

Réponses individuelles aux enquêtes de conjoncture et prévision de la production manufacturière

Olivier Biau, Hélène Erkel-Rousse et Nicolas Ferrari*

Nous comparons les performances de soldes d'opinion et d'indicateurs proposés par Mitchell, Smith et Weale pour la prévision à un trimestre du taux de croissance de la production manufacturière. Les sources utilisées sont l'enquête sur la situation et les perspectives dans l'industrie et les comptes trimestriels publiés par l'Insee. Les indicateurs se réfèrent aux questions portant sur les productions passée et prévue des unités de production enquêtées.

Contrairement aux soldes d'opinion, les indicateurs de Mitchell, Smith et Weale ont comme particularité de tenir compte de l'hétérogénéité des comportements de réponse des entrepreneurs à l'enquête de conjoncture. Les réponses des entrepreneurs qui sont les plus en phase avec le taux de croissance de la production manufacturière sont celles qui contribuent le plus à la variabilité de ces indicateurs. Il s'agit de vérifier si cette propriété se traduit par une capacité prédictive supérieure à celle d'indicateurs plus classiques, comme le solde d'opinion. Les applications de Mitchell, Smith et Weale sur des données britanniques et allemandes le suggèrent, mais pas leurs applications sur des données suédoises et portugaises.

Dans cette étude effectuée sur des données françaises, les performances prédictives des indicateurs de Mitchell, Smith et Weale s'avèrent inférieures ou, au mieux, équivalentes à celles des soldes d'opinion, selon les modèles utilisés. Ce résultat paraît robuste en raison de la grande taille du panel de données françaises et de la méthode d'évaluation des indicateurs qui est retenue, les qualités prédictives de ces derniers étant testées en dehors de leur période d'estimation.

** Au moment de la réalisation de cette étude, Olivier Biau et Nicolas Ferrari étaient membres de la division des Enquêtes de Conjoncture de l'Insee. Hélène Erkel-Rousse est chef de la division Croissance et Politiques Macroéconomiques de l'Insee. Les traitements statistiques de cette étude ont été réalisés courant avril 2005. Les auteurs remercient Éric Dubois, Dominique Ladiray, Philippe Scherrer, Fabien Toutlemonde et deux rapporteurs anonymes pour leurs commentaires et suggestions. Ils sont aussi redevables aux participants de la conférence internationale du Ciret tenue à Rome en septembre 2006 pour leurs questions et réactions.*

Du fait de leur publication très rapide, vers la fin du mois de leur réalisation, les résultats des enquêtes de conjoncture permettent l'élaboration d'indicateurs précoces d'activité économique en amont de la sortie des comptes nationaux trimestriels. Sur leur base, des prévisions à court terme des principaux agrégats des comptes trimestriels (Pib, production et emploi par secteur, investissement, consommation, etc.) peuvent être effectuées.

La plupart des questions posées à ces enquêtes de conjoncture sont qualitatives et appellent une réponse à trois modalités, positive (« en hausse » ou « supérieur(e) à la normale »), intermédiaire (« stable » ou « proche de la normale ») et négative (« en baisse » ou « inférieur(e) à la normale »). Les principaux indicateurs résumés calculés à partir des réponses individuelles des enquêtés à ces questions sont des combinaisons des pourcentages de réponses positives, intermédiaires et négatives. Ainsi, le solde d'opinion, l'indicateur le plus couramment utilisé par les conjoncturistes pour résumer les réponses à une question d'une enquête, est défini comme la différence entre le pourcentage de réponses positives et le pourcentage de réponses négatives à cette question. De tels indicateurs sont qualifiés d'*agrégés* au sens où ils sont obtenus à partir d'informations agrégées sur les réponses des enquêtés.

Les indicateurs agrégés sont parfois critiqués au motif qu'ils n'exploitent pas l'hétérogénéité des comportements de réponses aux enquêtes. Mitchell, Smith et Weale (désignés, dans la suite de cet article, par MSW) (2002, 2004 et 2005) présentent des indicateurs concurrents, qualifiés de *désagrégés* au sens où le mode de traitement des réponses individuelles sous-jacent à leur élaboration tient compte de cette hétérogénéité. En effet, il assure que les réponses des entreprises reflétant les fluctuations d'activité globalement les plus en phase avec celles de la production manufacturière dans son ensemble contribuent davantage à la variabilité de leurs indicateurs que les réponses des entreprises à activité plus acyclique. L'objectif est d'aboutir ainsi à des indicateurs plus performants que les indicateurs agrégés pour la prévision à court terme de la production manufacturière globale.

