Temporal Aggregation and Causality in Multiple Time Series Models*

Jörg Breitung
Humboldt University

Norman R. Swanson
Texas A&M University

January 26, 2001

Abstract

Large aggregation interval asymptotics are used to investigate the relation between Granger causality in disaggregated vector autoregressions (VARs) and associated contemporaneous correlation among innovations of the aggregated system. Our approach allows us to better understand the informational content in non-diagonal error covariance matrices, which play an important role in structural VAR analysis, for example. From a theoretical perspective, we outline various conditions under which the informational content of error covariance matrices yields insight into the causal structure of the VAR. Tis allows us, for example, to summarize the types of empirical questions which can be answered by evaluating the correlation structure of error covariance matrices in aggregated VARs. Monte Carlo results suggest that our asymptotic findings are applicable even when the aggregation interval is small, as long as the time series are not characterized by high levels of persistence.

JEL Classification: C32, C43, C51.

KEYWORDS: instantaneous causality, Granger causality, contemporaneous correlation, temporal aggregation, stock and flow variables.

---

* *Corresponding Author:* Joerg Breitung, Institute of Statistics and Econometrics, Humboldt University, Berlin, Spandauer Str. 1, D-10178 Berlin, Germany. The authors wish to thank Maike Burda, Clive W.J. Granger, Helmut Lütkepohl, Michael Neumann and an anonymous referee for helpful comments. The first author gratefully acknowledges financial support from *Sonderforschungsbereich 373*, Humboldt University. The second author thanks the Private Enterprise Research Center and the Bush Program in the Economics of Public Policy at Texas A&M University for financial support. Parts of this paper were written while the second author was visiting the University of California, San Diego.

1 Introduction

Temporal aggregation poses many interesting questions which have been explored in time series analysis and which yet remain to be explored. An early example of research in this area is Quenouille (1957), where the temporal aggregation of ARMA processes is studied. Amemiya and Wu (1972), and Brewer (1973) refine and generalize Quenouille's result by including exogenous variables. Zellner and Montmarquette (1971) discuss the effects of temporal aggregation on estimation and testing. Engle (1969) and Wei (1978) analyze the effects of temporal aggregation on parameter estimation in a distributed lag model. Granger (1987) discusses the implications of aggregation on systems with common factors. Other important contributions include Tiao (1972), Stram and Wei (1986), Lütkepohl (1987), Weiss (1994), and Marcellino (1999), to name but a few. The findings of these studies can be summarized by quoting Tiao (1999): *"So the causality issue is muddled once the data are aggregated. The problem is that if the data are observed at intervals when the dynamics are not working properly, then we may not get any kind of causality."*

In this paper, we re-examine the impact of temporal aggregation on Granger causal relations in vector autoregressions (VARs). In particular, we use large aggregation interval asymptotics to investigate the relation between Granger causality in the original variables and contemporaneous correlation among the residuals of a temporally aggregated system. From a theoretical perspective, we outline various conditions under which the informational content of error covariance matrices yields insight into the underlying causal structure of the VAR. This allows us to characterize the extent of information loss due to aggregation.

To illustrate the type of problem which we consider, assume that one is interested in analysing a system of three aggregated variables X, Y and Z. Suppose that we observe contemporaneous correlation between X and Y conditional on Z. In this case one may conclude that there is a causal relationship between X and Y (e.g. Dawid (1979)). However, it may be the case that the corresponding disaggregated variables x and y do not exhibit any causal relationship, so that the contemporaneous correlation spuriously indicates causality between the variables. Therefore, it is important to know under which conditions contemporaneous correlation is merely due to temporal aggregation (i.e. to quantify the risk of infering that there is Granger causality when in fact the original variables do not possess any causal linkage. In this paper asymptotic theory for large aggregation intervals is used to derive sufficient conditions for ruling out "spurious causality" stemming from temporal aggregation.

The rest of the paper is organized as follows. In Section 2, we review the concepts of Granger causality and contemporaneous correlation. Section 3 presents our asymptotic results. Section 4 contains the results from a Monte Carlo investigation, where it is found that our large sample results are applicable even when the aggregation interval is small, as long as the time series are not characterized by high levels of persistence. Section 5 concludes. All proofs are gathered in the appendix.

2 Basic definitions

Following Granger (1969) we consider a conditional distribution with respect to two information sets which are available at time t, say $\mathcal{I}_t$ and $\mathcal{I}_t^+ = \{\mathcal{I}_t, x_t, x_{t-1}, \ldots\}$, where x_t denotes a (possibly causal) variable. In the following, we use a conditional mean definition of causality. Specifically, we define a variable x_t to be a Granger cause for the variable y_t (or $x \to y$) if

$$E(y_{t+h}|\mathcal{I}_t) \neq E(y_{t+h}|\mathcal{I}_t^+) \quad \text{for some } h = 1, 2, \ldots \tag{1}$$

Note also that in a bivariate VAR, (1) holds for all h if it holds for $h = 1$ (e.g. see Lütkepohl (2000)). In the sequel, we simplify our discussion by setting $h = 1$ in (1), although our asymptotic results are not limited to the case $h = 1$.

Granger (1969) also gives a definition of *instantaneous causality* (see also Pierce and Haugh (1977)). For a multivariate system with $\xi_t = [x_t, y_t, z_t']'$, where z_t is a m^*-dimensional vector of time series with $m^* \geq 1$, we say that instantaneous causality $x \Rightarrow y$ (or by reasons of symmetry $y \Rightarrow x$) occurs if

$$E(y_t|\mathcal{J}_t) \neq E(y_t|\mathcal{J}_t^+) \ , \tag{2}$$

where $\mathcal{J}_t = \{\mathcal{I}_{t-1}, z_t\}$ and $\mathcal{J}_t = \{\mathcal{I}_{t-1}, z_t, x_t\}$. This definition can be seen as a dynamic version of the causality concept used by Dawis (1979) and Pearl (2000), among others. For example, if ξ_t is white noise, we find that there is no instantaneous causality between x_t and y_t if $E(y_t|z_t, x_t) = E(y_t|z_t)$. This condition is satisfied if x_t and y_t are *conditionally independent* given a sufficient set of variables z_t (see Dawis (1979)). Furthermore, conditional independence implies a causal relationship that can be represented by using directed graphs (see Swanson and Granger (1997) and Pearl (2000) for more details).

