2008.1 and 2011) proposes a decomposition of the mean-square filter error 7

---

[5]The proposed frequency-domain representation remains valid as long as the filter error $y_t - \hat{y}_t$ is stationary - even if x_t happens to be integrated - see the appendix for reference. We can therefore substitute t to T in the above criterion.

[6]For notational simplicity we here assume that T is an even number.

based on a straightforward application of the law of the cosine in the complex plane:

$$\frac{2\pi}{T} \sum_{k=-T/2}^{T/2} |\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)|^2 \left|\Xi_{TX}^0(\omega_k)\right|^2$$

$$= \frac{2\pi}{T} \sum_{k=-T/2}^{T/2} (A(\omega_k) - \hat{A}(\omega_k))^2 \left|\Xi_{TX}^0(\omega_k)\right|^2 \tag{9}$$

$$+ \frac{2\pi}{T} \sum_{k=-T/2}^{T/2} 4A(\omega_k)\hat{A}(\omega_k) \sin(\hat{\Phi}(\omega_k)/2)^2 \left|\Xi_{TX}^0(\omega_k)\right|^2 \tag{10}$$

In this decomposition $A(\omega_k) = \Gamma(\omega_k)$ is the real transfer function of the symmetric filter[7]. The first summand 9 measures that part of the total (mean-square) filter error which is attributable to the amplitude mismatch: it summarizes distortions in the passband and noise-rejection in the stopband of the real-time filter. The second summand 10 measures the contribution by the non-vanishing phase function: it is restricted to the passband since the dimensionless phase-term $\sin(\hat{\Phi}(\omega_k)/2)^2$ is scaled by the product of amplitude functions. We are now in a position to tackle a famous real-time filter dilemma addressing *timeliness* and *reliability* (noise suppression) issues. Specifically, we introduce additional 'weights' λ and $W(\omega_k)$ in the above expressions:

$$\frac{2\pi}{T} \sum_{k=-T/2}^{T/2} (A(\omega_k) - \hat{A}(\omega_k))^2 W(\omega_k) \left|\Xi_{TX}^0(\omega_k)\right|^2 \tag{11}$$

$$+ (1+\lambda)\frac{2\pi}{T} \sum_{k=-T/2}^{T/2} 4A(\omega_k)\hat{A}(\omega_k) \sin(\hat{\Phi}(\omega_k)/2)^2 W(\omega_k) \left|\Xi_{TX}^0(\omega_k)\right|^2 \tag{12}$$

$$= \frac{2\pi}{T} \sum_{k=-T/2}^{T/2} |\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)|^2 W(\omega_k) \left|\Xi_{TX}^0(\omega_k)\right|^2$$

$$+ 4\lambda\frac{2\pi}{T} \sum_{k=-T/2}^{T/2} A(\omega_k)\hat{A}(\omega_k) \sin(\hat{\Phi}(\omega_k)/2)^2 W(\omega_k) \left|\Xi_{TX}^0(\omega_k)\right|^2 \rightarrow \min \tag{13}$$

For $\lambda = 0$ and $W \equiv \text{Id}$ the mean-square criterion 7 is replicated; for $\lambda > 0$ phase-artifacts are artificially magnified and *timeliness* of the optimal filter must improve (the delay decreases); noise suppression or *reliability* can be improved if $W(\cdot)$ emphasizes the fit in the stop-band. Obviously, timeliness and reliability issues can be addressed simultaneously: a customized real-time filter can outperform traditional mean-square designs in both dimensions at costs of poorer mean-square performances. This result confirms that the classical mean-square error norm is not a 'panacea': in particular it is not congruent with the early detection of turning-points (of a trend or a cycle signal). Illustrative examples/exercises and R-scripts addressing the proposed customization of 5 are maintained at disposition on SEFBlog, see the appendix for reference. A generalization to

---

[7]For notational ease we assumed $\Gamma(\cdot) \geq 0$ such that $\Gamma(\omega_k) = |\Gamma(\omega_k)| =: A(\omega_k)$.

non-stationary integrated processes is provided in Wildi (2008.1), chapter 6.

Let us briefly conclude this section by commenting the 'interface' linking the user to the statistical algorithm: the user can supply the type of signal (in 3) and the type of spectral estimate (in 5) and he can customize the optimization criterion 13 by supplying λ and $W(\cdot)$. The user can replicate real-time performances of classical model-based approaches (TRAMO/SEATS, X-12-ARIMA or STAMP, for example) or classical filter designs (as for example Hodrick-Prescott, Christiano-Fitzerald or Baxter-King) by supplying corresponding (pseudo) spectral densities $\hat{dS}(\omega_k)$ and he can customize either of these designs, according to his particular research priorities, by relying on 13. Alternatively, he can rely on the (pseudo-) DFT and possibly improve RTSE-performances even further, see Wildi (2008.1) chapters 4,5, 7 and 8 for corresponding results.

3.2 Multivariate Direct Filter Approach (MDFA)

Wildi (2008.2) extends the above results to a multivariate framework, the so-called Multivariate Direct Filter Approach (MDFA). We here briefly summarize the main results by considering the stationary case (the case of integrated and/or cointegrated releases is discussed in section 4.3 as well as in the appendix). Let y_t be the target signal 3 and assume the existence of m additional explaining variables w_{jt}, $j = 1, ..., m$. We can consider the following generalization of 8:

$$\hat{\Gamma}_X(\omega_k)\Xi_{TX}(\omega_k) + \sum_{n=1}^{m}\hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k) \tag{14}$$

where

$$\hat{\Gamma}_X(\omega_k) = \sum_{j=0}^{L} b_{Xj}\exp(-ij\omega_k)$$

$$\hat{\Gamma}_{W_n}(\omega_k) = \sum_{j=0}^{L} b_{w_nj}\exp(-ij\omega_k)$$

are the (one-sided) transfer functions applying to the 'explaining' variables and $\Xi_{TX}(\omega_k)$, $\Xi_{TW_n}(\omega_k)$ are the corresponding DFT's. Theorem 7.1 in Wildi (2008.2) establishes that the following (generalized) optimization criterion

$$\frac{2\pi}{T}\sum_{k=-T/2}^{T/2}\left|\left(\Gamma(\omega_k) - \hat{\Gamma}_X(\omega_k)\right)\Xi_{TX}(\omega_k) - \sum_{n=1}^{m}\hat{\Gamma}_{W_n}(\omega_k)\Xi_{TW_n}(\omega_k)\right|^2 \to \min_{\mathbf{B}} \tag{15}$$

inherits all efficiency properties of the (univariate) DFA and therefore the whole customization principle can be carried over to a general multivariate framework (the parameter matrix $\mathbf{B}$ is

defined in 19 below). Specifically, we can re-write the above criterion as:

$$\frac{2\pi}{T} \sum_{k=-T/2}^{T/2} \left| \left(\Gamma(\omega_k) - \hat{\Gamma}_X(\omega_k) \right) \Xi_{TX}(\omega_k) - \sum_{n=1}^{m} \hat{\Gamma}_{W_n}(\omega_k) \Xi_{TW_n}(\omega_k) \right|^2$$

$$= \frac{2\pi}{T} \sum_{k=-T/2}^{T/2} \left| \left(\Gamma(\omega_k) - \hat{\Gamma}_X(\omega_k) \right) - \sum_{n=1}^{m} \hat{\Gamma}_{W_n}(\omega_k) \frac{\Xi_{TW_n}(\omega_k)}{\Xi_{TX}(\omega_k)} \right|^2 |\Xi_{TX}(\omega_k)|^2$$

$$= \frac{2\pi}{T} \sum_{k=-T/2}^{T/2} \left| \Gamma(\omega_k) - \tilde{\Gamma}(\omega_k) \right|^2 |\Xi_{TX}(\omega_k)|^2 \tag{16}$$

where

$$\tilde{\Gamma}(\omega_k) := \hat{\Gamma}_X(\omega_k) + \sum_{n=1}^{m} \hat{\Gamma}_{W_n}(\omega_k) \frac{\Xi_{TW_n}(\omega_k)}{\Xi_{TX}(\omega_k)} \tag{17}$$

Note that potential singularities affecting $\tilde{\Gamma}(\omega_k)$ (due to a vanishing denominator DFT) will be canceled in 16[8]. A 'customization' of the (multivariate) mean-square error norm is straightforwardly obtained in the multivariate case by applying the decomposition in the previous section to 16:

$$\frac{2\pi}{T} \sum_{k=-T/2}^{T/2} |\Gamma(\omega_k) - \tilde{\Gamma}(\omega_k)|^2 W(\omega_k) |\Xi_{TX}(\omega_k)|^2$$

$$+4\lambda \frac{2\pi}{T} \sum_{k=-T/2}^{T/2} A(\omega_k) \tilde{A}(\omega_k) \sin(\tilde{\Phi}(\omega_k)/2)^2 W(\omega_k) |\Xi_{TX}(\omega_k)|^2 \to \min \tag{18}$$

where amplitude and phase functions address expression 17.

