



BANCA NAȚIONALĂ A ROMÂNIEI

CAIETE DE STUDII
Nr. 33

CAIETE DE STUDII

Nr. 33

Octombrie 2014

N o t ă

*Opiniile prezentate în această lucrare aparțin în întregime autorilor
și nu implică sau angajează în vreun fel Banca Națională a României.*

Toate drepturile rezervate.

*Reproducerea informațiilor în scopuri educative și necomerciale
este permisă numai cu indicarea sursei.*

ISSN 1584-0883 (versiune *online*)

ISSN 1584-0883 (versiune e-Pub)

**MODELAREA INCERTITUDINII
ASOCIATE DATELOR STATISTICE
ÎN CONTEXTUL REVIZUIRILOR ISTORICE**

Valeriu Nalban*

* Banca Națională a României, Direcția Modelare și Prognoze Macroeconomice

Cuprins

Rezumat	7
1. Introducere	9
2. Tiparul revizuirilor seriilor PIB real	10
3. Filtrarea seriei PIB real ajustate sezonier.....	14
4. Analiza robusteții modelului	20
4.1. Senzitivitatea modelului la lungimea orizontului revizuirilor	20
4.2. Includerea unei variabile observabile suplimentare.....	21
5. Concluzii	22
Bibliografie	24

Rezumat

Datele statistice oficiale privind dinamica seriilor de timp din conturile naționale trimestriale sunt supuse revizuirilor în comunicatele ulterioare, deoarece acestea reprezintă estimări și nu măsurări exacte, fapt care indică incertitudinea asociată variabilelor macroeconomice. Revizuirile sunt cauzate de apariția unor noi informații (privind evoluțiile unor variabile cu putere explicativă), dar și de îmbunătățirea metodologiilor utilizate de autoritatea responsabilă cu statistica oficială. Revizuirile succesive ar trebui să aducă seriile estimate către valorile „adevărate”, însă probabilitatea ridicată ca estimările recente să fie revizuite induce obstacole și provocări în procesul de analiză și prognoză economică. În studiul de față este cuantificată și analizată incertitudinea asociată ultimei versiuni a seriei PIB real ajustate sezonier (aferentă comunicatului Institutului Național de Statistică din octombrie 2013), adoptând ipoteza că tiparul revizuirilor trecute reprezintă o aproximare optimă a incertitudinii prezente, rezultatul estimării fiind distribuția de probabilitate a potențialelor revizuri viitoare. Reprezentarea incertitudinii estimate prin intermediul unor grafice de tip evantai și construirea intervalelor de incertitudine pentru valorile istorice relevă natura probabilistică asociată variabilei analizate.

Cuvinte-cheie: incertitudinea datelor, revizuirea datelor, modele *state-space*

Coduri de clasificare JEL: C10, C81, E01

1. Introducere

O mare parte a datelor statistice, în particular majoritatea variabilelor macroeconomice (cum ar fi cele din conturile naționale) nu sunt direct observate sau măsurate exact, ci sunt estimări, fiind astfel caracterizate de un anumit grad de incertitudine. Un indiciu relevant pentru această incertitudine este reprezentat de revizuirile aferente datelor statistice comunicate anterior. În general, autoritățile responsabile cu statistica oficială trebuie să facă un compromis între acuratețea datelor comunicate și raportarea cât mai operativă a acestora. Odată cu trecerea timpului, sunt recepționate noi informații și sunt implementate tehnici mai performante de estimare. În aceste condiții, revizuirile devin inevitabile, acestea trebuind să imprime o dinamică convergentă către valorile „adevărate” neobservabile, care caracterizează cel mai bine conceptul economic pe care îl reprezintă variabila respectivă (cum ar fi produsul intern brut). Susceptibilitatea datelor statistice de a fi revizuite în viitor complică procesul de formulare și luare a deciziilor economice de către autorități, inclusiv banca centrală, și agenții privați, ale căror acțiuni optime sunt condiționate de cunoașterea cu exactitate a condițiilor economice trecute și curente și prin urmare implică provocări suplimentare pentru procesul de analiză și prognoză economică.

Literatura de specialitate identifică multiple cauze ale revizuirilor succesive ale seriilor statistice¹. Printre aceste, conform Jacobs și van Norden (2011), cele mai importante sunt: înlocuirea estimărilor inițiale derivate din sondaje incomplete (de exemplu 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; modificarea metodologiei statistice, a conceptelor, definițiilor și clasificărilor (acestea sunt mai greu de anticipat și nu sunt considerate a fi indicative pentru revizuirile viitoare) etc.

Analiza prezintă modelează seria filtrată² aferentă PIB real – cea care aproximează într-un mod fundamental activitatea economică³ și care în urma unui anumit număr de revizuiți succesive nu ar trebui să sufere modificări ulterioare – ca variabilă neobservabilă în cadrul unui model reprezentat în spațiul stărilor (engl. *state-space*). Utilizând filtrul Kalman pentru estimarea variabilei neobservabile, se poate calcula și eroarea standard asociată, care a fost utilizată pentru construirea distribuției de probabilitate a revizuirilor potențiale. Jacobs și van Norden (2011) susțin utilitatea seriei filtrate în cadrul estimării PIB potențial, deoarece seriile PIB real publicate succesiv de autoritățile pentru statistică prezintă o incertitudine ridicată. De asemenea, în același studiu se susține că *gap*-ul PIB (deviația PIB real de la nivelul său potențial) asociat seriei PIB filtrate explică mai bine evoluția prognozată a ratei inflației decât *gap*-ul PIB asociat seriei oficiale. Cunningham et al. (2007) sugerează chiar posibilitatea de a substitui seria oficială cu cea filtrată pentru analize și prognoze economice cu caracter recurent.

¹ Pentru fiecare comunicat (engl. *release*) sursele de date sunt reprezentate preponderent de sondaje, care inevitabil conțin zgomot asociat erorilor de eșantionare.

² În literatura de specialitate conceptul respectiv este utilizat sub sintagma *true data*.

³ În general se consideră că activitatea economică este o variabilă neobservabilă, PIB fiind cea mai bună aproximare (engl. *proxy*) pentru aceasta.

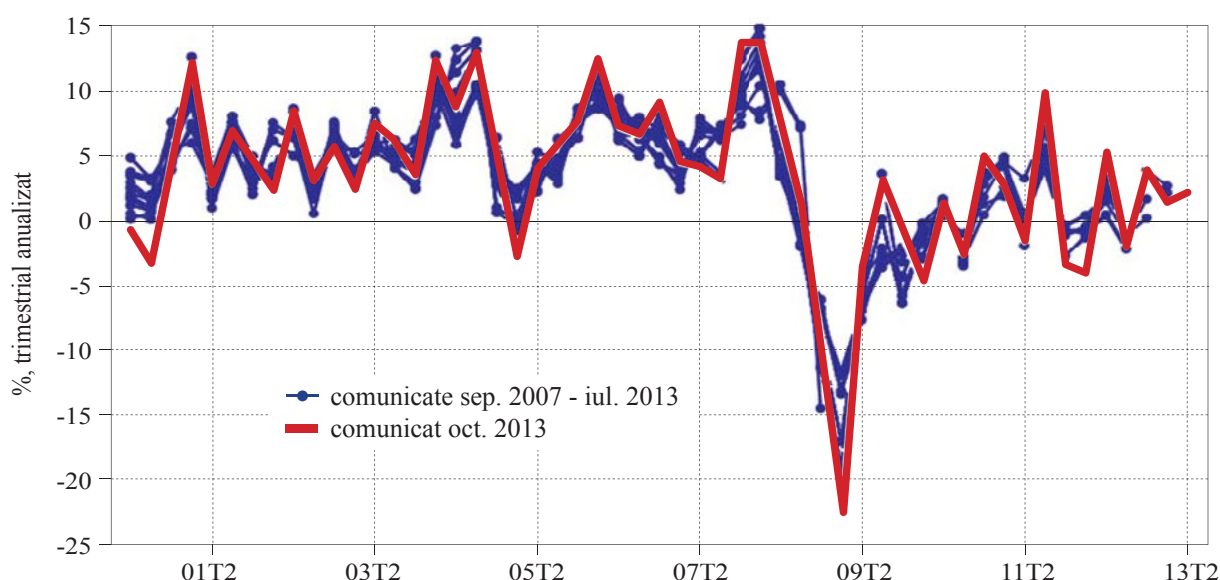
În continuare, secțiunea 2 analizează tiparul revizuirilor succesive operate asupra seriilor PIB real (exprimate în rate de creștere trimestriale anualizate). Descrierea și estimarea modelului *state-space*, împreună cu vizualizarea seriei PIB filtrate și a intervalelor de încredere asociate, sunt prezentate în secțiunea 3. Analiza sensibilității modelului la variația unor parametri și specificații este descrisă în secțiunea 4, în timp ce concluziile sunt prezentate în secțiunea 5.

