

Construction d'intervalles de confiance et relecture du passé avec le modèle Mésange

Journées de Méthodologie Statistique de l'Insee 2022

Alexandre Bourgeois et Benjamin Favetto

31 mars 2022

INSEE - Département des Études Économiques & Direction Générale du Trésor



Institut national de la statistique
et des études économiques

Mesurer pour comprendre



Direction générale du Trésor

1. Introduction et problématique
2. Construction d'intervalles de confiance pour les fonctions de réponse
3. Retour sur le passé : l'exemple de la Grande Récession
4. Conclusion et perspectives

Introduction et problématique

Mésange : un modèle macroéconométrique de l'économie française

- Modèle trimestriel permettant de simuler les comportements agrégés des agents économiques (ménages, SNF, APU ...)
- Une quarantaine d'équations de comportement estimées sur le passé à partir des comptes trimestriels, de type modèle à correction d'erreur (ECM)
- Plus de 1700 équations comptables assurent la cohérence des résultats dans un cadre comptable simplifié
- Depuis sa dernière ré-estimation (2017) : deux types de travailleurs, ré-estimation après la crise de 2008, désagrégation sectorielle plus riche

[Bardaji et al., 2017] : un modèle utilisé **en variante** pour simuler *ex ante* la réponse de l'économie française à un choc ou à une réforme de politique économique, à l'Insee et à la DG Trésor.

BREF SURVOL DU MODÈLE MÉSANGE (2/2)

Un exemple de fonction de réponse (IRF)

Le modèle permet de simuler les effets dans le temps d'une hausse permanente l'investissement public, par rapport à un scénario dit *central* (sentier de croissance équilibré).

Hausse de l'investissement public de 1 % du PIB en volume

	% en écart au compte central					
	1 an	2 ans	3 ans	5 ans	10 ans	LT
PIB en volume	1,31	1,38	1,22	0,93	0,57	0,22
VA du secteur marchand	1,71	1,75	1,51	1,10	0,63	0,17
Consommation des ménages	0,41	0,90	1,13	1,21	0,85	0,40
Investissement	6,58	6,98	6,48	5,70	5,09	4,84
Investissement des SNF et EI	3,06	3,54	2,62	1,17	0,33	0,26
Exportations	-0,07	-0,36	-0,68	-1,15	-1,63	-2,07
Importations	1,62	1,92	1,73	1,49	1,29	1,18
Revenu disponible réel des ménages	0,59	0,97	1,27	1,34	0,79	0,40
Taux d'épargne	0,15	0,06	0,12	0,11	-0,05	0,00

Figure 1 – Exemple de fonction de réponse (d'après [Bardaji et al., 2017])

Équation de consommation des ménages

$$\begin{aligned} \Delta c_t = & \underbrace{-0.02}_{(-2.9)} + \underbrace{0.11}_{(3.1)} \Delta_4 r d b r_t - \underbrace{0.001}_{(-1.6)} \Delta e u r 3 m r_{t-2} - \underbrace{0.005}_{(-2.2)} \Delta U_t + P A C \\ & - \underbrace{0.16}_{(-3.1)} \underbrace{\left(c - r d b r + 0.04 \text{ libfin} \right)}_{\substack{(***) \\ \text{long terme}}} \Big|_{t-1} + \varepsilon_t \end{aligned} \quad (1)$$

$(R^2_{adj} = 53.4\%, DW = 1.88, SER = 0.4\%, \text{période d'estimation : } 1987T4 - 2014T4).$

Intérêts :

- bonnes performances empiriques
- réconciliation des approches court terme / long terme (CT estimé librement et cadre théorique pour le LT qui se traduit par une relation de cointégration)

