

# MODELAREA SI EVALUAREA IMPACTULUI INVESTITIILOR DIRECTE, NATIONALE SI INTERNATIONALE ASUPRA PIETEI MUNCII SI EVOLUTIEI MACROECONOMICE DIN ROMANIA<sup>1</sup>

-metode VAR și VEC-

Bianca Păuna<sup>2</sup>

## Abstract

*Metodele VAR și VEC sunt niște tehnici foarte răspândite datorită flexibilității dar și a avantajelor legate de absența necesității introducerii unor restricții arbitrare pentru identificarea sistemului, așa cum se întâmplă în cazul ecuațiilor simultane. În acest tip de metode, nu există variabile exogene, setul de variabile endogene sunt estimate prin lag-urile lor. Intervenția modelatorului se face în alegerea variabilelor care sunt introduse în model precum și a numărului de lag-uri.*

## Introducere

Metodologia vector autoregresiv (VAR) este o metodă foarte întâlnită în analiza seriilor de timp, în principal datorită flexibilității și a ușurinței de folosire. Ea poate fi considerată o generalizare atât a modelului univariat autoregresiv, pentru că variabilele dependente sunt lag-uri ale variabilelor explicative dar și a ecuațiilor simultane, pentru că se estimează simultan un sistem de ecuații. În plus, în fiecare ecuație se includ și laguri ale celorlalte variabile endogene. În general, în model nu sunt incluse variabile exogene, așa cum se procedează în modelare.

În modelarea ecuațiilor simultane unele variabile sunt considerate variabile endogene iar altele variabile exogene. O condiție pentru estimarea acestui tip de model este condiția ca modelul să fie identificat. Acest lucru se realizează prin includerea unor variabile exogene numai în anumite ecuații, o procedură care introduce subiectivismul în model, decizie care este criticată de Sims<sup>3</sup>.

Conform lui Sims, dacă un set de variabile sunt simultane, ar trebui tratate egal, nu ar trebui să se facă împărțiri apriori a variabilelor în endogene sau exogene. Alternativa popularizată de Sims este modelul VAR în care toate variabilele sunt tratate ca endogene, evitând astfel coruperea modelului cu restricții de identificare stabilite arbitrar. În acest tip de modele,

---

<sup>1</sup> Prezentată în cadrul programului: "Modelarea si evaluarea impactului investitiilor directe nationale si internationale asupra pietii muncii si evolutiilor macroeconomice din Romania", contract MEC 91-052/ 10 sept 2007, faza decembrie 2007: Etapa I: Documentare stiintifica, crearea bazei informationale, rafinare si analiza exploratorie a datelor.

<sup>2</sup> Institutul Național de Cercetări Economice, Academia Română, București

<sup>3</sup> C.A. Sims, "Macroeconomics and Reality" *Econometrica*, vol. 48, 1980.

intervenția modelatorului se face numai în stabilirea variabilelor care să intre în model, și a numărului de lag-uri care trebuie considerate.

Una din caracteristicile modelelor VAR este faptul că surprind structura dinamică a mai multor variabile simultan, iar funcțiile de impuls răspund surprind propagarea unui șoc al unei variabile dependente asupra sistemului.

Odată cu descoperirea importanței trendurilor stohastice ale variabilelor economice și dezvoltarea analizelor de cointegrare, au fost dezvoltate modele care să permită separarea relației pe termen lung de dinamica pe termen scurt. Relația pe termen lung numită și relația de cointegrare este asociată cu o dependență economică, în timp ce dinamica pe termen scurt reprezintă ajustarea modelului către relația pe termen lung. Modelele VEC (vector error correction – modelul de corectare a erorii) permit separarea componentelor pe termen lung de cele pe termen scurt din procesul de generarea a datelor.