2. Tiparul revizuirilor seriilor PIB real

În cadrul literaturii de specialitate, datele primare utilizate în analize de acest tip sunt reprezentate sub forma ratelor de creștere trimestriale anualizate sau anuale ale seriilor PIB real ajustate sezonier disponibile la diferite momente de timp⁴. Pentru exemplificarea metodologiei econometrice, în cadrul analizei au fost colectate seriile de PIB real ajustate sezonier publicate în perioada septembrie 2007 – octombrie 2013 (un total de 29 serii) și au fost construite seriile de rate de creștere trimestriale anualizate asociate (Figura 1). De asemenea, este important a se menționa faptul că seriile PIB real au fost ajustate sezonier în cadrul BNR până în 2009, deoarece INS nu publica și serii ajustate sezonier în acea perioadă.

Conform Figurii 1, evoluția PIB a cunoscut revizuiți semnificative și variate ca direcție și magnitudine. De asemenea, au existat cazuri în care a fost schimbat și semnul ratei de creștere trimestriale anualizate: de exemplu, comunicatele inițiale plasau creșterea economică pentru 2009 T3 între -3 la sută și -2 la sută, însă cel mai recent comunicat (octombrie 2013) indică o creștere trimestrială de circa 4 la sută pentru acea perioadă.

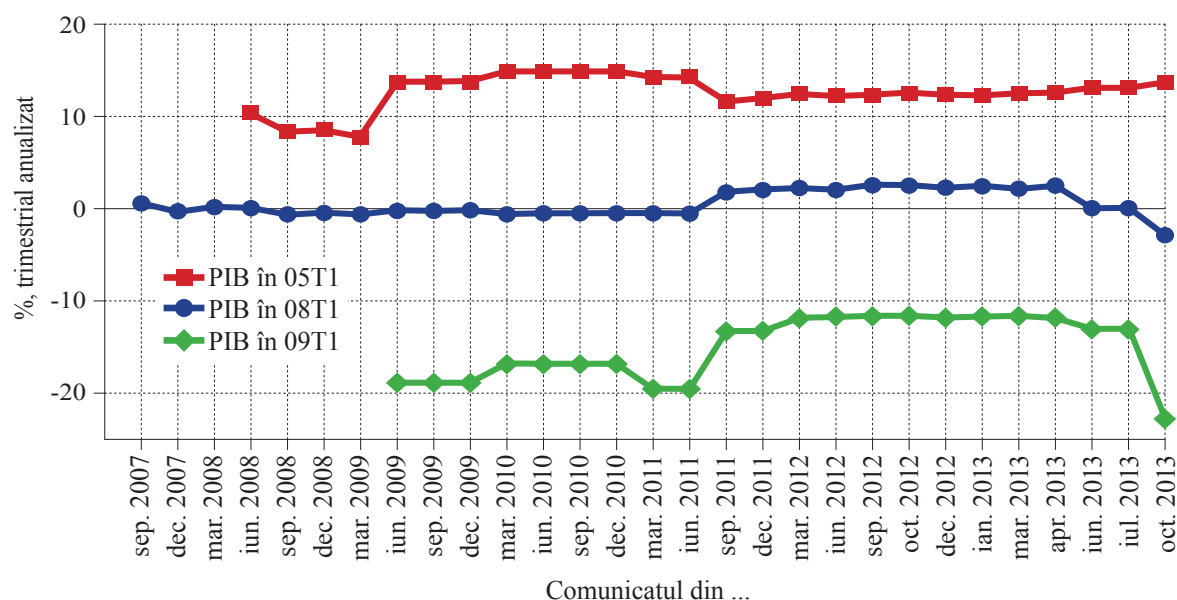
Figura 1. Seriile PIB real ajustate sezonier, variații trimestriale anualizate, %



⁴ Pentru statele cu o experiență statistică mai bogată, cum ar fi Marea Britanie și SUA, pentru fiecare trimestru sunt disponibile mai multe versiuni ale seriilor. De exemplu, conform Mankiw și Shapiro (1986), pentru Conturile Naționale ale SUA există estimări semnal, preliminar, provizorii I, provizorii II și finale. În cazul INS estimările semnal vizează doar trimestrul în curs (fără a publica eventualele revizuiți istorice), iar datele provizorii II au început a fi publicate abia începând cu anul 2012. Astfel, conform metodologiei actuale aplicate de INS, există trei versiuni ale seriei PIB pentru fiecare trimestru: „date semnal”, publicată aproximativ la 45 zile după încheierea trimestrului, „date provizorii I”, publicată la 70 zile după încheierea trimestrului, și „date provizorii II”, disponibilă după circa 100 de zile.

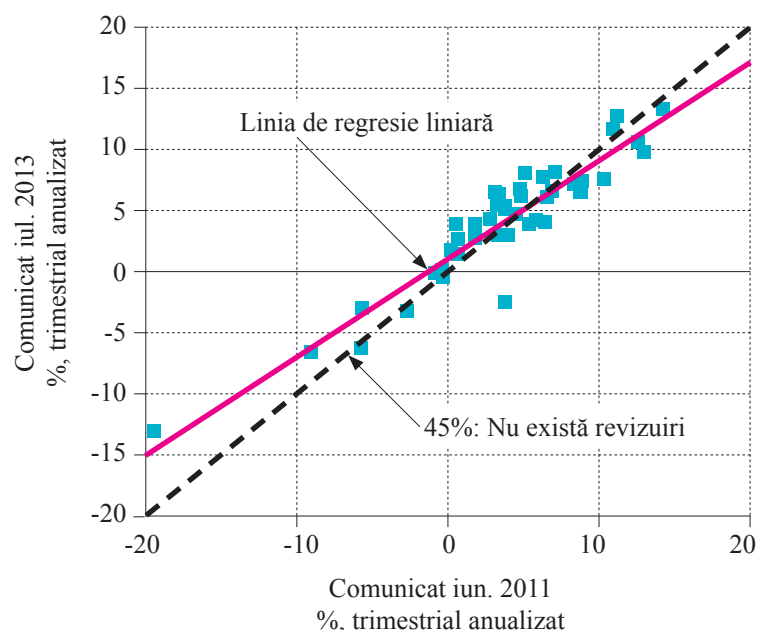
Analiza detaliată a magnitudinii revizuirilor trecute este un prim pas în interpretarea incertitudinii asociate seriei PIB. Conform Figurii 2, creșterea trimestrială anualizată pentru 2009 T1 (linia verde), publicată pentru prima dată în comunicatul din iunie 2009, a suferit revizuirii substanțiale, luând valori între aproximativ -23 la sută (aferentă comunicatului din octombrie 2013) și aproximativ -11 la sută (aferentă comunicatelor de pe parcursul anului 2012). Astfel, magnitudinea revizuirilor este mare în valoare absolută, dar și în raport cu nivelul efectiv al seriei de rate de creștere, implicând revizuirii caracterizate de abateri standard mari și, prin urmare, un grad ridicat de incertitudine. Suplimentar, revizuirile pot avea loc chiar și după o perioadă îndelungată de timp de la publicarea inițială, cum a fost cazul creșterii economice pentru 2005 T1 (linia albastră).

