



**Banca Națională a României**

# **Modelarea incertitudinii asociate datelor statistice în contextul revizuirilor istorice**

**Valeriu Nalban**

Direcția Modelare și Prognoze Macroeconomice

# Objective

- Analiza **tiparului revizuirilor** datelor statistice având ca exemplu seria PIB real ajustată sezonier.
- Evaluarea caracterului revizuirilor: generate de "**zgomot**" sau "**informație nouă**" (engl. *noise* vs. *news*).
- Filtrarea seriei PIB și estimarea **incertitudinii asociate revizuirilor viitoare** prin intermediul distribuției de probabilitate.
- Construirea unui **grafic de tip evantai** (engl. *fan chart*) pentru seria istorică PIB real (asemănător cu cel din *Bank of England Inflation Report*).

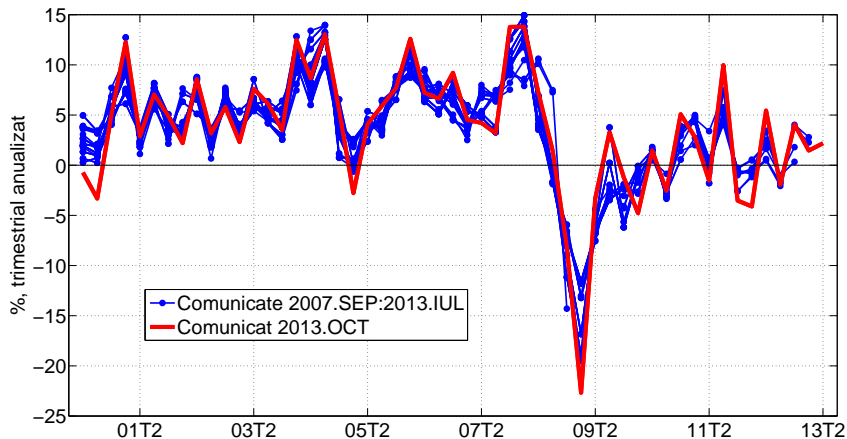
# Cauze ale revizuirilor

- înlocuirea estimărilor inițiale derivate din sondaje incomplete (ex.: respondenții întârziați), judecăți expert și ajustări statistice cu informații din surse cu un grad crescut de încredere, complete și actualizate;
- corectarea erorilor de sondaj și de procesare;
- actualizarea factorilor sezonieri;
- modificarea anului de bază → nu este de natură să producă revizuiți ale ratelor de creștere;
- modificarea metodologiei statistice, a conceptelor, definițiilor și clasificărilor, etc.

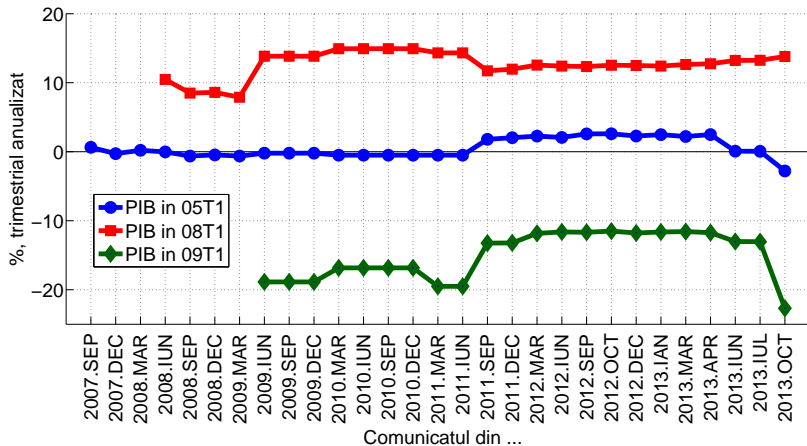
# Seriile PIB disponibile în timp real

$y_t^j$ : seria de rate de creștere trimestrială anualizată a PIB real ajustat sezonier\* aferentă comunicatelor INS din perioada  $j$ ,  $j = \{\text{Septembrie 2007, ..., Octombrie 2013}\}$  (29 serii).

\*ajustate sezonier intern (BNR) până în 2009.