As already noted by Granger (1969), an important problem with the definitions is the choice the sampling interval. For example, variables which are Granger causal according to (1) and based on daily data, may not be Granger causal based on monthly data, and vice-versa. In the sequel, we examine two types of temporal aggregation (e.g. see Lütkepohl (1987)). For a flow variable, say y_t, observations are cumulated (or averaged) at k successive time periods in order to form:

$$\bar{y}_t = k^{-1/2} \sum_{j=0}^{k-1} y_{t-j},$$

where the factor $k^{-1/2}$ is introduced in order to obtain a limiting process with finite variance. The aggregated series, $\bar{Y}_T$, results from applying *skip*-sampling (i.e. $\bar{Y}_T = \bar{y}_{kT}$, for $T = 1, 2, \ldots$), where it is assumed that the time series starts at the beginning of the aggregation period. Stock data are aggregated by directly applying the skip-sampling scheme to the data, so that $Y_T = y_{kT}$ for $T = 1, 2, \ldots$.

As a simple illustration of the problems that occur when analyzing causality in aggregated data, assume that the 3×1 vector time series of flow variables $y_t = [y_{1t}, y_{2t}, y_t]'$ has a stationary VAR(1) representation: $y_t = Ay_{t-1} + \varepsilon_t$, where $E(\varepsilon_t \varepsilon_t') =$ - . In Section 3 we show that for $k \to \infty$ the aggregated process, $\bar{Y}_T$, can be represented as a white noise vector with covariance matrix $E(\bar{Y}_T \bar{Y}_T') = (I - A)^{-1}$- $(I - A')^{-1}$. It is important to notice that the aggregated variables no longer have the Granger causality feature. Instead, the causal information in the disaggrated VAR manifests itself as *contemporaneous* correlation among the aggregated variables. Hence, some of the Granger causal information in the contemporaneous correlations can be extracted if sufficient structure is imposed on Σ (see below).

3 Results

Assume that $n = kN$, where N and n are the sample sizes of the aggregated and disaggregated variables. Since $k \to \infty$ implies $n \to \infty$, it is not necessary to assume that N also tends to infinity. Our findings are summarized in the following propositions.[1]

Proposition 1 (Stationary Variables): *Let y_t be generated by an m dimensional linear process $y_t = C_0\varepsilon_t + C_1\varepsilon_{t-1} + C_2\varepsilon_{t-2} + \cdots$, where $C_0 = I_m$, $E(\varepsilon_t\varepsilon_t') = -$, and y_t is one-summable such that $\sum_{j=0}^{\infty} j|C_j| < \infty$, where $|C_j| = max_{n,m}|C_{j,(n,m)}|$ and $C_{j,(n,m)}$ denotes the (n,m)-element of C_j. As $k \to \infty$, the processes for the aggregated vectors Y_T and $\bar{Y}_T$ are such that for stock variables:*
(i) $\lim_{k\to\infty} E(Y_T Y_T') = \sum_{j=0}^{\infty} C_j - C_j'$
(ii) $\lim_{k\to\infty} E(Y_T Y_{T+j}') = 0 \quad$ for $j \geq 1$
For flow variables we have:
(iii) $\lim_{k\to\infty} E(\bar{Y}_T \bar{Y}_T') = 2\pi f_y(0)$
(iv) $\lim_{k\to\infty} k \cdot E(\bar{Y}_T \bar{Y}_{T+1}') = \sum_{j=1}^{\infty} \left(\sum_{i=0}^{j} C_i\right) - \left(\sum_{i=j+1}^{\infty} C_i\right)'$
(v) $\lim_{k\to\infty} k \cdot E(\bar{Y}_T \bar{Y}_{T+j}') = 0 \quad$ for $j \geq 2$,
where $f_y(\omega)$ denotes the spectral density matrix of y_t at frequency ω.

The result of Proposition 1 is intuitive, since it states that as the sampling interval increases, short-run dynamics disappear. Furthermore, for moderate k, aggregated flow variables are well approximated by a vector MA(1) processes. The reason for this is that from *(iv)* we know that the first order autocorrelation is $O(k^{-1})$, while *(v)* implies that higher order autocorrelations are $o(k^{-1})$.

Proposition 2 (Difference Stationary Variables): *Let Δy_t be generated by*

[1] Our asymptotic framework follows closely that used by Tiao (1972). An alternative asymptotic framework which could in principle be applied in the current context is that used in Christiano and Eichenbaum (1982) and Renault and Szafarz (1991). In particular, it may be assumed that the data are generated by a stationary continuous process such as $y(t) = \int f(\tau)\varepsilon(t-\tau)\mathrm{d}\tau$, where $\varepsilon(t)$ is continuous white noise.

an m dimensional linear process $y_t = \varepsilon_t + C_1\varepsilon_{t-1} + C_2\varepsilon_{t-2} + \cdots$, where it is assumed that $E(\varepsilon_t\varepsilon_t') =$ - , $\sum_{j=1}^{\infty} j|C_j| < \infty$ and the matrix $\bar{C} = \sum_{j=0}^{\infty} C_j$ has full rank. As $k \to \infty$, the processes for the aggregated vectors Y_T and $\bar{Y}_T$ are such that for stock variables:

(i) $\lim_{k\to\infty} \frac{1}{k} E(Y_T - Y_{T-1})(Y_T - Y_{T-1})' = 2\pi f_{\Delta y}(0)$

(ii) $\lim_{k\to\infty} \frac{1}{k} E(Y_T - Y_{T-1})(Y_{T+j} - Y_{T+j-1})' = 0 \quad \text{for } j \geq 1$.

