Nowcasting causality in mixed frequency vector autoregressive models

Maastricht University

Thomas Götz, Alain Hecq

Nowcasting Causality in Mixed Frequency Vector Autoregressive Models

RM/13/050

GSBE

Maastricht University School of Business and Economics
Graduate School of Business and Economics

P.O Box 616
NL- 6200 MD Maastricht
The Netherlands

Nowcasting Causality in Mixed Frequency Vector Autoregressive Models

Thomas B. Götz*
Alain Hecq

Maastricht University, SBE, Department of Quantitative Economics

September 3, 2013

Abstract

This paper introduces the notion of *nowcasting causality* for mixed-frequency VARs as the mixed-frequency version of instantaneous causality. We analyze the relationship between nowcasting and Granger causality in the mixed-frequency VAR setting of Ghysels (2012) and illustrate that nowcasting causality can have a crucial impact on the significance of contemporaneous or lagged high-frequency variables in standard MIDAS regression models.

JEL Codes: C12, C22, C32
JEL Keywords: Instantaneous Causality, Granger Causality, Mixed Frequency VAR, Mixed Data Sampling, MIDAS

*Corresponding author: Thomas Götz, Department of Quantitative Economics, School of Business and Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Email: t.gotz@maastrichtuniversity.nl, Tel.: +31 43 388 3578, Fax: +31 43 388 2000.

1 Mixed-frequency VARs

Economic time series are available at different frequencies. For two variables,[1] let us denote by y_t, $t = 1, \ldots, T$ the low-frequency variable and by $x^{(m)}_{t-i/m}$ the high-frequency variables with m high-frequency observations per low-frequency period t. The value of i indicates how many high-frequency observations, starting from the end of the low-frequency period, i.e., $i = 0$, we take into account. For the quarter/month-scenario $m = 3$ and for $t = 2012Q4$, $i = 1$ refers to November 2012; $x^{(3)}_{2012Q4-2/3}$ to October 2012 etc. While it has been the standard approach to address this issue by simply aggregating the high-frequency series, it potentially leads to a loss of information. Mi(xed) Da(ta) S(ampling) regressions (Ghysels et al., 2004) have been developed as means to preserve the information embedded in the higher frequencies without sacrificing parsimony of the model. Until recently, mixed-frequency problems were limited to a simple regression framework, in which one of the low-frequency variables is chosen as the dependent variable. Since the work of Ghysels (2012) for stationary series and the extension of Götz et al. (2013) or Miller and Ghysels (2013) for the non-stationary and possibly cointegrated case, we can analyze the links between high- and low-frequency series in a VAR system treating all variables as endogenous. Letting $X_t = (x^{(m)}_t, x^{(m)}_{t-1/m}, \ldots, x^{(m)}_{t-(m-1)/m})'$, a dynamic structural equations model for $Z_t = (y_t, X_t')'$ is given by $A_c Z_t = A_1 Z_{t-1} + \ldots + A_p Z_{t-p} + \varepsilon_t$, with a diagonal covariance matrix Σ_ε and where A_c pertains to contemporaneous relationships between the series.[2] For a quarter/month-example and with $p = 1$ the model reads as

$$\begin{pmatrix} 1 & \beta_1 & \beta_2 & \beta_3 \\ \delta & 1 & -\rho_1 & -\rho_2 \\ 0 & 0 & 1 & -\rho_1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_t \\ x^{(3)}_t \\ x^{(3)}_{t-1/3} \\ x^{(3)}_{t-2/3} \end{pmatrix} = \begin{pmatrix} \rho_y & \phi_1 & \phi_2 & \phi_3 \\ \pi_1 & \rho_3 & 0 & 0 \\ \pi_2 & \rho_2 & \rho_3 & 0 \\ \pi_3 & \rho_1 & \rho_2 & \rho_3 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x^{(3)}_{t-1} \\ x^{(3)}_{t-1-1/3} \\ x^{(3)}_{t-1-2/3} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{pmatrix}, \quad (1)$$

where an AR(3) structure for the high-frequency process is assumed in this example;[3] the relationship between matrices A_c and A_1 (Ghysels, 2012) is seen in (1). The matrix A_c links contemporaneous values of y and x: $\beta_j \neq 0$ implies that y_t is affected by incoming observations of X_t, whereas $\delta \neq 0$ implies that the last monthly observation, i.e., March, June and so on, of x is influenced by y_t. The latter becomes interesting for studying policy analysis, where the high-frequency policy variable may react to current low-frequency conditions. Note that the policy variable could as well be $x^{(3)}_{t-1/3}$ or $x^{(3)}_{t-2/3}$, although y_t may not yet be observed, such that the entire first column of A_c is non-zero (see Ghysels, 2012 for details).

