

Étalonnages à l'aide d'enquêtes de conjoncture : de nouveaux résultats

Éric Dubois^(*)

Emmanuel Michaux^(*)

Les étalonnages sont des outils-clés pour l'analyse conjoncturelle et la prévision à court terme. Parmi ces étalonnages, ceux se rapportant à la production manufacturière font l'objet depuis longtemps d'une attention toute particulière en France. Il est toutefois encore possible d'apporter des améliorations méthodologiques à la construction de ces étalonnages. Cette note en propose deux.

Tout d'abord, la méthode usuelle qui consiste, pour prévoir la production d'un trimestre, à utiliser l'ensemble des enquêtes du trimestre paraît inutilement restrictive : elle impose, soit d'attendre que toutes les enquêtes d'un trimestre soient parues pour faire une prévision, soit – c'est ce qui est fait en pratique – de prévoir les enquêtes manquantes. Mais cette dernière méthode conduit à ajouter, de manière non transparente, l'incertitude liée à l'outil de prévision des enquêtes à celle inhérente aux étalonnages.

On se propose donc dans ce travail d'estimer pour chaque trimestre autant d'étalonnages qu'il y a de mois dans le trimestre, en utilisant à chaque fois la seule information disponible dans les enquêtes parues jusqu'alors. Cette méthode a l'avantage d'une plus grande transparence et d'une utilisation statistiquement plus précise de l'information disponible. Elle a l'inconvénient de contraindre à changer d'étalonnage chaque mois, mais, pour limiter la portée de cet inconvénient, nous proposons une méthode qui permet d'analyser les modifications de prévisions entre deux enquêtes en fonction d'innovations, correctement définies, apportées par ces enquêtes.

Ensuite, alors que les étalonnages utilisés jusqu'à présent reposent sur une sélection "artisanale" des variables, qui conduit en pratique à ne retenir qu'un petit nombre de données d'enquête – par rapport à la totalité des soldes d'opinion disponibles – nous mettons en œuvre une méthode d'estimation plus "systématique", reposant sur un algorithme proposé récemment par Krolzig et Hendry permettant de reproduire "automatiquement" la méthodologie économétrique de la London School of Economics. Cette méthode nous permet de tester l'apport en prévision de nombreux soldes d'enquête, à un moindre coût et de manière là encore totalement transparente.

Les résultats obtenus sont globalement satisfaisants. On constate tout d'abord, à mesure que l'on avance dans le trimestre, une amélioration de la précision des prévisions. On constate également que cette méthode se compare avantageusement avec d'autres formes de modélisation (VAR) et qu'elle a, au moins à l'horizon du trimestre en cours, des performances très proches de celles des conjoncturistes de l'Insee.

(*) Direction de la Prévision et de l'Analyse Économique au moment de la rédaction de l'article.
Emmanuel Michaux était économiste au sein de la DGTPE. Il est désormais stratégeste chez HSBC Halbis Partners.
E-mail : eric.dubois@ensae.fr

Nous tenons à remercier M. Bodier, F. Sédillot et un rapporteur anonyme pour leurs précieux commentaires.

Les étalonnages à l'aide d'enquêtes de conjoncture constituent un outil essentiel de prévision à court terme. Pourtant, les résultats disponibles en la matière suscitent un certain nombre d'interrogations :

– les enquêtes sont mensuelles, alors que la variable à prévoir – la production manufacturière ou le PIB – est trimestrielle. Selon le mois du trimestre où l'on se place, les enquêtes disponibles ne sont donc pas les mêmes et les étalonnages devraient donc différer selon le mois du trimestre. Pourtant, les résultats publiés (cf. Scherrer, 1996, ou plus récemment encore Bouton et Erkel-Rousse, 2002) ne présentent pas d'étalonnages différents selon le mois du trimestre auquel on se réfère. Est-ce à dire qu'il n'est pas possible de prévoir correctement la production manufacturière avant de disposer de l'ensemble des enquêtes du trimestre ? ;

– les étalonnages publiés utilisent un très petit nombre de variables disponibles dans l'enquête. Est-ce que les autres variables n'apportent aucune information ? ;

– enfin, même s'il est acquis que les enquêtes de conjoncture dans l'industrie sont corrélées avec l'évolution de la production, quelle est leur réelle valeur ajoutée ?

Le présent travail apporte des éléments de réponse à ces trois questions à partir des étalonnages de la production manufacturière utilisant les enquêtes de conjoncture dans l'industrie. Dans un premier temps, on effectue, pour chaque mois du trimestre, des régressions reliant l'évolution de la production manufacturière aux données mensuelles d'enquêtes disponibles le mois considéré : ces données concernent l'opinion sur la production passée et les perspectives personnelles de production, mais aussi l'opinion sur les carnets de commandes et le jugement sur les stocks. La technique économétrique utilisée pour mener ces régressions est nouvelle : elle repose sur un algorithme proposé par Krolzig et Hendry (2000) pour reproduire la méthodologie économétrique généralement associée à la *London School of Economics* (LSE).

À l'issue de cette investigation, on apporte les réponses suivantes aux trois questions qu'on s'était posées : premièrement, il est possible de prévoir avec une précision satisfaisante la production manufacturière avant de disposer de l'ensemble des enquêtes du trimestre et sans faire d'hypothèse sur les mois d'enquêtes encore inconnus ; deuxièmement, l'opinion sur les carnets de commandes apporte de l'information utile à la prévision de la production manufacturière ; troisièmement, comparé à des modèles VAR, l'apport des enquêtes de conjoncture est incontestable.

Enfin, une stratégie de modélisation axée sur la production dans le secteur manufacturier et non sur le PIB peut paraître, à première vue, de portée limitée. Néanmoins, malgré son champ réduit – le

secteur manufacturier ne représente que 20% du PIB –, ce secteur contribue à lui seul à plus de 90%⁽¹⁾ des fluctuations (variance) du PIB. De plus, le recours à cette stratégie garantit une cohérence maximale entre le champ couvert par la variable à prévoir et celui relatif aux variables exogènes utilisées.

Les limites de la méthode usuelle

On se propose ici d'illustrer les limites de la méthode usuelle à partir de l'exemple de Scherrer (1996), qui estime, comme dans ce travail, des étalonnages de la production manufacturière à partir de l'enquête mensuelle de conjoncture dans l'industrie de l'Insee. Au travers de cette enquête, les industriels sont interrogés tous les mois par l'Insee sur l'évolution de leur production au cours des trois derniers mois, leurs perspectives d'activité au cours des trois ou quatre prochains mois, le niveau de la demande, globale et étrangère, qui leur est adressée, le niveau de leurs stocks de produits fabriqués, l'évolution prévue de leurs prix de vente, ainsi que les évolutions de la production et des prix concernant l'ensemble de l'industrie.

Pour prévoir l'évolution de la production manufacturière des trimestres passé et à venir, Scherrer (1996) propose deux étalonnages :

– le premier, "retardé", repose sur l'opinion des industriels quant à l'évolution de leur production passée, en moyenne sur le premier mois du trimestre et les deuxième et troisième mois du trimestre précédent ;

– le deuxième, "coïncident", utilise à la fois l'indicateur précédent et l'opinion des industriels sur leur production passée et leur production prévue issue de l'enquête de conjoncture du premier mois du trimestre ; cet étalonnage est utilisé pour prévoir la production du trimestre à venir.

Le problème des données d'enquête manquantes.

Ces étalonnages supposent donc connues toutes les enquêtes se rapportant à un trimestre. Au-delà de l'exemple traité ici, c'est une pratique assez générale en France : par exemple, Buffeteau et Mora (2000) ou Bouton et Erkel-Rousse (2002) présentent des étalonnages de PIB trimestriel utilisant l'ensemble des données d'enquêtes mensuelles d'un trimestre donné. Lorsque des données d'enquêtes sont manquantes, par exemple si l'on souhaite faire en mai une prévision pour le deuxième trimestre, l'utilisation de l'étalonnage suppose donc de prévoir les soldes d'enquêtes correspondants pour les mois de juin et juillet.

Cette manière de procéder présente de sérieux inconvénients. Tout d'abord, rien ne garantit que les

hypothèses retenues pour prévoir les soldes manquants utilisent au mieux l'information disponible. Prenons l'exemple de la production manufacturière et supposons – ce qui ne paraît pas absurde – que, le premier mois de chaque trimestre, les perspectives personnelles de production représentent la meilleure prévision que puissent faire les industriels (et donc l'économètre), mais qu'à la fin du trimestre, ce soit le solde d'opinion sur la production passée qui constitue le meilleur indicateur de la production du trimestre. Dans ces conditions, tenter de prévoir, en début de trimestre, la production à l'aide d'un étalonnage basé sur la production passée revient au mieux à prendre un détour inutile mais, plus vraisemblablement, à ne pas utiliser correctement l'information pertinente.

Exprimé en termes plus statistiques, la meilleure prévision, au sens de l'erreur quadratique moyenne, qui puisse être faite à partir des enquêtes à un moment donné est l'espérance conditionnelle de la production manufacturière relativement aux soldes d'enquête disponibles à ce moment-là. Par conséquent, la meilleure prévision qui puisse être faite, par exemple fin mai, de la production du trimestre, doit être basée sur une régression reliant la production trimestrielle à la seule information disponible dans les enquêtes de mai.

Les problèmes liés au choix des variables explicatives

De nombreux étalonnages reposent sur un nombre restreint de variables exogènes. C'est le cas par exemple de ceux retenus par Scherrer, dont les seules variables explicatives sont la tendance passée de la production et les perspectives personnelles de production. Or on ne peut exclure que l'information contenue par exemple dans l'opinion sur les carnets de commandes soit plus fiable, ou plus objective, parce qu'elle nécessite de la part des industriels une évaluation plutôt qu'une anticipation, que l'information retracée par les perspectives personnelles de production.

Par ailleurs, l'opinion sur la production passée “ au cours des trois mois qui précèdent ” est formulée de manière suffisamment vague pour ne pas correspondre nécessairement au taux de croissance

trimestriel (c'est-à-dire la moyenne des trois mois du trimestre rapportée à la moyenne des trois mois du trimestre précédent). Intuitivement elle correspond plutôt au glissement trimestriel (évolution du troisième mois du trimestre par rapport au troisième mois du trimestre précédent). Cette considération incite à utiliser des éléments des enquêtes précédentes pour capter les évolutions plus anciennes de la production industrielle et à laisser l'économétrie décider du poids à accorder à chaque enquête.

Aussi convient-il d'examiner si un usage plus systématique des informations d'enquête, n'imposant pas *a priori* de contraintes fortes, peut permettre d'améliorer la qualité de l'estimation et si les estimations ainsi obtenues ont un comportement plus satisfaisant au cours des années récentes.