## 1. Metodologia VAR

Presupunem că procesul care a generat seriile de timp  $\mathbf{Y}_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$  este o sumă dintre un trend deterministic și o parte pur stohastică:

$$\mathbf{Y}_t = \boldsymbol{\mu}_t + \mathbf{x}_t \tag{1.1}$$

Unde  $\boldsymbol{\mu}_t$  este partea deterministică și  $\mathbf{x}_t$  este un proces stohastic de medie zero. Componenta deterministică poate fi zero ( $\boldsymbol{\mu}_t = 0$ ) poate fi o constantă ( $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0$ ) sau poate avea o tendință liniară ( $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$ ). Componenta pur stohastică a procesului  $\mathbf{x}_t$  include tendințe stohastice sau relații de cointegrare. Se presupune că are media zero și o reprezentare VAR. Procesul observabil  $y_t$  preia proprietățile deterministice și stohastice de la  $\boldsymbol{\mu}_t$  și  $\mathbf{x}_t$ . Ordinul de integrare și cointegrare este dat de  $\mathbf{x}_t$ .

Pentru început se presupune că partea stohastică  $\mathbf{X}_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$  este generată de un proces VAR de ordinul  $p$ , (VAR( $p$ )) conform formulei:

$$\mathbf{Y}_t = \mathbf{C} + \mathbf{A}_1 \mathbf{Y}_{t-1} + \mathbf{A}_2 \mathbf{Y}_{t-2} + \dots + \mathbf{A}_p \mathbf{Y}_{t-p} + \mathbf{u}_t, \tag{1.2}$$

unde  $t = 1, \dots, T$ ; este lungimea seriei de timp;

$n$  este numărul de variabile endogene;

$\mathbf{C} = (c_1, c_2, \dots, c_n)'$  este un vector de constante;

$\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p$  sunt matricile coeficienților variabilelor lag și au dimensiunea  $(n \times n)$ ;

$\mathbf{u}_t$  are dimensiunea  $(n \times 1)$  și este un vector de erori necorelate și cu media zero (zgomot alb).

Un model bivariat VAR cu două lag-uri are următoarea formă:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} a_{11}^1 & a_{12}^1 \\ a_{21}^1 & a_{22}^1 \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} a_{11}^2 & a_{12}^2 \\ a_{21}^2 & a_{22}^2 \end{pmatrix} \begin{pmatrix} y_{1t-2} \\ y_{2t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad 1.3$$

sau scris sub formă de sistem de ecuații:

$$\begin{aligned} y_{1t} &= c_1 + a_{11}^1 y_{1t-1} + a_{12}^1 y_{2t-1} + a_{11}^2 y_{1t-2} + a_{12}^2 y_{2t-2} + \varepsilon_{1t} \\ y_{2t} &= c_2 + a_{21}^1 y_{1t-1} + a_{22}^1 y_{2t-1} + a_{21}^2 y_{1t-2} + a_{22}^2 y_{2t-2} + \varepsilon_{2t} \end{aligned} \quad 1.4$$

Se poate observa că fiecare ecuație conține aceleași variabile explicative, lag-urile variabilelor de explicat.

Procesul descris în ecuația (1.2) este stabil dacă  $\det(I_n - A_1 Z - \dots - A_p Z^p) \neq 0$  pentru oricare  $|z| \leq 1$ . Implicațiile acestei condiții de stabilitate este că toate rădăcinile determinantului polinomial sunt în exteriorul cercului unitate. În acest caz,  $x_t$  este integrat de ordinul zero  $I(0)$ . Dacă condiția de mai sus este respectată, procesul  $x_t$  este independent de timp, iar structura variație și a covariației este constantă.

Dacă determinantul polinomului are o rădăcină unitate  $z = 1$  și toate celelalte rădăcini în afara cercului unitate, atunci o parte din variabile sunt integrate și ar putea să existe și variabile cointegrate.