[1] We can straightforwardly extend the methodology to a multivariate setup as long as we only add low-frequency variables to the system. As soon as we add further high-frequency variables (with possibly different m), another framework is required due to the implicit multi-step nature of testing concomitant with the mixed-frequency characteristic of the time series involved. This is, however, left for further research.

[2] The analysis can straightforwardly be adapted to the presence of deterministic terms.

[3] Note that an AR(m) structure for the high-frequency process is often considered for finance applications, where, e.g., a daily volatility variable may depend on its value on the previous day, during the previous week and the entire previous month (Corsi, 2009).

2 Nowcasting causality

Starting from (1) we get the reduced form VAR model $Z_t = A_1^* Z_{t-1} + u_t$

$$Z_t = \begin{pmatrix} \rho_y^* & \phi_1^* & \phi_2^* & \phi_3^* \\ \pi_1^* & a_{2,2}^* & a_{2,3}^* & a_{2,4}^* \\ \pi_2^* & a_{3,2}^* & a_{3,3}^* & a_{3,4}^* \\ \pi_3^* & a_{4,2}^* & a_{4,3}^* & a_{4,4}^* \end{pmatrix} Z_{t-1} + u_t, \tag{2}$$

where $A_1^* = A_c^{-1} A_1$ and $u_t = A_c^{-1}\varepsilon_t$ and $\Sigma_u = A_c^{-1}\Sigma_\varepsilon A_c^{-1\prime}$. Without giving a formal definition of Granger (non-)causality (GC hereafter), it is clear that testing for X not Granger causing y corresponds to $\phi_1^* = \phi_2^* = \phi_3^* = 0$ jointly (Ghysels et al., 2013). Importantly, it is defined in terms of the low frequency, i.e., in terms of index t. Given the mixed-frequency nature of the variables under consideration, it may be of interest to analyze whether knowing the values of X_t helps to predict y_t. In the common-frequency setup this is referred to as instantaneous causality (Lütkepohl, 2005, p.42), because one tests for a causality pattern between y_t and X_t, where t refers to the common frequency under consideration. Since we intend to predict y_t using values of the high-frequency variables *within* period t (Giannone et al., 2008), we refer to instantaneous causality in the mixed-frequency case as *nowcasting causality*.

Formally, let Ω_t represents the set of information available at moment t such that $x_t^{(m)} \in \Omega_t$, but, e.g., $x_{t+1/m}^{(m)} \notin \Omega_t$, and let Ω_t^W be the corresponding information set containing the information for all stochastic processes except W. Now, we denote by $P[y_{t+1}|\Omega_t^W]$ and $P[X_{t+1}|\Omega_t^W]$ the best linear forecasts of y_{t+1} and X_{t+1}, respectively, based on Ω_t^W. We can then define nowcasting causality (NC hereafter) as follows:

Definition 1 *y does not nowcasting cause X if*

$$P[X_{t+1}|\Omega_t \cup \Omega_{t+1}^X] = P[X_{t+1}|\Omega_t].$$

Similarly, X does not nowcasting cause y if

$$P[y_{t+1}|\Omega_t \cup \Omega_{t+1}^y] = P[y_{t+1}|\Omega_t].$$

In other words, knowing y_{t+1} does not help in predicting X_{t+1} and vice versa.

Considering (1), testing for NC in both directions corresponds to $\beta_1 = \beta_2 = \beta_3 = 0$ and $\delta = 0$. Because of the interconnection between the matrices A_c and A_1 the question arises in what sense the presence or absence of NC has an impact on GC. Let us use our quarter/month-example and consider various specifications of the matrices A_c and A_1 in (1). *(i)* If $\phi_j \neq 0\ \forall j$ we will end up with GC in the VAR in (2), no matter whether NC is present or absent. *(ii)* In the absence of NC, i.e., $\beta_j = \delta = 0\ \forall j$, and if $\phi_j = 0\ \forall j$, then $\phi_j^* = 0\ \forall j$ in (2). *(iii)* Assuming, however, that $\beta_j \neq 0$ and $\phi_j = 0\ \forall j$, it turns out that $\phi_j^* \neq 0\ \forall j$ in (2). In other words, the presence of NC from X to y implies GC in the same direction. *(iv)* If $\delta \neq 0$ and $\beta_j = 0 \ \forall j$, and if $\phi_j = 0 \ \forall j$, then there is no GC from X to y in the reduced-form VAR parameters:

$$A_1^* = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \pi_1 - \delta + \pi_2\rho_1 + \pi_3\widetilde{\rho} & \rho_3 + \rho_1\rho_2 + \rho_1\widetilde{\rho} & \rho_1\rho_3 + \rho_2\widetilde{\rho} & \rho_3\widetilde{\rho} \\ \phi_2 + \pi_3\rho_1 & \widetilde{\rho} & \rho_3 + \rho_1\rho_2 & \rho_1\rho_3 \\ \pi_3 & \rho_1 & \rho_2 & \rho_3 \end{pmatrix}, \tag{3}$$

where $\widetilde{\rho} = \rho_1^2 + \rho_2$.