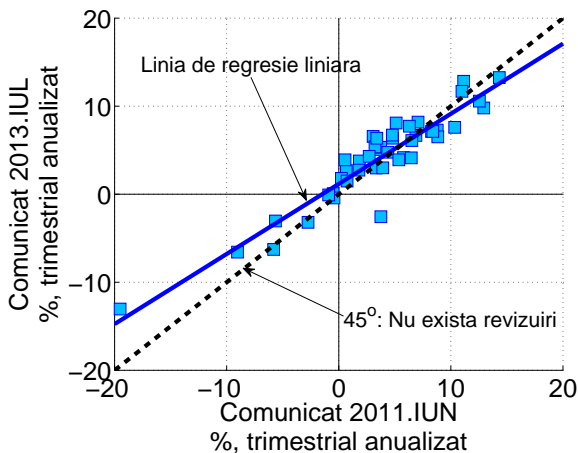


# Tiparul revizuirilor (1)



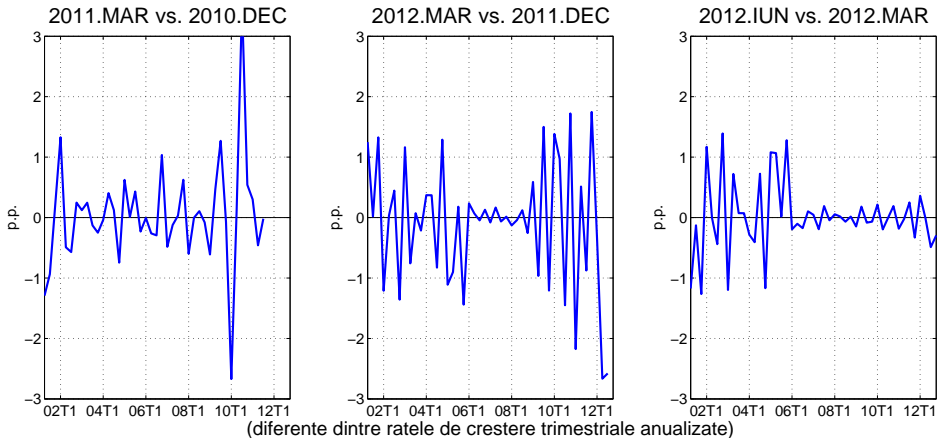
- Magnitudinea revizuirilor ratelor de creștere este ridicată atât în valoare absolută, cât și raportată la nivelul seriei.
- Revizuri semnificative au loc chiar și după o perioadă îndelungată de timp.

## Tiparul revizuirilor (2)



- Direcția revizuirilor (pentru orizontul de 8 trimestre) indică o tendință de netezire a ratelor de creștere trimestrială, prin reducerea în valoare absolută a ratelor de creștere publicate inițial.

# Tiparul revizuirilor (3)



- Natura revizuirilor nu corespunde întotdeauna ipotezei cu privire la magnitudinea superioară a acestora pentru perioadele mai recente.

# "zgomot" vs. "informație nouă"

Conform Mankiw și Shapiro (1986), există două ipoteze privind revizuirile:

- **"zgomot" / noise**: fiecare comunicat reprezintă o serie care conține **erori de măsură**, deci datele publicate sunt caracterizate de **ineficiență** și se recomandă tehnici de **filtrare a seriei publicate** (filtrul Kalman); implică  $\text{corr}(y_t^j - y_t^{j-1}, y_t^j) = 0$  și  $\text{corr}(y_t^j - y_t^{j-1}, y_t^{j-1}) \neq 0$
- **"informație nouă" / news**: fiecare comunicat reprezintă o **prognoză eficientă** a comunicatului următor, încorporând într-un **mod optim** toate informațiile disponibile la momentul publicării; implică  $\text{corr}(y_t^j - y_t^{j-1}, y_t^j) \neq 0$  și  $\text{corr}(y_t^j - y_t^{j-1}, y_t^{j-1}) = 0$

■ Media coeficienților de corelație calculați pentru datele PIB real:  
 $\text{corr}(y_t^j - y_t^{j-1}, y_t^{j-1}) = -0.06$  și  $\text{corr}(y_t^j - y_t^{j-1}, y_t^j) = 0.11$ .

■ Coeficienții de corelație calculați sunt apropiați de zero (și ne semnificativi statistic) → **nu se poate formula o concluzie privind ipoteza dominantă**.