Figura 2. Seriile PIB real ajustate sezonier pentru 2005 T1, 2008 T1 și 2009 T1, variații trimestriale anualizate, %



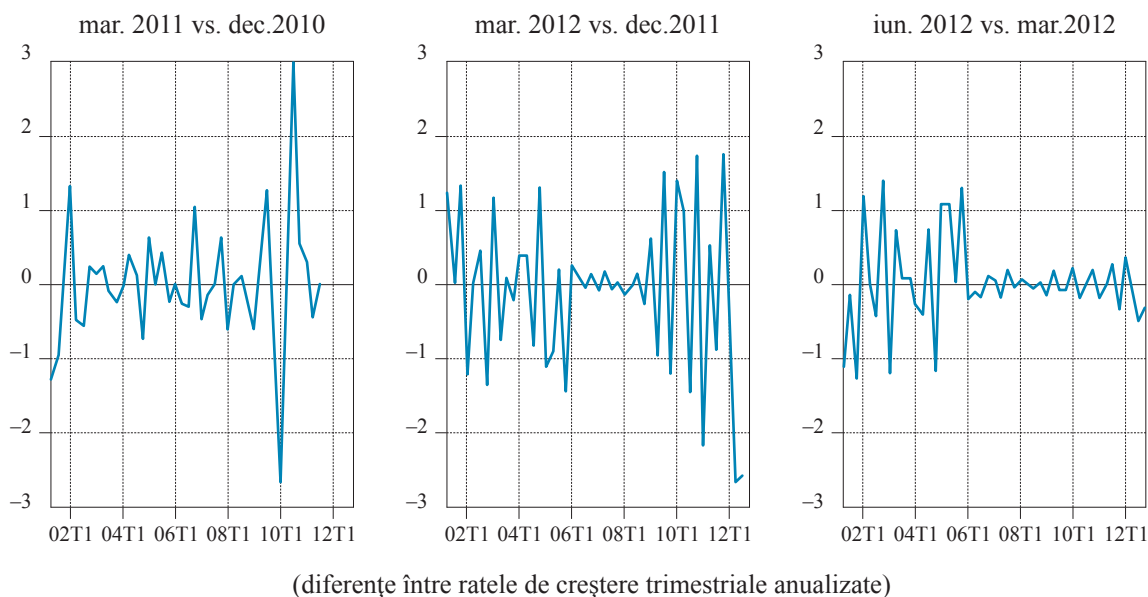
Referitor la direcția revizuirilor, datele statistice inițiale au fost, în general, reevaluate la valori mai mici în termeni absoluți (pozitive mai mici și negative mai mici), indicând o tendință de netezire a ratelor de creștere. Pentru exemplificare, în Figura 3 sunt prezentate seriile de variații trimestriale anualizate aferente comunicatelor din iunie 2011 (axa orizontală) și iulie 2013 (axa verticală), împreună cu dreapta de regresie liniară aferentă legăturii dintre cele două. Panta acesteia este mai mică decât bisectoarea cadranelor I și IV, care descrie lipsa revizuirilor (caz în care cele două comunicate ar fi identice) și implică revizuirii ulterioare către valori mai mici în termeni absoluți.

Figura 3. Seriile PIB real ajustate sezonier aferente comunicatelor din iunie 2011 (axa orizontală) și iulie 2013 (axa verticală), variații trimestriale anualizate, %



O ipoteză fundamentală a modelului de estimare a incertitudinii prezentat în continuare este presupunerea că observațiile cele mai recente suferă revizuirile cele mai mari ca amplitudine, în timp ce observațiile mai îndepărtate sunt revizuite mai puțin. În cazul României, ratele de creștere trimestriale anualizate relevă situații în care această ipoteză poate fi validată (Figura 4, stânga), dar și cazuri mai dificil de interpretat, în care doar observațiile cele mai recente și cele mai îndepărtate cunosc revizuirii importante (Figura 4, centru) sau în care doar observațiile cele mai îndepărtate au fost revizuite semnificativ (Figura 4, dreapta). Prin urmare, natura revizuirilor operate asupra seriilor istorice nu se încadrează întotdeauna în ipoteza modelului teoretic.

Figura 4. Revizuirile ratelor trimestriale anualizate de creștere a PIB între comunicate selective, puncte procentuale



În privința analizei tiparului revizuirilor, Mankiw și Shapiro (1986) propun două abordări. Prima presupune că fiecare comunicat reprezintă o serie care conține erori aleatoare de măsură (engl. *noise*), iar fiecare versiune ulterioară a seriei reduce sau elimină aceste erori, beneficiind de eșantioane cu un grad crescut de reprezentativitate. În acest caz, autorii recomandă modelarea seriei PIB ca variabilă neobservabilă în cadrul unui model reprezentat în spațiul stărilor, în vederea utilizării tehnicilor de filtrare (cum ar fi filtrul Kalman), deoarece datele comunicate de oficialități sunt caracterizate de „iraționalitate”. A doua abordare presupune că fiecare comunicat reprezintă o prognoză eficientă a seriei revizuite (adică a comunicatelor viitoare), care reflectă într-un mod optim toate informațiile disponibile la acel moment, iar fiecare comunicat ulterior reduce sau elimină eroarea de prognoză, încorporând noile informații (engl. *news*)⁵. În acest caz autoritatea responsabilă pentru statistica oficială se comportă rațional, iar seriile comunicate de aceasta sunt eficiente.

Formal, dacă seria aferentă unui comunicat preliminar diferă de cea dintr-un comunicat viitor prin eroare de măsură (ipoteza *noise*), atunci revizuirile sunt necorelate cu estimările revizuite, dar corelate cu estimările preliminare:

$$\text{corr}(y_t^j - y_t^{j-1}, y_t^j) = 0 \text{ și } \text{corr}(y_t^j - y_t^{j-1}, y_t^{j-1}) \neq 0,$$

unde y_t^j reprezintă seria PIB real ajustată sezonier (exprimată în variații trimestriale anualizate) aferentă comunicatului j pentru perioada t . Cu alte cuvinte, se pot folosi informațiile curente, y_t^{j-1} , pentru a prognoza estimările viitoare, y_t^j . Dacă în schimb revizuirile sunt caracterizate de erori de prognoză (ipoteza *news*), atunci acestea sunt corelate cu estimările revizuite, dar necorelate cu estimările preliminare:

$$\text{corr}(y_t^j - y_t^{j-1}, y_t^j) \neq 0 \text{ și } \text{corr}(y_t^j - y_t^{j-1}, y_t^{j-1}) = 0.^6$$

Mediile coeficienților de corelație dintre revizuri și seriile PIB real aferente calculate astfel pe datele INS sunt:

$$\text{corr}(y_t^j - y_t^{j-1}, y_t^{j-1}) = -0,07 \text{ și } \text{corr}(y_t^j - y_t^{j-1}, y_t^j) = 0,11.$$

Dat fiind faptul că ambii coeficienți de corelație calculați sunt apropiați de zero ca magnitudine și ne semnificativi din punct de vedere statistic, nu se poate formula o concluzie privind ipoteza dominantă.