Les applications de MSW à des données d'enquête issues de la CBI (Confederation of British Industry) et de l'Ifo, institut d'analyse économique allemand (Institut für Wirtschaftsforschung), suggèrent effectivement que ces indicateurs désagrégés permettraient d'obtenir des estimations précoces du taux de croissance de la production industrielle plus précises que celles

tirées d'un certain nombre d'indicateurs agrégés classiques. Toutefois, leurs applications à des données d'enquêtes portugaises et suédoises concluent, à l'opposé, en faveur de ces indicateurs agrégés. Ainsi, MSW ne parviennent pas à établir clairement la supériorité en prévision de leurs indicateurs, ce qui appelle des applications complémentaires sur d'autres données, telles que celle proposée ici sur données françaises.

Le but est ici de comparer les performances prédictives à un trimestre de quatre types d'indicateurs désagrégés élaborés en utilisant la méthodologie de MSW (2002 et 2004) à celles du solde d'opinion, concernant la prévision du taux de croissance de la production manufacturière.

Agrégation et quantification des réponses aux enquêtes de conjoncture : les principales approches

La majorité des indicateurs résumés des réponses aux questions qualitatives posées aux enquêtes de conjoncture résulte de l'application de méthodes standard d'agrégation et de quantification. Les pourcentages de réponses positives, intermédiaires et négatives, dans le cas le plus fréquent des réponses individuelles à une question à trois modalités, sont rarement utilisés comme tels. Un ensemble de trois séries n'est en effet pas aisé à suivre dans le temps. Dès lors, la plupart des indicateurs agrégés portant sur une question donnée sont fondés sur des combinaisons de ces trois pourcentages.

C'est le cas de l'indicateur le plus utilisé en analyse conjoncturelle, le solde d'opinion. Celui-ci est calculé comme la différence entre les pourcentages (généralement pondérés) de réponses positives et négatives à la question à laquelle il se réfère. Les soldes d'opinion sont simples à calculer, plus synthétiques, donc plus faciles à suivre dans le temps que les pourcentages de réponses positives, intermédiaires et négatives, au prix d'une perte d'information considérée en général comme acceptable. Ils sont peu révisés dans le temps, ou pas du tout, selon les méthodes pratiquées par les instituts producteurs des enquêtes sources. Surtout, les soldes d'opinion, notamment ceux qui portent sur l'activité, captent bien les grandes fluctuations de l'agrégat économique auquel ils se réfèrent, tout en étant plus lisses que cet agrégat, donc plus lisibles. C'est le cas, par exemple, des soldes d'opinion portant sur la production passée et les perspectives personnel-

les de production dans l'enquête de l'Insee sur la situation et les perspectives dans l'industrie, désignée dans la suite comme l'enquête *Industrie* de l'Insee (cf. graphique I). Theil (1952) puis d'autres parmi lesquels Fansten (1976) et Biau et Ferrari (2006) ont établi des fondements théoriques à l'usage des soldes d'opinion, mais assortis de conditions de validité relativement restrictives. Les soldes d'opinion sont en outre bien adaptés aux conditions de production et de publication rapides des enquêtes de conjoncture. Ceci n'empêche cependant pas de tenter d'élaborer en parallèle des indicateurs optimisés au regard d'un usage spécifique.