The matrix A_c also determines $\Sigma_u = A_c^{-1}\Sigma_\varepsilon A_c^{-1\prime}$. Actually, in the VAR, X_t is not nowcasting causal for y_t if and only if the corresponding errors are uncorrelated (Lütkepohl, 2005). In the absence of NC Σ_u is indeed block diagonal:

$$A_c = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -\rho_1 & -\rho_2 \\ 0 & 0 & 1 & -\rho_1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \Rightarrow A_c^{-1}\Sigma_\varepsilon A_c^{-1\prime} = \begin{pmatrix} \sigma_L & 0 & 0 & 0 \\ 0 & \sigma_H(1+\rho_1^2+\widetilde{\rho}^2) & \rho_1\sigma_H(1+\widetilde{\rho}) & \sigma_H\widetilde{\rho} \\ 0 & \rho_1\sigma_H(1+\widetilde{\rho}) & \sigma_H(\rho_1^2+1) & \rho_1\sigma_H \\ 0 & \sigma_H\widetilde{\rho} & \rho_1\sigma_H & \sigma_H \end{pmatrix},$$

where σ_L and σ_H are the variances of ε_{1t} and $\varepsilon_{jt}, j = 2,3,4$, respectively, i.e., the entries of the diagonal matrix Σ_ε.

The presence of NC can be tested using a standard Wald test on the $(1, j)$-elements of $\widehat{\Sigma_u}$, i.e., $\widehat{\sigma}_{1,j} = \frac{1}{T}\sum_{t=1}^T \widehat{u}_{1t}\widehat{u}_{jt}, j = 2,3,4$, where $\widehat{u}_{jt}$ corresponds to the residual of equation j in the corresponding VAR. Note that the test is asymptotically distributed as χ^2_m (see, e.g., Hamilton, 1994, p. 301 or Lütkepohl, 2005, p. 104) and that the concept is fully symmetric in the VAR, i.e., if y_t is nowcasting causal for X_t, then X_t is nowcasting causal for y_t. This implies that in practice we cannot distinguish between NC in one or the other direction, i.e., whether $\delta \neq 0$ or $\beta_j \neq 0$, we can merely check for its presence or absence.

3 Mixing GC and NC in MIDAS regressions

Let us now investigate the relationship between the mixed-frequency VAR and univariate MIDAS regressions. Partitioning Σ_u in the quarter/month VAR(1) example as

$$\Sigma_u = \begin{pmatrix} \sigma_{1,1} & \sigma_{\cdot 1}' \\ \sigma_{\cdot 1} & \Sigma_{2:4} \end{pmatrix}, \text{ where } \sigma_{\cdot 1} = (\sigma_{2,1}, \sigma_{3,1}, \sigma_{4,1})' \text{ and } \Sigma_{2:4} = \begin{pmatrix} \sigma_{2,2} & \cdot & \cdot \\ \sigma_{3,2} & \sigma_{3,3} & \cdot \\ \sigma_{4,2} & \sigma_{4,2} & \sigma_{4,4} \end{pmatrix}$$

we can factorize the VAR in (2) into the conditional model for y_t given X_t and the remaining marginal models:

$$\begin{aligned} y_t = \; & \sigma'_{\cdot 1}\Sigma_{2:4}^{-1}X_t + \left[\rho_y^* - \sigma'_{\cdot 1}\Sigma_{2:4}^{-1}\begin{pmatrix} \pi_1^* \\ \pi_2^* \\ \pi_3^* \end{pmatrix}\right] y_{t-1} \\ & + \left[\underbrace{\begin{pmatrix} \phi_1^* \\ \phi_2^* \\ \phi_3^* \end{pmatrix}'}_{\phi^*} - \sigma'_{\cdot 1}\Sigma_{2:4}^{-1}\underbrace{\begin{pmatrix} a_{2,2}^* & a_{2,3}^* & a_{2,4}^* \\ a_{3,2}^* & a_{3,3}^* & a_{3,4}^* \\ a_{4,2}^* & a_{4,3}^* & a_{4,4}^* \end{pmatrix}}_{A_{2:4}^*}\right] X_{t-1} + \left[u_{1t} - \sigma'_{\cdot 1}\Sigma_{2:4}^{-1}\begin{pmatrix} u_{2t} \\ u_{3t} \\ u_{4t} \end{pmatrix}\right] \end{aligned}$$