Rezultatele empirice din literatura de specialitate prezintă evidențe mixte. Mankiw și Shapiro (1986) acceptă ipoteza *news*, susținând că estimările curente ale PIB al SUA sunt prognoze eficiente ale seriei viitoare – la momentul apariției unui comunicat nou, seria aferentă este cea mai bună

⁵ Mankiw și Shapiro (1986) susțin că dacă revizuirile reprezintă erori de prognoză eficiente, atunci alte date disponibile la momentul publicării inițiale ar trebui să fie necorelate cu revizuirile, în sensul că informațiile suplimentare (cum ar fi seriile aferente unor comunicate anterioare) nu sunt de natură să îmbunătățească prognoza seriei din comunicatele viitoare.

⁶ Suplimentar, dacă abordarea *news* este dominantă, varianța seriilor din comunicate succesive trebuie să fie crescătoare, pentru că prognoza este în mod natural mai puțin volatilă decât seria istorică (abordarea *news* consideră că seria y^{j-1} este prognoza eficientă pentru seria din comunicatul următor, y^j). Datele INS nu prezintă un tipar clar (descendent sau ascendent) pentru varianța revizuirilor.

estimare disponibilă a seriei viitoare. În cazul Italiei, Frale și Raponi (2012) favorizează ipoteza *noise*, concluzionând că variațiile PIB nu pot fi considerate prognoze raționale ale seriilor viitoare și recomandând tehnici specifice pentru filtrarea seriei oficiale.

3. Filtrarea seriei PIB real ajustate sezonier

Pentru estimarea distribuției de probabilitate aferentă incertitudinii asociate revizuirilor viitoare și obținerea seriei filtrate, a fost aplicat modelul *state-space* descris în Cunningham et al. (2007) și Cunningham și Jeffery (2007) pentru Banca Angliei și Anderson și Gascon (2009) pentru Fed. Modelul ia în considerare tiparul revizuirilor anterioare și se bazează pe ipoteza că acestea reprezintă cu acuratețe magnitudinea și direcția revizuirilor viitoare, datele statistice devenind mai exacte și mai relevante cu fiecare comunicat succesiv, totodată ținând cont și de unii indicatori alternativi, dar și de proprietățile statistice ale seriei PIB real.

În particular, ultima versiune a seriei PIB real ajustate sezonier, exprimată în rate de creștere trimestriale anualizate, este descompusă în sumă a trei componente: (i) eroarea sistematică (sau revizuirii regulate nenule într-o anumită direcție), descrescătoare asimptotic în funcție de maturitatea⁷ observației, fiind mai mare pentru datele mai recente; (ii) PIB filtrat, definit ca seria neobservabilă care exprimă PIB real „net” de componentele reziduale asociate posibilelor revizuirii; (iii) eroarea de măsură (componenta aleatoare a revizuirilor), posibil autocorelată, heteroscedastică, adică cu varianță variabilă în timp, și care descrește monoton în funcție de maturitatea comunicatului.

Modelul descris anterior este reprezentat în spațiul stărilor, fiind format din relația (1.1), reprezentând ecuația de măsură, și relația (1.2), reprezentând ecuația de tranziție (t se referă la momentul observației, j se referă la maturitatea observației):

$$y_t = c^j + \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \bar{y}_t \\ v_t \end{bmatrix}, \quad (1.1)$$

$$\begin{bmatrix} \bar{y}_t \\ v_t \end{bmatrix} = \begin{bmatrix} \mu \\ 0 \end{bmatrix} + \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} \bar{y}_{t-1} \\ v_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix}, \text{ cu matricea de covarianță } \begin{bmatrix} \sigma_\varepsilon^2 & \rho_{\varepsilon\eta}\sigma_\varepsilon\sigma_{\eta^j} \\ \rho_{\varepsilon\eta}\sigma_\varepsilon\sigma_{\eta^j} & \sigma_{\eta^j}^2 \end{bmatrix}, \quad (1.2)$$

unde y_t este ultima versiune a seriei PIB real (rate de creștere trimestriale anualizate aferente comunicatului din octombrie 2013), c^j este vectorul de erori sistematice dependente de maturitatea comunicatului j , $\bar{y}_t$ este seria PIB filtrată, v_t este eroarea de măsură (modelată ca variabilă neobservabilă), ε_t și η_t reprezintă termeni de eroare, caracterizați de varianțele σ_ε^2 și $\sigma_{\eta^j}^2$ (ultima depinde de maturitatea observației) și coeficientul de corelație $\rho_{\varepsilon\eta}$. Modelul descrie seria PIB real (creșteri trimestriale anualizate) ca sumă a erorii sistematice, valorii „adevărate” și erorii de măsură. La rândul lor, seria „adevărată” și eroarea de măsură urmează procese autoregresive

⁷ În continuare, referirile privind maturitatea comunicatului au în vedere perioada parcursă de la publicarea acestuia.

de ordinul 1. A se vedea Harvey (1989) pentru aspecte tehnice privind formularea și estimarea modelelor *state-space*.

Pentru variabilele/parametrii dependenți de maturitate (j), se presupun următoarele procese, caracterizate de diminuarea treptată a erorii sistematice și a varianței erorii de măsură:

$$c^j = c^1(1 + \lambda)^{j-1}, \quad -1 \leq \lambda \leq 0, \quad (2.1)$$

$$\sigma_{\eta^j}^2 = \sigma_{\eta^1}^2(1 + \delta)^{j-1}, \quad -1 \leq \delta \leq 0. \quad (2.2)$$

Estimarea se efectuează în două etape: inițial revizuirile trecute sunt utilizate pentru a estima parametrii care guvernează dinamica erorii sistematice și varianței erorii de măsură, conform ecuațiilor (2.1) și (2.2). Optimizările numerice aferente sunt exogene modelului *state-space* propriu-zis și valorifică tiparul revizuirilor trecute. În etapa a doua, utilizând cel mai recent comunicat privind evoluția PIB se recuperează valorile filtrate asociate prin estimarea sistemului (1.1)-(1.2). Un neajuns al acestei abordări îl reprezintă faptul că pentru fiecare serie PIB real se poate estima o serie filtrată, iar revizuirile seriei filtrate de la un trimestru la altul este proporțională (dar de o magnitudine mai redusă) cu revizuirea suferită de seria oficială⁸.

Pentru a defini „revizuirea”, s-a considerat că valorile publicate pentru observațiile mai vechi de 2 ani ($N=8$ trimestre) reprezintă valori finale, care nu vor mai suferi modificări⁹. Astfel, primul vector de revizuirii este calculat ca diferență între seriile publicate în comunicatele din septembrie 2009 (09.SEP) și septembrie 2007 (07.SEP), iar ultimul vector de revizuirii este reprezentat de diferența dintre seriile aferente comunicatelor 13.OCT și 11.SEP (formula generală are forma $y_t^j - y_t^{j-N}$, $N = 8$)¹⁰. Având la dispoziție 29 de serii, se pot construi $29 - 8 = 21$ serii de revizuirii istorice, această construcție fiind denumită în literatura de specialitate „triumghiul revizuirilor”:

$$\begin{bmatrix} y_{00Q2}^{09.SEP} - y_{00Q2}^{07.SEP} & \dots & y_{00Q2}^{13.JUL} - y_{00Q2}^{11.JUN} & y_{00Q2}^{13.OCT} - y_{00Q2}^{11.SEP} \\ \dots & \dots & \dots & \dots \\ N/A & \dots & y_{11Q1}^{13.JUL} - y_{11Q1}^{11.JUN} & y_{11Q1}^{13.OCT} - y_{11Q1}^{11.SEP} \\ N/A & \dots & N/A & y_{11Q2}^{13.OCT} - y_{11Q2}^{11.SEP} \end{bmatrix}$$

⁸ Cu alte cuvinte, revizuirile imputate comunicatelor succesive ale seriei PIB real se regăsesc parțial în revizuirii ale seriilor PIB filtrate asociate, estimate în cadrul modelului.