For flow variables we have:

(iii) $\lim_{k\to\infty} \frac{1}{k^2} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_T' - \bar{Y}_{T-1})' = \frac{4\pi}{3} f_{\Delta y}(0)$

(iv) $\lim_{k\to\infty} \frac{1}{k^2} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+1} - \bar{Y}_T)' = \frac{\pi}{3} f_{\Delta y}(0)$

(v) $\lim_{k\to\infty} \frac{1}{k^2} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+j} - \bar{Y}_{T+j-1})' = 0 \quad \text{for } j \geq 2$,

where $f_{\Delta y}(\omega)$ denotes the spectral density matrix of Δy_t at frequency ω.

Given Proposition 2, it follows that as k tends to infinity, the vector of aggregated flow variables has a vector MA(1) representation. Namely,

$$k^{-1}(\bar{Y}_T - \bar{Y}_{T-1}) = U_T + (2 - \sqrt{3})U_{T-1} \ , \tag{3}$$

where

$$E(U_T U_T') = \frac{2\pi}{1 + (2 - \sqrt{3})^2} f_{\Delta y}(0) \ .$$

Note that for the special case where $m = 1$ (a single time series), our results correspond to the result of Working (1960), who shows that the first order autocorrelation of the increments from an aggregated random walk is $(2 - \sqrt{3})/[1 + (2 - \sqrt{3})^2] = 0.25$.

Using the limiting process for large aggregation intervals we are able to analyse the relationship between Granger causality among the original variables and the implied contemporaneous correlation of the aggregated variables. Following Granger (1988) we exclude "true instantaneous causality" and assume that the innovation of the VAR process for the disaggregated vector of time series y_t are mutually uncorrelated, that is, - $= E(\varepsilon_t\varepsilon_t')$ is diagonal. In the words of Granger (1988, p. 206): *"The true causal lag may be very small but never actually zero. The observed or apparent instantaneous causality can then be explained by either temporal aggregation or missing causal variables"*. In what follows we rule out the case that a contemporaneous correlation is due to missing causal variables and assume that we are able to condition on all relevant information.

To define the situation of "spurious instantaneous causality" we write $x \not\rightarrow y$ if x is not a (Granger) cause of y, and $x \not\leftrightarrow y$ if there is no causal relationship between x and y. Furthermore, temporally aggregated variables are indicated by uppercase letters, no matter whether they are flow or stock variables. To the aggregated data we apply the concept of instantaneous causality $X \Rightarrow Y$ (see Section 2 for a definition).

Definition 1: *Let $\xi = [x, y, z']'$ be a $m \times 1$ vector with $m > 3$. We say that there is spurious instantaneous causality between x_t and y_t if we have $y \not\leftrightarrow x$ for the original variables and $X \Rightarrow Y$ (resp. $Y \Rightarrow X$) for the aggregated variables.*

Of course, it is important to know whether an observed instantaneous causality between two variables is due to an underlying causal relationship between the original variables x_t and y_t or whether it is an artifact due to temporal aggregation. The latter situation arises if there is spurious instantaneous causality according to Definition 1. In the following definition we give sufficient conditions to rule out spurious instantaneous causality between two aggregated variables, X and Y.

Proposition 3: *Let $\xi_t = [x_t, y_t, z_t']'$, where z_t is an m^*–dimensional vector. Assume that either:*
(i) ξ_t is a vector of stationary flow variables, or
(ii) ξ_t is a vector of difference stationary flow variables, or
(iii) ξ_t is a vector of difference stationary stock variables.
If (a) the innovations of the VAR representation for ξ_t have a diagonal covariance matrix
(b) $x \not\leftrightarrow y$ and
(c) $\quad x_t \not\rightarrow z_{j,t}$ or (c') $\quad y_t \not\rightarrow z_{j,t}$ for all $j = 1, \ldots, m-2$,
then, as $k \to \infty$:
for case (i): $\rho(\bar{X}_T, \bar{Y}_T | \bar{Z}_T) = 0,$

for case (ii): $\rho(\Delta\bar{X}_T, \Delta\bar{Y}_T|\Delta\bar{Z}_T) = 0$
for case (iii): $\rho(\Delta X_T, \Delta Y_T|\Delta Z_T) = 0$,
where $\rho(a, b|c)$ *denotes the partial correlation between* a *and* b *conditional on* c.

From this proposition it follows that the two sets of conditions $\{(a), (b), (c)\}$ or $\{(a), (b), (c')\}$ are sufficient to rule out spurious causality for the cases (i) – (iii).[2]

An important special case implied by Proposition 3 is considered in

Corollary 1: *For cases (i) – (iii) in Proposition 3, and assuming that there is no feedback Granger causality among the variables, it follows that as* $k \to \infty$ *there is no spurious contemporaneous causality among the aggregated variables.*

In order to illustrate the implications of our results in this section, it is useful to consider some simple examples in which the vector of innovations $\varepsilon_t = [\varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t}]'$ is assumed to be white noise with a diagonal covariance matrix.

Example A: Assume that $\xi_t = [x_t, y_t, z_t]'$ is stationary and has a causal structure given by $x_t \to y_t$ and $y_t \to z_t$. The VAR(1) process is given by

$$\begin{aligned} x_t &= \varepsilon_{1,t} \\ y_t &= a x_{t-1} + \varepsilon_{2,t} \\ z_t &= b y_{t-1} + \varepsilon_{3,t} \ . \end{aligned}$$

so that $x_t \not\to z_t$ (x_t is non Granger causal for z_t) and $z_t \not\to y_t$. From Proposition 3 it follows that $\rho(\bar{X}_T, \bar{Z}_T|\bar{Y}_T) = 0$, and that there is no contemporaneous causality between $\bar{X}_T$ and $\bar{Z}_T$. Note also that as there is no feedback causality among the variables at $k = 1$, the above result also follows directly from Corollary 1.