La démarche suivie

Les arguments qui précèdent invitent à mettre en œuvre une approche plus systématique qui incorpore, chaque mois, l'information apportée par la nouvelle enquête de conjoncture et sélectionne *a posteriori* les variables pertinentes parmi l'ensemble des informations alors disponibles. La méthode consiste à découper, en fonction de la place du mois dans le trimestre, chaque série mensuelle en trois nouvelles séries trimestrielles. Ainsi, nous obtenons une première série trimestrielle composée des valeurs prises par la série initiale lors de chaque premier mois dans le trimestre, une deuxième série trimestrielle regroupant les valeurs pour le deuxième mois⁽²⁾, etc.

L'examen de la séquence des publications des Comptes Trimestriels et des enquêtes de conjoncture révèle (cf. tableau 1) que, si l'on se limite à l'horizon du trimestre à venir⁽³⁾, on peut faire, grâce au découpage des données, les 7 prévisions qui suivent⁽⁴⁾ :

– à la fin du premier mois du trimestre, on peut prévoir, avec l'enquête de conjoncture de fin de mois et celles qui précèdent, la croissance de la production

Tableau 1 : calendrier des publications des Comptes Trimestriels et de l'enquête mensuelle de conjoncture de l'Insee (exemple du 1^{er} trimestre de 2002)

| | |
|----------------|---|
| Janvier | |
| 11 | Comptes trimestriels – Résultats détaillés du 3 ^{ème} trimestre de 2001 (*) Enquête mensuelle de conjoncture dans l'industrie de janvier 2002 |
| 30 | |
| Février | |
| 22 | Comptes trimestriels – Premiers résultats du 4 ^{ème} trimestre de 2001 Enquête mensuelle de conjoncture dans l'industrie de février 2002 |
| 27 | |
| Mars | |
| 28 | Enquête mensuelle de conjoncture dans l'industrie de mars 2002 |

(*) depuis juin 2003, les résultats détaillés des comptes trimestriels sont publiés une dizaine de jours plus tôt ; ceci est sans incidence sur les résultats présentés ici, dans la mesure où les enquêtes sont utilisées pour prévoir des trimestres pour lesquels les comptes trimestriels ne sont pas encore publiés (donc pour lesquels manquent les premiers résultats).

manufacturière du trimestre précédent, du trimestre en cours et du trimestre à venir ;

– à la fin du deuxième mois du trimestre, on peut prévoir, avec l'enquête de conjoncture de fin de mois et celles qui précèdent, la croissance de la production manufacturière des trimestres en cours et à venir (la production du trimestre précédent est désormais connue) ;

– de même, à la fin du troisième mois du trimestre, on peut faire de nouvelles prévisions des trimestres en cours et à venir.

Les informations utilisées pour chacune de ces prévisions sont les suivantes :

– taux de croissance trimestriels passés de la production manufacturière ;

– opinion des industriels sur leur production passée, leurs perspectives personnelles de production, leur opinion sur leurs carnets de commande globaux et leur opinion sur leur niveau de stocks de produits finis.

Les informations suivantes, disponibles dans l'enquête de conjoncture dans l'industrie manufacturière, n'ont en revanche pas été utilisées :

– les perspectives générales de production, qui sont réputées représenter plus un indicateur " d'ambiance " qu'une véritable prévision ;

– l'opinion sur les carnets de commandes étrangers qui, de par son champ, est *a priori* un indicateur moins pertinent que les carnets de commande globaux.

Ces informations ont été utilisées avec les retards suivants :

– 3 retards pour le taux de croissance de la production manufacturière ;

– 5 retards pour chacune des variables de l'enquête de conjoncture retenues.

Ces retards ont été choisis pour préserver un nombre de degrés de liberté suffisant. Que le nombre de retards sur l'enquête de conjoncture soit constant quel que soit le mois considéré peut surprendre. Une stratégie *a priori* plus naturelle aurait consisté à ajouter, à l'ensemble d'information utilisé le premier mois, les résultats de l'enquête du deuxième, puis du troisième mois. La nécessité de conserver au troisième mois un nombre suffisant de degrés de liberté aurait cependant conduit à retenir un ensemble d'information relativement réduit pour le premier mois. Le gain de précision qu'on constatera en général en passant du premier au deuxième mois, puis du deuxième au troisième mois, permettra de confirmer que la perte d'information consécutive au retrait successif des résultats de deux enquêtes de conjoncture au profit des résultats de l'enquête la plus récente n'est pas importante.

Partant de cet ensemble d'information large, la démarche habituelle, *artisanale*, consiste à éliminer successivement des variables non significatives jusqu'à obtenir une spécification où ne restent que des variables significatives, quitte à rajouter le cas échéant des variables éliminées si la spécification obtenue ne passe pas les tests de spécification.

Plutôt que cette démarche artisanale, on a préféré recourir ici à une démarche plus systématique, proposée par Hoover et Perez (1999) et améliorée par Krolzig et Hendry (2000)⁽⁵⁾. Cette démarche consiste, à partir d'un modèle général, à procéder, selon plusieurs chemins possibles, à l'élimination successive de variables non significatives, à condition qu'un certain nombre de tests de spécification soient à chaque étape acceptés. Lorsque cette démarche conduit à plusieurs modèles possibles, ceux-ci sont sélectionnés selon leur degré d'acceptation vis-à-vis de leur union et, le cas échéant, en fonction d'un critère de sélection de type AIC, BIC ou HQ (cf. diagrammes 1 à 3 pour un exposé schématique de la méthode et l'annexe pour une présentation détaillée de cette méthode). Dans le cas où le modèle final est obtenu à l'aide d'un critère d'information, cela veut toutefois dire que plusieurs modèles ont une qualité d'ajustement très proche : s'il est toujours possible de se limiter à une approche totalement automatisée et de s'en tenir au modèle final sélectionné par le critère d'information choisi, un regard d'économiste peut toutefois être utile et éventuellement conduire à en choisir un autre parmi ces modèles de qualité d'ajustement très proche.

Cette démarche permet de tester un nombre important de modèles potentiellement pertinents, sans avoir à estimer l'ensemble des modèles possibles à partir du modèle général. Si le nombre de variables qu'on souhaite inclure dans l'estimation est de 20 (ce qui est *a priori* relativement peu), le nombre de modèles qui peuvent être construits à partir de ces 20 variables est en effet de 2^{20} , soit plus d'un million : le coût d'estimation de tous ces modèles et du choix entre ceux-ci serait prohibitif.

La démarche proposée permet donc tout à la fois de limiter le coût de sélection du modèle final, donc de bénéficier d'un gain de temps appréciable par rapport aux méthodes plus artisanales, tout en explorant un grand nombre de modèles possibles, plus grand que celui qu'on obtient par des méthodes artisanales. Krolzig et Hendry (2000) ont montré que, à condition de choisir correctement les tests de spécification et leur seuil d'acceptation, et de partir d'un ensemble d'information qui contienne le véritable modèle, cette démarche permet de retrouver le " bon modèle " à un seuil proche d'une valeur définie *a priori*⁽⁶⁾.

Diagramme 1 : estimation et réduction du modèle initial

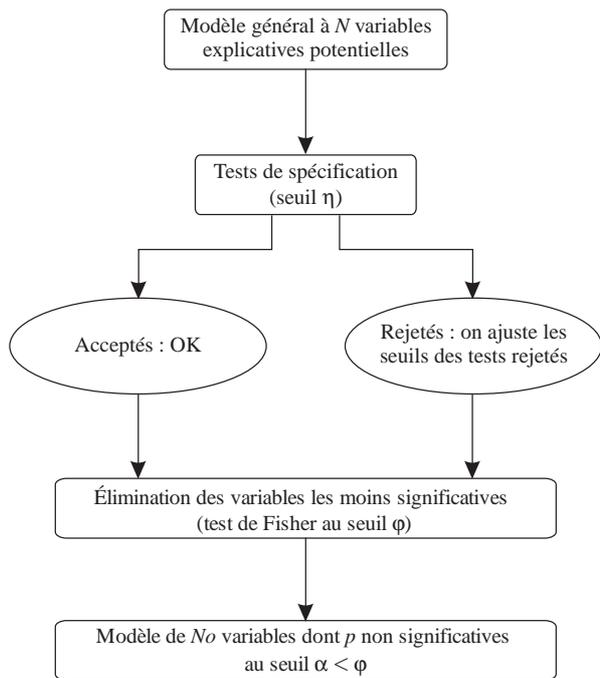
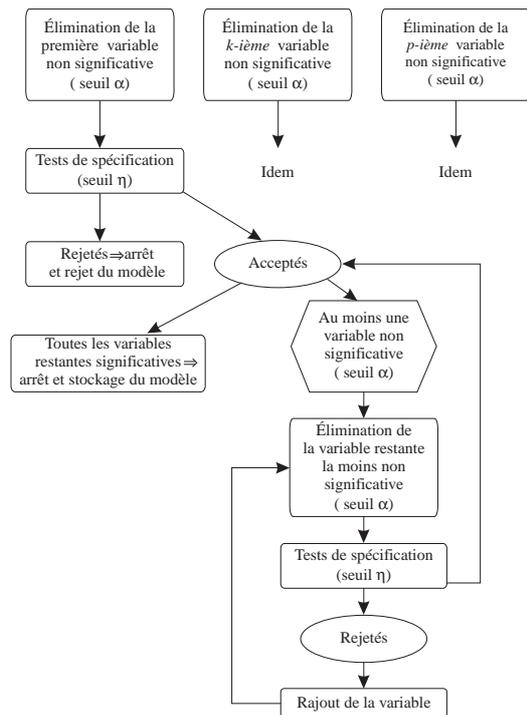
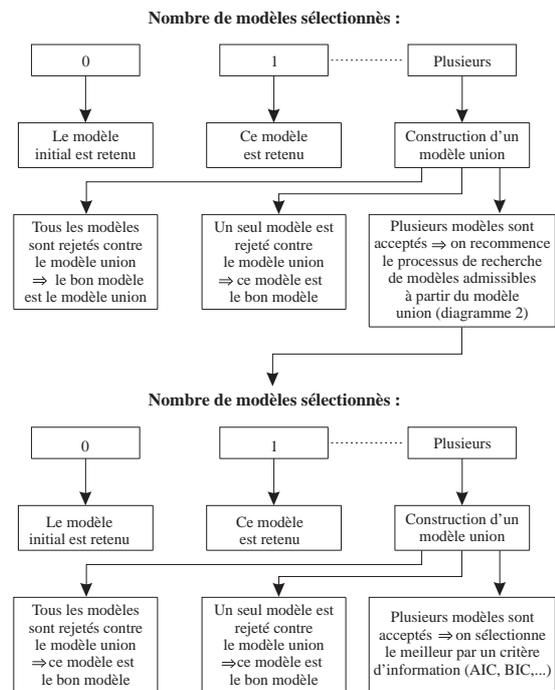


Diagramme 2 : recherche des modèles réduits qui sont admissibles par simplifications successives



En présence de multicollinéarité importante, l'algorithme aura toutefois du mal à détecter le bon modèle : il s'agit là d'une difficulté bien connue, qui n'est pas liée à l'algorithme lui-même, mais aux données. C'est un problème que nous n'avons toutefois pas rencontré ici : à chaque étape du processus de sélection des variables, les modèles ne présentaient pas de multicollinéarité forte et celle-ci est totalement absente de l'ensemble des spécifications obtenues.