⁹ Conform metodologiilor INS, versiunea definitivă a Conturilor Naționale Nefinanciare se publică după 26 de luni de la comunicatul inițial. Acest orizont este specific și INSEE, Franța, conform Lequiller și Blades (2006). Din păcate, revizuirile operate de INS presupun de multe ori și reexaminarea observațiilor cu maturitatea mai mare de 8 trimestre. În *Bank of England Inflation Report* august 2007 este menționat faptul că metodologiile statistice presupun un orizont de 2 ani până la publicarea datelor finale, însă Conturile Naționale sunt supuse revizuirilor chiar și după doi ani de la publicarea inițială, din diferite motive. În studiile de referință se menționează că lungimea orizontului revizuirilor este arbitrară și este de obicei aleasă în urma inspecției vizuale a seriilor: atât Anderson și Gascon (2009), cât și Cunningham et al. (2007) au ales $N = 20$. În cazul seriilor aferente economiei românești, lungimea redusă a acestora implică necesitatea alegerii unui orizont mai scurt.

¹⁰ Au fost folosite, de asemenea, comunicatele „date provizorii II” din iulie 2012, octombrie 2012, ianuarie 2013 și aprilie 2013, seriile aferente fiind considerate a fi semnificativ diferite de cele din comunicatele „date provizorii I” corespunzătoare. În celelalte cazuri au fost utilizate comunicatele „date provizorii II”. Astfel, ultima serie de revizuirii este definită ca diferență între publicațiile din octombrie 2013 (și nu septembrie 2013, aferentă „date provizorii I”) și septembrie 2011 (nu are asociat un comunicat „date provizorii II”, acestea fiind publicate începând cu aprilie 2012).

Fie W matricea cu 21 de coloane și 29 de rânduri (lungimea celui mai scurt comunicat: 2007 T2 – 2000 T2 = 29 de trimestre) care conține ultimele 29 de valori disponibile din fiecare coloană a matricei revizuirilor (ținând cont de forma acesteia descrisă anterior), astfel:

$$W = \begin{bmatrix} y_{00Q2}^{09.SEP} - y_{00Q2}^{07.SEP} & \dots & y_{04Q1}^{13.IUL} - y_{04Q1}^{11.IUN} & y_{04Q2}^{13.OCT} - y_{04Q2}^{11.SEP} \\ \dots & \dots & \dots & \dots \\ y_{07Q1}^{09.SEP} - y_{07Q1}^{07.SEP} & \dots & y_{10Q4}^{13.IUL} - y_{10Q4}^{11.IUN} & y_{11Q1}^{13.OCT} - y_{11Q1}^{11.SEP} \\ y_{07Q2}^{09.SEP} - y_{07Q2}^{07.SEP} & \dots & y_{11Q1}^{13.IUL} - y_{11Q1}^{11.IUN} & y_{11Q2}^{13.OCT} - y_{11Q2}^{11.SEP} \end{bmatrix}$$

Fiecare coloană a matricei W reprezintă revizuirile operate unui singur comunicat (cel din momentul s față de cel din momentul $s - 8$), și fiecare rând al matricei W reprezintă revizuirile în observațiile de o anumită maturitate (ultimul rând reprezintă revizuirile cu maturitatea 1, respectiv ale celor mai recente observații din cadrul fiecărui comunicat).

Pentru estimarea ecuației (2.1) au fost folosite ca și observații $\hat{c}^j$ mediile celor 29 de rânduri ale matricei W (media revizuirilor în funcție de maturitate). În cadrul estimării a fost minimizată suma pătratelor distanțelor dintre specificarea modelului, respectiv ecuația (2.1), și revizuirile empirice $\hat{c}^j$. Rezultatele estimărilor sunt prezentate în Tabelul 1.

Tabelul 1. Rezultatele estimării ecuației (2.1). Abateri standard în ()

$\min \sum [\hat{c}^j - c^j(1 + \lambda)^{j-1}]^2$ $c^j = \mathbf{0,02} \text{ (0,71),}$ $\lambda = \mathbf{-0,12} \text{ (0,33).}$	Anderson și Gascon (2009): $c^1 = \mathbf{0,58}$, interval de încredere 95%: [0,33; 0,83], $\lambda = \mathbf{-0,28}$, interval de încredere 95%: [-0,45; -0,11]. Cunningham et al. (2007): nesemnificative statistic, nu sunt raportate.
---	--

Coeficienții estimați sunt caracterizați de erori standard foarte mari, sugerând faptul că aceștia nu sunt semnificativi din punct de vedere statistic. Astfel, media revizuirilor nu este diferită de zero pentru fiecare maturitate, sugerând lipsa unor evidențe privind erori sistematice de subevaluare/supraevaluare a creșterii economice. Având în vedere aceste rezultate, s-a decis renunțarea la parametrul c^j din ecuația de măsură aferentă modelului *state-space*. În rapoartele trimestriale asupra inflației, Banca Angliei publică în secțiunea de sinteză un grafic de tip evantai pentru intervalul de proiecție, dar și pentru ultimele valori istorice ale PIB trimestrial (variații anuale). Acest grafic prezintă o asimetrie evidentă pe intervalul de date realizate, în sensul că mijlocul intervalului de încredere este superior seriei de date efective, implicând o potențială revizuire în sus a seriei¹¹.

Pentru a estima coeficientul de corelație dintre valorile filtrate și eroarea de măsură variabilă, $\rho_{\varepsilon\eta}$, Cunningham et al. (2007) propun aproximarea acestuia cu media corelațiilor dintre revizuirile aferente fiecărei maturități (fiecare rând al matricei W) și seria PIB corespunzătoare, respectiv $\hat{\rho}_{\varepsilon\eta} = \text{corr}(W(j,:), y_t^{j+8})$. De asemenea, acesta poate fi estimat și în cadrul modelului

¹¹ A se vedea, de exemplu, *Bank of England Inflation Report* august 2012, p. 6.

state-space. Estimările raportate de Cunningham et al. (2007) și Anderson și Gascon (2009) pentru acest parametru sunt opuse ca semn (-0,31 și, respectiv, 0,32), diferențele semnificative relevate în literatura de specialitate plasând o incertitudine suplimentară asociată parametrului.

Pentru estimarea $\sigma_{\eta^1}^2$, β și δ Cunningham et al. (2007) sugerează minimizarea

$$\left(\text{vec} V - \text{vec} \hat{V} \right)' \left(\text{vec} V - \text{vec} \hat{V} \right),$$

unde $\hat{V} = J^{-1}WW'$ este matricea de varianță-covarianță a matricei W , J este numărul de rânduri ale lui W ($J = 29$), și $V = V(\sigma_{\eta^1}^2, \delta, \beta)$ este matricea analitică de varianță-covarianță a revizuirilor, de forma:

$$V_{ij} = \frac{\sigma_{\eta^1}^2}{1 - (1 + \delta)\beta^2} \cdot \left[(1 + \delta)^{i-1} \beta^{j-i} \right], i = 1, \dots, 29, j = 1, \dots, 29, j \geq i. \quad (2.2.A)$$

Rezultatele optimizării numerice sunt prezentate în Tabelul 2.

