

POS tests for predictive regressions

July 9, 2020

- Framework
- The test statistic in the presence of stochastic regressors
- The problem
- Dependence modelling using copulas

Framework

- Predictive regressions assume the form

$$y_t = \beta x_{t-1} + \varepsilon_t \quad t = 1, \dots, n \quad (1)$$

- x_t is usually a highly persistent regressor such that the disturbances of the regressor are correlated with the error terms of the predictive regression. In a framework considered by, x_t is assumed to follow a stationary AR(1) process, such that

$$x_t = \theta x_{t-1} + u_t$$

where (u_t, ε_t) are bivariate normal distribution, with the covariance matrix

$$\begin{bmatrix} 1 & \sigma_{\varepsilon u} \\ \sigma_{\varepsilon u} & \sigma_u^2 \end{bmatrix}$$

- Therefore, u_t and ε_t are contemporaneously correlated with $\text{corr}(\varepsilon_t, u_t) \neq 0$, which implies that there is feedback from u_t to x_t , through ε_t , and as such $\text{corr}(u_t, x_{t+j}) \neq 0$ for $j \geq 0$.

- To build our POS-based tests in the context of predictive regressions. Let y_t be explained by a vector variable x_t , such that

$$y_t = \beta x_{t-1} + \varepsilon_t, \quad t = 1, \dots, n \quad (2)$$

where x_t is a $k \times 1$ vector of stochastic variables and $\beta \in \mathbb{R}^k$ is a $k \times 1$ vector of unknown parameters. Furthermore, following, define the information vector $\mathcal{I}_t = [x'_0, \dots, x'_t]$ defined for $t = 1, \dots, T-1$, with the convention that $\mathcal{I}_0 = x_0$.

- . Suppose we wish to test the null hypothesis

$$H_0 : \beta = 0$$

against the alternative

$$H_1 : \beta = \beta_1$$

- To build their tests assume that the error terms has a conditional zero median, such that

$$P[\varepsilon_t \geq 0 \mid \mathcal{I}_{t-1}] = P[\varepsilon_t < 0 \mid \mathcal{I}_{t-1}] \quad (3)$$

- Tests derived under (3) are interpreted as tests of whether the conditional median of y_t is predictable using x_{t-1} .
- To build our POS-bases tests, we follow by further assuming that the error terms follow a mediangale process conditional on \mathcal{I}_{t-1}

$$P[\varepsilon_t \geq 0 \mid \varepsilon_1, \dots, \varepsilon_{t-1}, \mathcal{I}_{t-1}] = P[\varepsilon_t < 0 \mid \varepsilon_1, \dots, \varepsilon_{t-1}, \mathcal{I}_{t-1}] = \frac{1}{2}$$

- This assumption allows for serial (non-linear) dependence between the error terms.

Theorem

Under the assumption (3) and given the model (2), the signs $s(\varepsilon_1), \dots, s(\varepsilon_n)$ are i.i.d conditional on \mathcal{I}_{t-1} , according to the distribution

$$P[s(\varepsilon_t) = 1 \mid \mathcal{I}_{t-1}] = P[s(\varepsilon_t) = 0 \mid \mathcal{I}_{t-1}] = \frac{1}{2}, \quad t = 1, \dots, n$$

- We define the vector of signs $U(n) = (s(y_1), \dots, s(y_n))'$, where $s(y_t) = \mathbb{1}_{\mathbb{R}^+ \cup 0}\{y_t\}$. Thus, the likelihood function of the sample in terms of signs under the null hypothesis is

$$\begin{aligned} L(U(n), 0) &= \prod_{t=1}^n P[s(y_t) = s_t \mid \mathcal{I}_{t-1}] \\ &= \prod_{t=1}^n P[y_t \geq 0 \mid \mathcal{I}_{t-1}]^{s_t} P[y_t < 0 \mid \mathcal{I}_{t-1}]^{1-s_t} \\ &= \prod_{t=1}^n P[\varepsilon_t \geq 0 \mid \mathcal{I}_{t-1}]^{s_t} P[\varepsilon_t < 0 \mid \mathcal{I}_{t-1}]^{1-s_t} \\ &= \left(\frac{1}{2}\right)^n \end{aligned}$$

- Under the alternative, the signs $s(y_1), \dots, s(y_n)$ are no longer independent

$$\begin{aligned}
 L(U(n), \beta_1) &= \prod_{t=1}^n P[s(y_t) = s_t \mid \mathbf{S}_{t-1}, \mathcal{I}_{t-1}] \\
 &= \prod_{t=1}^n P[y_t \geq 0 \mid \mathbf{S}_{t-1}, \mathcal{I}_{t-1}]^{s(y_t)} \times \\
 &\quad P[y_t < 0 \mid \mathbf{S}_{t-1}, \mathcal{I}_{t-1}]^{1-s(y_t)}
 \end{aligned}$$

- where $\mathbf{S}_{t-1} = \{s(y_{t-1}) = s_{t-1}, \dots, s(y_1) = s_1\}$ with $\mathbf{S}_0 = \{\emptyset\}$ and $P[s(y_1) = s_1 \mid \mathbf{S}_0] = P[s(y_1) = s_1]$.