Cu ajutorul operatorului lag  $L^j y_t = y_{t-j}$ , ecuația (1.2) poate fi scrisă mai compact:

$$A(L)X_t = u_t \quad 1.5$$

Unde  $A(L) = I_n - A_1 L - \dots - A_p L^p$  este o matrice polinomială în operatorul lag de ordinul  $p$ . Dacă  $\mu_t = \mu_0 + \mu_{1t}$  și dacă expresia lui  $X_t$  se înmulțește cu  $A(L)$  după simplificări se ajunge la următoarea expresie:

$$Y_t = v_0 + v_1 t + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + u_t \quad 1.6$$

Dacă este folosit modelul din ecuația (1.6), iar valorile parametrilor  $v_0$  și  $v_1$  sunt diferite de zero, în cazul în care  $Y_t \sim I(1)$  trendul deterministic ar putea să fie pătratic. Deci modelul este restrictiv impunând restricții asupra parametrilor deterministici. În general, parametrizarea trendul deterministic ar trebui să fie făcută la începutul analizei prin alegerea ordinului adecvat al polinomului deterministic. O altă practică este și extragerea trendului deterministic înaintea analizării procesului stohastic, deoarece acesta este în principal partea care se dorește modelată în analiza econometrică.

## 2. Metodologia VEC

Dacă variabilele  $Y_t \sim I(1)$  iar anumite variabile sunt cointegrate atunci modelul (1.6) nu este cea mai bună reprezentare a procesului, deoarece nu conține relația de cointegrare în mod explicit. În acest caz, modelul poate fi re-parametrizat prin scăderea lui  $X_{t-1}$  din ambele părți ale ecuației și se obține ecuația:

$$\Delta X_t = \Pi X_{t-1} + \Gamma_1 \Delta X_{t-1} + \dots + \Gamma_{p-1} \Delta X_{t-p+1} + u_t \quad 2.1$$

unde  $\Pi = -(I_n - A_1 - \dots - A_p)$  și  $\Gamma_j = -(A_{j+1} + \dots + A_p)$  pentru  $j=1, \dots, p-1$ . Această formă funcțională este cunoscută sub numele de forma de model de corecție a erorilor (VEC) a VAR-ului de ordin  $p$ .

Se poate observa că  $\Delta X_t$  nu conține trend stohastic pentru că s-a presupus că  $X_t \sim I(1)$ . Deci  $\Pi X_{t-1}$  este singurul termen care include variabile  $I(1)$  și în consecință  $\Pi X_{t-1}$  trebuie să fie  $I(0)$ . Pentru îndeplinirea condiției, trebuie să existe o relație de cointegrare. Termenul  $\Pi X_{t-1}$  este numit termenul de corecție a erorii. De asemenea, mișcările pe termen scurt ale variabilelor sunt determinate de  $\Gamma_j$  ( $j = 1, \dots, p-1$ ) care sunt numiți parametri pe termen scurt.

Dacă  $y_t \sim I(1)$  și polinomul  $\det(I_n - A_1 Z - \dots - A_p Z^p)$  are o rădăcină unitară, adică  $\det(I_n - A_1 Z - \dots - A_p Z^p) = 0$  pentru  $z = 1$ , matricea  $\Pi$  este singulară. Dacă rangul matricei  $\Pi$  este  $r$  cu  $0 < r < n$ , atunci din teoria matriceală se știe că există matricile  $\alpha$  și  $\beta$  de dimensiune  $(n \times r)$  cu rangul  $r$  astfel încât  $\Pi = \alpha \beta'$ . Înmulțind expresia  $\Pi X_{t-1}$  cu  $(\alpha' \alpha)^{-1} \alpha'$  se demonstrează că  $\beta' X_{t-1}$  este  $I(0)$ , deci că există  $r$  relații de cointegrare între componentele lui  $X_t$ . Rangul matricei  $\Pi$  este numit rangul de cointegrare al procesului.