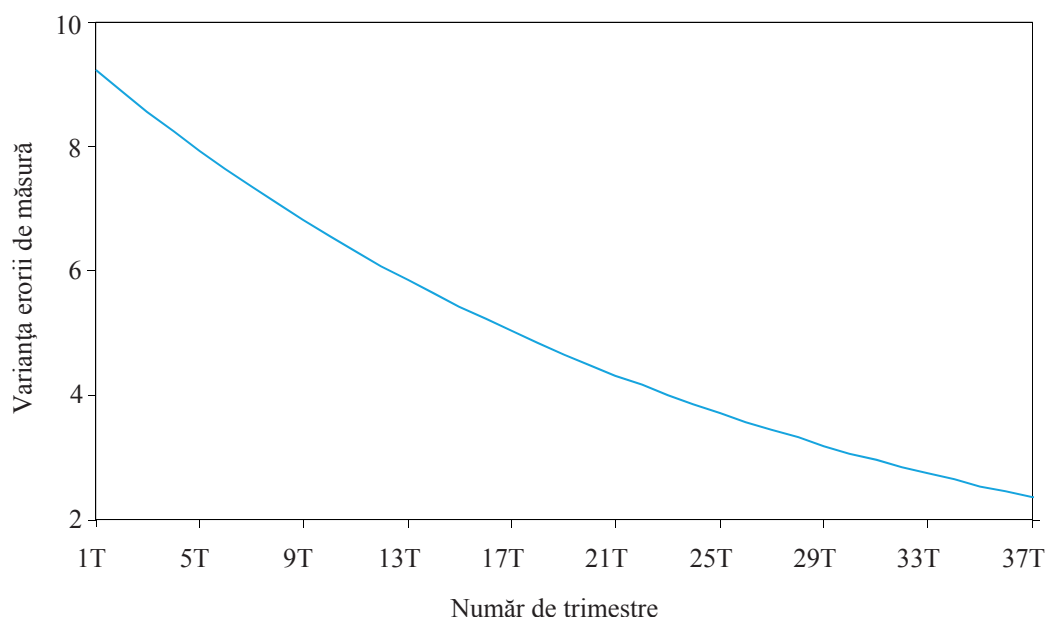
Tabelul 2. Rezultatele estimării. Abateri standard în ()

$\min \left(\text{vec} V - \text{vec} \hat{V} \right)' \left(\text{vec} V - \text{vec} \hat{V} \right).$ $\sigma_{\eta^1}^2 = \mathbf{9,23} \text{ (0,38)}, \quad \sigma_{\eta^1} = 3,04,$ $\delta = \mathbf{-0,07} \text{ (0,00)},$ $\beta = \mathbf{-0,03} \text{ (0,02)}.$	Anderson și Gascon (2009): $\sigma_{\eta^1}^2 = 2,70, \text{ interval de încredere 95\%: } [2,40; 3,00],$ $\delta = -0,08, \text{ interval de încredere 95\%: } [-0,09; -0,06],$ $\beta = 0,20, \text{ interval de încredere 95\%: } [0,15; 0,26].$ Cunningham et al. (2007): $\sigma_{\eta^1}^2 = 3,58 \text{ (0,30)},$ $\delta = -0,06 \text{ (0,01)},$ $\beta = -0,22 \text{ (0,06)}.$
---	---

Primii doi estimatori sunt semnificativi din punct de vedere statistic (abaterea standard este mică în raport cu valoarea estimată a coeficientului). Față de studiile de referință, parametrul care determină rata de descreștere a varianței erorii de măsură odată cu maturitatea (δ) prezintă magnitudini similare celor estimate în cazul PIB al SUA și al Marii Britanii. În cazul varianței inițiale a erorii de măsură ($\sigma_{\eta^1}^2$) însă, rezultatele indică o valoare mult mai mare, rezultat susținut parțial și de natura și magnitudinea mare ale revizuirilor operate de INS în runde succesive de publicare a seriilor PIB real, prezentate în secțiunea 2. Varianța erorii de măsură aferentă ecuației (2.2) este reprezentată în Figura 5. Aceasta este mai mare pentru observațiile mai recente (partea din stânga a graficului), pentru care sunt asociate probabilități superioare de a înregistra revizuri și, implicit, un grad mai mare de incertitudine. Chiar dacă parametrul β apare și în cadrul modelului *state-space*, acesta va fi fixat la valoarea estimată în cadrul optimizării numerice, facilitând în acest mod estimarea celorlalți coeficienți (și evitând eventuale probleme de identificare).

¹² Matricea V este simetrică, în formulă fiind reprezentată doar forma generală a elementelor situate pe diagonala principală și deasupra acesteia.

Figura 5. Varianța erorii de măsură $\sigma_{\eta^j}^2$ conform ecuației (2.2)



Coeficienții aferenți modelului cu componente neobservabile (1.1)-(1.2) sunt estimați prin metoda verosimilității maxime ajustate cu distribuții *a priori* (engl. *Regularized Maximum Likelihood*). Distribuțiile *a priori* pentru parametri au fost construite pe baza rezultatelor obținute în literatura de specialitate, dar ținând cont de specificul seriei de rate de creștere trimestriale anualizate ale PIB real (autocorelație ridicată și lungime scurtă a eșantionului). Rezultatele estimării sunt prezentate în Tabelul 3.

Tabelul 3. Rezultatele estimării sistemului (1.1)-(1.2). Abateri standard în ()

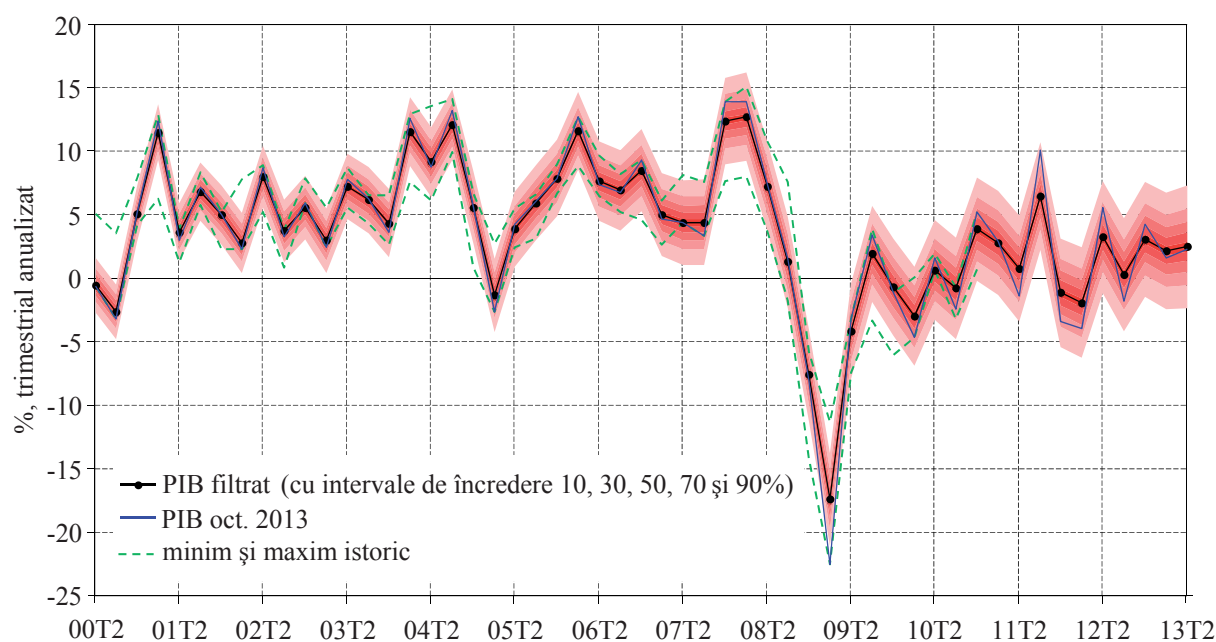
$\mu = 1,49$ (0,24), $\alpha = 0,59$ (0,10), $\sigma_{\varepsilon} = 4,07$ (0,84), $\rho = 0,02$ (0,10).	Anderson și Gascon (2009): nu raportează Cunningham et al. (2007): $\mu = 1,18$ (0,24), $\alpha = 0$ (fixat), $\sigma_{\varepsilon} = 3,22$ (0,67).
---	--

Cu excepția parametrului ρ , erorile standard sunt mici în raport cu valorile estimate, sugerând coeficienți semnificativ diferiți de zero din punct de vedere statistic. Valorile estimate pentru ecuația de tranziție implică o creștere trimestrială anualizată la echilibru de $\mu / (1 - \alpha) = 3,59$, foarte apropiată de media de 3,54 la sută pentru ultima versiune a seriei PIB real ajustate sezonier. Totodată, valorile coeficienților susțin ipoteza *noise* pentru tiparul revizuirilor suferite de seria PIB, în condițiile în care coeficientul ρ nu este semnificativ statistic (implică lipsa corelației dintre revizuri și ultimul comunicat disponibil, $\text{corr}(y_t^j - y_t^{j-1}, y_t^j) = 0$). Aceasta, după cum am menționat anterior (a se vedea secțiunea 2), susține necesitatea utilizării tehnicilor de filtrare (în particular, filtrul Kalman) pentru extragerea componentelor neobservabile.