or

$$y_t = \pi^* y_{t-1} + \underbrace{\sum_{i=0}^{2}\theta_{i,0}x_{t-i/3}^{(3)}}_{\sigma'_{\cdot 1}\Sigma_{2:4}^{-1}X_t} + \underbrace{\sum_{i=0}^{2}\theta_{i,1}x_{t-1-i/3}^{(3)}}_{(\phi^* - \sigma'_{\cdot 1}\Sigma_{2:4}^{-1}A_{2:4}^*)X_{t-1}} + u_t^*, \tag{4}$$

using straightforward substitutions and rearrangements. (4) is a standard mixed-frequency regression, the parameters of which being estimated unrestrictedly (what Foroni et al., 2012 call U-MIDAS) or using MIDAS restrictions (e.g., Ghysels et al., 2004). Equation (4) reveals that with a mixed-frequency VAR as the underlying data generating process, the univariate mixed-frequency model can mix GC and NC.

Indeed, Granger non-causality from X to y in the mixed-frequency VAR implies $\phi^* = \mathbf{0}$, whereas nowcasting non-causality corresponds to $\sigma'_{\cdot 1} = \mathbf{0}$. Assuming that the VAR in (2), derived from (1), generates the data, let us consider an analyst who works with the mixed-frequency regression model in (4) instead, i.e., who assumes the high-frequency variables X_t to be weakly exogenous for the parameters of interest. Three cases, in which the analyst may draw different conclusions from the parameter estimates of (4) than from the ones of (2), can be distinguished:

Case 1: In the absence of GC from X to y in the mixed-frequency VAR (i.e., $\phi^* = \mathbf{0}$) it is still possible to obtain non-zero coefficients on X_{t-1} in (4), namely if y is nowcasting causal for X (and hence $\sigma'_{\cdot 1} \neq \mathbf{0}$). This case corresponds to the autoregressive matrix in (3) and is analyzed in detail via Monte Carlo simulations.

Case 2: As indicated above, the parameters in (4) are often estimated after MIDAS restrictions (e.g., Ghysels et al., 2004) have been imposed. Without going into too much detail, instead of estimating all θ-coefficients unrestrictedly, which could lead to parameter proliferation issues (especially for large m), the polynomial lag structure is hyper-parameterized to yield

$$y_t = \pi^* y_{t-1} + \beta \sum_{j=0}^{1}\sum_{i=0}^{2} w_{3j+i+1}(\gamma)x_{t-j-i/3}^{(3)} + \epsilon_t, \tag{5}$$

where $w(\gamma)$ is a weight function, for which different specifications are proposed in the literature (see, e.g., Ghysels et al., 2007). In the sequel we employ the two-dimensional exponential Almon lag polynomial. Contrary to estimating one weight function for all high-frequency variables as done in (5), one can also estimate a separate weight function for each set of high-frequency variables per t-period:

$$y_t = \pi^* y_{t-1} + \beta_1 \sum_{i=0}^{2} w^1_{i+1}(\gamma^1) x^{(3)}_{t-i/3} + \beta_2 \sum_{i=0}^{2} w^2_{i+1}(\gamma^2) x^{(3)}_{t-1-i/3} + \upsilon_t. \tag{6}$$

Let us come back to the estimation of (4) instead of (2). In the absence of NC, but presence of GC from X to y, i.e., $\sigma'_{\cdot 1} = \mathbf{0}$ and $\phi^* \neq \mathbf{0}$, (6) is probably more appropriate than (5): Due to nowcasting non-causality, coefficients on X_t should equal zero, whereas the ones on the respective observations in period $t-1$ should be non-zero due to existing GC. This implies an abrupt jump in the coefficients from zero to a non-zero value. In (6) this is easily achieved when $\widehat{\beta_1} = 0$, while (5) probably yields non-zero coefficients for some observations in X_t. This case is illustrated by means of an application in Section 5, and in particular Figure 2.

Case 3: Suppose X Granger causes y in the mixed-frequency VAR, i.e., $\phi^* \neq \mathbf{0}$, and NC is present as well, i.e., $\sigma'_{\cdot 1} \neq \mathbf{0}$, with the additional restriction that $\phi^* = \sigma'_{\cdot 1}\Sigma^{-1}_{2:4}A^*_{2:4}$. In this unlikely case $\widehat{\theta_{i,1}}, i = 0, 1, 2$, from (4) could be statistically indistinguishable from zero.