If we postulate that the revision-sequence is of finite length h_0, i.e. $x_t^{h_0-1} = x_t^{\infty}$, see section 4, then the data-revision problem can be transposed straightforwardly into the above MDFA by defining:

$$x_t := x_t^{h_0-1}; \ w_{nt} := x_t^n; \ \Xi_{TX}(\omega_k) := \Xi_{TX}^{h_0-1}(\omega_k); \ \Xi_{TW_n}(\omega_k) := \Xi_{TX}^n(\omega_k), \ n = 0, ..., h_0 - 2$$

$$b_{Xj} := b_{h_0-1,j}; \ b_{w_nj} := b_{nj}, n = 0, ..., h_0 - 2$$

The n-th column of the *release* triangle 2, denoted by $\mathbf{x}^n := (x_1^n, ..., x_{T-n}^n, 0, ..., 0)'$ is an 'explaining' variable and b_{nj} are the filter weights applied to x_{T-j}^n. Accordingly, the structure of the *release*

---

[8]From a numerical perspective a better solution which avoids singularities is provided in Wildi (2011.2).

triangle swaps over to the parameter space of filter coefficients:

$$
\begin{pmatrix}
b_{0,L-1} & b_{1,L-1} & ... & b_{h_0-1,L-1} & 0 \\
... & ... & ... & & \\
b_{0,h_0} & b_{1,h_0} & ... & b_{h_0-1,h_0} & 0 \\
b_{0,h_0-1} & b_{1,h_0-1} & ... & b_{h_0-1,h_0-1} & 0 \\
... & ... & ... & & \\
b_{01} & b_{11} & & & \\
b_{00} & & & &
\end{pmatrix}
\tag{19}
$$

We call this particular arrangement 'parameter *release* triangle': it reflects 'structure' imposed by the revision-background on the general multivariate filter framework. Recall that our preference for the *release* triangle relies on a statistical congruency argument.

The data-revision problem and the RTSE problem can now be tackled in a unified framework. In this multivariate context the user could replicate classical multivariate model-based approaches (VEC-ARIMA, State-Space) by plugging corresponding (pseudo- and/or cross-) spectral distribution functions into 15 and he could customize these designs - enhancing timeliness and/or noise suppression - by supplying λ and $W(\cdot)$ in 18. Alternatively, he could rely on the (pseudo-) DFT - as we are assuming implicitly in our notation -. Additional interesting 'structures' imposed by the revision process on the parameter space of the (real-time) filter are discussed in the following section.

4   Combining RTSE and Data-Revisions

We here combine the revision filtering task and the RTSE-task in the proposed MDFA-framework. For this purpose we briefly analyze the estimation of the 'true' data. The corresponding vintage filter/smoother is then plugged into the RTSE-filter.

4.1   Vintage-Filtering/Smoothing

Estimation of the 'true' latent data x_t^∞ can be obtained in the following general form

$$
\hat{x}_{T-j}^\infty \;\; = \;\; \sum_{l=-j}^{T-j-1} \sum_{h=0}^{j+l} \gamma_{j,j+l}^{j+l-h} x_{T-(j+l)}^{j+l-h}
\tag{20}
$$

The three-dimensional index-space spanned by j, l, h fixes the target variable x_{T-j}^∞ and identifies longitudinal and lateral contributions of the *vintage* triangle 1: for given j, the remaining indices l and h allow for movement along columns and rows, respectively. In a model-based perspective the above filter/smoother weights could be derived by imposing particular identifying (structural)

10

constraints, see section 6.2 for reference. In contrast, we here propose to consider relevant 'structures' imposed on the filter/smoother-weights 'directly'.

If the latest vintage $x_{T-(j+l)}^{j+l}$ encompasses all earlier releases $x_{T-(j+l)}^0, ..., x_{T-(j+l)}^{j+l-1}$ then $\gamma_{j,j+l}^{j+l-h} = 0$ for $0 < h \leq j + l$ and 20 simplifies to

$$\hat{x}_{T-j}^\infty = \sum_{l=-j}^{T-j-1} \gamma_{j,j+l}^{j+l} x_{T-(j+l)}^{j+l} \tag{21}$$

This setting is nearly - though not strictly - identical to the 'pure news' case analyzed in section 6.2[9]. We feel free to label this derived 'structure' by the term 'extended pure news' case: it will play an important role on the top RTSE-layer by imposing 'structure' upon the parameter triangle 19.

Practitioners often have an idea about the extent or the usefulness of information supported by x_{T-j}^{j-h}, $0 < h \leq j$, beyond that conveyed by the latest release x_{T-j}^j, $h = 0$. Assuming a fair amount of rationality of data providers would restrict h to typically 'small' values. We would like to suggest that it may be easier for an experienced practitioner to impose constraints on the vintage-filter/smoother weights 'directly' than to formulate a suitable structure for an underlying model. As an example, the user could specify

$$\gamma_{j,j+l}^0 := \alpha^h \gamma_{j,j+l}^h \tag{22}$$

where $0 \leq \alpha \leq 1$ thus reflecting the idea that the importance of earlier releases for the determination of the true value decreases at an exponential rate. The extended pure news case corresponds to $\alpha = 0$.

In the following we assume that revisions beyond a certain horizon h_0 are negligible, see for example Jacobs and van Norden (2011), Aruoba (2005) and Swanson and van Dijk (2006) (who coined the term "remaining revision"); not least because the effect of the RTSE-filter contributes in damping any "remaining revision". Formally, we assume that $x_t^{h_0+h} = x_t^{h_0-1}$ for $h \geq 0$ i.e. $x_t^{h_0-1}$ is the final data. This upper-limit h_0 of the number of 'relevant releases' corresponds to the dimension l of the observation vector in the Jacobs and van Norden model, see the appendix.

---

[9]The case $h = 0$ may be relevant for non-'pure news' cases. Consider the following (artificial) situation: $x_t^\infty = x_1^\infty$ (the noise variance in the state transition of the 'true' values vanishes) and $x_1^{T-1} = x_1^\infty$ i.e. the last release x_1^{T-1} is the true value. Then $x_{T-j}^\infty = \hat{x}_{T-j}^\infty = x_1^{T-1}$ for all j i.e. $h = 0$ applies, independently of the precise structure of the revision process.

4.2 Setting-up MDFA-Designs: the Stationary Case

In the absence of data revisions, the 'optimal' real-time filter for estimating the signal in $t = T$ expands according to

$$\hat{y}_T = \sum_{j=0}^{L} b_j x_{T-j}^0 = \sum_{j=0}^{L} b_j x_{T-j}^\infty$$

where L is the filter length. In the presence of revisions we plug-in estimates of the true data:

$$\hat{\hat{y}}_T := \sum_{j=0}^{L} b_j \hat{x}_{T-j}^\infty \tag{23}$$

It appears that RTSE adds an extra aggregation layer to the vintage filtering/smoothing 20 of the data. We now briefly analyze the consequence of this additional complexity layer.

Formally, one can plug equation 20 into 23 and rearrange terms in two different ways: the outer summand in 25 refers to vintages (for fixed h the data-point $x_{T-(j+h)}^j$ wanders along the h-th column of the *vintage* triangle) whereas the outer-index of 26 emphasizes releases (for fixed h the data-point x_{T-j}^h wanders along the h-th column of the *release* triangle):

$$
\begin{aligned}
\hat{y}_T &= \sum_{j=0}^{T-1} b_j \left(\sum_{l=-j}^{T-j-1} \sum_{h=0}^{j+l} \gamma_{j,j+l}^{j+l-h} x_{T-(j+l)}^{j+l-h} \right) \tag{24} \\
&= \sum_{h=0}^{T-1} \sum_{j=0}^{T-1-h} \left(\sum_{l=0}^{j} b_l \gamma_{l,j+h}^j \right) x_{T-(j+h)}^j \\
&= \sum_{h=0}^{T-1} \sum_{j=0}^{T-1-h} b_{j,j+h} x_{T-(j+h)}^j \tag{25} \\
&= \sum_{h=0}^{T-1} \sum_{j=h}^{T-1} \left(\sum_{l=0}^{j} b_l \gamma_{l,j}^h \right) x_{T-j}^h \\
&= \sum_{h=0}^{T-1} \sum_{j=h}^{T-1} b_{hj} x_{T-j}^h = \sum_{h=0}^{T-1} \sum_{j=0}^{T-1-h} b_{h,j+h} x_{T-(j+h)}^h \tag{26}
\end{aligned}
$$

where we defined

$$b_{hj} := \sum_{l=0}^{j} b_l \gamma_{l,j}^h \tag{27}$$

for $j \geq h$ (and similarly for $b_{j,j+h}$). The first representation 25 is convenient when working with *vintages* (and possibly vintage restrictions). The second one 26 emphasizes *releases*: the index h of its outer summand fixes the h-th data-release. It is in some sense a natural choice because releases are (assumed to be) statistically congruent and because cointegration constraints are more meaningfully applied to releases than to vintages, see the discussion in Hecq and Jacobs (2009)

as well as section 6.1 in the appendix. Interestingly, the reduced-form expression 27 implies that the convolution literally 'integrates out' the l-dimension of the three-dimensional index (l, j, h) of $\gamma_{l,j}^h$. Thus, if we focus on a 'direct' determination of the filter weights b_{hj}, then a whole section of the 'structure' determining the three-dimensional index-space l, j, h of the revision-filter/smoother collapses[10].