Limites : pas d'anticipations, modèle à agent représentatif

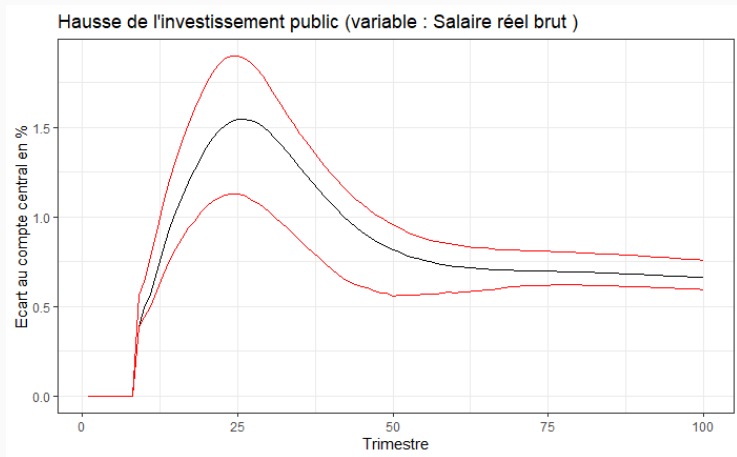
Enjeux de la construction d'intervalles de confiance ?

- Des sources d'aléa dans le modèle prises en compte dans les équations économétriques (estimateurs des coefficients, distribution des résidus)
- Intérêt de la construction d'intervalles de confiance : quantifier l'effet de cet aléa sur les simulations
- Pour avoir une indication de la significativité des effets à moyen terme

Remarque : une pratique courante dans les simulations de modèles VAR / DSGE, mais un discours plus délicat à tenir avec un modèle institutionnel!

EXEMPLE DE FONCTION DE RÉPONSE

Exemple : courbe de réponse du salaire réel brut lors d'une hausse de l'investissement public de 1% du PIB en volume

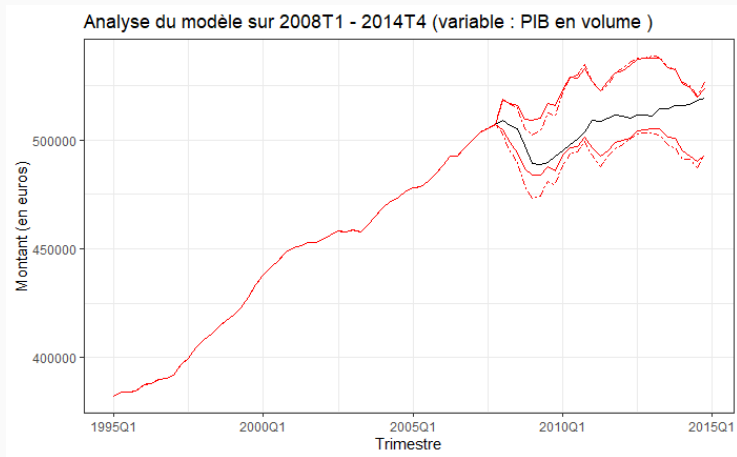


Une analyse rétrospective des résultats du modèle ?

- But : confronter les simulations du modèle aux observations sur une période du passé récent (incluse dans la période d'estimation du modèle)
- En particulier : étudier la réaction à des chocs observés, et non plus calibrés pour l'analyse
- Application : relecture de la crise de 2008 et construction de scénarios contrefactuels

En pratique : questionner *globalement* la capacité du modèle à reproduire le passé, et construire un intervalle de fluctuation autour d'une trajectoire observée (mesure de la volatilité du modèle *ex post*)

Exemple : construction d'un intervalle de fluctuation pour le PIB



Construction d'intervalles de confiance pour les fonctions de réponse

La source de variabilité du modèle

- L'estimation des coefficients des ECM donne lieu au calcul des **résidus**.
- Ces résidus représentent l'écart entre ce qui est expliqué par le modèle, et les valeurs prises par les variables endogènes du modèle.
- Pour une équation donnée, ces résidus sont supposés être **indépendants, identiquement distribués et centrés**.

Principes généraux

- Calculer pour chaque date de la simulation et pour les principales variables du modèle un intervalle de confiance
- Utiliser la distribution empirique des résidus économétriques de chaque ECM (calculée lors de la phase d'estimation) ...
- ... pour simuler un grand nombre de trajectoires du modèle légèrement différentes qui intègrent cet aléa → **bootstrap!**

Idée : utiliser les capacités de simulation du modèle, sans avoir besoin d'autre information que celle contenu dans les séries de résidus initialement calculées, pour quantifier la sensibilité des résultats aux modifications de certaines variables

Description formalisé du modèle Mésange

Un ensemble d'équations f_i de la forme :

$$f_i(y_t, y_{t-1}, \dots, y_{t-p}, x_t, \dots, x_{t-q}, \alpha_i) = \varepsilon_{i,t}, \quad i = 1, \dots, n, \quad t = 1, \dots, T. \quad (2)$$

avec :


- y_t le vecteur des n endogènes à la date t
- x_t le vecteur des exogènes à la date t
- α_i le vecteur des coefficients (estimés) de l'équation. On note $\alpha = (\alpha'_1, \dots, \alpha'_m)'$ l'ensemble des coefficients du modèle.
- $\varepsilon_{i,t}$ le résidu économétrique (ou terme d'erreur), supposé centré et i.i.d.