4 Monte Carlo study

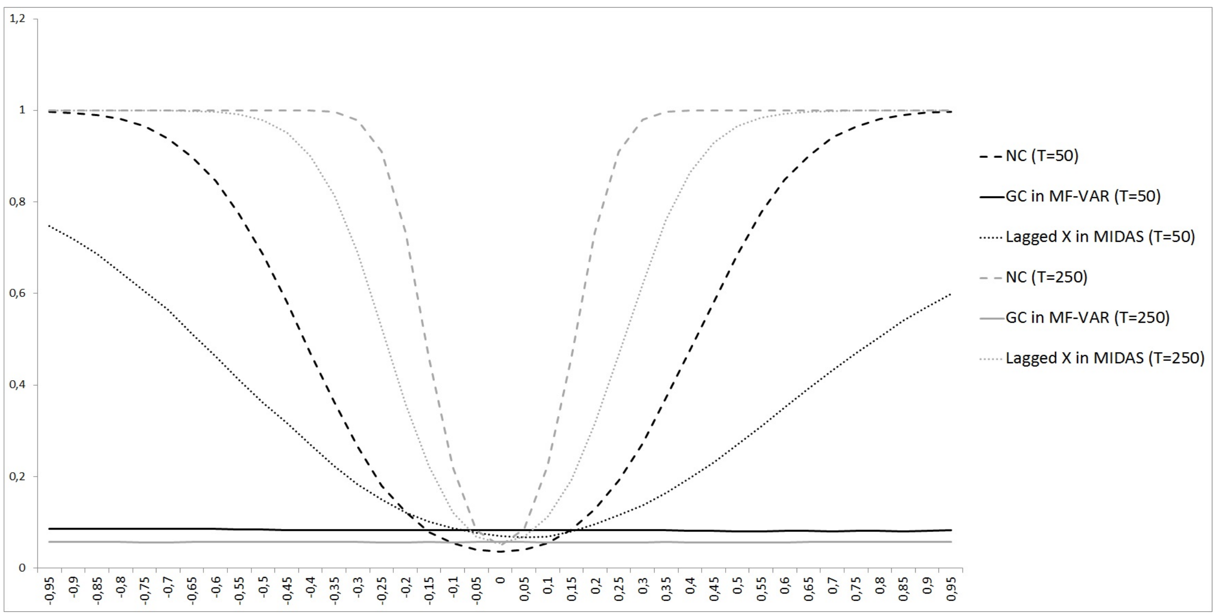
We analyze Case 1 presented in the previous section by generating the VAR(1) in (2) from the following dynamic structural equations model:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ \delta & 1 & -0.4 & 0.2 \\ 0 & 0 & 1 & -0.4 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_t \\ x^{(3)}_t \\ x^{(3)}_{t-1/3} \\ x^{(3)}_{t-2/3} \end{pmatrix} = \begin{pmatrix} 0.5 & 0 & 0 & 0 \\ 0.3 & 0.6 & 0 & 0 \\ 0.3 & -0.2 & 0.6 & 0 \\ 0.3 & 0.4 & -0.2 & 0.6 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x^{(3)}_{t-1} \\ x^{(3)}_{t-1-1/3} \\ x^{(3)}_{t-1-2/3} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{pmatrix}, \tag{7}$$

where $\sigma_L = \sigma_H = 1$. Since $\beta_j = \phi_j = 0 \ \forall j$, A^*_1 has the same structure as in (3). We test for Granger and nowcasting non-causality in the mixed-frequency VAR using a standard Wald test on $\widehat{\phi^*_j} \ \forall j$ for the former and the one described at the end of Section 2 for the latter. Furthermore, we test $\theta_{0,1} = \theta_{1,1} = \theta_{2,1} = 0$ in (4), i.e., the significance of X_{t-1} in the univariate mixed-frequency regression. We do so for $\delta \in [-0.95, 0.95]$ and $T = 50, 250$. Figure 1 depicts the various rejection frequencies at a 5% significance level and are all based on 10,000 replications using GAUSS12.

Due to the absence of GC independent from δ the corresponding rejection frequencies for $T = 250$ almost equal the nominal size of 5%, whereas the test is oversized for $T = 50$. For

Figure 1: Rejection frequencies of the various tests



Note: This figure shows the rejection frequencies for Granger non-causality (solid lines) and nowcasting non-causality (dashed lines) in the mixed-frequency VAR as well as for the insignificance of X_{t-1} in the univariate mixed-frequency regression (dotted lines) at the 5% level. The data are generated from (7) with $\delta \in [-0.95, 0.95]$ (horizontal axis). The sample sizes considered are $T = 50$ (black lines) and 250 (grey lines).