- We may assume that process of signs $\{s(y_t)\}_{t=0}^{\infty}$ follows a first order Markov process - i.e.

$$\begin{aligned}
 P[y_t \geq 0 \mid \mathbf{S}_{t-1}, \mathcal{I}_{t-1}] &= \\
 P[y_t \geq 0 \mid y_{t-1} \geq 0, \mathcal{I}_{t-1}]^{s(y_{t-1})} P[y_t \geq 0 \mid y_{t-1} < 0, \mathcal{I}_{t-1}]^{1-s(y_{t-1})} \\
 P[y_t < 0 \mid \mathbf{S}_{t-1}, \mathcal{I}_{t-1}] &= \\
 P[y_t < 0 \mid y_{t-1} \geq 0, \mathcal{I}_{t-1}]^{s(y_{t-1})} P[y_t < 0 \mid y_{t-1} < 0, \mathcal{I}_{t-1}]^{1-s(y_{t-1})}
 \end{aligned}$$

- Thus, it can be shown that the test statistic is given by

$$\tilde{S}L_n(\beta_1) = \sum_{t=1}^n \tilde{a}_t(\beta_1)s(y_t) + \sum_{t=1}^n \tilde{b}_t(\beta_1)s(y_t)s(y_{t-1}),$$

where

$$\tilde{a}_1(\beta_1) = \ln \left\{ \frac{1 - P[\varepsilon_1 < -\beta'_1 \mathbf{x}_0 \mid \mathcal{I}_0]}{P[\varepsilon_1 < -\beta'_1 \mathbf{x}_0 \mid \mathcal{I}_0]} \right\}, \quad \tilde{b}_1(\beta_1) = 0$$

- and for $t = 2, \dots, n$,

$$\tilde{a}_t(\beta_1) = \ln \left\{ \frac{1 - \frac{P[\varepsilon_t < -\beta'_1 x_{t-1}, \varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]}{P[\varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]}}{\frac{P[\varepsilon_t < -\beta'_1 x_{t-1}, \varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]}{P[\varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]}} \right\}$$

$$\begin{aligned} \tilde{b}_t(\beta_1) = & \ln \left\{ \frac{1 - \left(\frac{P[\varepsilon_t < -\beta'_1 x_{t-1} | \mathcal{I}_{t-1}]}{1 - P[\varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]} - \frac{P[\varepsilon_{t-1} < -\beta'_1 x_{t-2}, \varepsilon_t < -\beta'_1 x_{t-1} | \mathcal{I}_{t-1}]}{1 - P[\varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]} \right)}{\frac{P[\varepsilon_t < -\beta'_1 x_{t-1} | \mathcal{I}_{t-1}]}{1 - P[\varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]} - \frac{P[\varepsilon_{t-1} < -\beta'_1 x_{t-2}, \varepsilon_t < -\beta'_1 x_{t-1} | \mathcal{I}_{t-1}]}{1 - P[\varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]}} \right\} \\ & - \ln \left\{ \frac{1 - \frac{P[\varepsilon_t < -\beta'_1 x_{t-1}, \varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]}{P[\varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]}}{\frac{P[\varepsilon_t < -\beta'_1 x_{t-1}, \varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]}{P[\varepsilon_{t-1} < -\beta'_1 x_{t-2} | \mathcal{I}_{t-1}]}} \right\} \end{aligned}$$

The problem

The weights

The problem consists of evaluation the univariate and bivariate probabilities

$$P [\varepsilon_t < -\beta'_1 x_{t-1}, \varepsilon_{t-1} < -\beta'_1 x_{t-2} < 0 \mid \mathcal{I}_{t-1}]$$

and

$$P [\varepsilon_t < -\beta'_1 x_{t-1} \mid \mathcal{I}_{t-1}]$$

for $t = 1, \dots, T$

Dependence modelling using copulas

- As the mediangale assumption allows for non-linear serial dependence, testing assumption A1 by considering linear correlation is inappropriate.
- we suggest fitting copula models, which provide the means of separating the marginal distribution of the process from their respective dependence structure.
- For instance, the marginals can be assumed to possess standard normal distributions, while the nonlinear dependency is modeled using rank correlation measures (e.g. Kendall Tau) and copulas, where the latter are invariant under monotonic transformations; hence, they are not affected by the marginal distributions [see].