We now consider the case of 'extended pure news', as discussed in the previous section, for which $h = 0$ in 25:

$$\hat{y}_T = \sum_{j=0}^{L} b_{\min(h_0-1,j),j} x_{T-j}^{\min(h_0-1,j)} \tag{28}$$

$$\text{where} \quad b_{\min(h_0-1,j),j} = \sum_{l=0}^{j} b_l \gamma_{l,j}^{\min(h_0-1,j)} \tag{29}$$

As shown in the appendix, expression 28 can be transposed into the frequency-domain

$$\sum_{j=0}^{L} b_{\min(h_0-1,j),j} \exp(-ij\omega_k) \Xi_{TX}^{\min(h_0-1,j)}(\omega_k) \tag{30}$$

where $\Xi_{TX}^{\min(h_0-1,j)}(\omega_k)$ is the DFT of the corresponding column $\mathbf{x}^{\min(h_0-1,j)}$ of the *release* triangle[11]. Defining DFT's on the releases, instead of the vintages, makes sense because of our statistical congruency assumption. Accordingly, criterion 7 is generalized to

$$\frac{2\pi}{T} \sum_{k=-T/2}^{T/2} \left| \Gamma(\omega_k) \Xi_{TX}^{h_0-1}(\omega_k) - \sum_{j=0}^{L} b_{\min(h_0-1,j),j} \exp(-ij\omega_k) \Xi_{TX}^{\min(h_0-1,j)}(\omega_k) \right|^2 \rightarrow \min_{b_{\min(h_0-1,j),j}} \tag{31}$$

We refer to the appendix for details about the derivation of this expression which is valid for stationary data. It is easily seen that criterion 31 ('extended pure news' case) is a generalization of 7 by noting that $\Xi_{TX}^j(\omega_k) = \Xi_{TX}^0(\omega_k)$ for all j in the absence of revisions. Customization of the mean-square criterion can be obtained by plugging

$$\tilde{\Gamma}(\omega_k) := \sum_{j=0}^{L} b_{\min(h_0-1,j),j} \exp(-ij\omega_k) \frac{\Xi_{TX}^{\min(h_0-1,j)}(\omega_k)}{\Xi_{TX}^{h_0-1}(\omega_k)} \tag{32}$$

into 16 and by applying 18. The structure imposed by the extended pure news case spills over the

---

[10]Trivially, a projection must occur because the parameter triangle is two-dimensional 'only': 27 then assigns a closed-form expression to this projection.

[11]For ease of exposition and notation we neglected the fact that the time series upon which $\Xi_{TX}^{\min(h_0-1,j)}(\omega_k)$ are based are not of equal length: this issue can be tackled easily in practice and therefore we here omit a discussion on a bland topic.

parameter *release* triangle:

$$
\begin{pmatrix}
 & & 0 & 0 & \dots & 0 \\
 & & b_{h_0-1,L} & 0 & \dots & 0 \\
 & & \dots & \dots & \dots & \dots \\
 & & b_{h_0-1,h_0} & 0 & \dots & 0 \\
 & & b_{h_0-1,h_0-1} & 0 & \dots & 0 \\
 & \dots & & 0 & \dots & 0 \\
 & b_{11} & & 0 & \dots & 0 \\
 b_{00} & & & 0 & \dots & 0
\end{pmatrix}
$$

In contrast to standard multivariate frameworks the observed 'single-lag' arrangement of filter parameters is pretty uncommon.

Next up, we consider general revision dynamics by adopting the data arrangement entailed by the *release* triangle 2. Specifically, we rewrite the time-domain expression 26 by accounting for the upper-limit h_0 of the number of useful releases as well as for a finite filter-length L

$$
\sum_{h=0}^{h_0-1} \sum_{j=0}^{L-h} b_{h,j+h} x_{T-(j+h)}^h
$$

Transposed into the frequency-domain this expression becomes (see the appendix for further details):

$$
\sum_{h=0}^{h_0-1} \sum_{j=0}^{L-h} b_{h,j+h} \exp(-i(h+j)\omega_k) \Xi_{TX}^h(\omega_k)
$$
$$
= \sum_{h=0}^{h_0-1} \hat{\Gamma}^h(\omega_k) \Xi_{TX}^h(\omega_k) \tag{33}
$$

where

$$
\hat{\Gamma}^h(\omega_k) := \exp(-ih\omega_k) \sum_{j=0}^{L-h} b_{h,j+h} \exp(-ij\omega_k) \tag{34}
$$

(we assume $L \geq h_0$ as usual). The left/upper-triangular shape of the parameter *release* triangle 19 is accounted for by the back-shift operator $\exp(-ih\omega_k)$ in this expression. The associated generalized optimization criterion becomes

$$
\frac{2\pi}{T} \sum_{k=-T/2}^{T/2} \left| \Gamma(\omega_k) \Xi_{TX}^{h_0-1}(\omega_k) - \sum_{h=0}^{h_0-1} \hat{\Gamma}^h(\omega_k) \Xi_{TX}^h(\omega_k) \right|^2 \to \min_{b_{h,j+h}} \tag{35}
$$

($L \geq h_0$ is assumed). An extension to non-stationary processes is proposed in the appendix, see 58.

14

If later vintages dominate earlier releases we may use the parametrization introduced in the previous section and plug 22 into 27 to obtain:

$$b_{h,j+h} \; := \; \alpha^j b_{j+h,j+h} \tag{36}$$

For $h < h_0$. Note that for $j + h > h_0 - 1$ we formally identify the (fictive) parameter $b_{j+h,j+h}$ with $b_{h_0-1,j+h}/\alpha^{j+h-(h_0-1)}$. As a result

$$\hat{\Gamma}^h(\omega_k) := \exp(-ih\omega_k) \sum_{j=0}^{L-h} \alpha^j b_{j+h,j+h} \exp(-ij\omega_k) \tag{37}$$

and the number of freely determined filter parameters would match the 'extended pure news' case 31, up to the additional weighting α linking *releases*. The atypical structure imposed by the revision framework on the multivariate filter design implies that the filter weights in $\hat{\Gamma}^h(\omega_k)$ in 37 are an (uncommon) combination of cross-sectional exponential decay α^j and of proper longitudinal signal-extraction $b_{j+h,j+h}$ weights, as summarized in the parameter *release* triangle:

$$
\begin{pmatrix}
& & & 0 & 0 & \dots & 0 \\
& & & \dots & 0 & \dots & 0 \\
& & & 0 & 0 & \dots & 0 \\
\alpha^L b_{L,L} & \alpha^{L-1} b_{L,L} & & \alpha^{L+1-h_0} b_{L,L} & 0 & \dots & 0 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots \\
\alpha^{h_0+1} b_{h_0+1,h_0+1} & \alpha^{h_0} b_{h_0+1,h_0+1} & \dots & \alpha^2 b_{h_0+1,h_0+1} & 0 & \dots & 0 \\
\alpha^{h_0} b_{h_0,h_0} & \alpha^{h_0-1} b_{h_0,h_0} & \dots & \dots & \alpha b_{h_0,h_0} & 0 & \dots & 0 \\
\alpha^{h_0-1} b_{h_0-1,h_0-1} & \alpha^{h_0-2} b_{h_0-1,h_0-1} & \dots & \dots & b_{h_0-1,h_0-1} & 0 & \dots & 0 \\
\dots & \dots & \dots & & 0 & \dots & 0 \\
\alpha b_{11} & b_{11} & & & 0 & \dots & 0 \\
b_{00} & & & & 0 & \dots & 0
\end{pmatrix} \tag{38}
$$

For notational ease we substituted the (fictive) expression $\alpha^k b_{h_0-1+k,h_0-1+k}$ to the (existing) parameter b_{h_0-1,h_0-1+k} in all the above expressions. The generalization of the previous mean-square criteria to the customization in 18 is straightforward.

4.3   An Extension to Non-Stationary Time Series

Hecq and Jacobs (2009) argue "This paper clarifies the link between different ways of modeling cointegration in nonstationary real-time data within a multivariate dynamic time series framework, using VAR, VECM and restricted VAR representations. We emphasize the alternative ways to deal with the *diagonals* of the *vintage triangle* and observe the pros and cons of the different approaches" and, further on p.15 "The presence of cointegration implies that a linear combination

between *diagonals* is stationary". We position our approach in continuation of this discussion and consider diagonals of the *vintage* triangle - columns of the *revision* triangle - as well. However, we propose a generalization of the traditional cointegration concept which tackles RTSE through specific filter constraints in the unit root frequency/ies (rather than through model-constraints). We briefly sketch the topic: a comprehensive technical treatment is provided in Wildi (2008.2).