În urma aplicării filtrului Kalman utilizând valorile estimate anterior, seria PIB filtrată obținută ($\bar{y}_t$) reprezintă o variantă ușor mai netedă a ultimei serii PIB real; a se vedea Figura 6. Modelul indică o direcție a revizuirilor așteptate compatibile cu observațiile din Figura 3, respectiv o tendință către serii revizuite care prezintă rate de creștere trimestriale mai mici (în modul).

Pentru vizualizarea distribuției de probabilitate estimate, aferentă revizuirilor potențiale, a fost construit un grafic de tip evantai pentru intervalele de încredere de 10, 30, 50, 70 și 90 la sută din jurul seriei filtrate¹³; a se vedea Figura 6. Britton et al. (1998) prezintă aspecte generale și specifice privind utilizarea graficelor de tip evantai în cadrul Băncii Angliei. Astfel, seria filtrată obținută este caracterizată pentru perioada recentă de un interval de încredere de 90 la sută de o lățime comparabilă cu cea aferentă canalului calculat ca diferență între valorile maxime și minime istorice din intervalul 2000 T2 : 2010 T4, sugerând o aproximare adecvată a incertitudinii aferente revizuirilor potențiale. De asemenea, intervalul de încredere asociază o incertitudine mai mare valorilor recente ale ratelor de creștere trimestriale anualizate ale PIB real, ultimele observații prezentând probabilități semnificative de fi revizuite la valori apropiate de zero.

Figura 6. Seria PIB filtrată și intervalele de încredere



Estimarea alternativă a intervalelor de încredere utilizând simulări Monte Carlo din distribuția implicită a modelului a rezultat în lățimi similare, fără semne evidente privind prezența asimetriei acestora¹⁴. Pentru un număr suficient de mare de simulări (în estimarea de bază au fost aplicate 10 000 de trageri), seria filtrată coincide cu mijlocul intervalelor de încredere – situație identică cu estimarea bazată pe ipoteza de distribuție normală a seriilor (Figura 6) –, chiar dacă în acest caz nu mai există o interdependență între acestea. Prin urmare, cele două metode de construire a intervalelor de încredere dau rezultate foarte apropiate.

¹³ Intervalele de încredere au fost construite adoptând ipoteza că seriile urmează o distribuție normală. Procedul implică intervale de încredere simetrice în jurul seriei filtrate.

¹⁴ Lipsa asimetriei intervalelor de încredere este parțial datorată renunțării la seria de erori sistematice c^j în cadrul modelului *state-space*.

4. Analiza robusteții modelului

4.1. Senzitivitatea modelului la lungimea orizontului revizuirilor

În modelul construit și estimat în secțiunea precedentă orizontul revizuirilor (distanța dintre două comunicate care definește revizuirea) a fost fixat la $N = 8$ trimestre. Însă alegerea orizontului revizuirilor este arbitrară și necesită examinarea robusteții modelului la variația acestui parametru, conform Anderson și Gascon (2009). În vederea analizei sensibilității coeficienților și a seriei filtrate la modificarea orizontului revizuirilor, modelul prezentat anterior a fost estimat pentru $N = 6$ și $N = 10$ trimestre. În cadrul estimării ecuației aferente varianței erorii de măsură $\sigma_{\eta_i}^2$, apar unele diferențe în magnitudinea coeficienților: cu cât este mai mare orizontul revizuirilor, cu atât sunt mai mari (în modul) varianța inițială a erorii de măsură ($\sigma_{\eta_i}^2$) și rata de descreștere (δ) a acesteia (însă relația nu este una liniară); a se vedea partea superioară a Tabelului 4.

Tabelul 4. Senzitivitatea coeficienților la lungimea orizontului revizuirilor (N), abateri standard în ()

	$N = 6$	$N = 8$	$N = 10$
$\sigma_{\eta_i}^2$	8,246 (0,38)	9,227 (0,38)	10,408 (0,52)
δ	-0,072 (0,01)	-0,073 (0,00)	-0,122 (0,01)
β	-0,050 (0,02)	-0,028 (0,02)	-0,027 (0,02)
μ	1,487 (0,25)	1,486 (0,24)	1,486 (0,24)
α	0,585 (0,10)	0,586 (0,10)	0,586 (0,10)
σ_e	4,100 (0,73)	4,069 (0,84)	4,067 (0,85)
ρ	0,023 (0,10)	0,024 (0,10)	0,024 (0,10)

Utilizând aceste seturi de coeficienți, parametrii estimați în cadrul modelului *state-space* nu se modifică semnificativ; a se vedea partea inferioară a Tabelului 4. Ca urmare, seriile PIB filtrate aferente, împreună cu intervalele de încredere asociate, nu sunt sensibile la variații ușoare ale lungimii orizontului revizuirilor.

4.2. Includerea unei variabile observabile suplimentare

Cunningham et al. (2007) propun augmentarea modelului cu o variabilă observabilă care nu suferă revizuirii și care prezintă o legătură puternică cu seria PIB real (și implicit cu seria PIB filtrată). Majoritatea studiilor susțin că cel mai bun candidat este reprezentat de datele privind sondajele economice. Ashley et al. (2005) analizează relația dintre datele din sondaje și incertitudinea asociată seriilor din conturile naționale în cazul Marii Britanii. În cadrul analizei a fost selectată includerea deviației viitoare a indicatorului de încredere în economie pentru România (ESI, engl. *Economic Sentiment Indicator*) de la media pe termen lung: $x_t = ESI_{t+1} - 100$ (acesta prezintă un coeficient de corelație de 0,66 cu ultima serie PIB real)¹⁵. Ecuația de tranziție (1.2) rămâne aceeași, însă ecuația de măsură (1.1) devine:

$$\begin{cases} y_t = c^j + \bar{y}_t + v \\ x_t = \varphi + \omega \bar{y}_t + \zeta_t. \end{cases} \quad (1.1.A)$$

Parametrii estimați exogen (prima etapă) rămân nemodificați. Modelul multivariat determină un număr de 3 parametri suplimentari: φ , ω și σ_ζ (ultimul este abaterea standard a termenului de eroare ζ_t). Rezultatele estimării utilizând *Regularized Maximum Likelihood* pentru parametrii modelului *state-space* (1.1.A)-(1.2) sunt prezentate în Tabelul 5, alături de cele aferente modelului univariat.