Se observă că matricile  $\alpha$  și  $\beta$  nu sunt unice. Alegând orice matrice  $P$  nesingulară de dimensiune  $(r \times r)$  se obține o altă descompunere  $\Pi = \alpha^* \beta^{*'} cu  $\alpha^* = \alpha P^{-1}$  și  $\beta^{*'} = \beta P'$ . De cele mai multe ori, pentru identificarea relațiilor de cointegrare (pentru a se obține relații de cointegrare unice) se recurge la teoria economică.$

Relațiile de cointegrare reprezintă relații pe termen lung sau relații de echilibru. Această înțelegere poate ajuta în specificarea restricțiilor de identificare. De exemplu, dacă există numai o relație de cointegrare, este suficient să se normalizeze un coeficient la 1 și se obține identificarea parametrilor ecuației de cointegrare.

### 3. Estimarea ecuațiilor VAR

Se consideră modelul VAR (p) discutat anterior:

$$Y_t = C + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + u_t$$

Cea mai simplă metodă de estimare a ecuațiilor este metoda celor mai mici pătrate (ordinary least squares – OLS). Deoarece variabilele explicative sunt numai lag-uri ale variabilelor endogene, simultaneitatea nu este o problemă pentru acest tip de modele, și estimatorul OLS este consistent. A fost demonstrat în literatură că acest estimator este identic cu estimatorul metodei celei mai mici pătrate generalizată (generalized least squares – GLS), deoarece ecuațiile au aceleași variabile explicative. Atunci când procesul este distribuit normal acest estimator este identic cu estimatorul verosimilității maxime. În concluzie nu se pierde din eficiența estimării prin folosirea OLS-ului.

În cazul în care procesul este stabil ( $I(0)$ ) estimatorul metoda celor mai mici pătrate produce estimatori care tind asimptotic către distribuția normală. Matricea variație-covariație poate fi estimată consistent, și testele statistice sunt asimptotic corecte, deci pot fi folosite pentru testarea ipotezelor statistice.

În cazul în care  $Y_t \sim I(1)$  deci procesul nu este stabil și pot exista relații de cointegrare între variabile, rezultatul că estimatorul OLS tinde asimptotic către distribuția normală este încă respectat. În acest caz, matricea variație covariație este singulară, pentru că există relații liniare între parametri estimați. În această situație statisticele  $t$ ,  $\chi$  și  $F$  privind parametri VAR nu mai sunt valabile asimptotic. Chiar în aceste condiții, sunt cazuri când inferențele rămân valabile, și anume cazul în care testele statistice se referă numai la coeficienții uneia din matrici și nu pun condiții asupra coeficienților mai multor matrici  $A$ .

Pentru estimarea unui model VAR trebuie parcurse mai multe etape. Se începe cu specificarea modelului, prin alegerea variabilelor dependente care vor fi modelate. Acestea pot fi, din puncte de vedere teoretic, oricât de multe dar în practică suntem limitați de numărul de observații disponibile. Trebuie remarcat că în cazul aplicării unui VAR(p) la  $n$  variabile dependente, numărul de coeficienți conținuți în model este  $n^2 \times p$ . Dependența numărului de coeficienți de numărul de variabile endogene nu este liniară, ci pătratică, și de aceea din punct de vedere al gradelor de libertate această alegerea a numărului de variabilelor dependente este foarte importantă.

Ca exemplificare se consideră 3 lag-uri și 2 variabile endogene, acest model va necesita 12 coeficienți de estimat, câte 6 în fiecare ecuație, dacă însă se dorește includerea a 3 variabile exogene cu 3 lag-uri, fiecare ecuație va avea acum 9 coeficienți de estimat, fără considerarea trendului, în total 27. Din acest motiv, numărul de variabile endogene care pot fi modelate simultan este limitat de numărul de observații disponibile.

Un al doilea pas în specificarea modelului este alegerea numărului de lag-uri. Cel mai simplu, alegerea numărului de lag-uri se face cu ajutorul criteriilor informaționale. Criteriile informaționale cuantifică practic partea din variabila endogenă care nu este explicată de model. Cele mai răspândite criterii informaționale sunt Akaike și Schwartz.