In the following we assume that x_t in 3 is integrated with a (single) unit root in frequency zero; we also assume that $\Gamma(0) > 0$ such that y_t and x_t are indeed (co)integrated[12]. Then any of the proposed univariate minimization criteria in section 3.1 would (attempt to) establish cointegration between the real-time estimate $\hat{y}_t$ and the signal y_t - implicitly - because the filter mean-square error is minimized. Besides this implicit built-in mechanism we here propose a formal framework for working with the relevant cointegration structure through suitable filter-constraints. To start, we consider the univariate criterion 7. Theorem 10.18 in Wildi (2008.1) implies that the simple filter-constraint

$$\hat{\Gamma}(0) = \Gamma(0)$$

imposed upon the one-sided filter $\hat{\Gamma}(\cdot)$ in the unit-root frequency zero ensures that $\hat{y}_t$ 'tracks' the level of y_t: both series are cointegrated. An intuitive - informal - argument goes as follows

$$
\begin{aligned}
E[(y_t - \hat{y}_t)^2] &= \int_{-\pi}^{\pi} |\Gamma(\omega) - \hat{\Gamma}(\omega)|^2 dS(\omega) \\
&= \int_{-\pi}^{\pi} |\Gamma(\omega) - \hat{\Gamma}(\omega)|^2 \frac{\tilde{S}(\omega)}{|1 - \exp(-i\omega)|^2} d\omega \\
&= \int_{-\pi}^{\pi} \frac{|\Gamma(\omega) - \hat{\Gamma}(\omega)|^2}{|1 - \exp(-i\omega)|^2} \tilde{S}(\omega) d\omega
\end{aligned}
$$

where $S(\cdot)$ is the pseudo-spectral distribution of the non-stationary process generating x_t and $\tilde{S}(\cdot)$ is the spectral distribution of its stationary first differences. The last equality illustrates that traditional spectral factorizations apply, as expected, but one should consider $\frac{|\Gamma(\omega) - \hat{\Gamma}(\omega)|^2}{|1 - \exp(-i\omega)|^2}$ in place of $|\Gamma(\omega) - \hat{\Gamma}(\omega)|^2$. Theorem 10.18 in Wildi (2008.1) provides a rigorous derivation of this significant notational trick. Note that the latter expression above emphasizes the importance of the proposed filter restriction (a simple first order regularity argument is needed to ensure that the singularity is removed effectively). A general test for verifying (falsifying) the pertinence of the proposed real-time filter restriction is proposed in chapter 6 in Wildi (2008.1).

In the multivariate revision context of the previous section we assume that all pairwise cointegration vectors linking the columns of the *release* triangle are $(1, -1)$, see Hecq and Jacobs (2009).

---

[12]Under standard regularity assumptions about $\Gamma(\omega)$ the time series $y_t - \Gamma(0)x_t$ is stationary.

Then, as shown in Wildi (2008.2), the pertinent filter restriction is

$$\sum_{h=0}^{h_0-1} \hat{\Gamma}^h(0) = \Gamma(0) \tag{39}$$

for the transfer functions defined in 34 (a sketch of the proof is provided in the appendix). In this case, the stationary 'form' 35 of the optimization criterion has to be modified in order to account for the non-stationarity as well as for the long-term equilibrium linking releases, see 58 in the appendix.

5 Conclusion

Combining RTSE and revision dynamics in a unified framework is an exigent statistical estimation problem entailing simultaneous control of one- and multi-step ahead forecasting performances in an unusual (triangularly-shaped) multivariate time series framework. We here propose to unify RTSE and revision topics in a reduced-form form approach which addresses filter parameters 'directly'. Thereby we avoid to specify 'redundant' structure which is integrated out by the additional (signal extraction) aggregation layer. Moreover, our reduced-form approach allows for a variety of practically relevant customizations, by emphasizing a fundamental timeliness-reliability dilemma: the latter generalizes the classical mean-square paradigm. This type of optimization criterion offers a natural interface which allows the user to operationalize a wide range of research priorities. Our approach is generic in the sense that the user could plug any signal specification and any spectral estimate into the proposed optimization criteria. In particular he could replicate classical uni- and multivariate model-based approaches and/or classical filter designs - by providing the corresponding spectral contents - and he could customize performances of such designs. We mark our own (subjective) preference for the (pseudo-) DFT for which open source R-script and illustrative examples are available and maintained on SEFBlog, see the appendix for reference.

6 Appendix

6.1 Vintage vs. Release Triangle: Statistical Congruency

The following assumption about the *release* triangle will be useful when setting-up our approach.

Hypothesis: The columns of the *release* triangle 2 are statistically congruent time series. To be more specific we assume that the columns are either stationary in levels or in first (eventually higher order) differences.

From a subject matter perspective, we assume that data providers process data 'identically' over time i.e. they aim at a reliable publication policy. Our hypothesis seems to be in accordance

17

with the recent approaches proposed by Jacobs and van Norden (2011) and by Hecq and Jacobs (2009). There is, however, a singular occurrence which eventually conflicts with the above assumption, namely so-called benchmark revisions. Hecq and Jacobs consider benchmark revisions in the context of cointegration and they compare two well-known model-structures (observation balanced and vintage balanced systems) with respect to co-breaking. With regards to cointegration, specifically, the authors argue on p.15: "Indeed, it is not very interesting to investigate the presence of long-run relationship between the verticals of the real-time data matrix, i.e. between vintages ... because the series are radically identical except for the last two or three years" and further "The presence of cointegration implies that a linear combination between *diagonals* is stationary". Patterson and Heravi (1991, 2004.1) consider *vintages* of UK GNP and they argue "Was it possible, therefore, to relate GNP on the most recent basis back to previous published series? ... By using the estimated long-run relationship amongst the different vintages, all the time series can be put onto a single and comparable base". According to their findings, *vintages are cointegrated* and the cointegration rank is one. Pain (1994) analyzes forecasts produced by the National Institute of Economic and Social Research and argues "Our results also *fail to reject non-cointegration* between different *vintages* of data, suggesting that considerable care should be exercised in both the choice of realisation data used and in the means by which efficiency is tested". Despite the technically subtle difference between 'not falsifying non-cointegration' and 'falsifying cointegration', located on the level of the respective null-hypotheses, we can infer that the latter author seems to disagree with the formerly cited. Going further, Knetsch and Reimers (2009) consider "The Case of German Production and Orders Statistics" and they argue: "The distinct *vintage* transformation functions found for industrial production and orders imply that the *benchmark revision* alters the estimates of the cointegrating relation. This corollary is rather unpleasant because cointegrating relations are typically interpreted as long-run economic relationships and, thus, of central interest in many empirical applications". Addressing benchmark revisions specifically, Patterson and Heravi (2004.2) document difficulties and pitfalls associated with linear rebiasing. In this context, Knetsch and Reimers (2009) argue "Moreover, hypothesis tests check whether benchmark revisions are innocuous with regard to the parameters of cointegrating relations and whether differencing and rebasing are inadequate methods for adjusting real-time data for *benchmark revisions*. According to theoretical arguments as well as the empirical evidence from the application at hand, *vintage* transformation functions estimated by cointegrating regressions have been proven to be flexible means of creating congruent real-time data sets". Let us now attempt to distil a pragmatic strategy from these findings.

The above partially conflicting outcomes reflect the diversity of application fields (type of national time series), tests (is the null-hypothesis cointegration or non-cointegration) or data arrangements (vintages or diagonals). Unsurprisingly, no fixed pattern emerges from our brief review. As

a result, we here take on a pragmatic perspective and assume that our hypothesis (statistical congruency of releases) is satisfied up to benchmark revisions and that the latter can be tackled either by 'standard' rebasing or by more sophisticated techniques as proposed by Knetsch and Reimers (2009), for example. In any case, we assume that our hypothesis is satisfied after a suitable transformation of the data. This choice dictates cointegration across *releases* (columns in 2) rather than across *vintages* (columns in 1): we here follow the convincing argument proposed by Hecq and Jacobs (2009), as cited above.