Example B: Assume that a vector of flow variables is generated by a stationary

[2] Necessary *and* sufficient conditions for ruling out spurious contemporaneous causality in aggregated time series can in principle be derived from the relationship between the VAR with $k = 1$ and the limiting VAR (i.e. $k \to \infty$). However, such conditions are complicated nonlinear functions of the VAR parameters with $k = 1$. If only aggregated data are available with $k > 1$, then the conditions cannot be evaluated in practice.

process given by:

$$\begin{aligned} x_t &= ay_{t-1} + bz_{t-1} + \varepsilon_{1,t} \\ y_t &= \varepsilon_{2,t} \\ z_t &= \varepsilon_{3,t} \ . \end{aligned}$$

Applying Granger's concept of causality, there is no causality between y_t and z_t. Further, a simple calculation shows that for the limiting process, $\rho(\bar{Y}_T, \bar{Z}_T|\bar{X}_T) = -ab/(a^2 + b^2 + 1)$. Thus, a necessary and sufficient condition for the aggregated variables $\bar{Y}_T$ and $\bar{Z}_T$ to have zero partial correlations is that either a, b, or both parameters are equal to zero. This result also follows from Proposition 3, which states that there is no contemporaneous causality if either y_t or z_t is not Granger causal for x_t.

Example C: To illustrate why Proposition 3 does not extend to aggregated stock variables, consider the a stationary process given by:

$$\begin{aligned} x_t &= \varepsilon_{1,t} \\ y_t &= ax_{t-1} + \varepsilon_{2,t} \\ z_t &= by_{t-1} + \varepsilon_{3,t}. \end{aligned}$$

In this system, $x_t \to y_t$ and $y_t \to z_t$. For $k \geq 3$ the aggregated process becomes white noise with:

$$\begin{aligned} X_T &= U_{1,T} \\ Y_T &= U_{2,T} \\ Z_T &= abX_T + U_{3,T} \end{aligned}$$

For $ab \neq 0$ there exists *spurious* contemporaneous causality between X_T and Z_T, as there is no Granger causality between x_t and z_t. Stated another way, the indirect causal relationship between x_t and z_t via y_t becomes a direct causal link under aggregation.

4 Monte Carlo Experiments

In this section, the asymptotic implications of Proposition 3 and Corollary 1 are examined via a simple Monte Carlo experiment. In particular, we begin with the following VAR(1) model:

$$\begin{bmatrix} \Delta^d x_t \\ \Delta^d y_t \\ \Delta^d z_t \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \\ b & a & 0 \\ 0 & b & a \end{bmatrix} \begin{bmatrix} \Delta^d x_{t-1} \\ \Delta^d y_{t-1} \\ \Delta^d z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix}, \tag{4}$$

where $d \in \{0, 1\}$ and $\varepsilon_{i,t}$ is an i.i.d. vector of standard normal random variables. For $b \neq 0$, the Granger causal structure of this system is: $x_t \to y_t$ and $y_t \to z_t$. From Proposition 3 and Corollary 1 it follows that as $k \to \infty$, the limiting process has a partial correlation structure such that $E(\widehat{u}_{1,T}, \widehat{u}_{3,T} | \widehat{u}_{2,T}) = 0$ and all other partial correlations are nonzero, where the $\widehat{u}_{j,T}$ $(j = 1, 2, 3)$ are the residuals from an estimated VAR(4) model using data generated according to (4) and aggregated appropriately.[3] Evaluation of this restriction allows us to assess the empirical size of the empirical tests proposed by Swanson and Granger (1997).

Empirical level figures for 5% nominal size tests and for various parameterizations of the VAR are reported in Table 1. We also report results for stationary stock variables, which are not treated in Proposition 3. In all experiments, b is set equal to 0.5. The values of the parameter a are taken from $\{0, 0.2, 0.4, 0.6, 0.8\}$. Not surprisingly, the magnitude of the parameter a is crucial when k is small, as a determines the roots of the autoregressive polynomial in our model. Thus, our asymptotic results may be a poor guide to finite sample behavior for small k and $|a|$ close to unity.[4] The entries in the tables are based on 10,000 Monte Carlo replications for $k \leq 100$ and 2,000 replications for $k = 200$ and $k = 500$. All tests are based on 100 observations of appropriately aggregated data.

Tables 1a-c contain results for cases (i) – (iii) in Proposition 3. For small and moderately sized values of a the empirical sizes converge quickly to their

[3]Swanson and Granger (1997) propose tests for assessing whether the above partial correlation restriction holds which are based on Fisher's z-statistics or alternatively on t-statistics from regressions involving the residuals. Here we use the t-statistic approach.

[4]Recall also that the aggregated processes which we construct are VARMA processes, in general. Thus, lower order VAR approximations may not yield good estimates of the errors of the process.

limiting value of 0.05 when k increases. For $a = 0.8$, however, the empirical sizes tend very slowly to the limiting value of 0.05. This is due to the fact that for a persistent processes the dynamics remain important for short aggregation intervals. However, if the aggregation interval is as large as $k = 500$, then the short-run dynamics becomes negligible and the asymptotic results are applicable.

5 Concluding Remarks

In this paper we examine the asymptotic effects of temporal aggregation on causal inference by examining the concept of Granger causality in the context of aggregated systems, using the framework of Swanson and Granger (1997). We argue that as Granger causal findings are aggregation dependent, understanding the relationship between aggregation and causality is important. In particular we consider the relationship between Granger causality among disaggregated variables and instantaneous causality found among temporally aggregated data. Conditions are derived that are sufficient to rule out the case where instantaneous causality of the aggregated data is a pure artifact of temporal aggregation. Our results are illustrated via three simple examples and via a series of Monte Carlo experiments which indicate that our asymptotic results are reliable in finite samples as long as the time series are not characterized by high level of persistence.