- A special case is where $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{T-1}, \varepsilon_T$ are distributed according to $N(0, 1)$. As suggested before, since the form of the serial dependence of the errors is non-linear, we may estimate the bivariate probabilities using copulas.
- The latter stem from a Theorem brought forward by, which states that there is a copula C such that

$$F_\varepsilon(-\beta'x_0, \dots, -\beta'x_{n-1}) = C(F_1(-\beta'x_0), \dots, F_n(-\beta'x_{n-1})).$$

- Therefore, as noted earlier, the bivariate probabilities can be calculated using copulas, which separate the marginal distribution of the error terms from their dependence structure. We may consider a rank correlation measure (e.g. Kendall Tau) as the measure of non-linear dependency and choose the Gaussian copula for evaluating the bivariate probabilities $P[\varepsilon_{t-1} < \cdot, \varepsilon_t < \cdot | \mathcal{I}_{t-1}]$.

- First let us forget about the Theorem mentioned earlier, where $s(\varepsilon_1), \dots, s(\varepsilon_n)$ are i.i.d conditional on \mathcal{I}_{t-1} . Let us explore this from an earlier point and what led to this Theorem in . We know from the mediangale assumption that

$$P[s(\varepsilon_t) = s_t \mid s(\varepsilon_{t-1}) = s_{t-1}, \dots, s(\varepsilon_1) = s_1] = P[\varepsilon_t \geq 0 \mid \varepsilon_{t-1}, \dots, \varepsilon_1]^{s(\varepsilon_t)} P[\varepsilon_t < 0 \mid \varepsilon_{t-1}, \dots, \varepsilon_1]^{1-s(\varepsilon_t)} = \frac{1}{2}$$

- As this corresponds only to the median, it is only true *iff* both sides of the inequality are zero. Since the probabilities $P[\varepsilon_t \geq 0 \mid \varepsilon_{t-1}, \dots, \varepsilon_1]$ and $P[\varepsilon_t < 0 \mid \varepsilon_{t-1}, \dots, \varepsilon_1]$ are the same and equal to $1/2$ for all $t = 1, \dots, n$. Thus, it can be argued that $s(\varepsilon_1), \dots, s(\varepsilon_n)$ which are defined by the aforementioned probabilities can be considered i.i.d, as their distribution is determined by the said probabilities, which are identical over time.

- And here is the proposition (pay careful attention to the last paragraph):

- Now let us instead assume that we are interested in the signs $\tilde{s}(\varepsilon + c_t)$, where c_t for $t = 1, \dots, n$ are some constants. Then the likelihood function in terms of the signs $\tilde{s}_1, \dots, \tilde{s}_n$

$$L(\tilde{s}_1, \dots, \tilde{s}_n \mid X) = \prod_{t=1}^n P[\varepsilon_t \geq -c_t \mid \varepsilon_1, \dots, \varepsilon_n]^{\tilde{s}_t} P[\varepsilon_t < -c_t \mid \varepsilon_1, \dots, \varepsilon_n]^{1-\tilde{s}_t}$$

- Now the joint p.m.f. depends on all the past signs and thus is no longer *i.i.d.*, as it violates the Mediangale assumption (particularly the point about permutations mentioned in Proposition 1 of). In other words, we no longer have that the p.m.fs $P[\varepsilon_t < -c_t \mid \varepsilon_{t-1}, \dots, \varepsilon_1]$ are constant and identical over time, so the distribution of the signs $\tilde{s}(\varepsilon_1 + c_1), \dots, \tilde{s}(\varepsilon_n + c_n)$, now depends on the joint p.m.fs.

Alternative proof

- Under the null hypothesis, it is evident that

$$P[\varepsilon_t \geq 0 \mid X] = P[\varepsilon_t < 0 \mid X], \quad t = 1, \dots, n$$

is equivalent to

$$P[y_t \geq 0 \mid X] = P[y_t < 0 \mid X], \quad t = 1, \dots, n$$

- Therefore, the mediantale assumption can be extended to y_1, \dots, y_n :

$$P[y_t \geq 0 \mid y_1, \dots, y_n, X] = P[y_t < 0 \mid y_1, \dots, y_n, X] = \frac{1}{2}$$

Alternative proof

- Then the variables $s(y_1), \dots, s(y_n)$ are i.i.d conditional on X according to the distribution

$$P[s(y_t) = 1 \mid X] = P[s(y_t) = 0 \mid X] = \frac{1}{2}, \quad t = 1, \dots, n$$

- Under the alternative hypothesis, however, the mediangale property does not hold for permutations $\pi : i \rightarrow j$ as noted in the Proposition 3.1 of , as the conditional distribution of the signs vary across observations.
- Therefore, the signs $s(u_1), \dots, s(y_n)$ can no longer be assumed to be i.i.d.