Diagramme 3 : détermination du modèle final unique



Notons enfin que cette démarche favorise la transparence : dès que l'ensemble d'information initial, les tests de spécification et les divers seuils statistiques utilisés pour sélectionner les différents modèles sont spécifiés, l'ensemble de la démarche conduisant au modèle finalement retenu peut être reproduite par le lecteur, ce qui n'est pas le cas lorsque l'estimation obtenue résulte de l'application raisonnée d'un *savoir-faire* spécifique.

Les résultats d'estimation

L'algorithme de sélection automatique de modèles a été ici mis en œuvre avec les paramètres suivants :

- les tests de spécification retenus : il s'agit du test du multiplicateur de Lagrange d'auto-corrélation des résidus à l'ordre 5 (Godfrey, 1978), du test de normalité de Doornik et Hansen (1994), du test d'hétéroscédasticité quadratique entre les régresseurs (Nicholls et Pagan, 1983), enfin du test d'échec prédictif de Chow sur respectivement 50% et 90% de la période (Hendry, 1979). Il s'agit de la batterie des tests préconisée par Krolzig et Hendry (2000) ;

- les seuils des tests : ils ont été posés à 5% pour le test de significativité des coefficients (seuil de l'étape 1), à 1% pour les tests de spécification (seuil de l'étape 1 ; il s'agit du seuil recommandé par Krolzig et Hendry, qui se traduit par un rejet à tort du vrai modèle environ 5% du temps), enfin à 5% pour les tests de Fischer de sélection des modèles (seuil de l'étape 4).

La manière dont l'ensemble d'informations a été traité est également importante. Même s'il s'agit à chaque fois de projeter, *au début du processus*, la variable expliquée sur le même espace, la procédure d'élimination des variables non significatives, et donc de sélection des modèles, peut être différente selon que l'on aura par exemple retenu les variables χ_1 et χ_2 ou χ_1 et $\chi_2 - \chi_1$.

Ici, pour chaque *item* de l'enquête mensuelle de conjoncture, on a choisi systématiquement de retenir la donnée de l'enquête la plus récente en niveau et les données des enquêtes précédentes en écart à l'enquête du mois précédent⁽⁷⁾ : on a donc à chaque fois la dernière variable de l'enquête en niveau et les "informations" nouvelles apportées par chaque enquête. Concrètement, l'ensemble d'informations utilisé pour effectuer la régression du premier mois du trimestre est ainsi, avec les notations de l'encadré 1 :

$\{y(-1); y(-2); y(-3); ypa_m1; ypa_m1-ypa_m3(-1); ypa_m3(-1)-ypa_m2(-1); ypa_m2(-1)-ypa_m1(-1); ypa_m1(-1)-ypa_m3(-2); ppy_m1; ppy_m1-ppy_m3(-1); ppy_m3(-1)-ppy_m2(-1); ppy_m2(-1)-ppy_m1(-1); ppy_m1(-1)-ppy_m3(-2); ccg_m1; ccg_m1-ccg_m3(-1); ccg_m3(-1)-ccg_m2(-1); ccg_m2(-1)-ccg_m1(-1); ccg_m2(-1)-ccg_m3(-2); sto_m1; sto_m1-sto_m3(-1); sto_m3(-1)-sto_m2(-1); sto_m2(-1)-sto_m1(-1); sto_m1(-1)-sto_m3(-2); constante\}$

Encadré 1 : notations

Les variables tirées de l'enquête mensuelle dans l'industrie sont notées : xxx_mi où :

- $xxx = ypa$ (opinion sur la production passée)
- = ppy (perspectives personnelles de production)
- = sto (opinion sur les stocks de produits finis)
- = ccg (carnets de commande globaux)

- $i = n^\circ$ du mois considéré (1 = janvier, avril, juillet, octobre ; 2 = février, mai, août, novembre ; 3 = mars, juin, septembre, décembre).

Lorsque la production qu'on cherche à prévoir est celle du trimestre courant, elle sera appelée y ; si c'est celle du trimestre précédent, elle sera appelée y_p ; si c'est celle du trimestre à venir, elle sera appelée y_f .

Enfin, la période d'estimation va du deuxième trimestre de 1979 (première date possible compte tenu de la date de démarrage des comptes trimestriels en base 1995 et du nombre de retards retenu pour y) au quatrième trimestre de 2002.

Les estimations (tableaux 2 à 4) confirment que chaque nouvelle enquête apporte une information nouvelle pour la prévision du trimestre courant : au fur et à mesure que le temps passe, l'écart type des

résidus (mesure de la précision des estimations) passe de 0,86% à 0,77% et des variables de la nouvelle enquête deviennent significatives. Pour l'estimation du trimestre à venir, les résultats sont moins encourageants : l'écart type de la prévision est nettement plus élevé et augmente, passant de 0,97% à 0,99%, entre le premier et le deuxième mois, avant de diminuer de nouveau légèrement, de 0,99% à 0,96% entre le deuxième mois et le troisième mois.

Cette dégradation de la qualité de l'ajustement trouve son origine dans la suppression d'une variable lors du passage du premier au deuxième mois dans le trimestre. En effet, les étalonnages du premier et du deuxième mois ont des structures similaires : ils se différencient par (i) le remplacement de l'opinion sur les carnets de commandes du premier mois par ceux du deuxième et par (ii) la présence dans l'étalonnage, donc dans l'ensemble d'information de départ, du premier mois et non du deuxième, de la variable $ccg_m1(-1)-ccg_m3(-2)$. Le remplacement des carnets de commandes du premier mois par ceux du deuxième, outre qu'il paraît intuitif, ne peut pas conduire à une perte d'information dans la mesure où si les carnets du premier mois possédaient réellement une capacité prédictive supérieure, ils seraient sélectionnés par l'algorithme *via* la variable $ccg_m2 - ccg_m1$. C'est donc la "perte" de $ccg_m1(-1) - ccg_m3(-2)$ dans le passage d'un mois à l'autre qui est à l'origine de cette dégradation de la performance prédictive.

Lorsque cette variable est réintégrée dans l'ensemble d'information de départ du deuxième mois, on retrouve parmi les modèles *in fine* sélectionnés par le critère d'information, un modèle (que l'on appellera *modèle alternatif* dans la suite de l'article) qui inclut cette variable. Ce modèle possède un écart type des résidus plus faible (0,97%) que celui de l'étalonnage initialement retenu (0,99%). Avec le critère BIC retenu ici conformément aux préconisations d'Hendry et Krolzig, ce modèle n'est toutefois pas sélectionné comme modèle final mais il le serait si l'on avait recours au critère AIC ou HQ (cf. tableau 5). Le critère BIC est en effet le critère qui pénalise le plus l'accroissement du nombre de variables exogènes et le modèle *alternatif* comprend une variable de plus que le modèle retenu. Dans une perspective opérationnelle, il est d'ailleurs sans doute préférable de retenir ce modèle *alternatif* comme modèle final, en raison de sa plus grande proximité avec le modèle sélectionné pour le mois qui précède (cf. *infra* pour des développements supplémentaires sur ce thème) ; dans cette configuration il n'y a pas de détérioration de la performance prédictive des étalonnages lors du passage du premier au deuxième mois.

Tableau 2 : résultats de l'estimation du premier mois du trimestre pour le trimestre précédent

| | Modèle |
|--|-------------------|
| $y_p(-1)$ | -0,18 (-2,04)* |
| y_{pa_m1} | 0,05 (7,91) |
| $y_{pa_m2(-1)} - y_{pa_m1(-1)}$ | 0,11 (5,91) |
| $y_{pa_m1(-1)} - y_{pa_m3(-2)}$ | 0,07 (3,30) |
| Constante | 0,58 (6,46) |
| \bar{R}^2 | 0,60 |
| Écart type des résidus | 0,75 |
| Indice de conditionnement ⁽⁸⁾ | 24 |
| Durbin-Watson | 1,83 |
| Chow (50%) | 0,88 [0,67]** |
| Chow (90%) | 0,61 [0,79] |
| Autocorrélation (ordre 5) | 1,83 [0,11] |
| Normalité | 0,48 [0,78] |
| Hétéroscédasticité | 1,31 [0,24] |

(*) Entre parenthèses les t -statistiques

(**) Entre crochets les p -values

Ces développements illustrent bien le fait que la méthode de sélection, si elle facilite le travail de l'économètre, ne l'élimine pas totalement, notamment lorsque l'algorithme révèle plusieurs modèles statistiquement équivalents.

Par ailleurs, bien que la méthode de sélection de chaque modèle ne le garantisse pas *a priori*, on constate que tous les coefficients estimés ont le signe attendu : positif s'agissant des carnets de commandes, des perspectives personnelles de production ou de la tendance de la production passée. En revanche lorsque les soldes d'opinion relatifs aux stocks intervenaient dans les régressions, les coefficients qui leur étaient associés avaient le mauvais signe, c'est-à-dire positif⁽⁹⁾. Dans ces cas les variables de stocks ont été supprimées de l'ensemble d'information de départ et les modèles ré-estimés. Il est certes possible, pour le premier mois, de rationaliser ce résultat en notant qu'un niveau de stocks élevé en janvier est l'indice que les entreprises ont beaucoup produit au quatrième trimestre, mais cet argument n'est pas totalement convaincant : cette hausse de production devrait en effet se constater dans l'opinion sur la production passée et pas dans les stocks.

