Mixed Causal-Noncausal Models

Sean Telg

Before reading...

In this article, I want to familiarize you with mixed causal-noncausal autoregressive models: a topic that I have extensively studied in my PhD thesis and have continued to do - together with Sébastien Fries - here at the VU. The first section provides some intuition on how such models emerged from theory. For applied econometricians, the second section discusses some empirical applications and future research ideas for these models.

1 Some background

Consider the following difference equation:

$$y_t = a_1 y_{t-1} + \epsilon_t, \tag{1}$$

where (ϵ_t) is a sequence of white noise. You might be tempted to claim that this is an autoregressive (AR) process of order 1, but we have to be careful. In standard econometrics courses and books, you study this model mostly under the assumption of model stability: i.e. $|a_1| < 1$. In combination with (ϵ_t) being white noise, this ensures stationarity of the AR(1). Applying some limiting results and repeated substitution to (1), it can be shown that the convergent solution to this difference equation is the moving average (MA) representation $y_t = \sum_{j=0}^{\infty} a_1^j \epsilon_{t-j}$. It shows that y_t only depends on current and past values of the sequence (ϵ_t). You probably also studied the case where $a_1 = 1$, i.e. the unit root case. It is well-known that there is no stationary solution to this problem. However, what happens when $|a_1| > 1$? In that case, (1) is often indicated as an "explosive process", but note that it also has a convergent solution. Increase each time subscript by one and rewrite (1) as

$$y_t = \frac{1}{a_1} y_{t+1} - \frac{1}{a_1} \epsilon_{t+1}. \tag{2}$$

Using similar arguments as before, the convergent solution is given by $y_t = -\sum_{j=1}^{\infty} a_1^{-j} \epsilon_{t+j}$, i.e. y_t depends solely on future values of (ϵ_t).

1.1 Why were such solutions barely studied?

A quote (adjusted to our notation) in Brockwell and Davis (1991) clarifies: "The stationary solution is frequently regarded as unnatural since y_t is correlated with $\{\epsilon_s, s > t\}$. [...] It should be noted that every AR(1) process with $|a_1| > 1$ can be reexpressed as an AR(1) process with $|a_1| < 1$ and a new white noise sequence. **From a second-order point of view** therefore, nothing is lost by eliminating AR(1) processes with $|a_1| > 1$ from consideration."

1.2 Why should we study them anyways?

The bold-faced part in the last sentence reveals why one might be interested in studying - as they are called - noncausal processes. The two processes discussed above indeed have the same autocovariance function (and thus, spectral density) as this is a symmetric measure. This means that based on second moments they are identical. However, only the Normal distribution is **fully** characterized by second-order properties. Hence, if we assume *any other distribution* for the sequence (ϵ_t), there is information hidden in either lower or higher moments than the second, which makes the probabilistic structure of the two processes different from each other! In the next section, we discuss how this is useful, but let us first generalize the idea.

1.3 Generalization

Our starting point was an AR(1), but note that we can generalize our claim by considering an AR(p) process:

$$a(L)y_t = \epsilon_t, \tag{3}$$

where $a(L) = 1 - a_1L - ... - a_pL^p$ and L is a lag operator such that $L^iy_t = y_{t-i}$. Suppose that this process has r well-behaved roots for stationarity, (i.e. they lie outside the unit circle) an s ill-behaved roots (they lie inside the unit circle). Then we can rewrite $a(L) = \phi(L)\varphi^*(L)$, where the first polynomial contains the r "good" roots and $\varphi^*(L)$ the s "bad" roots. In order to solve the issue of the "bad" roots, we apply a procedure similar to (2). Lanne and Saikkonen (2011) apply this idea in their article and rewrite (3) as

$$\phi(L)\varphi(L^{-1})y_t = \varepsilon_t, \tag{4}$$

with (ε_t) a new white noise sequence and $\phi(L) = 1 - \phi_1L - ... - \phi_rL^r$ and $\varphi(L^{-1}) = 1 - \varphi_1L^{-1} - ... - \varphi_sL^{-s}$. Note that L^{-1} is a lead operator such that $L^{-i}y_t = y_{t+i}$. Finally, (4) is known as the mixed causal-noncausal autoregressive (MAR) model. But enough algebra for now, what can we actually do with these models?

2 Applications

The popularity of the MAR model has received a lot of attention in the last decade. This can mostly be explained by three reasons. Firstly, it has been shown that MAR models can generate processes exhibiting nonlinear features such as speculative bubbles and asymmetric cycles (see e.g., Gouriéroux and Zakoïan, 2017 or Fries and Zakoïan, 2019). Secondly, the MAR can be seen as "nonfundamental"[1] solutions of rational expectations models (see e.g. Lanne and Luoto,

[1] A situation in which the econometrician has less information to his/her disposal than the economic agents.

2013 for an application to the Hybrid New Keynesian Phillips Curve). Thirdly, MAR models may have forecasting advantages over the conventional AR model (Lanne et al., 2012). Let us focus on the first reason in this research article.