6.2 State-Space Approach

The generic state-space model in Jacobs and van Norden (2011) starts as

$$\mathbf{x}_t = \mathbf{Z}\mathbf{a}_t$$
$$\mathbf{a}_{t+1} = \mathbf{T}\mathbf{a}_t + \mathbf{R}\mathbf{o}_t$$

The observation vector is defined as $\mathbf{x}_t = (x_t^1, ..., x_t^l)'$[13] which corresponds to rows in either *vintage* or *release* forms 1, 2. The state vector $\mathbf{a}_t$ and the observation matrix $\mathbf{Z}$ are partitioned according to

$$\mathbf{a}_t = (x_t^\infty, \mathbf{f}_t, \mathbf{m}_t, \mathbf{n}_t)'$$
$$\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3, \mathbf{Z}_4)$$

where $\mathbf{f}_t$ is a b-dimensional vector describing time-dynamics of the 'true' x_t^∞, $\mathbf{m}_t$ is an l-dimensional news component, $\mathbf{n}_t$ is an l-dimensional noise component, $\mathbf{Z}_1$ is an l-dimensional vector of one's, $\mathbf{Z}_2$ is an $l*b$-matrix of zeroes, and $\mathbf{Z}_3$, $\mathbf{Z}_4$ are $l*l$-dimensional identities. Accordingly, the state-transition matrix $\mathbf{T}$ can be partitioned into

$$\mathbf{T} = \begin{pmatrix} T_{11} & \mathbf{T}_{12} & \mathbf{0} & \mathbf{0} \\ \mathbf{T}_{21} & \mathbf{T}_{22} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{T}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{T}_4 \end{pmatrix}$$

where the above blocks are conformably defined matrices corresponding to the partition of $\mathbf{a}_t$. Finally, $\mathbf{R}$ is partitioned into an $(1 + b + 2l)*r$ matrix

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 & \mathbf{R}_3 & \mathbf{0} \\ \mathbf{R}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{U}_1\mathrm{diag}(\mathbf{R}_3) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_4 \end{pmatrix}$$

---

[13]The authors allow for a publication lag of one time unit which explains that the first release is x_t^1 (whereas we assumed x_t^0 to be the first one).

where $\mathbf{U}_1$ is $l*l$ with zeroes below the main diagonal and ones everywhere else, $\mathbf{R}_3 = (\sigma_{m1}, ..., \sigma_{ml})$ where σ_{mi} is the standard error associated with x_t^i and $\mathbf{R}_4$ is $l*l$. Finally, the error term in the transition matrix is partitioned into $\mathbf{o}_t = (\mathbf{o}'_{et}, \mathbf{o}'_{mt}, \mathbf{o}'_{nt})'$ where the sub-vectors address errors associated to the 'true' values, the news and the noise components respectively. Jacobs and van Norden populate the system-matrices by discriminating various practically relevant cases.

The pure noise case is set-up by defining

$$\mathbf{a}_t = \begin{pmatrix} x_t^\infty \\ \mathbf{f}_t \\ \mathbf{n}_t \end{pmatrix} \;,\; \mathbf{T} = \begin{pmatrix} T_{11} & \mathbf{T}_{12} & \mathbf{0} \\ \mathbf{T}_{21} & \mathbf{T}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \;,\; \mathbf{R} = \begin{pmatrix} \mathbf{R}_1 & \mathbf{0} \\ \mathbf{R}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_4 \end{pmatrix}$$

and $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_4)$. The noise structure is imposed by setting

$$\mathbf{R}_4 = \begin{pmatrix} \sigma_{n1} & 0 & . & 0 \\ 0 & \sigma_{n2} & . & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & . & \sigma_{nl} \end{pmatrix}$$

i.e. the noise component is independently (but not necessarily identically) distributed. An increasing precision of observations over time could be ensured by imposing monotonically increasing noise variances $\sigma_{ni} < \sigma_{nj}$ for $i < j$.

The pure news case is accounted for by setting

$$\mathbf{a}_t = \begin{pmatrix} x_t^\infty \\ \mathbf{f}_t \\ \mathbf{m}_t \end{pmatrix} \;,\; \mathbf{T} = \begin{pmatrix} T_{11} & \mathbf{T}_{12} & \mathbf{0} \\ \mathbf{T}_{21} & \mathbf{T}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \text{ and } \mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)$$

Imposing the news structure is achieved through

$$\mathbf{R} = \begin{pmatrix} \mathbf{R}_1 & \mathbf{R}_3 \\ \mathbf{R}_2 & \mathbf{0} \\ \mathbf{0} & -\mathbf{U}_1\mathrm{diag}(\mathbf{R}_3) \end{pmatrix}$$

such that

$$x_{t+1}^\infty = T_{11}x_t^\infty + \mathbf{T}_{12}f_t + \mathbf{R}_1\mathbf{o}_{et} + \mathbf{R}_3\mathbf{o}_{mt}$$

$$\mathbf{m}_{t+1} = -\mathbf{U}_1\mathrm{diag}(\mathbf{R}_3)\mathbf{o}_{nt} = \begin{pmatrix} \sigma_{m1} & \sigma_{m2} & . & \sigma_{ml} \\ 0 & \sigma_{m2} & . & \sigma_{ml} \\ . & . & . & . \\ 0 & . & 0 & \sigma_{ml} \end{pmatrix}$$

The model implies that x_t^j encompasses x_t^i, in informational terms, if $j > i$ i.e. more recent vintages encompass earlier ones.

Spillovers correspond to relationships between measurement errors in different economic time periods. They are structurally unrelated to the previous two revision components and can be accounted for by $\mathbf{T}_3$ or $\mathbf{T}_4$ in the state transition matrix $\mathbf{T}$.

The most general revision process in this approach is obtained by a superposition of all three components.

Dynamics of true values are tackled by the system matrices $T_{11}, \mathbf{T}_{12}, \mathbf{T}_{21}, \mathbf{T}_{22}$ in the general state transition matrix $\mathbf{T}$, as well as by the variance covariance matrices $\mathbf{R}_1$ and $\mathbf{R}_2$. An ARMA process is obtained by setting

$$
\mathbf{n}_t = (x_{t-1}^\infty, ..., x_{t-p+1}^\infty, \epsilon_t, ..., \epsilon_{t-q+1})'
$$

$$
\begin{pmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{pmatrix} = \begin{pmatrix} \sigma_\epsilon \\ \mathbf{1}_{p-1} \\ \sigma_\epsilon \\ \mathbf{1}_{q-1} \end{pmatrix}
$$

$$
\begin{pmatrix} T_{11} & \mathbf{T}_{12} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{AR}' & \mathbf{MA}' \\ \mathbf{I}_{p-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{q-1}|\mathbf{0} \end{pmatrix}
$$

where $\mathbf{AR}$ and $\mathbf{MA}$ contain the autoregressive and moving averages coefficients. Jacobs and van Norden also propose a transposition of a structural time series model which is omitted here.

6.3 DFA and Revisions: Reduced-Form Criterion

We here focus on the link between 28 and 30 in section 4. We first note that the 'no-revision' case has been tackled in Wildi (2008.1), chapter 3. Let us briefly summarize the main ideas behind the original DFA before proceeding to revisions. For this purpose consider the following equality/approximation:

$$
\frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2 = \frac{2\pi}{T}\sum_{k=-T/2}^{T/2}\left|\Xi_{T,Y-\hat{Y}}(\omega_k)\right|^2 \tag{40}
$$

$$
\approx \frac{2\pi}{T}\sum_{k=-T/2}^{T/2}\left|\Gamma(\omega_k) - \hat{\Gamma}(\omega_k)\right|^2 |\Xi_{TX}(\omega_k)|^2 \tag{41}
$$

where y_t is the target signal 3, $\hat{y}_t$ is the output of the one-sided filter 4 (28 in the case of revisions) and $\Xi_{T,Y-\hat{Y}}(\omega_k)$ is the DFT of their difference. The first equality 40 is the standard (finite-sample) spectral factorization of the (finite sample) mean-square filter error. It is a 'number identity' and therefore it can pretend to generality, see proposition 10.4 in Wildi (2008.1): note that these 'numbers' are not observed in finite samples, though, and therefore our identity is in some sense fictive. The approximation in 41 results from the application of a standard finite-sample convolution result, see for example Brockwell and Davis (1993), theorem 10.3.1: now these 'numbers' are observed. A less-known result addresses the magnitude of the approximation error: Proposition 10.8 in Wildi (2008.1) asserts that 41 is a superconsistent estimate of the mean-square filter error i.e. the approximation error is of smaller magnitude than the traditional $1/\sqrt{T}$-order under fairly general assumptions about the DGP of x_t as well as the transfer functions $\Gamma(\cdot)$ and $\hat{\Gamma}(\cdot)$ (including, for example, traditional ARMA-models). An extension of the results to non-stationary integrated processes can be obtained very easily by noting that the right-hand side of 40 addresses the filter error: the latter is generally (supposed to be) stationary even if the DGP of x_t isn't. To be more precise: one can impose real-time filter constraints such that signal and real-time estimates are cointegrated, see section 4.3. Therefore, the validity of standard frequency-domain results is still ensured. However, some care is necessary when referring to 41 because the convolution unfolds the non-stationarity. As shown in Wildi (2008.1), chapter 6, the DFT of an integrated process may be affected by severe bias due to unit-root leakage over the whole frequency spectrum. Therefore, Wildi (2008.1) (chapter 6) proposes to use the pseudo-DFT

$$\frac{\Xi_{T\Delta^d X}(\omega_k)}{(1-\exp(-i\omega_k))^d}$$

(where $\Delta^d X$ are stationary differences and d is the integration order) in lieu of the ordinary DFT in 41 and to impose filter constraints in frequency zero such that the resulting singularities are removed (this proceeding extends straightforwardly to unit-roots in arbitrary frequencies). A rigorous derivation of these results is proposed in theorem 10.18 in the cited literature. Note that the mean-square filter error (left side of 40) is generally not observable, because the signal y_t isn't, whereas 41 is (observable): therefore filter parameters can be optimized by minimizing the latter expression. Efficiency arguments of the corresponding DFA rely on the fact that 41 is a (uniformly) superconsistent estimate of 40 and that the latter is an (asymptotically) efficient estimate of the true unknown mean-square filter error $E[(y_t - \hat{y}_t)^2]$, see propositions 10.16 and 10.17 in Wildi (2008.1). Moreover, 41 lends itself to powerful customization, as illustrated in 13 or 18. Taken together, these formal results underline the potential of the 'Direct' Filter-Approach.