Tableau 3 : résultats de l'estimation pour le trimestre courant

| | Mois 1 | Mois 2 | Mois 3 |
|-------------------------------|------------------|----------------|----------------|
| ppy_m1 | 0,05 (6,29)* | | |
| $ppy_m1 - ppy_m3(-1)$ | | 0,07 (3,05) | 0,06 (2,44) |
| y_{pa_m2} | | 0,04 (7,23) | |
| y_{pa_m3} | | | 0,04 (7,59) |
| $y_{pa_m2} - y_{pa_m1}$ | | 0,11 (5,23) | 0,10 (4,94) |
| $y_{pa_m1} - y_{pa_m3(-1)}$ | | 0,06 (3,08) | 0,07 (3,21) |
| $ccg_m1 - ccg_m3(-1)$ | 0,11 (4,10) | | |
| $ccg_m3(-1) - ccg_m2(-1)$ | 0,08 (2,93) | | |
| Constante | 0,49 (5,24) | 0,49 (5,73) | 0,48 (5,84) |
| \bar{R}^2 | 0,48 | 0,57 | 0,58 |
| Écart type des résidus | 0,86 | 0,78 | 0,77 |
| Indice de conditionnement | 13 | 17 | 17 |
| Durbin-Watson | 2,29 | 2,17 | 2,17 |
| Chow (50%) | 0,99 [0,51]** | 0,73 [0,85] | 0,76 [0,82] |
| Chow (90%) | 0,87 [0,55] | 0,49 [0,87] | 0,52 [0,85] |
| Autocorrélation (ordre 5) | 1,20 [0,31] | 0,84 [0,52] | 0,69 [0,63] |
| Normalité | 0,82 [0,66] | 0,44 [0,80] | 0,25 [0,88] |
| Hétéroscédasticité | 0,62 [0,71] | 1,00 [0,44] | 1,06 [0,40] |

(*) Entre parenthèses les t -statistiques

(**) Entre crochets les p -values

On notera enfin que les variables d'enquête utiles pour prévoir le trimestre courant et le trimestre à venir ne sont pas les mêmes et que les différences sont conformes à l'intuition : la production passée est utile pour prévoir le trimestre en cours et pas le trimestre à venir ; les perspectives personnelles de production sont utiles pour prévoir le trimestre à venir, mais pas ou peu pour prévoir le trimestre en cours (elles le sont le premier mois, date à laquelle la production passée se rapporte au trimestre précédent). Les carnets de commande jouent un rôle plus ambivalent, puisqu'ils sont utiles pour prévoir le trimestre courant comme le trimestre à venir et semblent donc jouer un rôle plus uniforme d'indicateur d'"état de la conjoncture".

Simulations hors échantillon

Afin d'appréhender qualitativement quels auraient été les comportements de ces étalonnages lors de plusieurs exercices successifs de prévisions, nous avons mené des simulations récursives sur la base de variables sélectionnées sur l'ensemble de l'échantillon. La première prévision est réalisée de la façon suivante : les modèles sont estimés de la période allant du deuxième trimestre de 1979 au premier trimestre de 1992 et des prévisions du trimestre passé (si nous sommes le premier mois), présent et à venir sont réalisées. Puis nous incrémentons d'un trimestre et recommençons

Tableau 4 : résultats de l'estimation pour le trimestre à venir

| | Mois 1 | Mois 2 | Mois 3 |
|--------------------------------|------------------|----------------|----------------|
| <i>ppy_m3</i> | | | 0,04 (4,75) |
| <i>ppy_m3 - ppy_m2</i> | | | 0,08 (3,97) |
| <i>ppy_m1 - ppy_m3(-1)</i> | 0,07 (2,37)* | 0,09 (3,07) | |
| <i>ppy_m3(-1) - ppy_m2(-1)</i> | 0,09 (3,04) | 0,09 (2,90) | |
| <i>ccg_m1</i> | 0,01 (2,16) | | |
| <i>ccg_m2</i> | | 0,01 (2,30) | |
| <i>ccg_m1 - ccg_m3(-1)</i> | | | 0,01 (3,19) |
| <i>ccg_m2(-1) - ccg_m1(-1)</i> | 0,09 (2,60) | 0,10 (2,91) | |
| <i>ccg_m1(-1) - ccg_m3(-2)</i> | 0,07 (2,15) | | |
| Constante | 0,84 (6,32) | 0,86 (6,34) | 0,49 (4,78) |
| \bar{R}^2 | 0,34 | 0,31 | 0,36 |
| Écart-type des résidus | 0,97 | 0,99 | 0,96 |
| Indice de conditionnement | 35 | 36 | 13 |
| Durbin-Watson | 1,91 | 1,92 | 2,27 |
| Chow (50%) | 1,37 [0,15]** | 1,51 [0,09] | 0,97 [0,54] |
| Chow (90%) | 1,33 [0,23] | 1,41 [0,19] | 0,95 [0,49] |
| Autocorrélation (ordre 5) | 0,58 [0,71] | 0,86 [0,51] | 1,02 [0,41] |
| Normalité | 0,43 [0,80] | 0,47 [0,79] | 0,01 [1,00] |
| Hétéroscédasticité | 0,79 [0,85] | 0,81 [0,59] | 0,80 [0,57] |

(*) Entre parenthèses les *t*-statistiques

(**) Entre crochets les *p*-values

Tableau 6 : MSE des modèles

| | Mois 1 | Mois 2 | Mois 3 |
|---------------------------------|--------|---------------------|--------|
| Prévisions du trimestre courant | 0,79 | 0,54 | 0,51 |
| Prévisions du trimestre futur | 1,33 | 1,39 ^(*) | 1,02 |

(*) Le MSE associé au modèle alternatif du mois 2 est de 1,31

l'expérience. Au total, nous avons à notre disposition 45 prévisions hors échantillon (graphiques 1 à 3). Nous pouvons dans ce cadre réaliser des tests de comparaisons des performances prédictives, basés sur les *Mean Squared Error* (MSE, reportés dans le tableau 6), entre les différents modèles (Diebold et Mariano, 1995, et Harvey *et alii*, 1997).

Concernant la prévision du trimestre courant, on constate (cf. tableau 7) qu'entre le premier et le deuxième mois la nouvelle information disponible permet de mieux prévoir la production manufacturière ; le test d'égalité des performances prédictives donne une hypothèse alternative (MSE1 < MSE2) significative au seuil de 5%. Mais l'écart

Tableau 5 : modèles terminaux pour la prévision du trimestre futur le deuxième mois

| | Modèle 1 | Modèle 2 | Modèle 3 |
|--------------------------------|----------------|----------------|----------------|
| <i>ppy_m1 - ppy_m3(-1)</i> | 0,07 (2,28) | | 0,08 (3,07) |
| <i>ppy_m3(-1) - ppy_m2(-1)</i> | 0,09 (3,04) | 0,08 (2,51) | 0,09 (3,07) |
| <i>ccg_m2</i> | 0,01 (2,08) | 0,01 (2,16) | 0,01 (2,31) |
| <i>ccg_m1 - ccg_m3(-1)</i> | | 0,07 (2,07) | |
| <i>ccg_m2(-1) - ccg_m1(-1)</i> | 0,08 (2,45) | 0,07 (2,21) | 0,10 (2,91) |
| <i>ccg_m1(-1) - ccg_m3(-2)</i> | 0,07 (2,03) | 0,07 (2,25) | |
| Constante | 0,84 (6,24) | 0,84 (6,25) | 0,86 (6,35) |
| \bar{R}^2 | 0,33 | 0,33 | 0,31 |
| AIC | 0,010* | 0,019 | 0,034 |
| BIC | 0,171 | 0,181 | 0,169** |
| HQ | 0,075*** | 0,085 | 0,089 |

Modèle retenu selon (*) le critère AIC, (**) le BIC, (***) le HQ

Tableau 7 : comparaisons des performances prédictives

| Prévision du trimestre : | Mois | Hypothèse nulle | Statistiques |
|--------------------------|------|---------------------------|----------------------|
| Futur | 1 | - | - |
| | 2 | $MSE_1^{fu} = MSE_2^{fu}$ | -1,03 ⁽¹⁾ |
| | 3 | $MSE_2^{fu} = MSE_3^{fu}$ | 1,75 ⁽²⁾ |
| Présent | 1 | $MSE_3^{pr} = MSE_1^{pr}$ | 2,21* |
| | 2 | $MSE_1^{pr} = MSE_2^{pr}$ | 2,71* |
| | 3 | $MSE_2^{pr} = MSE_3^{pr}$ | 1,06** |
| Passé | 1 | $MSE_3^{pa} = MSE_1^{pa}$ | -0,17 |

(*) Significatif au seuil de 5%

(**) Significatif au seuil de 15%

(1) La statistique s'élève à 0,84 et (2) à 1,44** dans le cas de l'utilisation du modèle *alternatif* du mois 2.

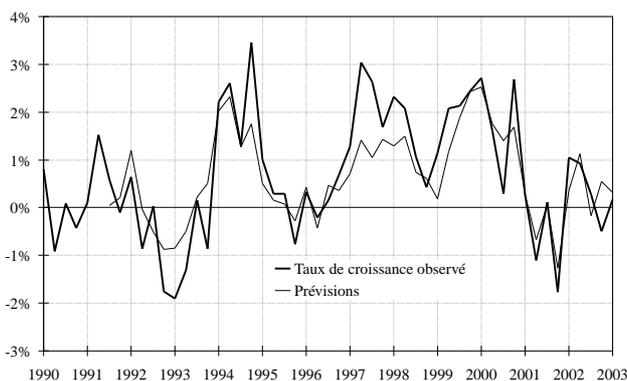
Lecture : les statistiques correspondant à chaque fois à l'hypothèse nulle que la prévision d'un mois ($i = 1, 2, 3$) donné a une capacité prédictive identique à celle effectuée le mois suivant ; *fu* désigne les prévisions faites pour le trimestre futur ; *pr* les prévisions faites pour le trimestre courant et *pa* les prévisions faites pour le trimestre passé. Par exemple, l'hypothèse nulle que la prévision de l'évolution de la production manufacturière pour le trimestre *T* faite en fin de trimestre précédent ($T-1$) est aussi précise que celle faite au début du trimestre (*T*) a une statistique de 2 ; 215. Avec une valeur critique de 1,65 au seuil de 5 %, cette hypothèse est donc nettement rejetée.

entre les performances prédictives des modèles estimés le deuxième et le troisième mois n'est pas statistiquement significatif. On enregistre également un gain important quand on passe du troisième mois du trimestre précédent au premier mois du trimestre à prévoir (ce dont atteste la valeur élevée de la statistique de test, à 2,21).