2.1 Bubbles

We start by considering the simplest two models within the class of MAR models, the purely causal and noncausal AR(1):

$$y_t = \phi_1 y_{t-1} + \varepsilon_t \Rightarrow y_t = \sum_{j=0}^{\infty} \phi_1^j \varepsilon_{t-j} \tag{5}$$

$$y_t = \varphi_1 y_{t+1} + \varepsilon_t \Rightarrow y_t = \sum_{j=0}^{\infty} \varphi_1^j \varepsilon_{t+j} \tag{6}$$

The two models and their corresponding MA representations look very similar, but there is an important difference. Suppose that there is a very large shock ε_t at $t = 10$. In (5), we will be surprised by the shock at $t = 20$, as y_t is only determined by the current and past errors. This is completely different from (6). Since y_t depends on its future errors, we are closely building towards the big shock. In other words, this process is anticipative. Figure 1(a)-(b) visualizes this by idea by simulating causal and noncausal AR(1) processes where (ε_t) follows a Cauchy distribution. Figure 1(c)-(d) shows examples of the MAR model, which combines a causal and noncausal part. You can observe that there is no sudden increase or crash: the noncausal component builds up the bubble, the causal part ensures the exponential decay after we reached the highest (or as in panel (d) - lowest) point. The rate at which this happens depends on the values of the coefficients. Note that the choice of a fat-tailed distribution, such as the Cauchy in this simulation, is necessary to draw a "large value" that drives the bubble. In case we had drawn the shocks (ε_t) from a Normal distribution, there would hardly be any difference between the four processes. Empirical implementations of MAR models on bubbles can be found in e.g.

Hencic and Gouriéroux (2015) for Bitcoin/USD exchange rates, Hecq et al. (2016) for Belgian solar panels and Fries and Zakoïan (2019) for financial price series. In case you are interested in playing around a bit with these models, we developed an R package "MARX" which makes it possible to simulate, estimate and forecast with MAR models (Hecq et al., 2017).

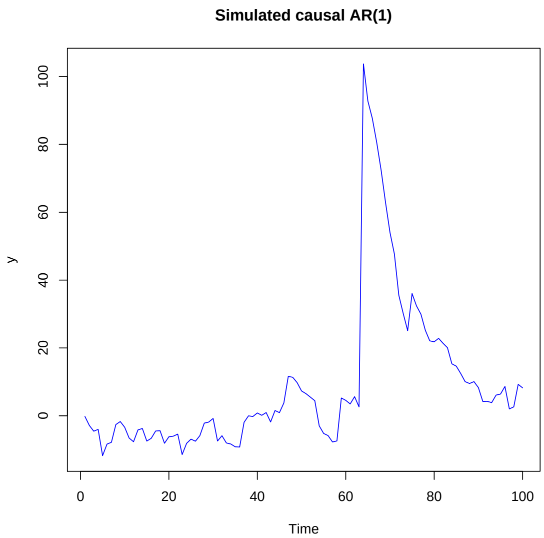
2.2 Challenges

This research article is an attempt to comprehensively explain the concept of MAR models and their main advantages for modelling economic and financial processes. However, there is still quite some room for research:

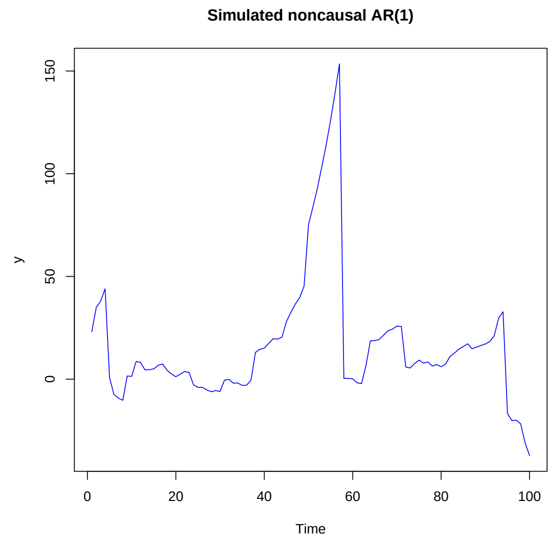
- Purely causal and noncausal models are typically nested, i.e. an AR(1) can be seen as a restricted version of an AR(2). This is not the case for MAR models in (4) as the multiplication of polynomials makes the model nonlinear in the parameters. This complicates model selection procedures.
- To estimate MAR models, the method of non-Gaussian maximum likelihood is typically employed. This implies that a distributional assumption is necessary to perform estimation. Are there other methods that are less stringent?
- In the multivariate case, things rapidly become more complicated. The MAR in (4) now has polynomial matrices which are generally not commutative (i.e. $AB \neq BA$). What are the implications of this? Can the models be used to evaluate policy measures (e.g. by impulse-response analysis)? Is it sensible to look into structural multivariate MAR models? Can we introduce a mixed-frequency component? And more....