Bootstrap non-paramétrique

1. À l'étape j de la boucle de rééchantillonnage, on tire avec remise dans $\hat{\varepsilon}$ un vecteur de résidus $\varepsilon_t^{*,j}$ pour chaque date $t = 1, \dots, T$. Les éléments de $\varepsilon_t^{*,j}$ proviennent donc tous d'une même date dans la période d'estimation (*block bootstrap*).
2. À partir de ces séries de résidus, on simule la trajectoire des variables endogènes du modèle sur la période $1 \leq t \leq T$.
3. À l'aide de ces nouvelles séries d'endogènes simulées, on estime α par la méthode utilisée lors de l'estimation initiale du modèle. On note $\hat{\alpha}^{*,j}$ l'estimateur des coefficients obtenu.
4. On répète les étapes précédentes pour $j = 1, \dots, J$, ce qui fournit J jeux de données et autant de version du modèle associées.

BOOTSTRAP : SCHÉMA EXPLICATIF

date	1	...	S	...	T
eq. 1	$\varepsilon_{1,1}$...	$\varepsilon_{1,S}$...	$\varepsilon_{1,T}$
\vdots	\vdots		\vdots		\vdots
eq. i	$\varepsilon_{i,1}$...	$\varepsilon_{i,S}$...	$\varepsilon_{i,T}$
\vdots	\vdots		\vdots		\vdots
eq. n	$\varepsilon_{n,1}$...	$\varepsilon_{n,S}$...	$\varepsilon_{n,T}$


 ε_t^*

Pour choisir les résidus ε_t^* à la date t lors du rééchantillonnage, on choisit une date s dans la période d'estimation (avec remise).

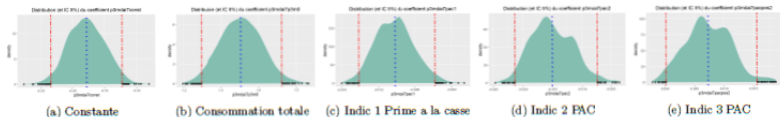
- Calculer une **distribution empirique** pour chaque coefficient
- Analyser la transmission de **l'aléa sur les coefficients** aux résultats de simulation du modèle (variabilité des IRF)
- Construire les bandes de confiance à partir d'un grand nombre de simulations

Intérêt : par rapport à la méthode de Monte-Carlo, pas d'hypothèse paramétrique sur les distributions des résidus, pas d'estimation nécessaire de leur covariance

Hypothèses de projection :

- 1000 jeux de coefficients obtenus par *bootstrap*
- Pour chaque jeu, on projette le modèle à long terme → calcul d'un compte central et d'un compte variantiel
- Résultats présentés en écart au compte central, avec un IC à 95 %

EXEMPLE : DISTRIBUTION EMPIRIQUE DES COEFFICIENTS

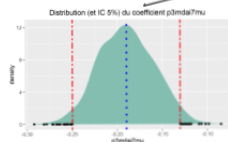


$$\Delta c_t^{DAL} = -0,17 + 1,37 \Delta c_t - 0,39 \Delta PxRel_t^{DAL} + PAC'$$

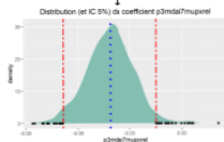
(-3,9) (17,3)
(-3,0)

$$-0,18 \left[c^{DAL} - c + 0,25 PxRel^{DAL} + 0,004 libfin^{DAL} \right]_{t-1}$$

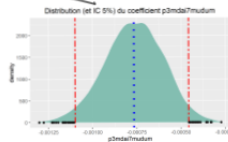
(-3,8)
(***)
(***)



(g) LT consommation



(h) LT prix relatif



(i) LT Tendence lib. fin.