# Modelul *state-space*

**Ecuția de măsură:**

$$\underbrace{y_t}_{\text{PIB publicat}} = \underbrace{c^j}_{\text{eroare sistematică}} + \underbrace{\bar{y}_t}_{\text{PIB filtrat}} + \underbrace{\nu_t}_{\text{eroare de măsură}}$$

**Ecuția de tranziție:**

$$\begin{cases} \bar{y}_t = \mu + \alpha \bar{y}_{t-1} + \varepsilon_t \\ \nu_t = \beta \nu_{t-1} + \eta_t \end{cases} \quad \Sigma = \begin{bmatrix} \sigma_\varepsilon^2 & \rho \sigma_\varepsilon \sigma_{\eta^j} \\ \rho \sigma_\varepsilon \sigma_{\eta^j} & \sigma_{\eta^j}^2 \end{bmatrix}$$

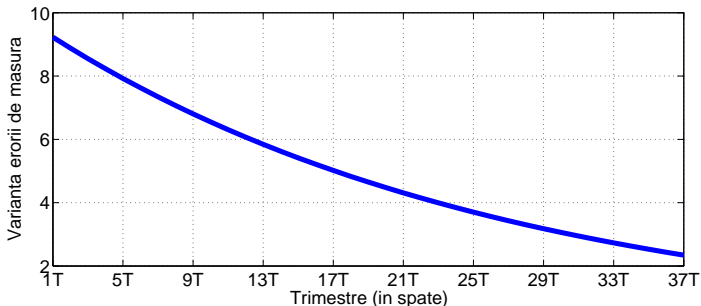
Eroarea sistematică și varianța erorii de măsură descresc în funcție de “maturitatea” observației:

$$\begin{cases} c^j = c^1 (1 + \lambda)^{j-1} \\ \sigma_{\eta^j}^2 = \sigma_{\eta^1}^2 (1 + \delta)^{j-1} \end{cases} \quad \left\{ \begin{array}{l} \text{Estimări realizate exogen, utilizând seturile} \\ \text{de date PIB real disponibile (29 serii)} \end{array} \right.$$

Sursa: Cunningham et. al (2007), Anderson și Gascon (2009)

# Proprietățile erorii de măsură

- Estimarea **abaterii standard variabile** pentru eroarea de măsură necesită o analiză detaliată a tiparului revizuirilor.
- Se consideră un orizont al revizuirilor de 8 trimestre. Conform metodologiei INS, datele definitive pentru conturile naționale se publică aproximativ în 21 luni de la comunicatul inițial.
- **Revizuirile** sunt definite ca diferență dintre seria de la un anumit moment de timp și seria disponibilă cu 8 trimestre în urmă:  $\phi_t = y_t^j - y_t^{j-8}$ .
- Variația erorii de măsură este **mai mare pentru observațiile mai recente**, acestora fiindu-le asociate probabilități superioare de a înregistra revizuriri.



# Rezultatele estimărilor

## I. Optimizări numerice exogene modelului *state-space*:

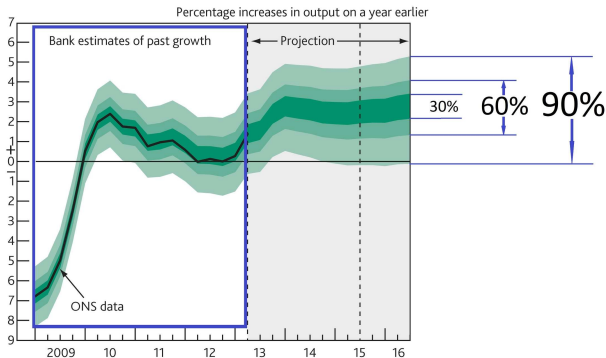
- Coeficienții aferenți **erorii sistematice** ( $c^j$ ) nu sunt semnificativi statistic, deci **revizuirile au medie zero**.
- **Varianța inițială a erorii de măsură** ( $\sigma_{\eta_1}^2$ ) este semnificativă statistic și mai mare decât în literatura de specialitate, indicând **revizuirile de mai mare amploare în cazul seriei PIB real al României**.

## II. Estimarea *Maximum Likelihood* (ML) a modelului *state-space*:

- Coeficientul de corelație dintre eroarea de măsură și PIB filtrat ( $\rho$ ) este ne semnificativ diferit de zero, implicând **ipoteza noise**.
- Estimarea ML ajustată cu distribuții *a priori* (*regularized ML*) confirmă valorile rezultate în urma estimării ML.
- Pentru estimarea seriilor neobservabile (PIB filtrat  $\bar{y}_t$  și eroarea de măsură  $\nu_t$ ) a fost utilizat **filtrul Kalman**.