6 Appendix: Proofs

Proposition 1:

(i) From $Y_T = y_{kT}$ and assuming stationarity we have $E(Y_T Y_T') = E(y_t y_t') = \sum_{i=0}^{\infty} C_i \cdot C_i'$
(ii) Since the process is assumed to be ergodic we have $\lim_{k\to\infty} E(Y_T Y_{T+j}') = \lim_{k\to\infty} E(y_{kT} y_{kT+jk}') = 0$ for all $j \geq 1$.
(iii) The vector of aggregated flow variables is given by $\bar{Y}_T = k^{-1/2} \sum_{j=0}^{k-1} y_{kT-j}$ and therefore $\bar{Y}_T$ behaves as a (normalized) vector partial sum. For partial sums it is known that $\lim_{k\to\infty} E(\bar{Y}_T \bar{Y}_T') = \Sigma + \Gamma + \Gamma'$, where $\Sigma = E(y_t y_t')$ and $\Gamma = \sum\limits_{j=1}^{\infty} E(y_1 y_{1+j}')$. In the frequency domain this result can be represented as

$$\lim_{k\to\infty} E(\bar{Y}_T \bar{Y}_T') = 2\pi f_y(0) = \left(\sum_{j=0}^{\infty} C_j\right) \cdot \left(\sum_{j=0}^{\infty} C_j'\right).$$

(iv) Let

$$k^{1/2}\bar{Y}_T = (I_m + L + L^2 + \cdots + L^{k-1})C(L)\varepsilon_t \equiv D(L)\varepsilon_t,$$

where $D(L) = I_m + D_1 L + D_2 L^2 + \cdots$ and $D_j = \sum_{i=0}^{\min(j,k-1)} C_{j-i}$.

It is convenient to decompose $\bar{Y}_T$ as

$$\begin{aligned} k^{1/2}\bar{Y}_T &= D_0\varepsilon_t + D_k\varepsilon_{t-k} + D_{2k}\varepsilon_{t-2k} + \cdots \\ &\quad + D_1\varepsilon_{t-1} + D_{k+1}\varepsilon_{t-k-1} + D_{2k+1}\varepsilon_{t-2k-1} + \cdots \\ &\quad \vdots \\ &\quad + D_{k-1}\varepsilon_{t-k+1} + D_{2k-1}\varepsilon_{t-2k+1} + D_{3k-1}\varepsilon_{t-3k+1} + \cdots \\ &\equiv u_{0t} + \cdots + u_{k-1,t}\ , \end{aligned}$$

where $u_{jt} = D_j\varepsilon_{t-j} + D_{j+k}\varepsilon_{t-j-k} + \cdots$.

From

$$\begin{aligned} k^{1/2}\bar{Y}_T &= u_{0t} + \cdots + u_{k-1,t} \\ k^{1/2}\bar{Y}_{T+1} &= u_{0,t+k} + \cdots + u_{k-1,t+k} \\ k^{1/2}\bar{Y}_{T+2} &= u_{0,t+2k} + \cdots + u_{k-1,t+2k} \end{aligned}$$

we obtain $k \cdot E(\bar{Y}_T \bar{Y}'_{T+1}) = \sum_{j=0}^{k-1} E(u_{jt} u'_{j,t+k})$. Next consider $E(u_{0t} u'_{0,t+k}) = D_0 - D'_k + D_k - D'_{2k} + \cdots$. For a summable sequence C_i we have

$$\lim_{k\to\infty} |D_{2k}| = \lim_{k\to\infty} |C_{k+1} + C_{k+2} + \cdots + C_{2k}| = 0$$

and, similarly, $\lim_{k\to\infty} |D_{jk}| = 0$ for $j \geq 2$. It follows that

$$\begin{aligned} \lim_{k\to\infty} E(u_{0t} u'_{0,t+k}) &= D_0 - D'_k \\ &= -(C_1 + C_2 + \cdots + C_k)'. \end{aligned}$$

Similarly we get:

$$\begin{aligned} \lim_{k\to\infty} E(u_{1t} u'_{1,t+k}) &= D_1 - D'_{k+1} \\ &= (I_m + C_1) - (C_2 + C_3 + \cdots + C_{k+1})' \end{aligned}$$

and

$$\lim_{k\to\infty} E(u_{k-1,t} u'_{k-1,t-k}) = (C_1 + \cdots + C_{k-1}) - (C_k + C_{k+1} + \cdots + C_{2k-1})'.$$

Adding these expressions gives the desired result.

It remains to show that $\sum_{j=0}^{\infty} \left(\sum_{i=0}^{j} C_i \right) - \left(\sum_{i=j+1}^{\infty} C_i \right)'$ is bounded. Let $\bar{c} = \sup_t \left\| \sum_{j=0}^{t} C_j \right\|$. Then:

$$\left\| \sum_{j=0}^{\infty} \left(\sum_{i=0}^{j} C_i \right) - \left(\sum_{i=j+1}^{\infty} C'_i \right) \right\| \leq \sum_{j=0}^{\infty} \left\| \sum_{i=0}^{j} C_i \right\| \| - \| \sum_{i=j+1}^{\infty} j \, \|C_i\|$$

which is finite by assumption.

(v) Consider $E(u_{0t} u'_{0,t-pk}) = D_0 - D'_{pk} + D_k - D_{(p+1)k} + \cdots$. Since

$$\lim_{k\to\infty} D_{(p+j)k} = \lim_{k\to\infty} \left[C_{(p+j-1)k+1} + \cdots + C_{(p+j)k} \right] 0 \qquad \text{for } p \geq 2 \text{ and } j = 0, 1, \ldots$$

it follows that the autocovariances disappear for $p \geq 2$.

Proposition 2

(i) The difference $Y_T - Y_{T-1} = y_{kT} - y_{kT-k} = \sum_{i=1}^{k} \Delta y_{(k-1)T+i}$ is a partial sum process with asymptotic covariance matrix

$$\lim_{k\to\infty} k^{-1} E(Y_T - Y_{T-1})(Y_T - Y_{T-1})' = \Sigma + \Gamma + \Gamma' = 2\pi f_{\Delta y}(0)$$

(ii) Let $S_1 = \sum\limits_{i=1}^{k} u_i$ and $S_2 = \sum\limits_{i=k+1}^{2k} u_i$, where u_t is stationary with covariance function Γ_j. The covariance between S_1 and S_2 is given by $E(S_1 S_2') = \Gamma_1 + 2\Gamma_2 + \cdots + k\Gamma_k + (k-1)\Gamma_{k+1} + \cdots + \Gamma_{2k-1}$ For $\sum_{j=1}^{\infty} j|\Gamma_j| < \infty$ we have

$$|E(S_1 S_2')| < |\sum_{j=1}^{\infty} j\Gamma_j| \leq \sum_{j=1}^{\infty} j|\Gamma_j| \ < \ \infty$$

and, thus, by letting $S_1 = Y_T - Y_{T-1}$ and $S_2 = Y_{T+1} - Y_T$ it follows that $E(Y_T - Y_{T-1})(Y_{T+1} - Y_T)$ is $O(1)$. A similar result is obtained for higher order autocovariances.