Illustration de la méthode : hausse de l'efficience du travail

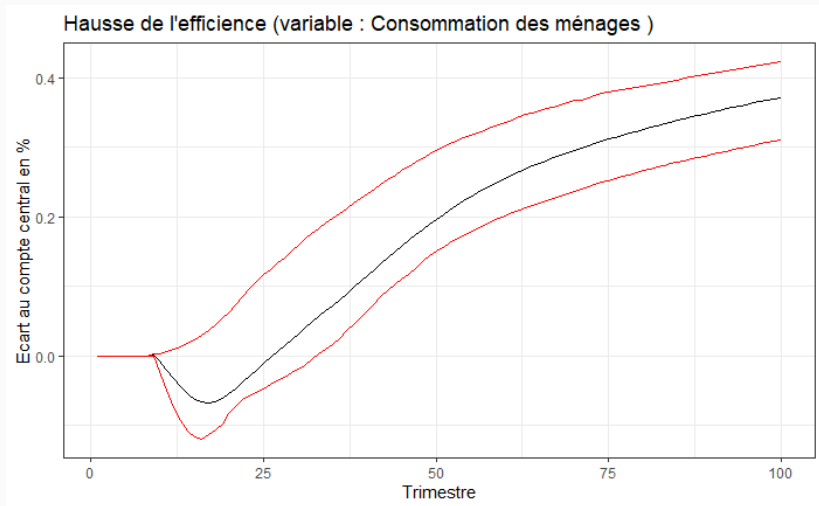
Variante correspondant à une **hausse permanente de l'efficience du travail de 1 %** → choc d'offre, spécifique à la France.

Mécanismes sous-jacents

À court terme, l'emploi s'ajuste à la baisse du fait de la plus grande productivité du travail, car la demande est rigide. L'augmentation du chômage limite la diffusion des gains de productivité aux salariés, et pèse sur la consommation en raison d'un mécanisme d'épargne de précaution.

À moyen terme, la baisse des prix, plus marquée que celle des salaires, conduit à une amélioration progressive de la consommation des ménages.

VARIANTE EFFICIENCE DU TRAVAIL (2/3) : INTERVALLE DE CONFIANCE



VARIANTE EFFICIENCE DU TRAVAIL (3/3)