$\delta = 0$, the test for NC has an actual size almost identical to the nominal one for both sample sizes. The rejection frequencies increase quickly (more rapidly for large T) as we turn away from the null. As expected, the test for $\theta_{0,1} = \theta_{1,1} = \theta_{2,1} = 0$ in (4) behaves similarly to the one for NC. It rejects close to 5% of the times for $\delta = 0$ and more often as $|\delta|$ increases. However, its rejection frequencies are lower than the ones for NC for both sample sizes: From (4), we have that $\sum_{i=0}^{2} \theta_{i,1} x_{t-1-i/3}^{(3)} = (\phi^* - \sigma'_{\cdot 1}\Sigma_{2:4}^{-1} A_{2:4}^*) X_{t-1}$, which, for a DGP characterized by an autoregressive matrix as in (3), boils down to $\begin{pmatrix} \delta\frac{\rho_3}{\delta^2+1} & 0 & 0 \end{pmatrix} X_{t-1}$. With $\rho_3 = 0.6$ fixed, we know that for reasonable values[4] of δ the coefficient of $x_{t-1}^{(3)}$ is going to be different from zero.

5 Applications

This section illustrates the three cases outlined in Section 3 using data from the Federal Reserve Bank of St. Louis. Whenever non-stationarity in the time series is detected, we compute growth rates in order to achieve stationarity. Cointegration issues are neglected here for simplicity, but the methodology can be extended to the non-stationary, possibly cointegrated, case (see Götz et al., 2013).

We first consider the relationship between the monthly growth rate of the seasonally adjusted industrial production index (IPI hereafter), and the weekly growth rate of the seasonally adjusted stock of money M2 (M hereafter). Estimated on the period from January 1991 to December 2010 the mixed-frequency VAR(1) model with $m = 4$ gives

$$\begin{pmatrix} \widehat{IPI_t} \\ \widehat{M_t^{(4)}} \\ \widehat{M_{t-1/4}^{(4)}} \\ \widehat{M_{t-2/4}^{(4)}} \\ \widehat{M_{t-3/4}^{(4)}} \end{pmatrix} = \widehat{\mu} + \begin{pmatrix} \underset{(0.069)}{0.284} & \underset{(0.205)}{0.092} & \underset{(0.194)}{-0.105} & \underset{(0.266)}{-0.469} & \underset{(0.182)}{0.213} \\ \underset{(0.025)}{-0.06} & \underset{(0.073)}{-0.019} & \underset{(0.069)}{0.008} & \underset{(0.094)}{-0.09} & \underset{(0.064)}{-0.168} \\ \underset{(0.026)}{-0.021} & \underset{(0.076)}{0.153} & \underset{(0.071)}{0.017} & \underset{(0.098)}{-0.014} & \underset{(0.067)}{0.021} \\ \underset{(0.017)}{0.013} & \underset{(0.05)}{0.137} & \underset{(0.047)}{-0.044} & \underset{(0.065)}{-0.014} & \underset{(0.044)}{0.077} \\ \underset{(0.025)}{-0.015} & \underset{(0.073)}{0.01} & \underset{(0.069)}{-0.171} & \underset{(0.095)}{0.033} & \underset{(0.065)}{0.107} \end{pmatrix} \begin{pmatrix} IPI_{t-1} \\ M2_{t-1}^{(4)} \\ M2_{t-5/4}^{(4)} \\ M2_{t-6/4}^{(4)} \\ M2_{t-7/4}^{(4)} \end{pmatrix},$$

where standard errors are displayed in brackets beneath the least squares coefficient estimates. The Wald test for $H_0 : \phi_{1,1} = \phi_{1,2} = \phi_{1,3} = \phi_{1,4} = 0$, i.e., ξ_{GC_VAR}, has a p-value of 0.2484. The Wald statistic for testing nowcasting non-causality à la Hamilton (1994), i.e., $\sigma'_{\cdot 1} = \mathbf{0}$, turns out to be $\xi_{NC} = 36.55$ with a p-value of practically 0%. Thus, Granger non-causality from M to IPI in the mixed-frequency VAR as well as NC between the two series is detected. A mixed-frequency regression in (4) yields:

$$\begin{aligned} \widehat{IPI}_t = {} & \underset{(0.0007)}{0.0038} + \underset{(0.065)}{0.196} IPI_{t-1} - \underset{(0.179)}{1.069} M_t^{(4)} - \underset{(0.175)}{0.667} M_{t-1/4}^{(4)} + \underset{(0.249)}{0.27} M_{t-2/4}^{(4)} \\ & - \underset{(0.169)}{0.441} M_{t-3/4}^{(4)} + \underset{(0.192)}{0.142} M_{t-1}^{(4)} - \underset{(0.18)}{0.148} M_{t-5/4}^{(4)} - \underset{(0.244)}{0.557} M_{t-6/4}^{(4)} + \underset{(0.171)}{0.073} M_{t-7/4}^{(4)}. \end{aligned}$$

[4] $\lim_{|\delta|\to\infty} \delta\frac{\rho_3}{\delta^2+1} = 0$ such that for δ large enough the corresponding coefficient converges to zero. However, for $|\delta| \le 5$ we have that $\theta_{0,1} > 0.1$.