(iii) Let

$$\begin{aligned} k(\bar{Y}_T - \bar{Y}_{T-1}) &= y_{kT} - y_{kT-k} + y_{kT-1} - y_{kT-k-1} + \cdots + y_{kT-k+1} - y_{kT-2k+1} \\ &= S_k(L)\Delta y_{kT} + S_k(L)\Delta y_{kT-1} + \cdots + S_k(L)\Delta y_{kT-k+1} \\ &= S_k(L)^2 \Delta y_{kT}, \end{aligned}$$

where $S_k(L) = 1 + L + L^2 + \cdots + L^{k-1}$ and

$$\begin{aligned} S_k(L)^2 &= 1 + 2L + 3L^2 + \cdots + kL^{k-1} + (k-1)L^k + \cdots + L^{2k-2} \\ &= w_0 + w_1 L + w_2 L + \cdots + w_{2k-2} L^{2k-2} \end{aligned}$$

is a symmetric filter with triangular weights.

The covariance matrix is given by

$$\begin{aligned} k \cdot E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_T - \bar{Y}_{T-1})' &= E\left(\sum_{i=0}^{2k-2} w_i \Delta y_{kT-i}\right)\left(\sum_{i=0}^{2k-2} w_i \Delta y'_{kT-i}\right) \\ &= \sum_{p=-2k+2}^{2k-2} \sum_{i=1}^{2k-1-|p|} w_i w_{i+|p|} \Gamma_p \end{aligned}$$

where $\Gamma_p = E(\Delta y_t \Delta y'_{t-p})$.

Consider the odd values $p = \pm 1, \pm 3, \pm 5, \ldots$. We have

$$\sum_{i=1}^{2k-1-|p|} w_i w_{i+|p|} = 2 \sum_{i=1}^{k-(|p|+1)/2} i(i+p)$$

and as $k \to \infty$

$$\begin{aligned} \lim_{k\to\infty} \sum_{i=1}^{k-|(p+1)/2|} 2(i^2 - ip) &= 2(\sum_{i=1}^{\infty} i^2) - 2p(\sum_{i=1}^{\infty} i) \\ &= \frac{2}{3}k^3 + O(k^2). \end{aligned}$$

For even values $p = 0, \pm 2, \pm 4, \dots$ we have

$$\sum_{i=1}^{2k-1-|p|} w_i w_{i+|p|} = (k - |p|/2)^2 + 2 \sum_{i=1}^{k-|p|/2-1} i(i+p)$$

and, thus,

$$\lim_{k\to\infty} \sum_{i=1}^{k-|(p+1)/2|} 2(i^2 - ip) = \frac{2}{3}k^3 + O(k^2).$$

Using these results yields

$$\begin{aligned} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_T - \bar{Y}_{T-1})' &= \frac{2}{3}k^2(\Gamma_0 + \sum_{j=1}^{\infty} \Gamma_j + \Gamma_j') + o(k^2) \\ &= \frac{4\pi}{3}k^2 f_{\Delta y}(0) + o(k^2). \end{aligned}$$

(iv) The first order autocovariance matrix is given by

$$k \cdot E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+1} - \bar{Y}_T)' = \sum_{p=-2k+2}^{2k-2} \sum_{i=1}^{2k-1-|p|} w_{i+k} w_{i+k+|p|} \Gamma_p$$

where $\Gamma_p = E(\Delta y_t \Delta y_{t-p}')$.

For an odd value of p we have

$$\begin{aligned} \lim_{k\to\infty} \sum_{i=1}^{2k-1-|p|} w_{i+k} w_{i+k+|p|} &= \sum_{i=1}^{\infty} (k-i)(i+p) + O(k^2) \\ &= k(\sum_{i=1}^{\infty} i) - (\sum_{i=1}^{\infty} i^2) + k^2 p - p(\sum_{i=1}^{\infty} i) + O(k^2) \\ &= \frac{1}{6}k^3 + O(k^2) \end{aligned}$$

It follows that

$$\begin{aligned} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+1} - \bar{Y}_T)' &= \frac{1}{6}k^2(\Gamma_0 + \sum_{j=1}^{\infty} \Gamma_j + \Gamma_j') + o(k^2) \\ &= \frac{\pi}{3}k^2 f_{\Delta y}(0) + o(k^2). \end{aligned}$$

(v) To simplify the proof we assume that Δy_t has a vector MA(q) representation with $q < k$. Since $k \to \infty$ the proof is valid for $q \to \infty$ as long as k grows with a faster rate that q.

The second order autocovariance matrix is given by

$$
\begin{aligned}
k \cdot E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+2} - \bar{Y}_{T+1})' &= E\left(\sum_{i=0}^{2k-2} w_i \Delta y_{kT-i}\right)\left(\sum_{i=0}^{2k-2} w_i \Delta y'_{kT+2k-i}\right) \\
&= \sum_{p=1}^{k}\sum_{i=1}^{|p|} w_i w_{2k-i-|p|+1}(\Gamma_p + \Gamma'_p)
\end{aligned}
$$

There exist a constant $c < \infty$ such that for all p

$$
\sum_{i=1}^{|p|} w_i w_{2k-i-|p|+1} = \sum_{i=1}^{p} i(p-i+1) < cp^3.
$$

Thus, we have

$$
\begin{aligned}
\sum_{p=1}^{k}\sum_{i=1}^{|p|} w_i w_{2k-i-|p|+1}|\Gamma_p + \Gamma'_p| &< \sum_{p=1}^{k} 2cp^3|\Gamma_p| \\
&< 2ck^2 \sum_{p=1}^{k} p|\Gamma_p|,
\end{aligned}
$$

for $p < k$. From $\sum_{p=1}^{k} p|\Gamma_p| < \infty$ it finally follows that

$$
\lim_{k\to\infty} \frac{1}{k^2} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+2} - Y_{T+1})' = 0.
$$

Similarly it can be shown that the higher order autocorrelations converge to zero as well.