3 A brief conclusion

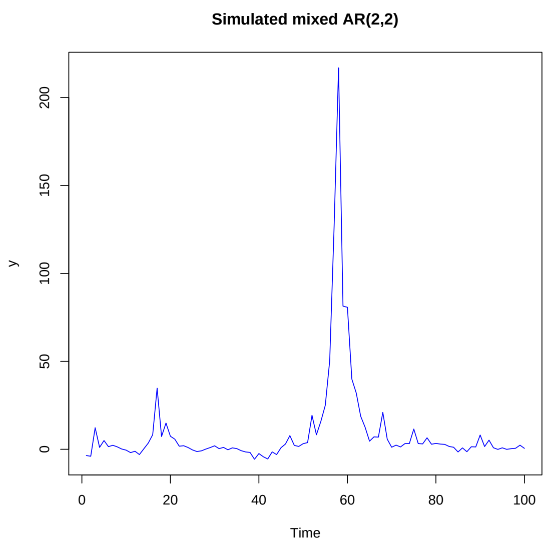
The MAR model resulted in a new stream of literature in the field of economics and finance over the last years. In case this topic piqued your interest, you can find some references below.



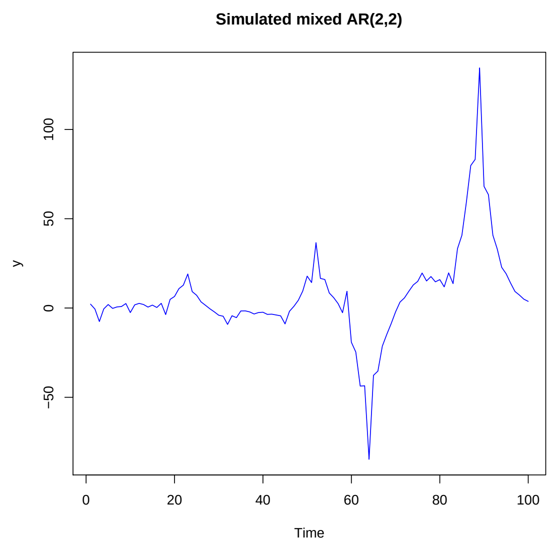
(a) Simulated causal MAR(1,0) process ($\phi_1 = 0.9$) with standard Cauchy errors



(b) Simulated noncausal MAR(0,1) process ($\varphi_1 = 0.9$) with standard Cauchy errors



(c) Simulated MAR(2,2) process ($\boldsymbol{\phi} = [0.2, 0.3]', \boldsymbol{\varphi} = [0.2, 0.1]'$) with standard Cauchy errors



(d) Simulated MAR(2,2) process ($\boldsymbol{\phi} = [0.3, 0.3]', \boldsymbol{\varphi} = [0.4, 0.4]'$) with standard Cauchy errors

Figure 1: Simulated processes from various MAR(r, s) specifications, $T = 100$

References

Brockwell, P. & Davis, R. (1991). *Time Series: Theory and Methods*. Springer-Verlag New York, Second Edition.

Fries, S. & J.-M. Zakoian (2019). Mixed Causal-Noncausal AR Processes and the Modelling of Explosive Bubbles. *Econometric Theory*, 35, 1234–1270. doi:10.1017/S0266466618000452

Gouriéroux, C. & Zakoïan, J.M. (2017). Local explosion modelling by non-causal process. *Journal of the Royal Statistical Society, Series B*, 79, 737–756. doi:10.1111/rssb.12193

Hecq, A., Lieb, L. & Telg, S. (2016). Identification of Mixed Causal-Noncausal Models in Finite Samples. *Annals of Economics and Statistics*, 123/124, 307–331. doi:10.15609/annaeconstat2009.123-124.0307

Hecq, A., Lieb, L. & Telg, S. (2017). Simulation, Estimation and Selection of Mixed Causal-Noncausal Autoregressive Models: The MARX Package. *Working Paper*, available at SSRN: `https://ssrn.com/abstract=3015797`.

Hencic A. & Gouriéroux C. (2015). Noncausal Autoregressive Model in Application to Bitcoin/USD Exchange Rates. In: Huynh VN., Kreinovich V., Sriboonchitta S., Suriya K. (eds). *Econometrics of Risk: Studies in Computational Intelligence*, 583, Springer.

Lanne, M., Luoto J. & Saikkonen, P. (2012). Optimal Forecasting of Noncausal Autoregressive Time Series. *International Journal of Forecasting*, 28, 623–631. doi:10.1016/j.ijforecast.2011.08.003

Lanne, M. & Luoto, J. (2013). Autoregression-Based Estimation of the New Keynesian Phillips Curve. *Journal of Economic Dynamics & Control*, 37, 561–570. doi:10.1016/j.jedc.2012.09.008

Lanne, M. & Saikkonen, P. (2011). Noncausal Autoregressions for Economic Time Series. *Journal of Time Series Econometrics*, 3(3), 1-32. doi:10.2202/1941-1928.1080