On constate donc que globalement la prise en compte systématique des nouvelles informations conjoncturelles constitue un gain dans la prévision. Si l'on compare à un étalonnage " usuel " utilisant des soldes d'enquête trimestrialisés⁽¹⁰⁾ (moyenne du trimestre) (proche de celui de Scherrer, 1996) on constate qu'il donne des prévisions de moins bonne qualité que ceux obtenus ici, *lorsqu'ils utilisent le même ensemble d'information* : si nous menons les mêmes estimations récursives sur l'étalonnage " usuel ", nous obtenons un MSE de 0,63, supérieur à celui obtenu ici avec les enquêtes du troisième mois du trimestre (0,51). Son MSE est également supérieur à celui obtenu ici avec les enquêtes du deuxième mois du trimestre : au deuxième mois du trimestre, l'étalonnage présenté ici, qui n'utilise que les enquêtes du deuxième mois du trimestre, constitue donc une amélioration incontestable par rapport à la méthode " usuelle ", qui nécessite elle de faire une prévision, nécessairement imprécise, des résultats de l'enquête du mois à venir.

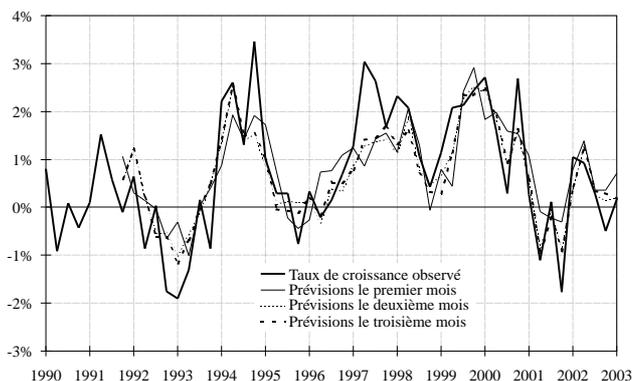
À l'inverse, les résultats obtenus pour la prévision du trimestre futur ne sont pas totalement satisfaisants. Comme pouvaient le laisser présager les résultats de l'estimation, le passage du premier mois dans le trimestre au deuxième mois se traduit par une légère dégradation des prévisions. Dès lors, la nette amélioration de la prévision entre le deuxième mois et le troisième mois doit être relativisée. Néanmoins, il est possible de remédier à ce défaut en utilisant, pour le deuxième mois, le *modèle alternatif* obtenu précédemment avec le critère AIC : on obtient alors une légère amélioration des performances prédictives entre le premier et le deuxième mois, tout en préservant l'amélioration statistiquement significative entre le deuxième et le troisième mois.

Graphique 1 : prévisions du trimestre passé le premier mois du trimestre

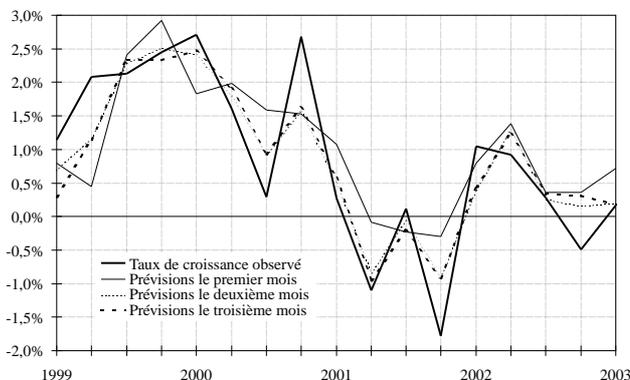


Mais, dans tous les cas avec un MSE de 1,02 le troisième mois du trimestre, l'écart type de l'erreur de prévision (*Root Mean Squared Error*) atteint 1%, à peine inférieur à celui de la série de variations trimestrielles de la production manufacturière (1,2%). Nous atteignons une limite, souvent mise en valeur, concernant le pouvoir prédictif des enquêtes à un horizon de plus d'un trimestre.

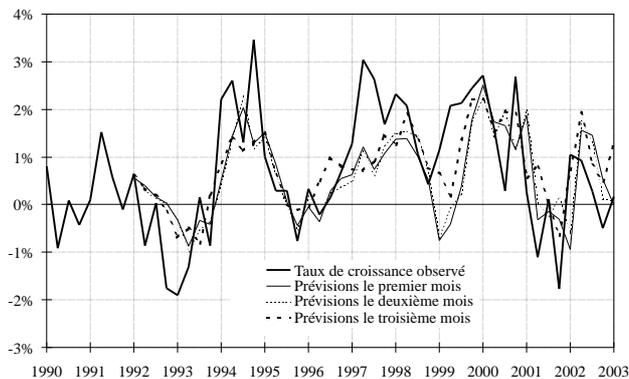
Graphique 2 : prévisions du trimestre courant selon le mois dans le trimestre



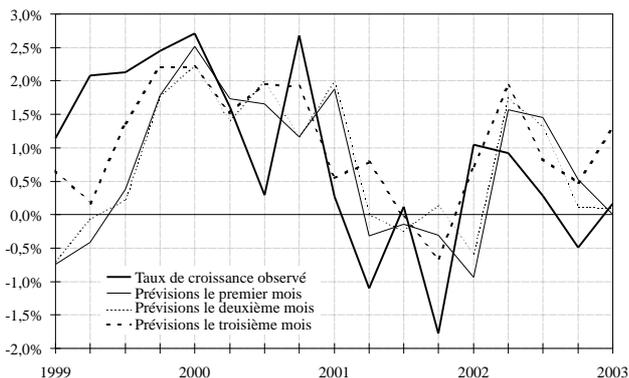
Graphique 2b : prévisions du trimestre courant selon le mois dans le trimestre Zoom sur la période 1999 - fin 2002



Graphique 3 : prévisions du trimestre futur selon le mois dans le trimestre



Graphique 3b : prévisions du trimestre futur selon le mois dans le trimestre
Zoom sur la période 1999 - fin 2002



Calcul de l'innovation conjoncturelle apportée par chaque nouvelle enquête

L'utilisation selon le mois dans le trimestre et selon l'horizon de prévision de variables différentes dans les étalonnages rend difficile l'interprétation des modifications de prévision d'un mois sur l'autre : lorsque la prévision change, comment séparer ce qui est une modification liée au changement de variable dans l'étalonnage de ce qui est imputable aux véritables chocs exogènes ? Nous proposons, ici, une méthode permettant de relier les changements de prévision d'un mois sur l'autre aux innovations conjoncturelles apportées par chaque nouvelle enquête.

On peut montrer que la différence entre deux prévisions successives peut se décomposer – approximativement – linéairement en fonction des innovations apportées par la nouvelle enquête à chaque exogène inclus dans l'étalonnage correspondant à la deuxième prévision. En effet, les étalonnages présentés ont la forme suivante :

$$(1) y_t = Z_t \alpha_i + X_{it} \beta_i + \varepsilon_{it}, \quad i=1,2$$

où Z_t de format $1 \times k$ désigne les variables communes à deux étalonnages successifs et X_{it} de format $1 \times l_i$ contient les variables utilisées uniquement pour l'étalonnage i du trimestre t ; \hat{y}_{it} est la prévision de y_t réalisée le mois i à partir de l'étalonnage correspondant.

Soit alors la régression de la variable j de X_{2t} sur X_{1t} et Z_t

$$(2) \hat{X}_{2t,j} = Z_t \hat{a} + X_{1t} \hat{b}$$

avec

$$(3) \hat{U}_{2t,j} = X_{2t,j} - \hat{X}_{2t,j}$$

et

$$(4) \hat{U}_{2t} = (\hat{U}_{2t,1} \dots \hat{U}_{2t,l_2})$$

Comme Z_t et X_{1t} sont exogènes, dans l'équation précédente (3), nous en déduisons

$$(5) \hat{U}_{2t} \perp Z_t, X_{1t}$$

De plus, dans la mesure où les variables contenues dans X_{1t} ne sont pas sélectionnées par l'algorithme lors de l'estimation de l'équation du mois 2, nous pouvons considérer que la relation

$$(6) \hat{\varepsilon}_{2t} \perp X_{1t}$$

est approximativement vérifiée.

Par ailleurs, par construction :

$$(7) \hat{\varepsilon}_{2t} \perp Z_t$$

Or nous avons

$$(8) y_t - Z_t \hat{\alpha}_2 - \hat{X}_{2t} \hat{\beta}_2 = y_t - Z_t \hat{\alpha}_2 - X_{2t} \hat{\beta}_2 + (X_{2t} - \hat{X}_{2t}) \hat{\beta}_2 = \varepsilon_{2t} + \hat{U}_{2t} \hat{\beta}_2$$

En raison des relations (5) et (6), on a donc approximativement

$$(9) \hat{\varepsilon}_{2t} + \hat{U}_{2t} \hat{\beta}_2 \perp X_{1t}, Z_t$$

D'où

$$(10) y_t - Z_t \hat{\alpha}_2 - \hat{X}_{2t} \hat{\beta}_2 \perp X_{1t}, Z_t$$

Donc, comme l'estimateur des MCO dans la régression de y_t sur X_{1t}, Z_t est tel que le résidu de la régression est orthogonal à X_{1t}, Z_t et que $Z_t \hat{\alpha}_2 - \hat{X}_{2t} \hat{\beta}_2$ est bien une combinaison linéaire de X_{1t}, Z_t (par définition, cf. l'équation (2)), nous obtenons l'approximation suivante

$$(11) \hat{y}_{1t} \approx Z_t \hat{\alpha}_2 + \hat{X}_{2t} \hat{\beta}_2$$

Dès lors, la différence entre deux prévisions successives, $\hat{y}_{2t} - \hat{y}_{1t}$, peut s'écrire :

$$(12) \hat{y}_{2t} - \hat{y}_{1t} \approx (X_{2t} - \hat{X}_{2t}) \hat{\beta}_2$$

Cette formulation nous permet donc décomposer la variation de la prévision entre les mois 1 et 2, $\hat{y}_{2t} - \hat{y}_{1t}$, comme une somme des innovations enregistrées sur chaque variable explicative j ,

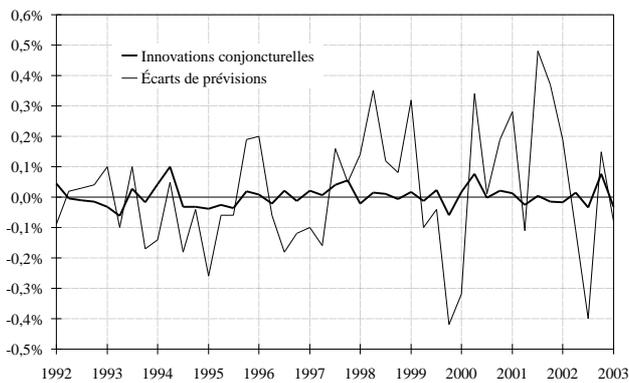
pondérée par le poids dans la régression de cette variable : $(X_{2t,j} - \hat{X}_{2t,j})\hat{\beta}_{2,j}$. La relation définie dans (12) n'est qu'approximative : elle serait exacte si X_{1t} était présent dans le deuxième étalonnage.