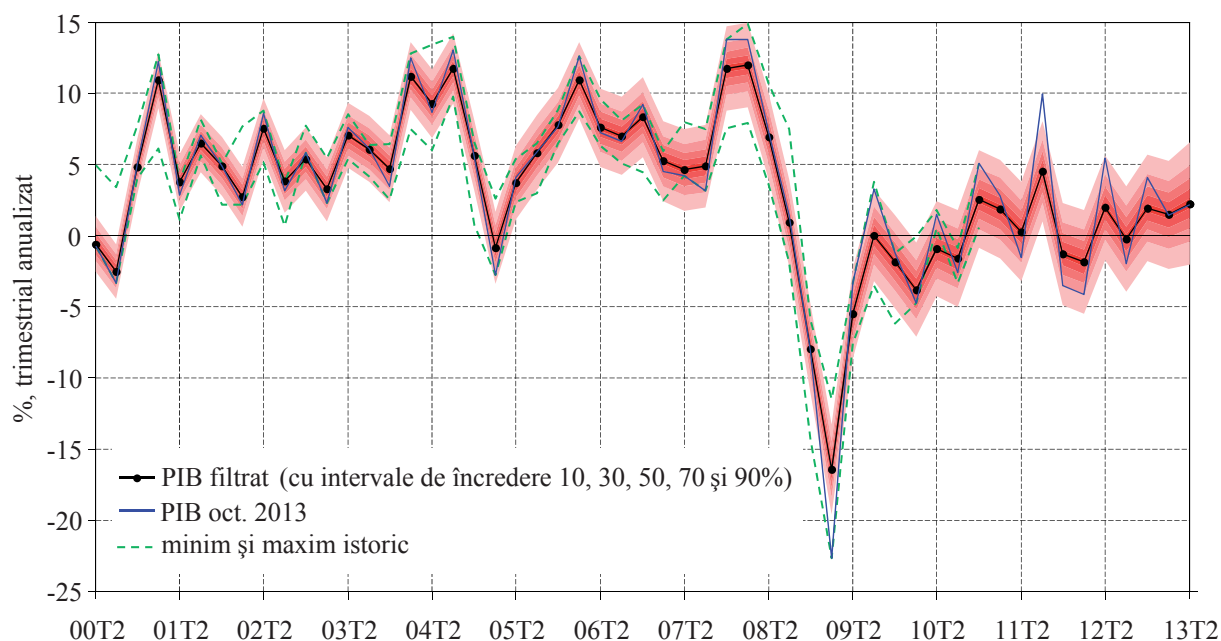
Tabelul 5. Rezultatele estimării pentru modelele univariat și multivariat. Abateri standard în ()

	Model univariat	Model multivariat
μ	1,486 (0,24)	1,398 (0,25)
α	0,586 (0,10)	0,664 (0,10)
σ_ε	4,069 (0,84)	3,498 (0,72)
ρ	0,024 (0,10)	0,013 (0,10)
φ	-	-0,418 (0,65)
ω	-	1,132 (0,47)
σ_ζ	-	7,970 (0,20)

Valorile estimate pentru parametrii comuni se modifică doar marginal. Legătura dintre indicatorul de încredere în economie și PIB real, exprimată prin coeficientul ω , implică o relație directă de magnitudine aproximativ unitară. În urma aplicării filtrului Kalman, seria PIB real filtrată obținută este similară celei estimate în cadrul modelului univariat (a se compara Figurile 6 și 7).

¹⁵ Comunicatele privind evoluția ESI pentru trimestrul t sunt disponibile la sfârșitul trimestrului respectiv, în timp ce datele privind evoluția PIB pentru trimestrul t sunt publicate la circa 70 zile de la sfârșitul trimestrului („date provizorii I”).

Figura 7. Seria PIB filtrată și intervalele de încredere, model multivariat



Tiparul seriilor filtrate și al intervalelor de încredere, cu excepții minore, sunt foarte similare pentru cele două specificații ale modelului *state-space*. Seria filtrată aferentă modelului multivariat (Figura 7) prezintă un grad de netezire ușor mai mare decât seria filtrată în cadrul modelului univariat (Figura 6), implicând potențiale revizuiri de magnitudini mai mari. De asemenea, intervalele de încredere sunt marginal mai înguste în cazul includerii ESI ca variabilă observabilă suplimentară: de exemplu, pentru 2009 T3 sau 2010 T4 seria actuală a PIB real este plasată în intervalul de încredere de 70 la sută în cazul estimării univariete, însă este la limita intervalului de încredere de 90 la sută pentru estimarea multivariată.

5. Concluzii

În cadrul studiului au fost analizate aspecte privind tiparul revizuirilor operate seriilor PIB real ajustate sezonier și a fost estimată seria PIB filtrată (netă de revizuirile potențiale) aferentă comunicatului INS din octombrie 2013, împreună cu intervalele de încredere asociate. Acestea permit o evaluare optimă a incertitudinii prin intermediul distribuției de probabilitate a revizuirilor viitoare.

Tiparul revizuirilor nu se încadrează întotdeauna în ipotezele modelului, uneori fiind revizuite și observațiile mai îndepărtate, nu doar cele recente. Conform estimărilor rezultate, revizuirile operate de INS prezintă efecte *noise* (zgomot) – creșterile PIB real sunt caracterizate de erori de măsură aleatoare și nu pot fi considerate prognoze eficiente ale seriilor viitoare –, indicând necesitatea construirii unui model pentru extragerea seriei neobservabile a PIB filtrat. În cadrul analizei nu a fost semnalată existența erorilor sistematice de subevaluare/supraevaluare (în medie, revizuirile sunt nule), însă față de studiile de referință, varianța erorii de măsură este mai mare, indicând o incertitudine sporită asociată ratelor de creștere trimestriale anualizate ale PIB în cazul României.

Seria PIB filtrată obținută (atât în cadrul modelului univariat, cât și în cazul modelului augmentat cu indicatorul de încredere în economie ca variabilă observabilă suplimentară) reprezintă o variantă mai netedă (rate de creștere trimestriale anualizate mai mici în valoare absolută) a seriei PIB real ajustate sezonier aferentă comunicatului INS din octombrie 2013, iar valorile estimate ale coeficienților determină pentru cele mai recente observații intervale de încredere de 90 la sută de lățimi aproximativ egale cu cea aferentă canalului calculat ca diferență între valorile maxime și minime istorice. Intervalele de încredere reprezintă plaja de valori ale revizuirilor viitoare suferite de seria PIB real cu anumite probabilități. De asemenea, modelul este robust la definiția orizontului revizuirilor – estimările alternative cu intervale de 6 și 10 trimestre (față de 8 trimestre în modelul de bază) au rezultat în variații minore ale coeficienților modelului și ale seriilor filtrate.

Bibliografie

- Anderson, R.,
Gascon, C. *Estimating U.S. Output Growth with Vintage Data in a State-Space Framework*, Federal Reserve Bank of St. Louis Review, 91(4), pp. 349-369, 2009
- Ashley, J.,
Driver, R.,
Hayes, S.,
Jeffrey, C. *Dealing with Data Uncertainty*, Bank of England Quarterly Bulletin: Spring 2005, pp. 23-29, 2005
- Britton, E.,
Fisher, P.,
Whitley, J. *The Inflation Report Projections: Understanding the Fan Chart*, Bank of England Quarterly Bulletin: February 1998, pp. 30-37, 1998
- Cunningham, A.,
Eklund, J.,
Jeffery, C.,
Kapetanios, G.,
Labhard, V. *A State Space Approach to Extracting Signal from Uncertain Data*, Bank of England WP No. 336, 2007
- Cunningham, A.,
Jeffery, C. *Extracting Better Signal from Uncertain Data*, Bank of England Quarterly Bulletin 2007 Q3, 2007
- Frale, C.,
Raponi, V. *Revisions in Official Data and Forecasting*, Ministry of the Economy and Finance – Italy, WP No. 3, 2012
- Harvey, A. *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, 1989
- Jacobs, J.,
van Norden, S. *Modeling Data Revisions: Measurement Error and Dynamics of “True” Values*, Journal of Econometrics, 161(2), pp. 101-109, 2011
- Lequiller, F.,
Blades D. *Understanding National Accounts*, OECD, 2006
- Mankiw, G.,
Shapiro, M. *News or Noise: An Analysis of GNP Revisions*, Survey of Current Business, 66, pp. 20-25, 1986.