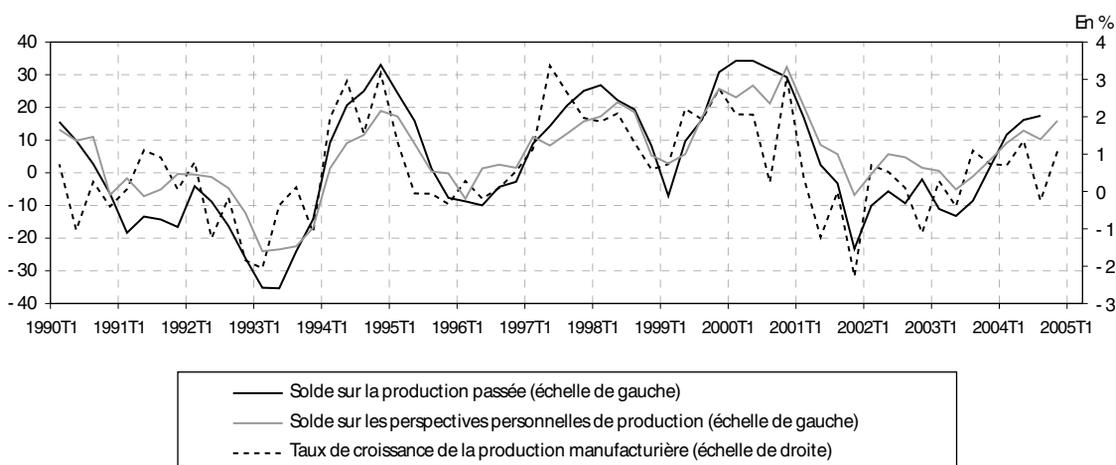
Plusieurs méthodes d'agrégation des réponses individuelles aux enquêtes de conjoncture sont envisageables dès que l'on souhaite obtenir des indicateurs répondant au mieux à un certain objectif : description de la position actuelle dans le cycle (on parle alors d'« indicateur coïncident »), évaluation de la probabilité d'une accélération ou d'un ralentissement à brève échéance (indicateur « de retournement ») ou prévision précoce d'un agrégat macroéconomique (indicateur « avancé »). La majorité des indicateurs existants sont de type agrégé. Les soldes d'opinion ou des combinaisons de soldes sélectionnés ou, encore, des dérivations plus complexes de soldes d'opinion (dans le cadre de l'élaboration d'indicateurs de retournement, particulièrement) sont souvent privilégiés.

De nombreux indicateurs *coïncidents* sont utilisés dans l'administration économique.

L'indicateur de sentiment économique publié par la Commission européenne, par exemple, est défini comme une moyenne pondérée de cinq composantes, les indicateurs de confiance dans l'industrie, les services, le commerce de détail, le BTP et des consommateurs. Ces composantes sont calculées à partir de soldes d'opinion tirés d'enquêtes de conjoncture auprès d'entrepreneurs de ces branches d'activité ou de consommateurs, sélectionnés pour leur corrélation élevée avec certains agrégats macroéconomiques ou sectoriels – cf. Commission européenne (2006). De même, les indicateurs de climat de l'Isae (Istituto di Studi e Analisi Economica) pour l'Italie et de l'Ifo pour l'Allemagne sont calculés comme des moyennes de transformations de trois (respectivement deux) soldes d'opinion. L'indicateur synthétique de climat dans l'industrie publié par l'Insee est également une combinaison linéaire de plusieurs soldes centrés réduits mais, dans ce cas, les pondérations des soldes sont estimées en utilisant des techniques d'analyse factorielle (1). Il en est de même pour les facteurs communs sectoriels dans les autres secteurs de l'économie construits par Bouton et Erkel-Rousse (2003) ainsi que pour l'indicateur synthétique pour la zone euro introduit par Lengart et Toutlemonde (2003) et publié régulièrement par l'Insee. Ce dernier indicateur combine des soldes d'opinion tirés des enquêtes de conjoncture harmoni-

1. Cf. Doz et Lengart (1999). Une définition rapide de l'indicateur est donnée sur le site Internet de l'Insee (http://www.insee.fr/fr/indicateur/indic_conj/donnees/doc_idconj_11.pdf).

Graphique I
Soldes d'opinion relatifs à la production manufacturière et taux de croissance trimestrielle de la production manufacturière



Lecture : le solde portant sur la production passée est décalé d'un trimestre par rapport aux deux autres séries afin que les périodes de référence de chaque série soient identiques pour tout trimestre *t*.
Sources : enquête Industrie de l'Insee et comptes nationaux trimestriels (situation au printemps 2005, moment où a été réalisée l'étude empirique).

sées des plus grands États membres de la zone. L'indicateur synthétique dans les services publié par l'Insee est estimé en ayant recours à des techniques d'analyse factorielle dynamique qui permettent de combiner un ensemble de soldes d'opinion de périodicités hétérogènes et renseignés sur des périodes différentes, cf. Cornec et Deperraz (2007). L'introduction d'indicateurs tirés de résultats d'enquêtes de conjoncture dans des modèles à facteurs est un moyen de lier les résultats de ces enquêtes à une variable économique latente reflétant le cycle des affaires dans un secteur donné ou au niveau macroéconomique, selon les indicateurs utilisés. L'indicateur synthétique qui en résulte, qualifié de facteur commun, est supposé refléter cette variable latente (2).