Cele două formule de calcul corespunzătoare celor două criterii sunt următoarele:

$$AIC = \left( \frac{VTR}{n} \right) \cdot e^{\frac{2(k+1)}{n}} = \frac{1}{n} \left( \sum_{t=1}^n u_t^2 \right) \cdot e^{\frac{2(k+1)}{n}}$$

$$SCHWARTZ = \left( \frac{VTR}{n} \right) \cdot n^{\frac{k+1}{n}} = \frac{1}{n} \left( \sum_{t=1}^n u_t^2 \right) \cdot n^{\frac{k+1}{n}}$$

Între cele două criterii, Akaike sugerează un număr de laguri mai mare de cât Schwartz. Valorile celor două criterii nu au nici o relevanță de sine, dar pentru decizia privind numărului de lag-uri, se estimează diverse modele de lag-uri diferite, iar pentru fiecare model se calculează cele două criterii informaționale. Alegerea unuia sau a altuia se face comparând criteriile informaționale și alegându-l pe cel care are valoarea cea mai mică. Ca regulă, atunci când cele două criterii dau rezultate contradictorii se recomandă selectarea modelului care dă rezultate mai bune la criteriul Schwartz.

#### 4. Estimarea ecuațiilor VEC

Considerăm un model VE fără termen deterministic:

$$\Delta Y_t = \alpha \beta' Y_{t-1} + \Gamma \Delta Y_{t-1} + \dots + \Gamma_{p-1} \Delta Y_{t-p+1} + u_t \quad 4.1$$

Modelul poate fi rescris în forma matriceală compactă astfel:

$$\Delta Y = \alpha \beta' Y_{-1} + \Gamma \Delta X + U \quad 4.2$$

Pentru rezolvarea modelelor de tip VEC se folosește metoda verosimilității maxime și metoda celor mai mici pătrate generalizate, datorită termenului de corecție erorii care este de forma:  $\alpha \beta' Y_{-1}$ , deci metoda uzuală de minimizare a pătratului erorilor nu mai poate fi folosită în mod direct. Dacă, produsul  $\alpha \beta'$  ar fi cunoscut, metoda celor mai mici pătrate ar putea fi folosită pentru estimarea parametrilor. În acest caz, estimatorul metodei celei mai mici pătrate pentru  $\Gamma$  are forma următoare:

$$\Gamma(\alpha \beta') = (\Delta Y - \alpha \beta' Y_{-1}) \Delta X' (\Delta X \Delta X')^{-1} \quad 4.3$$

Pentru obținerea estimatorilor coeficienților modelului VEC se recurge la estimarea în primă fază a lui  $\alpha$  și  $\beta$  după care aceștia sunt introduși în formula 4.3 și se obțin estimatori pentru coeficienții din matricea  $\Gamma$ . Coeficienții  $\alpha$  și  $\beta$  sunt obținuți aplicând metoda verosimilității maxime sau metoda celor mai mici pătrate generalizate.

Selecția numărului de lag-uri care trebuie incluse în modelarea VEC se face similar ca la VAR, prin încercări. Se estimează diverse modele și este ales cel care produce criteriul informațional cel mai mic.

Problemele pe care le ridică estimarea unui model VEC în plus față de VAR sunt legate de alegerea rangului de cointegrare. Deoarece numărul parametrilor în aceste modele crește cu pătratul numărului de variabile, este de dorit de cele mai multe ori să se impună restricții de egalitate cu zero asupra matricei parametrilor, cea ce duce la eliminarea unora din variabilele lag din o parte din sistemul de ecuații.

Există mai multe metode de determinare a numărului de ecuații de cointegrare. Metodele se rezumă la testarea mai multor ipoteze de genul:

$H_0(r_0)$ : rangul ( $\Pi$ ) =  $r_0$  în timp ce ipoteza alternativă este  $H_a(r_0)$ : rangul( $\Pi$ ) >  $r_0$  unde  $r_0 = 0, \dots, n - 1$ .