We now address the 'revision'-part by emphasizing the 'extended pure news' case, for simplicity:

$$
\frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2 = \frac{1}{T}\sum_{t=1}^{T}\left(y_t - \sum_{j=0}^{L} b_{\min(h_0-1,j),j} x_{t-j}^{\min(h_0-1,j)}\right)^2
$$

$$
= \frac{2\pi}{T}\sum_{k=-T/2}^{T/2}\left|\Xi_{TY-\hat{Y}}(\omega_k)\right|^2 \tag{42}
$$

$$
\approx \frac{2\pi}{T}\sum_{k=-T/2}^{T/2}\left|\Gamma(\omega_k)\Xi_{TX}^{h_0-1}(\omega_k) - \sum_{j=0}^{L} b_{\min(h_0-1,j),j}\exp(-ij\omega_k)\Xi_{TX}^{\min(h_0-1,j)}(\omega_k)\right|^2 \tag{43}
$$

In analogy to the standard 'no revision' case, criterion 31 (expression 43) minimizes the (unobservable) finite sample filter error up to the approximation error linking 42 and 43. In the absence of revisions, expression 43 simplifies to 41 since $\Xi_{TX}^{\min(h_0-1,j)}(\omega_k) = \Xi_{TX}^{0}(\omega_k)$ can be isolated as outer periodogram. If the different DFT's cannot be isolated as a single outer periodogram, then the original results in Wildi (2008.1) do not apply anymore. In this case we have to invoke the more general theorem 7.1 in Wildi (2008.2). Let us sketch the point linking 28 and 30. For this purpose we compare the DFT $\Xi_{T\hat{Y}}(\omega_k)$ of the real-time estimate $\hat{y}_t$ with the frequency domain expression 30 found in the optimization criterion: note that $\Xi_{TY-\hat{Y}}(\omega_k) = \Xi_{TY}(\omega_k) - \Xi_{T\hat{Y}}(\omega_k)$ and therefore $\Xi_{T\hat{Y}}(\omega_k)$ is a 'key-element' in 42, which is identified with $\sum_{j=0}^{L} b_{\min(h_0-1,j),j}\exp(-ij\omega_k)\Xi_{TX}^{\min(h_0-1,j)}(\omega_k)$ in 43 (whereas $\Xi_{TY}(\omega_k)$ is identified with $\Gamma(\omega_k)\Xi_{TX}^{h_0}(\omega_k)$). We here check the pertinence of this identification:

$$
\Xi_{T\hat{Y}}(\omega_k) = \frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T}\left(\sum_{j=0}^{L} b_{\min(h_0-1,j),j} x_{t-j}^{\min(h_0-1,j)}\right)\exp(-it\omega_k)
$$

$$
= \sum_{j=0}^{L} b_{\min(h_0-1,j),j}\exp(-ij\omega_k)\left(\frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T} x_{t-j}^{\min(h_0-1,j)}\exp(-i(t-j)\omega_k)\right) \tag{44}
$$

$$
= \sum_{j=0}^{L} b_{\min(h_0-1,j),j}\exp(-ij\omega_k)\left(\frac{1}{\sqrt{2\pi T}}\sum_{t=1}^{T} x_{t}^{\min(h_0-1,j)}\exp(-it\omega_k)\right) + R \tag{45}
$$

$$
= \sum_{j=0}^{L} b_{\min(h_0-1,j),j}\exp(-ij\omega_k)\Xi_{TX}^{\min(h_0-1,j)}(\omega_k) + R
$$

The error term R appearing in 45 comes from replacing the time-index $t-j$ by t in the inner summand: incidentally, $\Xi_{T\hat{Y}}(\omega_k)$ is not observable but $\sum_{j=0}^{L} b_{jj}\exp(-ij\omega_k)\Xi_{TX}^{\min(h_0-1,j)}(\omega_k)$ is. If L is small, then it is readily seen that most summands in the inner sums of 44 and 45 coincide and therefore the error-term is small (of order $L/\sqrt{T}$ in absolute mean, assuming a stationary process with finite first moment generating x_t). In fact R remains of order $1/\sqrt{T}$ if we assume that $b_{\min(h_0-1,j),j}$ converge to zero at a suitable rate, see proposition 7.4 in Wildi (2008.2). The

23

technical proof of the superconsistency in this proposition brings into play an additional aggregation layer, namely the outer-sum in the optimization criterion (the sum over the frequency-grid ω_k in expression 43 above). Superconsistency is obtained although R in 45 is of order $1/\sqrt{T}$ because the additional aggregation layer brings into play kind of a generalized law of large numbers argument which is evoked in Brockwell and Davis (1993), proposition 10.8.5. As an outcome, the error term in 43 is of smaller order than $1/\sqrt{T}$ (in absolute mean) if all 'filter-terms' (one-sided filter $b_{\min(h_0-1,j),j}$ and bi-infinite filter γ_k) decay sufficiently rapidly towards zero (classical ARMA models satisfy this restriction, for example). This 'technical' finite-sample multivariate convolution result links 28 and 30 formally, as intended.

6.4 Non-Stationarity: Cointegration and Filter Constraints

Consider the case of multiple (f) cointegration constraints linking x_t and w_{nt}, $n = 1, ..., m$, in a general multivariate framework:

$$X_t - \alpha_{11}W_{1t} - ... - \alpha_{1m}W_{mt} \sim I(0)$$
$$X_t - \alpha_{21}W_{1t} - ... - \alpha_{2m}W_{mt} \sim I(0)$$
$$...$$
$$X_t - \alpha_{f1}W_{1t} - ... - \alpha_{fm}W_{mt} \sim I(0)$$

Any linear combination

$$\sum_{j=1}^{f} w_j X_t - \sum_{j=1}^{f} w_j \alpha_{j1} W_{1t} - ... - \sum_{j=1}^{f} w_j \alpha_{jm} W_{mt}$$

is stationary (it is assumed from now on that $\sum_{j=1}^{f} w_j = 1$). Wildi (2008.2) then shows that

$$\hat{\Gamma}_{W_i}(0) = \left(\sum_{j=1}^{f} w_j \alpha_{ji} \right) (\Gamma(0) - \hat{\Gamma}_X(0)), i = 1, ..., n \tag{46}$$

are both necessary and sufficient filter restrictions in order to impose stationarity of the (real-time) filter error i.e. the real-time estimate and the signal are cointegrated. Transposed to the framework of section 4.3 we deduce:

$$\alpha_{ij} = \delta_{ij} , \ 0 \le i, j \le h_0 - 1$$

where the Kronecker-delta $\delta_{ij} = \begin{cases} 1, \ i = j \\ 0 \text{ otherwise} \end{cases}$ reflects the pairwise (1,-1) cointegration vectors.