Une comparaison entre les calculs *hors échantillon* de l'innovation conjoncturelle $(X_{2t} - \hat{X}_{2t})\hat{\beta}_2$ et la révision de prévision $\hat{y}_{2t} - \hat{y}_{1t}$ (graphiques 4 à 9) permet une évaluation empirique de la validité de cette approximation.

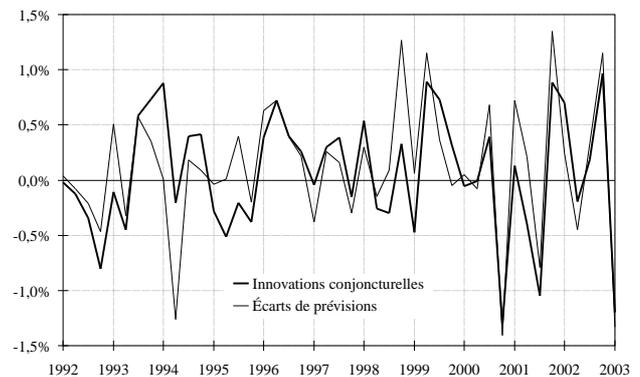
Les résultats paraissent dans l'ensemble satisfaisants, surtout si l'on tient compte du fait qu'il s'agit de prévisions effectuées hors période

d'estimation : les révisions effectuées sur les prévisions peuvent en général très bien s'expliquer par les innovations apportées par la nouvelle enquête à chaque exogène incluse dans le nouvel étalonnage. Il faut toutefois noter les mauvais résultats obtenus pour l'innovation conjoncturelle du premier au deuxième mois lors de la prévision du trimestre futur (graphique 4). Ceux-ci ne doivent pas surprendre dans la mesure où le passage du premier mois au deuxième mois ne se traduit pas par une amélioration de la prévision. De nouveau, le recours au modèle alternatif pour le mois 2 améliore grandement les choses (graphique 4b et 5b).

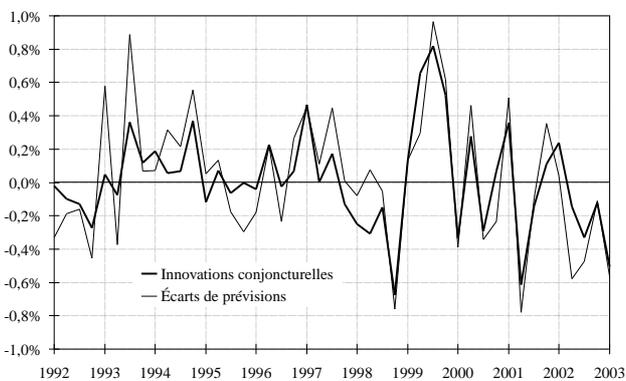
Graphique 4 : production manufacturière future
Écarts de prévisions et innovations conjoncturelles entre les premier et deuxième mois du trimestre



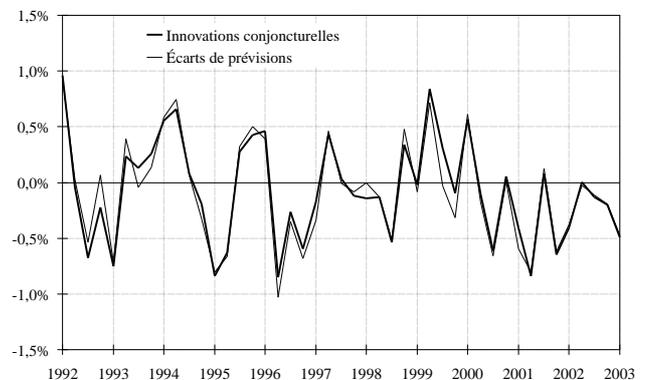
Graphique 5 : production manufacturière future
Écarts de prévisions et innovations conjoncturelles entre les deuxième et troisième mois du trimestre



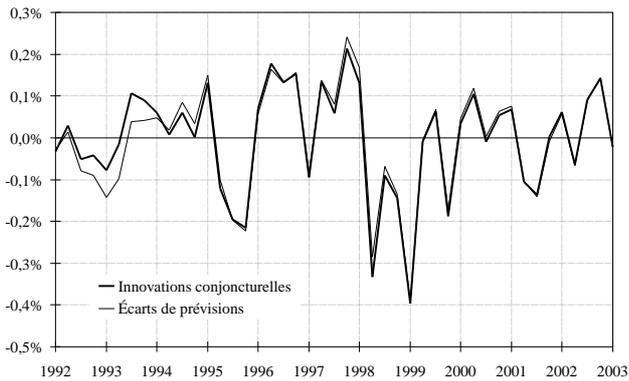
Graphique 6 : production manufacturière courante
Écarts de prévisions et innovations conjoncturelles entre le premier mois du trimestre et le troisième mois du trimestre précédent



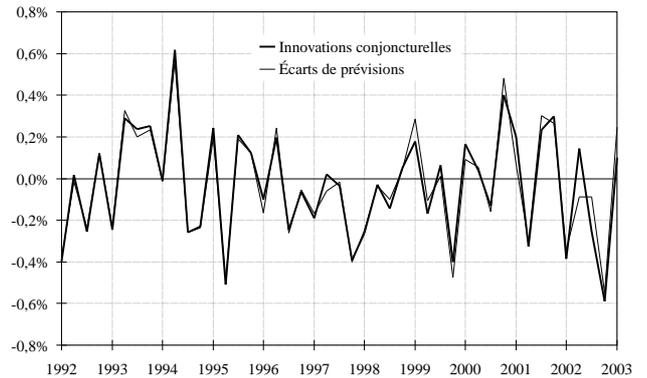
Graphique 7 : production manufacturière courante
Écarts de prévisions et innovations conjoncturelles entre les premier et deuxième mois du trimestre



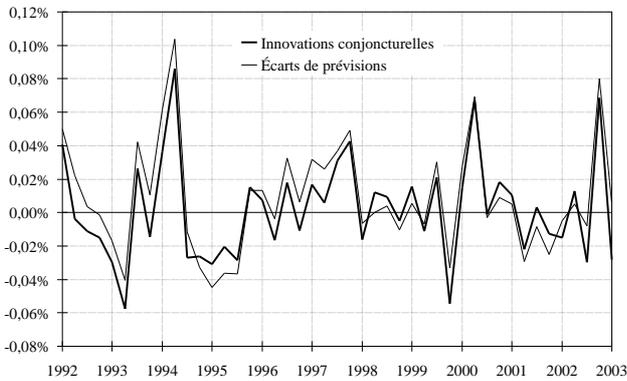
Graphique 8 : production manufacturière courante
Écarts de prévisions et innovations conjoncturelles
entre les deuxième et troisième mois du trimestre



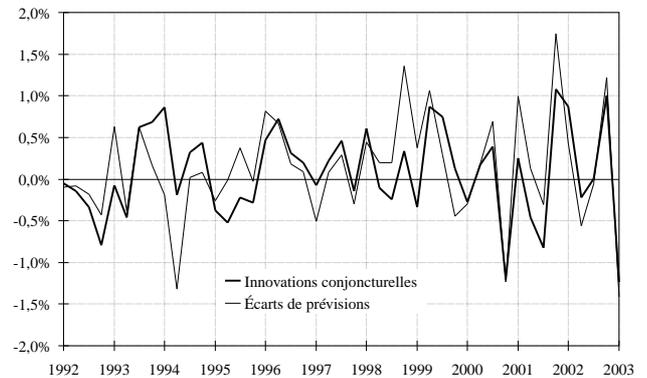
Graphique 9 : production manufacturière passée
Écarts de prévisions et innovations conjoncturelles
entre le premier mois du trimestre et le troisième
mois du trimestre précédent



Graphique 4b : production manufacturière future
Écarts de prévisions et innovations conjoncturelles
entre les premier et deuxième mois du trimestre



Graphique 5b : production manufacturière future
Écarts de prévisions et innovations conjoncturelles
entre les deuxième et troisième mois du trimestre



Comparaisons avec un modèle VAR

Pour appréhender l'apport des enquêtes de conjoncture, une démarche traditionnelle consiste à comparer les étalonnages estimés à des méthodes de prévision plus automatiques, les méthodes VAR constituant la référence en la matière. La littérature empirique sur le sujet ne nous fournissant pas de référence qui puisse nous guider, nous avons choisi d'estimer un modèle VAR à cinq variables (dans le secteur manufacturier) : la production, la consommation, la FBCF, les emplois intermédiaires produits et le ratio de couverture (exportations sur importations). Les variables sont prises en logarithme.

Ce modèle VAR comprend les principaux postes de l'équilibre ressources-emplois de la branche manufacturière. Seules manquent les variations de stocks. Les introduire aurait conduit à introduire une relation comptable, l'équilibre ressources-emplois, dans notre modèle VAR, ce qui aurait conduit à des problèmes économétriques d'identifiabilité⁽¹¹⁾. On

s'attend toutefois, sous l'hypothèse de stationnarité du ratio des variations de stocks au PIB, à retrouver au moins une relation de cointégration entre les variables du VAR, représentative de cet équilibre comptable. Les tests de la trace et de la valeur propre maximale (Johansen 1988, 1991) indiquent bien que le système représenté par ces variables comporte une (et une seule) relation de cointégration. Néanmoins, au sein de celle-ci, le coefficient relatif aux emplois intermédiaires en produits manufacturés a le mauvais signe (*i.e.* identique à celui de la production). Par conséquent, la relation de cointégration a peu de pertinence économique et ne peut en tout état de cause pas être assimilée à cet équilibre ressources-emplois. En conséquence, les simulations récursives présentées ont été obtenues à partir d'un VAR en différences⁽¹²⁾.

Notons que cela ne modifie pas notre diagnostic dans la mesure où les prévisions hors échantillons sont d'une qualité équivalente selon que l'on choisisse une modélisation VAR ou VECM (avec cette relation de cointégration contre intuitive) ; à titre d'exemple,

le MSE pour la prévision à un trimestre s'élève à 1,62 pour le modèle VAR et à 1,60 pour le VECM. En ce qui concerne les prévisions à deux trimestres, le VAR affiche un MSE de 1,84.

Il ressort de cet exercice (tableau 8) que les performances prédictives du VAR sont nettement inférieures à celles obtenues à partir des étalonnages (l'hypothèse d'égalité des performances prédictives entre le modèle VAR et les étalonnages est systématiquement rejetée, au bénéfice des étalonnages). L'information conjoncturelle contenue dans les données d'enquête est donc nettement supérieure à ce que permet d'obtenir une représentation plus structurelle du secteur manufacturier.