Now, for testing whether $\theta_{0,1} = \theta_{1,1} = \theta_{2,1} = \theta_{3,1} = 0$ the associated Wald test has a p-value of 0.0821. Hence, in the univariate mixed-frequency regression and at the 10% significance level we would conclude the set of lagged money-variables to be statistically different from zero.

The second pair of data consists of the monthly variation of the seasonally adjusted civilian unemployment rate (U hereafter), as regressand, and the aforementioned M, as regressor. The time period considered as well as m are the same as for the first set of data, i.e., $T = 240$ and $m = 4$. Testing for Granger non-causality in the mixed-frequency VAR(1) yields a p-value equal to 0.004, testing for nowcasting non-causality gives p-value= 0.678, and checking for non-causality in the mixed-frequency regression results in a p-value of 0.0059. In other words, there is no NC between the two series, but M continues to be statistically significant when conditioning on the high-frequency variables.

Here it is of interest to analyze what happens if MIDAS restrictions are imposed and, in particular, if we estimate only one or two weight functions, i.e., estimate (5) or (6). Given the estimates of $\gamma = (\gamma_1, \gamma_2)'$ in (5) with $m = 4$, we can compute the corresponding aggregate version of M as $\sum_{j=0}^{1}\sum_{i=0}^{3} w_{4j+i+1}(\widehat{\gamma})M^{(4)}_{t-j-i/4}$. This allows us to run a least squares regression and compute the standard error of the scale coefficient $\widehat{\beta} = 31.05$ as well as its t-statistic.[5] With $t_{\widehat{\beta}} = 4.01$ the corresponding coefficients on the whole set of M variables (t and $t-1$) are jointly statistically different from zero. These coefficients can be obtained by multiplying the γ-dependent weights by the scale coefficient and are plotted as solid line in Figure 2. Crucially and contrary to the missing NC feature, some M_t-variables possess a non-zero coefficient ($M^{(4)}_{t-3/4}$ and to a lesser extent $M^{(4)}_{t-2/4}$). If, however, (6) with $m = 4$ is estimated, $t_{\widehat{\beta_1}} = 1.34$ and $t_{\widehat{\beta_2}} = 3.84$ implying that $\widehat{\beta}_1$, corresponding to the instantaneous M variables, is not significantly different from zero. In other words, the nowcasting non-causality feature is preserved when estimating two separate weight functions. This is reflected by the associated coefficients plotted as dotted line in Figure 2.

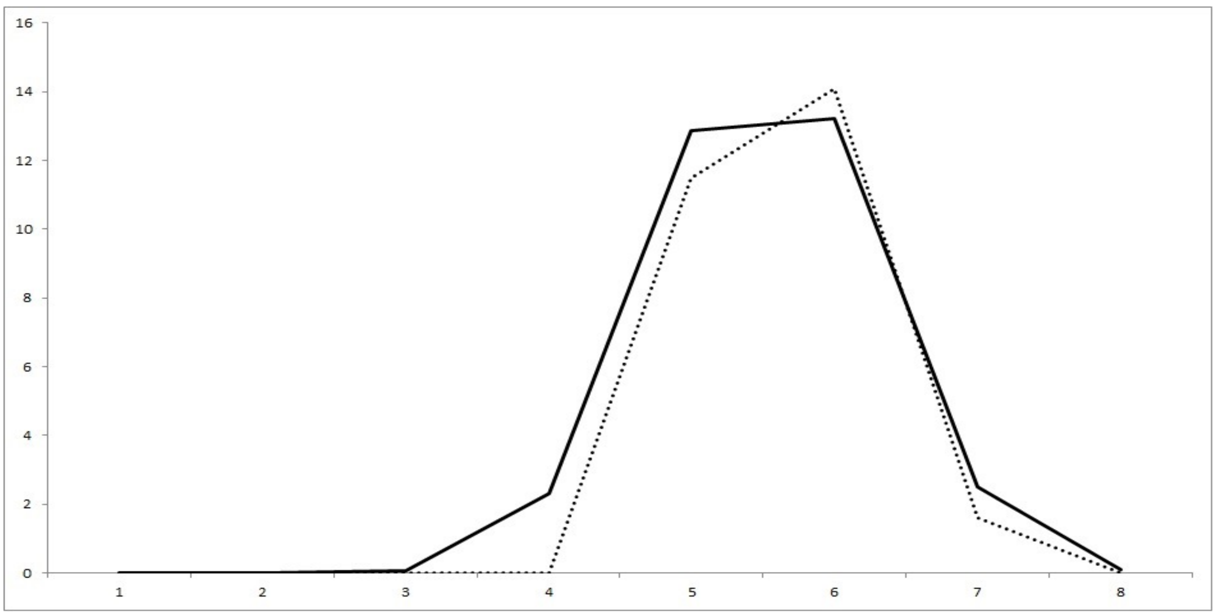
In the third application, the dependent variable is the quarterly growth rate of the seasonally adjusted real gross national product (GNP hereafter); the growth rate of IPI is the independent variable such that $m = 3$. This pair of time series is observed from 1948Q3 to 2011Q4. After estimating a VAR(1) we find that $\xi_{GC\_VAR} = 105.8$ leading to a clear rejection of the associated null hypothesis. Testing for nowcasting non-causality yields $\xi_{NC} = 78.18$ such that both GC in the mixed-frequency VAR and NC are detected. Estimation of the univariate mixed-frequency regression leads to