Using the notation introduced in section 3.2 and definition 34 we have $\hat{\Gamma}^h(\cdot) = \hat{\Gamma}_{W_i}(\cdot)$, $h = 1, ..., h_0 - 2$ and $\hat{\Gamma}^{h_0-1}(\cdot) = \hat{\Gamma}_X(\cdot)$. Inserting into 46 and integrating out w_j (summing-up over h) we then

obtain:

$$\sum_{h=0}^{h_0-2} \hat{\Gamma}^h(0) = \sum_{h=0}^{h_0-2}\left(\sum_{j=0}^{h_0-2} w_j \delta_{jh}\right)(\Gamma(0) - \hat{\Gamma}^{h_0-1}(0)) \tag{47}$$

$$= (\Gamma(0) - \hat{\Gamma}^{h_0-1}(0)) \tag{48}$$

confirming 39. We now generalize criterion 35, valid in the stationary case, to the postulated non-stationary setting, specifically: I(1)-DGP's and $(1, -1)$ cointegration vectors linking *releases* $x_t^{h_0-1}$ and x_t^h, for $h < h_0 - 1$. For this purpose we consider the filter error $r_T = y_T - \hat{y}_T$ (our selection $t = T$ is merely for notational convenience since we thereby emphasize the real-time aspect):

$$r_T = \sum_{l=-\infty}^{\infty} \gamma_l x_{T-l}^{h_0-1} - \left\{ \sum_{l=h_0-1}^{L} b_{h_0-1,l} x_{T-l}^{h_0-1} + \sum_{h=0}^{h_0-2}\left(\sum_{l=h}^{L} b_{hl} x_{T-l}^h\right)\right\}$$

Note that r_T cannot be observed because the signal y_T is unknown. We now decompose this filter error according to theorem 7.1 in Wildi (2008.2): in the following notation 'non-existing' coefficients $b_{h_0 j}$ are set to zero for j outside of the natural domain of definition of the index.

$$r_T = \sum_{l=-\infty}^{\infty} \gamma_l x_{T-l}^{h_0-1} - \left\{ \sum_{l=h_0-1}^{L} b_{h_0-1,l} x_{T-l}^{h_0-1} + \sum_{h=0}^{h_0-2}\left(\sum_{l=h}^{L} b_{hl} x_{T-l}^h\right)\right\}$$

$$= \sum_{l=-\infty}^{\infty} (\gamma_l - b_{h_0-1,l})\left(x_T^{h_0-1} + (x_{T-l}^{h_0-1} - x_T^{h_0-1})\right) - \sum_{h=0}^{h_0-2}\left(\sum_{l=h}^{L} b_{hl}\left(x_T^h + (x_{T-l}^h - x_T^h)\right)\right) \tag{49}$$

$$= (\Gamma(0) - \hat{\Gamma}^{h_0-1}(0))x_T^{h_0-1} - \sum_{h=0}^{h_0-2} \hat{\Gamma}^h(0)x_T^h \tag{50}$$

$$+ \sum_{l=-\infty}^{\infty} (\gamma_l - b_{h_0-1,l})(x_{T-l}^{h_0-1} - x_T^{h_0-1}) \tag{51}$$

$$- \sum_{h=0}^{h_0-2}\left(\sum_{l=h}^{L} b_{hl}(x_{T-l}^h - x_T^h)\right) \tag{52}$$

where we used

$$\hat{\Gamma}^h(0)x_T^h = \left(\sum_{l=h}^{L} b_{hl}\right)x_T^h$$

which are extracted from 49. The main intention behind this decomposition is that non-stationary variables in level appear in 50: we can thus impose 'cointegration' of y_t and $\hat{y}_t$ - equivalently: a finite mean-square filter error $E[r_t^2]$ - through suitable filter constraints in frequency zero. The link between 50 and the filter constraint 48 is straightforward. The remaining terms 51 and 52 put in evidence differences with variable lag (time-span): if all filter coefficients decay sufficiently rapidly, asymptotically, then the potential non-stationarity resulting from the increasing spread of the difference-lag is contained and r_t is stationary. This last result illustrates the importance

of 'regularity constraints' afforded by theorem 7.1 in Wildi (2008.2) (readers interested in formal derivations of the above claims are referred to the proof of the theorem). We now define

$$c_T := (\Gamma(0) - \hat{\Gamma}^{h_0-1}(0))x_T^{h_0-1} - \sum_{h=0}^{h_0-2} \hat{\Gamma}^h(0)x_T^h \tag{53}$$

Imposing 48 implies that the process c_t is stationary. Let $\Xi_{TC}(\omega_k)$ designate the corresponding DFT (of a finite realization of c_t). Theorem 7.1 shows that the DFT of r_t (which cannot be observed) can be estimated by the following frequency-domain expression, relying on 50-52:

$$\Xi'_{Tr}(\omega_k) \quad := \quad \left\{ \Delta\Gamma_X(0)\Xi_{TC}(\omega_k) \right. \tag{54}$$

$$- \exp(i\omega_k)\left[\frac{\Delta\Gamma_X(0) - \exp(-i\omega_k)\Delta\Gamma_X(\omega_k)}{1 - \exp(-i\omega_k)} - \Delta\Gamma_X(0)\right]^+ \Xi_{T\Delta X}^{h_0-1}(\omega_k) \tag{55}$$

$$+ \left[\frac{\Delta\Gamma_X(0) - \exp(-i\omega_k)\Delta\Gamma_X(\omega_k)}{1 - \exp(-i\omega_k)}\right]^- \Xi_{T\Delta X}^{h_0-1}(\omega_k) \tag{56}$$

$$\left. + \sum_{h=0}^{h_0-2} \exp(i\omega_k)\left[\frac{\hat{\Gamma}^h(0) - \exp(-i\omega_k)\hat{\Gamma}^h(\omega_k)}{1 - \exp(-i\omega_k)} - \hat{\Gamma}^h(0)\right] \Xi_{T\Delta X}^h(\omega_k) \right\} \tag{57}$$

where $\Delta\Gamma_X(\omega_k) := \Gamma(\omega_k) - \hat{\Gamma}^{h_0-1}(\omega_k)$, $\Xi_{T\Delta X}^h(\omega_k)$ is the DFT of $x_t^h - x_{t-1}^h$, $h = 0, ..., h_0 - 1$, and where the 'plus' and 'minus' superscripts in 55 and 56 mean the following positive and negative expansions:

$$\left[\frac{\Delta\Gamma_X(0) - \exp(-i\omega_k)\Delta\Gamma_X(\omega_k)}{1 - \exp(-i\omega_k)}\right]^- = \sum_{l=-\infty}^{-1}\left(\sum_{j=-\infty}^{l}\gamma_j\right)\exp(-il\omega_k)$$

$$\left[\frac{\Delta\Gamma_X(0) - \exp(-i\omega_k)\Delta\Gamma_X(\omega_k)}{1 - \exp(-i\omega_k)}\right]^+ = \sum_{l=0}^{\infty}\left(\sum_{j=l}^{\infty}(\gamma_j - b_{h_0-1,j})\right)\exp(-il\omega_k)$$

(note that $b_{h_0-1,j} = 0$ for $j < h_0 - 1$ and therefore it does not appear in the upper formula, right side). Finally

$$\frac{\hat{\Gamma}^h(0) - \exp(-i\omega_k)\hat{\Gamma}^h(\omega_k)}{1 - \exp(-i\omega_k)} = \sum_{l=0}^{\infty}\left(\sum_{j=l}^{\infty}b_{hj}\right)\exp(-il\omega_k)$$

It is again assumed that filter coefficients vanish if subscripts are outside of their domain of definition. The interested reader is referred to the proof of theorem 7.1 in Wildi (2008.2).

Let us briefly comment these results. The estimate $\Xi'_{Tr}(\omega_k)$ as defined by 54 is an efficient estimate of the (unobservable) DFT $\Xi_{Tr}(\omega_k)$ of the (unobservable) filter error r_T in the case of non-stationary (cointegrated) releases. Therefore, filter coefficients b_{hj} can be determined by minimizing

$$\sum_{|k|\leq T/2} \left|\Xi'_{Tr}(\omega_k)\right|^2 \rightarrow \min_{b_{hj}} \tag{58}$$

Criterion 58 is an extension of 35 to the non-stationary case[14]. It inherits all (uniform) efficiency properties applying to the univariate DFA and/or the multivariate stationary case 35: the resulting real-time filter minimizes an efficient estimate of the unobservable mean-square filter error. Customizations in view of tackling the timeliness-reliability dilemma could be obtained in the same vein as 18. Note, once again, that the user could supply model-based spectral representations instead of DFT's. Releases x_t^h enter in *levels* through 54 and in stationary *first differences* through 55-57. The latter expressions highlight the importance of the pseudo-DFT's which, in turn, emphasize the relevance of the pseudo-spectral densities and cross pseudo-spectral densities. Relaxing the intuitive filter constraint $\Delta\Gamma_X(0) = 0 = \Gamma(0) - \hat{\Gamma}^{h_0-1}(0)$ by imposing the less stringent restriction 48, instead, implies that the stationary cointegration residuum 53 enters into the optimization through 54. Note that the cointegration term c_t addresses the filter fit $\Delta\Gamma_X(0)$ in frequency zero, only. All other frequencies ω_k, $k > 0$, are entering the optimization through the pseudo-DFT's. Potential non-singularities in the pseudo-spectral representations 55-57 are removed by imposing regularity conditions ensuring a 'sufficiently fast' asymptotic decay (towards zero) of filter coefficients. Finite MA-filters satisfy these requirements trivially and stable ARMA filters satisfy them automatically. The above design relies heavily on our initial assumptions namely that the pairwise cointegration vectors of the first h_0 releases are $(-1, 1)$, recall the discussion in section 4.3 as well as section 6.1. In case of significant departures from these assumptions theorem 7.1 in Wildi (2008.2) provides a general framework for working with reduced-form 'filter constraints'.