Proposition 3

For convenience, we confine ourselves to a trivariate VAR(p) process. The proof can easily be generalized to systems with $m > 3$.

First consider a VAR process obeying the conditions $x_t \not\rightarrow y_t$, $y_t \not\rightarrow x_t$ and

$$
(c) \quad x_t \not\rightarrow z_t,
$$

that is, there is no causality between $y_{1,t}$ and $y_{2,t}$ and condition (a) is satisfied. From the Propositions 1 and 2 we know that the limiting processes for the cases (i) – (iii) is white noise with a covariance matrix proportional to the spectral density matrix of the original process. Thus, the limiting process for case (i), for example, has a representation of the form:

$$\left(\sum_{j=0}^{\infty} C_j\right)^{-1} \begin{bmatrix} \bar{X}_T \\ \bar{Y}_T \\ \bar{Z}_T \end{bmatrix} = \begin{bmatrix} U_{1,T} \\ U_{2,T} \\ U_{3,T} \end{bmatrix}$$

where $E(U_T U_T') = -$. A similar representation exists for the cases (ii) and (iii). We therefore confine ourselves to case (i). The proof for case (ii) and (iii) is straightforward.

Since we assume that the MA representation is invertible there exists an autoregressive representation with autoregressive polynomial

$$I_m - A_1 L - A_2 L^2 - \cdots = \left(\sum_{j=0}^{\infty} C_j L^j\right)^{-1}$$

and thus the limiting process can be written as

$$\begin{bmatrix} \bar{X}_T \\ \bar{Y}_T \\ \bar{Z}_T \end{bmatrix} = \bar{A} \begin{bmatrix} \bar{X}_T \\ \bar{Y}_T \\ \bar{Z}_T \end{bmatrix} + \begin{bmatrix} U_{1,T} \\ U_{2,T} \\ U_{3,T} \end{bmatrix}$$

where

$$\bar{A} = \sum_{j=1}^{\infty} A_j = \begin{bmatrix} \bar{a}_{11} & 0 & \bar{a}_{13} \\ 0 & \bar{a}_{22} & \bar{a}_{23} \\ 0 & \bar{a}_{32} & \bar{a}_{33} \end{bmatrix}.$$

The zero restrictions in the matrix $\bar{A}$ result from the assumptions on the causal relationship between the variables. Accordingly, we find

$$(1 - \bar{a}_{11})\bar{X}_T = \bar{a}_{13}\bar{Z}_T + U_{1,T}$$

$$(1 - \bar{a}_{11})\bar{Y}_T\bar{X}_T = \bar{a}_{13}\bar{Y}_T\bar{Z}_T + \bar{Y}_T U_{1,T} \ .$$

First, note that $\rho(\bar{Y}_T, \bar{Z}_T | \bar{Z}_T) = 0$. Furthermore, the system

$$(I - \bar{A}) \begin{bmatrix} \bar{X}_T \\ \bar{Y}_T \\ \bar{Z}_T \end{bmatrix} = \begin{bmatrix} U_{1,T} \\ U_{2,T} \\ U_{3,T} \end{bmatrix}$$

is block recursive so that $E(U_{1,T}|Z_T) = 0$ and $\rho(\bar{Y}_T, U_{1,T}|\bar{Z}_T) = 0$. It follows that $\rho(\bar{X}_T, \bar{Y}_T|\bar{Z}_T) = 0$. Note that the condition (c) is crucial for such a block recursive system.

Second, consider the condition

$$(c') \quad y_t \not\rightarrow z_t$$

instead of condition (c). In this case the limiting process can be represented as

$$\begin{bmatrix} \bar{X}_T \\ \bar{Y}_T \\ \bar{Z}_T \end{bmatrix} = \begin{bmatrix} \bar{a}_{11} & 0 & \bar{a}_{13} \\ 0 & \bar{a}_{22} & \bar{a}_{23} \\ \bar{a}_{31} & 0 & \bar{a}_{33} \end{bmatrix} \begin{bmatrix} \bar{X}_T \\ \bar{Y}_T \\ \bar{Z}_T \end{bmatrix} + \begin{bmatrix} U_{1,T} \\ U_{2,T} \\ U_{3,T} \end{bmatrix}$$

This gives:

$$\begin{aligned} (1-\bar{a}_{22})\bar{Y}_T &= \bar{a}_{23}\bar{Z}_T + U_{2,T} \\ (1-\bar{a}_{22})\bar{X}_T Y_T &= \bar{a}_{23}\bar{X}_T\bar{Z}_T + \bar{X}_T U_{2,T} \ . \end{aligned}$$

Obviously, $\rho(\bar{X}_T, \bar{Z}_T|\bar{Z}_T) = 0$. To show that $\rho(\bar{X}_T, U_{2,T}|\bar{Z}_T) = 0$ it is useful to re-arrange the system according to

$$\begin{bmatrix} 1-\bar{a}_{22} & \bar{a}_{23} & 0 \\ 0 & 1-\bar{a}_{33} & \bar{a}_{31} \\ 0 & \bar{a}_{13} & 1-\bar{a}_{11} \end{bmatrix} \begin{bmatrix} \bar{Y}_T \\ \bar{Z}_T \\ \bar{X}_T \end{bmatrix} + \begin{bmatrix} U_{2,T} \\ U_{3,T} \\ U_{1,T} \end{bmatrix} \ .$$

Since the re-arranged system is block recursive it follows that $E(U_{2,T}|\bar{Z}_T) = 0$ and, hence, $\rho(\bar{Y}_T, U_{2,T}|\bar{Z}_T) = 0$.

7 References

Amemiya, T. and R.Y. Wu (1972), The Effect of Aggregation on Prediction in the Autoregressive Model, *Journal of the American Statistical Association*, No. 339, 628-632.

Brewer, K.R.W. (1973), Some Consequences of Temporal Aggregation and Systematic Sampling for ARMA and ARMAX Models, *Journal of Econometrics*, 1, 133-154.