# Graficul de tip evantai (engl. *fan chart*)

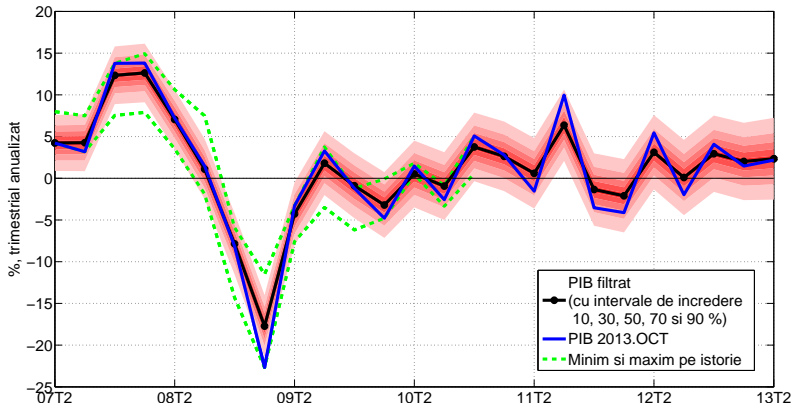
- Oferă o **representare probabilistică a incertitudinii** atât pe istorie, cât și pe intervalul de prognoză.
- Asimetria intervalelor de încredere indică **balanța riscurilor** asociată tendinței centrale.
- Pentru perioada istorică, intervalele de încredere indică **probabilitatea și magnitudinea revizuirilor potențiale**.



Sursa: *Bank of England Inflation Report, August 2013.*

# PIB filtrat și intervalele de încredere

- Intervalul de încredere de 90% pentru cele mai recente observații are o lățime aproximativ egală cu cea aferentă canalului calculat ca diferență între valorile maxime și minime istorice.
- PIB filtrat reprezintă o **variantă mai netedă a seriei PIB real**.
- Simulările Monte Carlo confirmă **validitatea ipotezei privind distribuția normală a erorilor de măsură**.



# Analiza robusteții modelului

## I. Lungimea orizontului revizuirilor:

- Alegerea orizontului revizuirilor este **arbitrară** și necesită examinarea robusteții modelului la variația acestui parametru, conform Anderson și Gascon (2009).
- Reestimarea și filtrarea modelului pentru lungimi ale orizontului revizuirilor de **6 și 10 trimestre** a rezultat în **dinamici similare** ale variabilelor neobservabile.

## II. Model multivariat:

- Includerea în model a unei variabile **corelate cu dinamica PIB**, care să nu fie revizuită și să fie disponibilă înainte de comunicatele oficiale privind PIB.
- Literatura de specialitate sugerează utilizarea **datelor din sondaje**.
- Reestimarea modelului *state-space* suplimentat cu o ecuație de măsură care leagă PIB filtrat de **indicatorul de încredere ESI** (DG-ECFIN).
- Rezultatele indică **o distribuție de probabilitate a revizuirilor potențială similară** celei estimate în cadrul modelului univariat.

# Concluzii

- Pentru o evaluare optimă a incertitudinii este necesară **estimarea distribuției de probabilitate a revizuirilor potențiale**.
- Tiparul revizuirilor operate nu se încadrează întotdeauna în ipotezele modelului.
- Revizuirile sunt compatibile cu **ipoteza *noise***.
- Nu a fost semnalată existența erorilor sistematice de subevaluare/supraevaluare, însă față de studiile de referință **eroarea de măsură tinde să fie mai mare**.
- Seria PIB filtrat reprezintă **o variantă mai netedă a seriei PIB real publicată**.
- Intervalul de încredere de 90% asociat revizuirilor viitoare are o lățime aproximativ egală cu diferența dintre valorile maxime și minime istorice.
- Evaluarea incertitudinii prin estimarea distribuției de probabilitate a revizuirilor potențiale este vizualizată printr-un **grafic evantai**, care oferă o **descriere detaliată a posibilelor revizurii viitoare, cu probabilitățile atașate**.

# Bibliografie selectivă

- Anderson, R., și Gascon, C. (2009). *Estimating U.S. Output Growth with Vintage Data in a State-Space Framework*. Federal Reserve Bank of St. Louis Review, 91(4), pp. 349-69.
- Cunningham, A., Eklund, J., Jeffery, C., Kapetanios, G., și Labhard, V. (2007). *A state space approach to extracting signal from uncertain data*. Bank of England WP No. 336.
- Cunningham, A., și Jeffery, C. (2007). *Extracting better signal from uncertain data*. Bank of England Quarterly Bulletin 2007 Q3.
- Frale, C., și Raponi, V. (2012). *Revisions in official data and forecasting*. Ministry of the Economy and Finance – Italy, WP No. 3.
- Jacobs, J., și van Norden, S. (2011). *Modeling data revisions: Measurement error and dynamics of “true” values*. Journal of Econometrics, 161(2), pp. 101-109.
- Mankiw, G., și Shapiro, M. (1986). *News or Noise: An Analysis of GNP Revisions*. Survey of Current Business, 66, pp. 20-25.