Tableau 8 : performance comparée des étalonnages et des modèles VAR

| | Mois 1 | Mois 2 | Mois 3 | VAR |
|-----------------------|------------------|------------------|-----------------|------|
| MSE trimestre courant | 0,80 (3,51)* | 0,55 (4,42)* | 0,51 (4,12)* | 1,62 |
| MSE trimestre à venir | 1,32 (1,65)** | 1,36 (1,64)** | 1,05 (3,14)* | 1,84 |

Entre parenthèses les statistiques du test d'égalité des performances prédictives entre le modèle VAR et les étalonnages $H_0: MSE_{VAR} = MSE_i$, où l'indice $i = 1,2,3$ désigne le mois dans le trimestre.

(*) Significatif au seuil de 5%

(**) Significatif au seuil de 10%

Comparaisons avec les prévisions de l'Insee

Nous proposons, dans cette section, de comparer les prévisions issues des étalonnages présentés précédemment avec celles publiées par l'Insee dans les *Notes de conjoncture*. Cette comparaison constitue un test sérieux de la qualité des étalonnages dans la mesure où : a) les conjoncturistes de l'Insee peuvent disposer d'informations privilégiées en raison de la position centrale de l'Insee au cœur du dispositif de production statistique b) pour bâtir leur prévision de production manufacturière, les conjoncturistes utilisent d'autres sources d'information, notamment la confrontation avec la prévision de demande. Chaque année, dans le courant des mois de mars, juin et décembre, l'Insee publie des prévisions qui ont pour horizon le trimestre courant (T), le trimestre $T+1$ et, pour les mois de juin et de décembre, le trimestre $T+2$. La *Note de conjoncture* de mars ne comporte de prévisions que depuis 1997. Si nous considérons la période de prévisions allant du premier trimestre de 1992 au quatrième trimestre de 2004, nous obtenons vingt-neuf prévisions de l'Insee, à la fois pour le trimestre courant et le trimestre futur, qui pourront être mises en rapport avec celles issues des étalonnages.

Néanmoins, les prévisions ne sont pas directement comparables car elles ne portent pas sur " la même variable " de production manufacturière. En effet, du fait du processus de révision des Comptes Annuels, les données sont définitives trois années après leur publication et sont donc soumises à révisions au cours de cette période. En outre, même les chiffres se rapportant à des années réputées définitives changent, avec la modification des coefficients saisonniers consécutive à l'incorporation de nouvelles données. Dès lors, lorsque l'Insee réalise ses prévisions, la dernière chronique de production manufacturière connue est composée, pour les données les plus récentes, des résultats provisoires puis, à mesure que l'on remonte dans le passé, de données semi-définitives et définitives : les données utilisées par l'Insee sont donc différentes, même sur leur plage commune, d'un exercice de prévision à l'autre. À l'inverse, la chronique de production manufacturière que nous utilisons pour tester nos prévisions est toujours la même (définitive jusqu'en 1999, puis semi-définitive et provisoire pour le passé plus récent), quel que soit l'exercice de prévision, et n'est donc pas celle dont disposaient les prévisionnistes, notamment ceux de l'Insee, en temps réel. En conséquence, la chronique utilisée par l'Insee n'est jamais identique à celle que nous avons utilisée pour estimer nos étalonnages, basée elle sur les données actuelles.

Afin de rendre les résultats de l'Insee comparables avec ceux obtenus à partir des étalonnages, nous avons calculé les erreurs de prévision de l'Insee en utilisant pour chaque date la production manufacturière telle que fournie par " les premiers résultats " (valeurs à la première publication des Comptes Trimestriels). Par cela, nous souhaitons tenir compte du fait (i) que lors de l'exercice de prévision les dernières valeurs de la production manufacturière que l'Insee avait à sa disposition étaient plus proches des résultats provisoires que des résultats définitifs, mais également (ii) que l'information conjoncturelle qu'il utilisait pour réaliser les prévisions ne permettait pas *a priori* d'anticiper les résultats définitifs des Comptes Trimestriels.

Tableau 9 : performance comparée des étalonnages et des prévisions de l'Insee

| | Mois 1 | Mois 2 | Mois 3 | Insee |
|-----------------------|-----------------|-----------------|-----------------|-------|
| MSE trimestre courant | 0,89 (-1,35) | 0,65 (-0,30) | 0,62 (-0,18) | 0,59 |
| MSE trimestre à venir | 1,42 (-1,38) | 1,47 (-1,53) | 1,02 (-0,52) | 0,80 |

Entre parenthèses les statistiques du test d'égalité des performances prédictives entre l'Insee et les étalonnages : $H_0: MSE_{INSEE} = MSE_i$, où l'indice $i = 1,2,3$ désigne le mois dans le trimestre.

(*) Significatif au seuil de 5%

(**) Significatif au seuil de 10%

Pour la prévision du trimestre courant, le tableau 9 montre que, à l'exception des prévisions réalisées le premier mois, les performances prédictives de l'Insee et des étalonnages sont très proches⁽¹³⁾. Pour la prévision du trimestre futur, nous constatons un écart nettement plus important dans les performances prédictives à la défaveur des étalonnages. Nous retrouvons la limite, déjà mise en évidence précédemment, concernant le faible pouvoir prédictif des enquêtes pour un horizon dépassant le trimestre courant : à un tel horizon, l'utilisation d'autres outils que les étalonnages d'enquête, reposant plus sur les raisonnements économiques, permet une meilleure prévision.

Conclusion

Nous avons proposé dans cet article une méthode alternative d'estimation et de prévision à court terme de la production manufacturière. Cette méthode a pour points de départ les constats suivants : (i) les étalonnages généralement utilisés sont construits à partir de données recouvrant l'ensemble du trimestre et (ii) ceux-ci tendent à intégrer un petit nombre de soldes disponibles dans l'enquête. Dans ces conditions, on a examiné s'il n'était pas possible d'obtenir une prévision de la production manufacturière avant de disposer de l'ensemble du trimestre et qui tienne compte de toute l'information pertinente et disponible dans l'enquête mensuelle dans l'industrie manufacturière.

Pour pouvoir mettre en relation l'information trimestrielle issue des Comptes Trimestriels et les données mensuelles issues de l'enquête dans l'industrie, nous avons donc utilisé une méthode d'étalonnage qui permette à la fois de préserver toute l'information conjoncturelle et qui soit opérationnelle même si le trimestre n'est pas encore achevé. Cette méthode consiste à découper chaque série mensuelle en trois nouvelles séries trimestrielles, puis d'estimer des étalonnages selon l'horizon de prévision et le mois dans le trimestre. Pour sélectionner les variables à intégrer dans les étalonnages (7 au total) nous avons mis en œuvre l'algorithme proposé par Krolzig et Hendry (2000).

Dans ce cadre, nous montrons que la prise en compte systématique de la nouvelle information conjoncturelle constitue un gain pour la prévision ; globalement la précision de celle-ci s'accroît à mesure que l'on avance dans le trimestre. De plus, nous montrons que les étalonnages ont de bonnes performances prédictives et cela particulièrement pour la prévision du trimestre courant. Concernant la prévision du trimestre futur, les données d'enquête se révèlent avoir un contenu prédictif nettement plus faible. Malgré cette dernière limitation, nous constatons que l'apport des enquêtes conjoncturelles, par rapport à une modélisation plus "structurelle" comme un VAR sur des données issues des Comptes Trimestriels, est incontestable à la fois pour la prévision du trimestre courant et futur. Enfin, pour la prévision du trimestre en cours, nous montrons que nos étalonnages ont des performances prédictives très proches des prévisions publiées par l'Insee dans les *Notes de conjoncture*, qui reposent à la fois sur des étalonnages et des raisonnements macroéconomiques

Deux prolongements sont envisageables à ce travail. D'une part, la comparaison avec d'autres méthodes d'estimation que les VAR mériterait d'être entreprise. D'autre part, la méthodologie d'estimation employée ici pourrait être étendue à d'autres enquêtes et sa validité réévaluée.

Notes

- (1) Cf. Bouton et Erkel-Rousse (2002).
- (2) Cette méthode n'est pas nouvelle, Fitzgerald et Miller (1996) la mettent en œuvre pour prévoir le PIB des États-Unis à partir des heures travaillées.
- (3) La démarche utilisée pourrait aisément être étendue plus avant dans le futur ; compte tenu de la dégradation continue de la prévision constatée sur les estimations menées ici, au fur et à mesure que l'horizon de prévision s'éloigne, cette extension semble peu pertinente.
- (4) On pourrait également rajouter deux prévisions : l'une lors de la publication des comptes trimestriels détaillés du début de premier mois (11 janvier 2002 au premier trimestre 2002), l'autre lors de la publication des premiers comptes trimestriels du trimestre précédent, après le milieu du deuxième mois du trimestre (22 février 2002 au premier trimestre 2002). Mais, comme nous le montrerons ultérieurement, les retards du taux de croissance de la production manufacturière sont non significatifs, de sorte que ces deux prévisions n'apporteraient rien.
- (5) Les estimations présentées ici ont été faites avec la « boîte à outils » Grocer pour Scilab (cf. Dubois, 2004).
- (6) Il peut arriver que le modèle de départ ne passe pas lui-même les tests de spécification. Cela tend alors à indiquer qu'il manque une variable explicative importante dans le modèle. L'utilisateur peut toutefois vouloir passer outre cette conclusion : il lui est alors nécessaire d'ajuster le seuil de départ du ou des tests de spécification rejetés (cf. diagramme 1)
- (7) Deux autres manières d'ordonner les données initiales ont également été testées : la première en n'introduisant que des variables en niveau, la seconde, sur la base d'une suggestion du rapporteur, en retenant à côté du solde en niveau de l'enquête la plus récente les écarts à ce solde dans les enquêtes passées. Dans chacun des cas, les résultats obtenus étaient, soit équivalents, à la fois en termes statistiques et de représentation obtenue, soit légèrement moins bons, ce qui justifie le choix effectué ici.
- (8) L'indice de conditionnement permet de détecter la présence de multicollinéarité (Besley *et alii*, 1980) : lorsque l'indice passe la valeur 300 cela indique la présence d'une colinéarité qui doit être considérée comme dangereuse.
- (9) Sont concernées : la prévision du trimestre passé et du trimestre futur le premier mois du trimestre, et la prévision du trimestre futur le deuxième mois du trimestre
- (10) L'étalonnage utilise la tendance passée de la production en niveau et les perspectives personnelles de production en différences premières contemporaines et retardées d'un trimestre comme variables exogènes. Les variables ont été sélectionnées à partir de l'algorithme de Krozlig et Hendry.

(11) Comme toutes les variables ont été prises en logarithme, en pratique, la relation comptable qui existe entre les variables en niveau n'est plus qu'approchée en logarithme ; introduire les variations de stocks aurait donc conduit plutôt à un problème de quasi-identifiabilité (donc de multicollinéarité) qui aurait rendu les estimations très imprécises.