	1 an	2 ans	3 ans	5 ans	10 ans
PIB (en volume)	0,04 [0,04; 0,07]	0,15 [0,14; 0,23]	0,27 [0,27; 0,37]	0,44 [0,43; 0,54]	0,62 [0,58; 0,71]
VA du secteur marchand	0,02 [0,01; 0,05]	0,09 [0,08; 0,19]	0,2 [0,2; 0,34]	0,37 [0,35; 0,5]	0,57 [0,51; 0,68]
Consommation des ménages	-0,01 [-0,04; 0,01]	-0,06 [-0,11; 0,02]	-0,06 [-0,1; 0,05]	0 [-0,04; 0,13]	0,17 [0,12; 0,28]
Investissement	0,02 [0,01; 0,05]	0,1 [0,08; 0,2]	0,23 [0,18; 0,37]	0,39 [0,25; 0,49]	0,48 [0,35; 0,54]
Investissement des SNF et EI	0,03 [0,02; 0,08]	0,16 [0,12; 0,32]	0,36 [0,3; 0,57]	0,65 [0,47; 0,78]	0,75 [0,58; 0,84]
Exportations	0,04 [0,03; 0,06]	0,16 [0,14; 0,23]	0,29 [0,27; 0,39]	0,46 [0,43; 0,58]	0,69 [0,6; 0,84]
Importations	0,03 [0,02; 0,06]	0,05 [0,03; 0,17]	0,09 [0,08; 0,24]	0,16 [0,11; 0,28]	0,21 [0,17; 0,28]
Revenu disponible réel des ménages	0,03 [0,03; 0,05]	0,02 [-0,01; 0,08]	-0,01 [-0,06; 0,08]	0 [-0,06; 0,14]	0,19 [0,14; 0,29]
Taux d'épargne	0,04 [0,03; 0,06]	0,07 [0,03; 0,11]	0,04 [0; 0,07]	0 [-0,02; 0,01]	0,02 [0,01; 0,03]
Prix de la consommation des ménages	-0,09 [-0,12; -0,09]	-0,27 [-0,36; -0,27]	-0,43 [-0,55; -0,4]	-0,61 [-0,78; -0,53]	-0,86 [-1,05; -0,7]
Prix de production	-0,11 [-0,14; -0,11]	-0,34 [-0,45; -0,34]	-0,54 [-0,69; -0,5]	-0,75 [-0,94; -0,65]	-0,98 [-1,2; -0,82]
Prix de production marchande	-0,1 [-0,14; -0,1]	-0,31 [-0,41; -0,31]	-0,49 [-0,63; -0,46]	-0,67 [-0,86; -0,59]	-0,91 [-1,13; -0,76]
Prix de VA marchande	-0,11 [-0,13; -0,1]	-0,31 [-0,39; -0,29]	-0,49 [-0,6; -0,43]	-0,67 [-0,84; -0,56]	-0,95 [-1,17; -0,76]
Prix des exportations	-0,06 [-0,09; -0,05]	-0,23 [-0,35; -0,21]	-0,39 [-0,56; -0,35]	-0,56 [-0,74; -0,48]	-0,74 [-0,93; -0,63]
Prix des importations	-0,09 [-0,15; -0,09]	-0,24 [-0,4; -0,26]	-0,38 [-0,59; -0,4]	-0,53 [-0,75; -0,51]	-0,71 [-0,89; -0,62]
Salaire nominal brut	-0,04 [-0,07; -0,03]	-0,16 [-0,28; -0,12]	-0,28 [-0,48; -0,18]	-0,45 [-0,74; -0,22]	-0,49 [-0,75; -0,24]
Coût réel du travail	0,07 [0,05; 0,09]	0,15 [0,07; 0,2]	0,2 [0,07; 0,29]	0,22 [0,05; 0,38]	0,46 [0,34; 0,57]
Emploi salarié (en milliers)	-15,08 [-18,49; -10,79]	-59,29 [-68,29; -40,33]	-77,4 [-84,2; -48,64]	-54,62 [-61,07; -25,02]	-33,37 [-46,91; -12,67]
Taux de chômage	0,06 [0,05; 0,08]	0,25 [0,17; 0,28]	0,32 [0,2; 0,35]	0,23 [0,1; 0,25]	0,14 [0,05; 0,2]
Balance commerciale (en pts de PIB)	0,01 [0; 0,03]	0,04 [0,01; 0,08]	0,07 [0,03; 0,11]	0,1 [0,06; 0,17]	0,16 [0,12; 0,22]
Taux de marge (EBE/VA)	0,03 [0,01; 0,06]	0,21 [0,17; 0,3]	0,34 [0,26; 0,46]	0,35 [0,23; 0,46]	0,22 [0,14; 0,3]
Déficit primaire (en pts de PIB)	0,04 [0,03; 0,05]	0,11 [0,06; 0,12]	0,09 [0,02; 0,09]	0,01 [-0,05; 0,02]	-0,05 [-0,09; -0,02]

Tableau 1 – Hausse de 1% de l'efficience du travail (en % d'écart au compte central)

Retour sur le passé : l'exemple de la Grande Récession

Une « relecture du passé » ?

- La possibilité de comparer sur le passé la trajectoire observée et des simulations historiques du modèle (sur la même période), par rapport à une incertitude analysée en projection
- Pendant la crise de 2008, des chocs de grande ampleur affectant l'ensemble des exogènes du modèle ...
- ... conduisant à une politique de relance massive en France et dans la Zone Euro.

Intérêt : la période 2008 – 2014 est incluse dans la plage d'estimation du modèle (contrairement à la période actuelle!)

Enjeu : construire des intervalles de *fluctuation* autour de la trajectoire historique afin de mesurer la robustesse des simulations du modèle (sensibilité aux aléas)

[Fair, 2003] : analyse systématique de l'incertitude des résultats d'un modèle macro-économique par *bootstrap*

Un matériel déjà dans notre boîte à outils!