Christiano, L.J. and M. Eichenbaum (1987), Temporal Aggregation and Structural Inference in Macroeconomics, *Carnegie-Rochester Conference Series on Public Policy*, 26, 63–130.

Dawid, A.P. (1979), Conditional Independence in Statistical Theory, Journal of the Royal Statistical Society B, 41, 1–31.

Engle, R.F. (1969), Biases from Time-Aggregation of Distributed Lag Models, Ph.D. Dissertation, Cornell University.

Geweke, J. (1986), Measurement of Linear Dependence and Feedback Between Multiple Time Series, *Journal of American Statistical Association,* 77, 304–313.

Granger, C.W.J. (1969), Investigating Causal Relations by Econometric Models and Cross Spectral Methods, *Econometrica*, 37, 428-438.

Granger, C.W.J. (1987), Implications of Aggregation with Common Factors, *Econometric Theory*, 32, 208-222.

Granger, C.W.J. (1988), Some Recent Developments in a Concept of Causality, *Journal of Econometrics*, 39, 199-211.

Lütkepohl, H. (1982), Non-Causality due to Omitted Variables, *Journal of Econometrics*, 19, 367–378.

Lütkepohl, H. (1987), *Forecasting Aggregated Vector ARMA Processes*, Berlin, Springer.

Lütkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, Berlin, Springer.

Lütkepohl, H. (2000), Vector Autoregressions, in: B. Baltagi (ed), Companion to Econometric Theory, Oxfort: Basil Blackwell.

Lütkepohl, H. and M.M. Burda (1997), Modified Wald Tests Under Nonregular Conditions, *Journal of Econometrics*, 78, 315–332.

Marcellino, M. (1999), Some Consequences of Temporal Aggregation in Empirical Analysis, Journal of Economic and Business Statistics, 17, 129–136.

Pierce, D.A. and L.D. Haugh (1977), Causality in Temporal Systems: Characterizations and a survey, *Journal of Econometrics*, 5, 265–293.

Renault, E. and A. Szafarz (1991), True Versus Spurious Instantaneous Causality, Working paper, GREMAQ and IDEI, Toulouse, France.

Pearl, J. (2000), Causality: Models, Reasoning, and Inference, Cambridge University Press.

Quenouille, M.H. (1957), The Analysis of Multiple Time Series, London: Griffin.

Stram, D.O. and W.W.S. Wei (1986), Temporal Aggregation in the ARIMA Process, *Journal of Time Series Analysis*, 7, 279-292.

Swanson, N.R. (1994), Essays in Forecasting Stationary and Nonstationary Stochastic Processes, Ph.D. Dissertation, University of California, San Diego.

Swanson, N.R. and C.W.J. Granger (1997), Impulse Response Functions Based on a Causal Approach to Residual Orthogonaliztion in Vector Autoregressions, *Journal of the American Statistical Association*, 92, 357-367.

Tiao, G.C. (1972), Asymptotic Behaviour of Temporal Aggregates of Time Series, Biometrika, 59, 525–531.

Tiao, G.C. (1999), The ET Interview: Professor George C. Tiao, Econometric Theory, 15, 389–424.

Wei, W.W.S. (1978), The Effect of Temporal Aggregation on Parameter Estimation in Distributed Lag Models, *Journal of Econometrics*, 8, 237-246.

Wei, W.W.S. (1981), Effect of Systematic Sampling on ARIMA Models, *Communications in Statistical-Theoretical Mathematics A*, 10, 2389-2398.

Weiss, A.A. (1984), Systematic Sampling and Temporal Aggregation in Time Series Models, *Journal of Econometrics*, 26, 271-281.

Working, H. (1960), Note on the Correlation of First Differences of Averages in a Random Chain, *Econometrica*, 26, 916–918.

Zellner, A. and C. Montmarquette (1971), A Study of Some Aspects of Temporal Aggregation Problems in Econometric Analyses, *Review of Economics and Statistics*, 335-342.

Table 1: Empirical sizes of the Swanson-Granger test procedure

a) stationary flow variables (case (i))					
k	a=0	a=0.2	a=0.4	a=0.6	a=0.8
2	0.10	0.10	0.11	0.11	0.11
5	0.07	0.08	0.16	0.33	0.61
10	0.07	0.07	0.10	0.35	0.95
20	0.07	0.07	0.08	0.19	0.98
50	0.07	0.07	0.07	0.09	0.80
100	0.07	0.07	0.07	0.07	0.47
200	0.07	0.07	0.07	0.07	0.22
500	0.07	0.07	0.07	0.07	0.11
b) Difference stationary flow variables (case (ii))					
k	a=0	a=0.2	a=0.4	a=0.6	a=0.8
2	0.11	0.11	0.11	0.11	0.11
5	0.08	0.10	0.17	0.34	0.61
10	0.07	0.07	0.10	0.31	0.93
20	0.07	0.07	0.07	0.14	0.95
50	0.07	0.07	0.07	0.07	0.53
100	0.07	0.07	0.07	0.07	0.17
200	0.07	0.07	0.07	0.07	0.08
500	0.07	0.07	0.07	0.06	0.06
c) Difference stationary stock variables (case (iii))					
k	a=0	a=0.2	a=0.4	a=0.6	a=0.8
2	0.10	0.11	0.11	0.11	0.11
5	0.07	0.09	0.16	0.33	0.61
10	0.07	0.07	0.10	0.35	0.95
20	0.07	0.07	0.08	0.19	0.98
50	0.07	0.07	0.07	0.09	0.80
100	0.07	0.07	0.07	0.07	0.47
200	0.07	0.07	0.07	0.07	0.23
500	0.07	0.07	0.07	0.07	0.11

Notes: Entries correspond to the empirical sizes of the Swanson-Granger test for a zero partial correlation between the residuals of the first and third equation conditional on the residuals of the second equation. The results for $k = 2$ to $k = 100$ are based on 10,000 Monte Carlo replications and for $k = 200$ and $k = 500$ 2,000 replications are used. Data are generated according to (4). The nominal size of the tests is 0.05.