$$\begin{aligned}\widehat{GNP}_t = \; & \underset{(0.0005)}{0.0053} - \underset{(0.064)}{0.12}\, GNP_{t-1} + \underset{(0.04)}{0.211} IPI^{(3)}_t - \underset{(0.062)}{0.045} IPI^{(3)}_{t-1/3} + \underset{(0.047)}{0.253} IPI^{(3)}_{t-2/3} \\ & + \underset{(0.06)}{0.042} IPI^{(3)}_{t-1} - \underset{(0.062)}{0.067} IPI^{(3)}_{t-4/3} + \underset{(0.039)}{0.07}\, IPI^{(3)}_{t-5/3}.\end{aligned}$$

Testing for $\theta_{0,1} = \theta_{1,1} = \theta_{2,1} = 0$, however, amounts to a p-value of 0.2997. Consequently,

[5] Due to non-identification of the weight-specifying parameter γ under the null hypothesis, we would have to follow the approach of Davies (1987) and compute Hansen (1996)'s p-value. Although not presented and reported here, the results do not differ qualitatively. The same holds for the estimation of (6).

Figure 2: Coefficients on M_t and M_{t-1} with one or two separate weight functions



Note: This figure shows the coefficients, i.e., their respective weights×scale coefficient, of $M_t^{(4)}, \ldots, M_{t-3/4}^{(4)}$ and $M_{t-1}^{(4)}, \ldots, M_{t-7/4}^{(4)}$, once when only one function is computed on the whole set of variables (solid line) and once when two separate weight functions, one for each set of 4 variables (dotted line), are computed. The weight functions are obtained using the two-dimensional Almon Lag Polynomial (see, e.g., Ghysels et al., 2007). The estimates are $\widehat{\beta} = 31.05$, $\widehat{\gamma} = (9.34, -0.75)'$ for the solid line and $(\widehat{\beta_1}, \widehat{\beta_2})' = (12.6, 27.2)'$, $\widehat{\gamma^1} = (1.4, -0.27)'$ and $\widehat{\gamma^2} = (3.75, -1.18)'$ for the dotted line. Due to the fact that the scale coefficient corresponding to $M_t^{(4)}, \ldots, M_{t-3/4}^{(4)}$, i.e., $\widehat{\beta_1}$, is statistically not different from zero, its estimate is set to zero.

estimating the univariate mixed-frequency regression instead of the mixed-frequency VAR leads to IPI_{t-1} being statistically insignificant.

Acknowledgements

We thank an anonymous referee, Michael Eichler, Daniela Osterrieder and Jean-Pierre Urbain for useful suggestions and comments on earlier versions of the paper.

References

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196.

Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternatives. *Biometrika*, 74(1):pp. 33–43.

Foroni, C., Marcellino, M., and Schumacher, C. (2012). U-midas: Midas regressions with unrestricted lag polynomials. CEPR Discussion Papers 8828, C.E.P.R. Discussion Papers.

Ghysels, E. (2012). Macroeconomics and the reality of mixed frequency data. Working paper, University of North Carolina.

Ghysels, E., Motegi, K., and Hill, J. (2013). Granger causality tests with mixed data frequencies. Discussion Paper.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2004). The midas touch: Mixed data sampling regression models. CIRANO Working Papers 2004s-20, CIRANO.

Ghysels, E., Sinko, A., and Valkanov, R. (2007). Midas regressions: Further results and new directions. *Econometric Reviews*, 26(1):53–90.

Giannone, D., Reichlin, L., and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4):665–676.

Götz, T. B., Hecq, A., and Urbain, J.-P. (2013). Testing for common cycles in non-stationary vars with varied frecquency data. *Advances in Econometrics Vol. 31*, forthcoming.

Hamilton, J. D. (1994). *Time series analysis.* Princeton Univ. Press, Princeton, NJ.

Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica*, 64(2):413–30.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis.* Springer Publishing Company, Incorporated.

Miller, J. I. and Ghysels, E. (2013). Testing for cointegration with temporally aggregated and mixed-frequency time series. Working paper, University of Missouri.