En pratique :

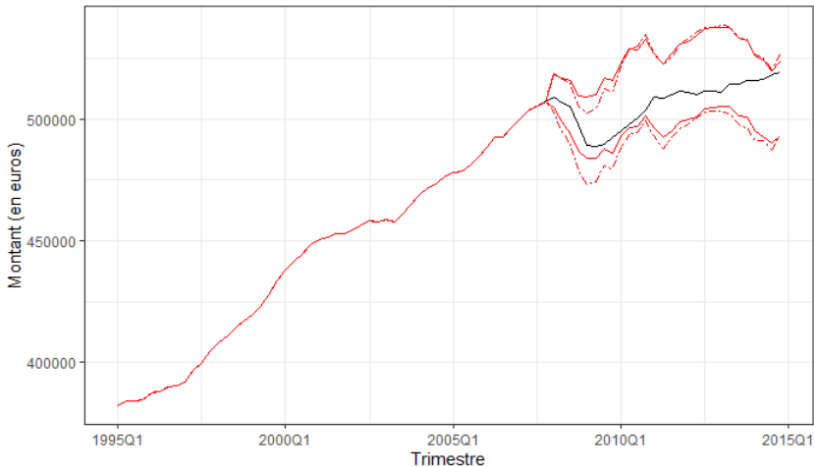
- on observe la trajectoire des **variables exogènes et endogènes du modèle** jusqu'en 2014T4,
- on estime l'**ensemble des coefficients du modèle**, présents dans les équations de comportement,
- on calcule les **résidus sur la période historique**, par inversion du modèle.

Une seconde application du bootstrap

- On dispose toujours des jeux de coefficients calculés lors de l'étape de rééchantillonnage
- On effectue 1000 simulations de trajectoires entre le T1 2008 et le T4 2014, en tirant avec remise les résidus dans cet intervalle
- Des intervalles de fluctuation à 95 % sont calculés pour les principales variables du modèle
- Les simulations peuvent être réalisées avec les coefficients « historiques » ou avec les jeux de coefficients réestimés (pour tenir compte de leur distribution)

EXEMPLE : INTERVALLE DE FLUCTUATION DU PIB SUR 2008 – 2014

Analyse du modèle sur 2008T1 - 2014T4 (variable : PIB en volume)



Plusieurs enseignements


- Des résultats **proches** pour les deux types d'intervalles (avec ou sans prise en compte de la variabilité des coefficients)
- Des résultats **hétérogènes** selon les variables : la volatilité des prix augmente de façon plus importante que celle des variables en volume
- La présence de trajectoires **hors des bandes de confiance** souligne le caractère exceptionnel de cette crise par rapport aux trajectoires simulées
- En fin de période, le rattrapage de l'économie **a placé la trajectoire observée dans le haut de la fourchette** pour le PIB et la consommation des ménages : est-il possible de distinguer l'effet des politiques publiques d'une évolution structurelle de l'économie ?

Conclusion et perspectives

1. Des résultats originaux relatifs à **la prise en compte de l'aléa** présent dans les simulations de Mésange
2. Une méthode simple à mettre en oeuvre fondée sur un algorithme de **bootstrap**
3. Une **analyse préliminaire** des résidus économétriques pour la validité de la méthode
4. Des **conclusions opérationnelles** : vitesse de diffusion de l'effet d'un choc à certaines variables, significativité des effets à moyen-terme
5. La construction d'**intervalles de fluctuation** pendant la crise analyse la robustesse du modèle


MERCI!



-  Bardaji, J., Campagne, B., Khder, M.-B., Lafféter, Q., Simon, O., Dufernez, A.-S., Elezaar, C., Leblanc, P., Masson, E., and Partouche, H. (2017).

Le modèle macroéconométrique Mésange : réestimation et nouveautés.

Technical report, INSEE.

-  Barlet, M., Clerc, M.-E., Garnero, M., Lapègue, V., and Marcus, V. (2011).

La nouvelle version du modèle mze, modèle macroéconométrique pour la zone euro.

Document de travail INSEE.

-  Bolatbayeva, A., Tolepbergen, A., and Abilov, N. (2020).

A macroeconomic model for russia.

Russian Journal of Economics, 6 :114.



Efron, B. (1979).

Bootstrap methods : Another look at the jackknife.

The Annals of Statistics, pages 1–26.



Fair, R. C. (2003).

Bootstrapping macroeconomic models.

Studies in Nonlinear Dynamics & Econometrics, 7(4).