6.5   List of Selected Blog-Entries

6.5.1   Background Information

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/142-Discrete-Fourier-and-Inverse-Fourier-Transforms-an-Atheoretical-Descriptive-Experimental-Approach.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/143-Orthonormal-Decompositions-in-Time-and-Frequency-Domains.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/144-The-Filter-Effect.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/146-Real-Time-Signal-Extraction-RTSE-the-Generic-Model-Based-Perspective.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/148-DFA-vs-Model-Based-Approach-Overfitting,-Richly-Parameterized-Designs-and-the-Virtues-of-Customized-Optimization-Criteria.html

---

[14]With assumed cointegration vectors $(1, -1)$ linking the first release x_t^0 with all successive releases.

6.5.2   Mean-Square Perspective

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/149-DFA-Customization-Minimizing-the-Mean-Square-Filter-Error.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/150-The-DFA-in-a-Mean-Square-Perspective-a-Wrap-Up.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/159-I-DFA-Exercises-Part-I-Mean-Square-Criterion.html

6.5.3   Customization

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/130-Analytical-DFA-Enhancing-Model-Based-Approaches-by-Accounting-for-Reliability-and-Speed-of-Early-Real-Time-Estimates.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/147-Speed-and-Reliability-Issues-an-Introduction.htm

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/160-I-DFA-Exercises-Part-II-Customization-SpeedReliability.html

6.5.4   Filter Constraints (Non-Stationarity)

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/131-Real-Time-Filter-Constraints-and-Unit-Roots.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/133-Some-More-Details-About-Real-Time-Filter-Constraints.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/131-Real-Time-Filter-Constraints-and-Unit-Roots.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/133-Some-More-Details-About-Real-Time-Filter-Constraints.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/137-Real-Time-Filter-Constraints-Revisited-a-Tale-About-the-Platonic-Idea-of-the-Data-Generating-Process.html

6.5.5   Tutorial

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/159-I-DFA-Exercises-Part-I-Mean-Square-Criterion.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/160-I-DFA-Exercises-Part-II-Customization-SpeedReliability.html

6.5.6 Real-Time Indicators

- Real-time indicator for the US: http://www.idp.zhaw.ch/usri

- A review of well-known monthly US-indicators is documented on SEFBlog. It can be acceded by selecting the category "Economic indicators" on the left margin.

6.5.7 R-Code

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/138-Analytical-DFA-a-Review-of-Code-versions,-Examples-and-Parameters.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/159-I-DFA-Exercises-Part-I-Mean-Square-Criterion.html

- http://blog.zhaw.ch/idp/sefblog/index.php?/archives/160-I-DFA-Exercises-Part-II-Customization-SpeedReliability.html

References

[1] Anderson R.G. and Gascon C.S. (2009) Estimating U.S. Output Growth with Vintage Data in a State-Space Framework. *Federal Reserve Bank of St. Louis Review, July/August 2009, 91(4), pp. 349-69.* http://research.stlouisfed.org/publications/review/09/07/Anderson_Gascon.pdf

[2] Aruoba B. (2008) Data Revisions are not Well Behaved. *Journal of Money, Credit and Banking, 40(2-3), 319-340.* Working paper: http://econweb.umd.edu/∼aruoba/research/paper10/well_behaved.pdf

[3] Croushore D. (2009) Frontiers of Real-Time Data Analysis. *FRB of Philadelphia, Working Paper No. 08-4* https://facultystaff.richmond.edu/ dcrousho/docs/Croushore%20Real%20Time%20Frontiers%20

[4] Cunningham, A., J. Eklund, C. Jeffery, G. Kapetanios, and V. Labhard (2009) A state space approach to extracting the signal from uncertain data. *Queen Mary, University of London, School of Economics and Finance in its series Working Papers with number 637.* http://www.econ.qmul.ac.uk/papers/doc/wp637.pdf

[5] McElroy, T., Wildi M. (2010) Signal Extraction Revision Variances as a Goodness-of-Fit Measure. *Journal of Time Series Econometrics.* Working paper available on http://www.census.gov/srd/papers/pdf/rrs2010-06.pdf

[6] McElroy, T., Wildi M. (2011) Multi-Step Ahead Estimation of Time Series Models. *Working paper presented at the 2010 NBER-NSF Time Series Conference:* http://econ.duke.edu/ brossi/NBERNSF/McElroy.pdf

[7] Garratt, Anthony, Kevin Lee, Emi Mise, and Kalvinder Shields (2008) Real Time Representations of the Output Gap. *Review of Economics and Statistics, 90, 792-804.*

[8] Garratt, Anthony, Kevin Lee, Emi Mise, and Kalvinder Shields (2009) Real Time Representations of the UK Output Gap in the Presence of Model Uncertainty. *International Journal of Forecasting, 25, 81-102.*

[9] Hamilton J.D.H. (1994) Time Series Analysis. *Princeton University Press* http://www.cirano.qc.ca/realtime/papers2009/Hecq.pdf

[10] Hecq A. and Jacobs J.P.A.M (2009) On the VAR-VECM Representation of Real Time Data. *Working paper* http://www.cirano.qc.ca/realtime/papers2009/Hecq.pdf

[11] Jacobs J.P.A.M and van Norden S. (2011) Modeling Data Revisions: Measurement Error and Dynamics of "True" Values. *Journal of Econometrics (forthcoming).*

[12] Knetsch T.A. and Reimers H.E. (2009) Dealing with Benchmark Revisions in Real-Time Data: The Case of German Production and Orders Statistics. *Oxford Bulletin of Economics and Statistics Volume 71, Issue 2, pages 209235.* http://www.eabcn.org/workshops/cambridge_2008/documents/Knetsch.pdf

[13] McKenzie, R. (2006) Undertaking revisions and real-time data analysis using the OECD main economic indicators original release data and revisions database. *Technical Report STD/DOC20062, Organisation for Economic Co-operation and Development.*

[14] Pain N. (1994) Cointegration and forecast evaluation: Some lessons from National Institute Forecasts. *Journal of Forecasting Volume 13, Issue 5, pages 481494.*

[15] K. D. Patterson and S. M. Heravi (1991) Are different vintages of data on the components of GDP co-integrated? : Some evidence for the United Kingdom. *Economics Letters Volume 35, Issue 4, April 1991, Pages 409-413.*

[16] K. D. Patterson and S. M. Heravi (2004.1) Revisions to Official Data on U.S. GNP: A Multivariate Assessment of Different Vintages. *Journal of Official Statistics, Vol. 20, No. 4, 2004, pp. 573602*

[17] K. D. Patterson and S. M. Heravi (2004.2) Rebasing, Common Cycles, and Some Practical Implications of Modelling Data Revisions. *Journal of Official Statistics, Vol. 20, No. 4, 2004, pp. 631644*

[18] Swanson, N. R. and Dijk, D. van (2006) Are Statistical Reporting Agencies Getting It Right? Data Rationality and Business Cycle Asymmetry. *Journal of Business and Economic Statistics, Vol. 24, pp. 24-42.*

[19] Wildi, M. (1998) Detection of Compatible Turning Points and Signal Extraction for Non-Stationary Time Series. *Operation Research Proceedings, Springer.*

[20] Wildi, M. (2004) Signal Extraction: How (In)efficient Are Model-Based Approaches? An Empirical Study Based on TRAMO/SEATS and Census X-12-ARIMA. *KOF-Working Paper Nr. 96*, ETH-Zurich.

[21] Wildi, M. (2005) Signal Extraction: Efficient Estimation, 'Unit-Root'-Tests and Early Detection of Turning-Points. *Lecture Notes in Economic and Mathematical Systems, 547*, Springer-Verlag Berlin Heidelberg.

[22] Wildi, M. (2008.1) Real-Time Signal-Extraction: Beyond Maximum Likelihood Principles. http://www.idp.zhaw.ch/fileadmin/user_upload/engineering/_Institute_und_Zentren/IDP/forschungsschwerp

[23] Wildi, M. (2008.2) Efficient Multivariate Real-Time Filtering and Cointegration. *IDP-working paper, IDP-WP-08Sep-01.* http://www.idp.zhaw.ch/fileadmin/user_upload/engineering/_Institute_und_Zentren/IDP/forschungsschwerp WP-08Sep-01.pdf

[24] Wildi, M. (2009) Real-Time US-Recession Indicator (USRI): A Classical Cycle Perspective with Bounceback. *IDP-working paper, IDP-WP-09Jun-06.* http://www.idp.zhaw.ch/en/engineering/idp-institute-of-data-analysis-and-process-design/research/finance-risk-management-and-econometrics/economic-indices/us-economic-recession-indicator.html

[25] Wildi, M. (2011) Real-Time Signalextraction: a Shift of Perspective. *Estudios de Economia Aplicada* [forthcoming]. Working paper: http://blog.zhaw.ch/idp/sefblog/uploads/customized_final.pdf

[26] Wildi, M. (2011.2) I-DFA and I-MDFA: Companion Paper to R-Code Published on SEFBlog. http://blog.zhaw.ch/idp/sefblog/uploads/working_paper1.pdf