Primul rang de cointegrarea din ipoteza nulă care nu poate fi respins este ales ca rangul de cointegrarea. Dacă  $H_0(0)$  adică prima ipoteză nulă nu poate fi respinsă, atunci procesul VAR este aplicat la diferențele de ordinul întâi. Dacă toate ipotezele nule sunt respinse inclusiv  $H_0(n-1)$  procesul este tratat ca un  $I(0)$  și se alege un proces VAR fără diferențierea variabilelor exogene.

Statistica care se aplică pentru testarea ipotezelor nule de mai sus este LR care are distribuție ne-standard. Această distribuție depinde de diferența  $K - r_0$  și de termenul determinist dar nu depinde de dinamica pe termen scurt. Valorile critice construite pentru diferite tipuri de trenduri deterministice au fost calculate și sunt disponibile.

## **5. Aplicarea metodologiei VAR/VEC la modelarea impactului investițiilor asupra pieței muncii și asupra variabilelor macroeconomice**

Metodologia VAR și VEC va fi aplicată pentru testarea impactului investițiilor în primul rând asupra pieței muncii, dar și asupra variabilelor macroeconomice.

### *Impactul asupra pieței muncii*

Variabilele avute în vedere sunt, bineînțeles, investițiile separat după tipul lor, investiții interne și investițiile străine directe. Variabilele de pe piața muncii avute în vedere sunt ocuparea sau alternativ șomajul, un indice al inegalității salariale, pentru a identifica dacă investițiile contribuie în mod diferențiat la creșterea salariilor în funcție de calificare forței de muncă. Modelul va fi estimat separat pentru investițiile interne și străine pentru a compara implicațiile celor două asupra pieței muncii. Cum o mare parte din variabilele trimestriale au devenit disponibile relativ recent, seria de care dispunem pentru analiză nu este foarte lungă, și va restricționa numărul de variabile care pot fi introduse în model. În măsura în care dispunem de serii de date lunare, modelul va fi estimat și pe date lunare.

Pentru stabilirea numărului de lag-uri se vor folosi criteriile informaționale, și se va testa staționaritatea seriilor. Se va testa și prezența cointegrării și în cazul în care se găsesc și relații de cointegrare se va aplica modelul VEC. Se lucrează cu serii de-sezonalizate.

### *Impactul asupra variabilelor macroeconomice*

Variabile macroeconomice care vor fi incluse sunt creșterea economică, importul, exportul și - depinzând de lungimea seriei de date - pot fi incluse și alte variabile, împreună cu investițiile separate pe investiții interne și investiții străine directe. Bineînțeles, se vor estima separat două modele pentru investițiile domestice și străine directe, și în felul ăsta se va putea identifica în ce măsură diferă impactul acestora asupra variabilelor macroeconomice.

Staționaritatea seriilor va fi testată, și vor fi diferențiate corespunzător pentru a obține staționaritatea. Se vor folosi criterii informaționale pentru a decide numărul de lag-uri care vor fi incluse în modelul VAR. În cazul existenței unor relații de cointegrare seriile vor fi modelate conform metodologiei VEC. Se lucrează cu serii de-sezonalizate.

## **Bibliografie**

**Clements, M. P., Hendry, D. F.** (1998) *Forecasting Economic Time Series*, Cambridge University Press, Cambridge.

**Eviews** (2005) *Eviews 5.1 User's Guide*, Quantitative Micro Software, Irvine, CA.

**Harvey** (1990) *The Econometric Analysis of Time Series*, ediția a doua, editor Philip Allan.

**Maddala, G.S.**(2000) *Unit Roots, Cointegration and Structural Change*, Cambridge University Press.

**Sims, C.A.**(1980): "Macroeconomics and Reality" *Econometrica*, vol. 48, 1980