Gonzalez-Astudillo, M. and Vilan, D. (2019).

A new procedure for generating the stochastic simulations in frb/us.

Technical report, Board of Governors of the Federal Reserve System (US).

Applications du bootstrap aux modèles macroéconométriques

- [Efron, 1979] Article fondateur sur la technique du *bootstrap*
- [Barlet et al., 2011] Mise en oeuvre du *bootstrap* afin de calculer des intervalles de confiance pour les IRF du modèle zone euro (MZE) de l'Insee
- [Fair, 2003] Description détaillée des application du *bootstrap* aux modèles macroéconomiques (y.c. retour sur le passé)
- [Bolatbayeva et al., 2020] Modèle structurel de l'économie russe comportant une analyse de la volatilité de ses résultats par *bootstrap*
- [Gonzalez-Astudillo and Vilan, 2019] Description de la construction des intervalles de confiance en prévision pour le modèle FRB/US

Du résidu à la cale

Forme générique d'une équation de comportement :

$$\Phi(L)y_t = \Psi(L)x_t + \varepsilon_t \quad (3)$$

Avec $cale_t = \Phi(L)^{-1}\varepsilon_t$ (fonction des valeurs passées des résidus) :

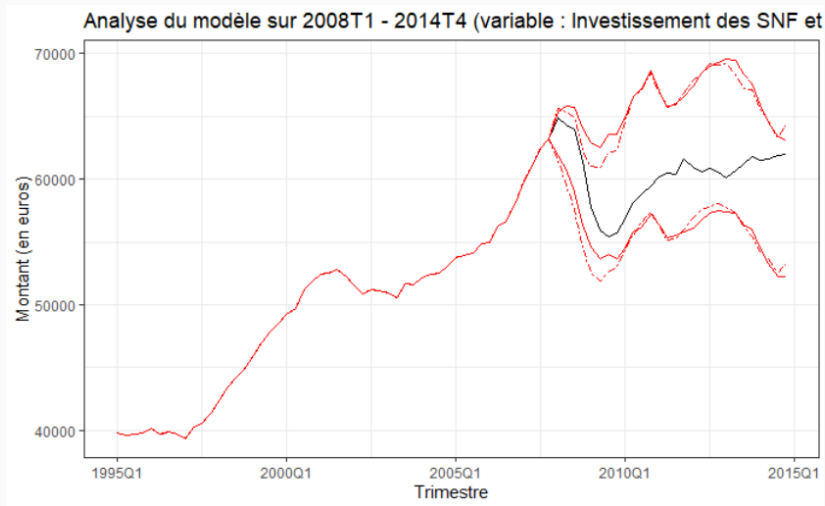
$$y_t = \Phi(L)^{-1}\Psi(L)x_t + cale_t \quad (4)$$

→ les variations de y_t peuvent être décomposées en fonction des seules valeurs – passées et présentes – des exogènes x_s et de la cale.

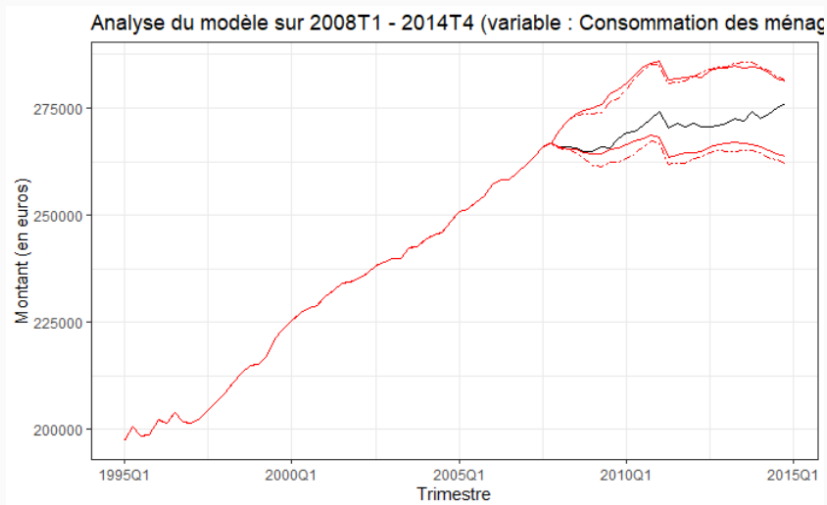
Construire un scénario contrefactuel

- **simuler sur le passé le modèle** en figeant l'ensemble des valeurs des cales des ECM à une date antérieure à la plage de relecture du passé (par exemple celle du T4 2007)
- → la contribution de l'inexpliqué est nulle. **Comparer la simulation et l'observation** permet de distinguer ce qui est expliqué par le modèle des effets qui échappent à la modélisation
- **analyser la volatilité** de ce scénario en construisant un intervalle de confiance, et le comparer avec la trajectoire historique