(12) Le critère BIC nous a conduit à retenir 1 retard.

(13) Les résultats ne peuvent, ici, être analysés que qualitativement. En effet, dans la mesure où les variables de production manufacturière utilisées par l'Insee et dans les étalonnages ne sont pas strictement identiques, les MSE ne peuvent pas, en toute rigueur, être comparés entre eux. De la même façon, la fonction d_t dans le test d'égalité des performances prédictives n'est plus, dans ce cadre, une fonction d'écart ; les hypothèses de construction du test ne sont plus vérifiées.

Bibliographie

- Belsley D.A., Kuh E. et Welsch R. E. (1980).** "Regression Diagnostics: Identifying Influential Data and Sources of Colinearity", ed. *John Wiley and Sons, Inc.* New-York.
- Bouton F. et Erkel-Rousse H. (2002).** "Conjonctures sectorielles et prévisions à court terme de l'activité : l'apport des enquêtes de conjoncture dans les services", *Économie et Statistique*, n°359-360, pp. 35-68.
- Buffeteau S. et Mora V. (2000).** "La prévision des comptes de la zone euro à partir des enquêtes de conjoncture dans les services", *Note de Conjoncture de l'Insee*, décembre, pp. 27-35.
- Diebold F. X. et Mariano R. S. (1995).** "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, 13, pp. 253-265.
- Doornik J.A. et Hansen H. (1994).** "A Practical Test for Univariate and Multivariate Normality", *Discussion paper*, Nuffield College.
- Dubois E. (2004).** "GROCER 1.0: an Econometric Toolbox for Scilab", *Document de travail*, disponible à l'adresse internet <http://dubois.ensae.net/grocer.html>
- Fitzgerald T. J. et Miller P. J. (1996).** "A Simple Way to Estimate Current-Quarter GNP", Federal Reserve Bank of Minneapolis, *Quarterly Review*, vol. 20, n°2.
- Godfrey L.G. (1978).** "Testing for Higher Order Serial Correlation in Regression Equations when the Regressors Include Lagged Dependent Variables", *Econometrica*, vol. 46, pp. 1303-1313.
- Harvey D.I., Leybourne S.J. et Newbold P. (1997).** "Testing the Equality of Prediction Mean Square d'Errors", *International Journal of Forecasting*, 13, pp. 273-281.
- Hendry D.F. (1979).** "Predictive Failure and Econometric Modelling in Macro-Economics: The Transactions Demand for Money". In *Economic Modelling*, Ormerod, P.(ed.), London: Heinemann, pp. 217-242.
- Hoover K. D. et Perez S. J. (1999).** "Data Mining Reconsidered: Encompassing and the General-to Specific Approach to Specification Search". *Econometrics Journal*, 2, pp. 167-191.
- Johansen S. (1988).** "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12, pp. 389-402.
- Johansen S. (1991).** "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, vol. 59, pp. 1551-1580.
- Krolzig H.-M. et Hendry D.F. (2001).** "Computer Automation of General-to-Specific Model Selection Procedures", *Journal of Economic Dynamics and Control*, 25 (6-7), pp. 831-866.
- Nicholls D.F. et Pagan A.R. (1983).** "Heteroscedasticity in Models with Lagged Dependent Variables", *Econometrica*, vol. 51, pp. 1233-1242.
- Scherrer S. (1996).** "Prévoir l'activité et les effectifs à partir des enquêtes de conjoncture", *Document de Travail*, n° 96-13, Direction de la Prévision.

Annexe : sélection automatique de modèles " du général au spécifique "

Pour une régression : $y_t = \xi_t \gamma + u_t$ où :

– $\xi_t = (\xi_{t1}, \dots, \xi_{kt})$ est la valeur à la date t du vecteur des exogènes ;

– $\gamma = (\gamma_1, \dots, \gamma_k)$ est le vecteur des coefficients à estimer ;

on note :

– $\hat{\gamma}$ l'estimation de γ ;

– $t\gamma_i$ le t de Student de γ_i ;

– $S = (s_1, \dots, s_p)$ le vecteur des p -values de p tests de spécification retenus *a priori* (tels que test de normalité, test d'hétéroscédasticité, ...).

On suppose que le processus de génération des données s'écrit : $E(y_t / X_t) = X_t \beta$ où X_t est un ensemble de variables, inconnues, appartenant à un ensemble $I_t = \{Z_{1t}, \dots, Z_{mt}\}$

La sélection du modèle s'obtient alors par application des étapes suivantes :

Étape 1 : estimation du modèle initial $y_t = Z_t \beta + u_t \Rightarrow \hat{\beta}$

On se fixe un seuil α de significativité des coefficients ($\alpha = 5\%$ en général) et un seuil η de significativité des tests de spécification ($\eta = 1\%$ en général). Soit $c_{1-\frac{\alpha}{2}}$ le $1-\frac{\alpha}{2}$ ème quantile de la loi de Student,

a) si tous les coefficients sont significatifs c'est-à-dire si $|t_{\beta_i}| > c_{1-\frac{\alpha}{2}}$, alors le modèle final est le modèle initial.

Arrêter le processus ;

b) si au moins un coefficient est non significatif, effectuer les p tests de spécification ; soit $S_1 = (s_1, \dots, s_p)$ le vecteur constitué des seuils de significativité de ces tests ; si $\forall j, s_j \geq \eta$ (tous les tests de spécification acceptés au seuil η) alors poser $S^* = [\eta, \dots, \eta]$; sinon, pour tous les j tels que $s_j \geq \eta$ (tests de spécification rejetés pour le modèle initial : cela peut signifier qu'il manque une variable explicative dans le modèle initial, mais le programme n'est pas capable d'en proposer de nouvelle ; pour continuer le processus, il faut donc ajuster le seuil d'acceptation des tests), poser $s_j^* = \eta$ et, pour tous les j tels que $s_j < \eta$, poser $s_j^* = s_j / k$ avec $k > 1$ imposé *a priori*.

Passer à l'étape 2.

Étape 2 : réduction du modèle initial

On se fixe un seuil ϕ de significativité (élevé, par exemple 50%),

a) classer les $\hat{\beta}_i$ par ordre croissant de significativité $\Rightarrow \hat{\beta}^*$;
poser $F=1$; $S = S_1$ (conditions d'amorçage du processus) ;

b) tant que $F > \phi$ et $S > \eta$ estimer le modèle : $y_t = Z_{t,-[1:F]} \beta + u_t$ où $Z_{t,-[1:F]}$ = ensemble des variables z_i classées par ordre croissant de significativité, les i premières variables étant alors exclues.

Calculer :

- le seuil de significativité F du test de $H_0: \beta_1^* = \dots = \beta_k^* = 0$;
- le vecteur S constitué des seuils de significativité des tests de spécification.

Étape 3 : recherche de modèles réduits admissibles par simplifications successives

Pour *tous* les coefficients initialement non significatifs du modèle obtenu à la fin de l'étape 2 :

a) enlever la variable correspondante, calculer le vecteur S_2 constitué des seuils de significativité des tests de spécification. Si $\exists i, S_2(i) \leq S^*(i)$ alors arrêter la recherche pour cette variable et passer au coefficient initialement non significatif suivant ;

b) sinon : poser $t = 0$; $S = S_2$ (conditions d'amorçage du processus). Tant que $|t\beta| < c_{1-\frac{\alpha}{2}}$, $S < S^*$, et que le modèle n'a pas déjà été estimé sur un chemin précédent :

– enlever la variable la moins significative restante ;

– estimer le modèle, calculer le t de Student du coefficient le moins significatif et le vecteur S constitué des seuils de significativité des tests de spécification ;

– stocker le chemin par lequel on est parvenu à ce modèle (c'est-à-dire l'ensemble des modèles intermédiaires estimés).

L'étape s'arrête, soit lorsqu'on rencontre un modèle déjà testé, soit lorsqu'il n'y a plus de variable non significative, soit parce qu'un test de spécification a été rejeté.

c) si l'étape b ne s'est pas arrêtée parce qu'on avait rencontré un modèle déjà testé, stocker le modèle obtenu à la fin de l'étape dans la liste des modèles terminaux.

Étape 4 : sélection d'un modèle terminal unique ou, si ce n'est pas possible, détermination d'un nouveau modèle initial

Soit $\Sigma = \{M_1, \dots, M_n\}$ la liste des modèles terminaux,

a) si la liste des modèles terminaux est vide ($n = 0$), le modèle final est le modèle estimé à l'étape 2. Arrêter le processus ;

b) si la liste des modèles terminaux ne comprend qu'un modèle ($n = 1$), c'est le modèle final. Arrêter le processus ;

c) sinon, former le modèle union U_1 comprenant l'ensemble des variables explicatives figurant dans les modèles terminaux Σ ;

d) pour chacun des modèles terminaux, tester par un test de Fisher au seuil ψ si ce modèle est accepté contre le modèle union :

– si un seul modèle est accepté contre le modèle union, ce modèle est le modèle final. Arrêter le processus ;

– si tous les modèles sont rejetés contre le modèle union et que tous les coefficients du modèle union sont significatifs, alors le modèle union est le modèle final. Arrêter le processus.

Sinon passer à l'étape 5.

Étape 5 : détermination du modèle final

a) refaire l'étape 3 en partant du modèle union U_1 ;

b) si la liste des modèles terminaux est vide, le modèle final est le modèle union. Arrêter le processus ;

c) si la liste des modèles terminaux ne comprend qu'un modèle, c'est le modèle final. Arrêter le processus ;

d) sinon, former le modèle union U_2 comprenant l'ensemble des variables explicatives figurant dans les modèles terminaux ;

e) pour chacun des modèles terminaux, tester par un test de Fisher si ce modèle est accepté contre le modèle union U_2 :

– si un seul modèle est accepté contre le modèle union, ce modèle est le modèle final. Arrêter le processus ;

– si tous les modèles sont rejetés contre le modèle union, alors le modèle union est le modèle final. Arrêter le processus ;

– sinon choisir le modèle final en appliquant un critère de sélection (Akaike, Schwarz, Hannan-Quinn, ...) à l'ensemble des modèles non rejetés contre le modèle union U_2 .

La méthode suppose donc :

– le choix de l'ensemble des variables initiales I ;

– le choix de p tests de spécification ;

– le choix du critère de sélection ;

– le choix des seuils suivants :

– le seuil ϕ du test de Fischer de réduction du modèle initial ;

– le seuil α de significativité des tests de simplification (t de Student et test de Fischer) ;

– le seuil η des tests de spécification ;

– le seuil ψ d'élimination des modèles terminaux.