Annexe : investissement des SNF



Annexe : consommation des ménages

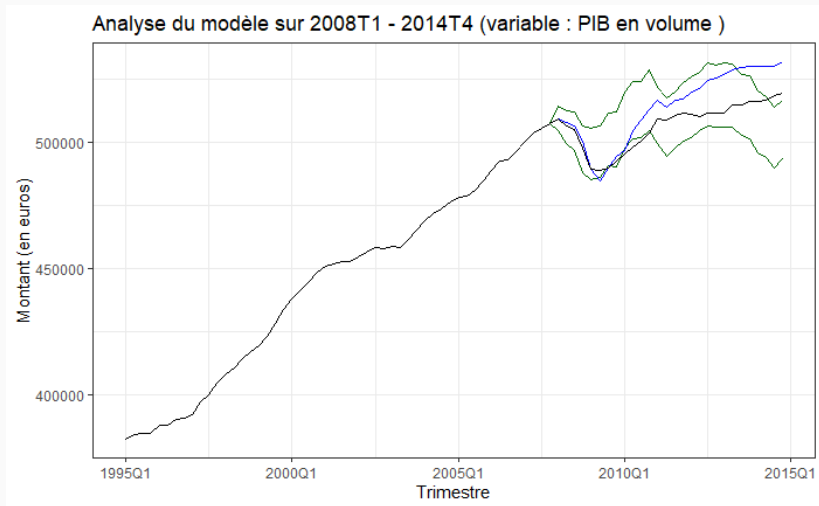


Une évolution uniquement déterminée par les exogènes du modèle

- **Construire un scénario contrefactuel** avec une contribution nulle de l'inexpliqué dans les ECM à partir du T1 2008. En d'autres termes, on postule que la crise n'affecte pas, dans ce scénario, les propriétés explicatives des équations économétriques.
- **Robustesse de cette approche** analysée par rééchantillonnage avec tirage des résidus dans la période 2001T1 – 2007T4 → ne pas utiliser les résidus des trimestres de crise.

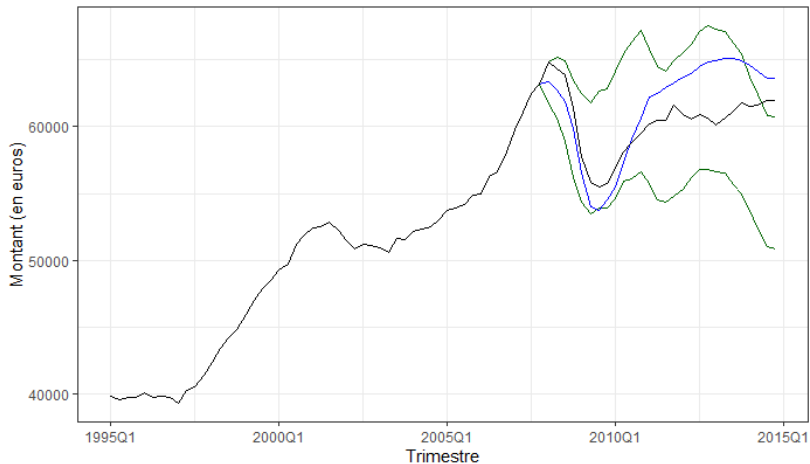
Remarque : dans cette analyse, la crise affecte uniquement les résultats des simulations à travers les exogènes (environnement international et monétaire ...)

Annexe : scénario contrefactuel et robustesse (PIB)



Annexe : Investissement des entreprises

Analyse du modèle sur 2008T1 - 2014T4 (variable : Investissement des SNF et



Annexe : exportations

Analyse du modèle sur 2008T1 - 2014T4 (variable : Exportations)