De nombreux indicateurs *de retournement* sont tirés de l'estimation de processus de Markov à variable cachée, suivant en cela Hamilton (1989, 1990) et Lahiri et Wang (1994). Par exemple, l'indicateur de retournement mensuel dans l'industrie publié par l'Insee est fondé sur une méthode en deux étapes proposée par Gregoir et Lengart (1998, 2000). Cette méthode consiste à extraire les innovations contenues dans six soldes d'opinion de l'enquête *Industrie* de l'Insee puis à déduire de leurs signes les probabilités d'une accélération ou d'un ralentissement conjoncturel en utilisant un modèle de Markov à variable cachée. L'indicateur de retournement est alors défini comme la différence entre les probabilités estimées d'une accélération et d'un ralentissement. La Commission européenne et la Banque d'Espagne ont développé un indicateur de retournement fondé sur une approche similaire, mais en mêlant des résultats d'enquêtes (les mêmes que ceux utilisés pour calculer l'indicateur de climat industriel de la Commission européenne) avec l'indice de production industrielle pour la zone euro. Leur indicateur de retournement est défini comme la probabilité estimée d'un ralentissement, cf. Bengoechea et Pérez-Quirós (2004).

Les méthodes d'élaboration d'indicateurs *avancés* (3) les plus courantes dans la littérature consistent à combiner les pourcentages de réponses positives et négatives en les liant à tel ou tel agrégat macroéconomique d'intérêt (observé ou latent). C'est le cas notamment des méthodes connues sous les dénominations d'approches « probabiliste » et « par la régression » (4). Esquissée par Theil (1952) et développée par Carlson et Parkin (1975), l'approche probabiliste repose sur l'hypothèse que la réponse donnée par un entrepreneur concernant

une variable d'intérêt comme la croissance de la production résulte d'une fonction de densité de probabilité subjective pour cette variable qui peut être spécifique à l'entreprise et est conditionnée à l'information disponible au sein de celle-ci. La moyenne de cette densité de probabilité subjective s'avère être un prédicteur sans biais de l'agrégat d'intérêt. On montre qu'elle peut être calculée à partir des pourcentages de réponses positives et négatives à la question portant sur la variable d'intérêt, sous certaines conditions techniques permettant d'identifier les fonctions de répartition de ces pourcentages (5). Plus simple, l'approche par la régression, introduite par Anderson (1952) et développée par Pesaran (1984), revient à régresser l'agrégat macroéconomique d'intérêt sur les pourcentages (pondérés de manière appropriée) de réponses positives et négatives portant sur cet agrégat (6). Citons aussi l'approche dite « du facteur latent » (D'Elia, 2005), qui considère les pourcentages de réponses positives, intermédiaires et négatives comme des fonctions d'une mesure latente commune de l'agrégat d'intérêt observée par les enquêtés, mais non par les statisticiens. Des techniques multivariées d'analyse factorielle permettent à D'Elia d'estimer la dynamique des variations de ce facteur latent. Enfin, dans l'approche dite de la régression inversée introduite par Cunningham, Smith et Weale (1998), les réponses qualitatives de chaque entrepreneur sont supposées déterminées par une variable latente spécifique à l'entreprise, liée à l'agrégat macroéconomique d'intérêt par un modèle linéaire, selon une règle commune à toutes les entreprises. Des liens ont été établis entre les différentes approches. On peut montrer que les indicateurs tirés de chacune d'entre elles peuvent être approchés par les soldes d'opinion sous certaines conditions techniques (Anderson, 1952 ; Theil, 1952 ; Lankes et Wolters, 1988 ; Mitchell, Smith et Weale, 2004 ; D'Elia, 2005).

2. Pour des fondements théoriques de ce genre de modèles (appliqués à des données d'enquêtes de conjoncture ou non), voir notamment Stone (1947), Sargent et Sims (1977), Stock et Watson (2002), Doz et Lengart (1999), Forni, Hallin, Lippi et Reichlin (2001), Cornec et Deperraz (2007) et cf. *infra*.

3. Sur les indicateurs avancés en général, on pourra se reporter par exemple à Lahiri et Moore (1991).

4. Pour des présentations plus détaillées, voir Nardo (2003), MSW (2002, 2004) ou D'Elia (2005).

5. Exemples d'applications de l'approche probabiliste : Carlson (1975), Watchel (1977), Fische et Lahiri (1981), Batchelor (1981), Batchelor et Orr (1987), Dasgupta et Lahiri (1992), Lee (1994), Balcombe (1996) et Berk (1999), parmi beaucoup d'autres.

6. L'approche par la régression est appliquée, par exemple, à la Banque d'Angleterre – voir Britton, Cutler et Wardlow (1999). Il y a eu de nombreuses extensions des approches probabilistes et par la régression – cf. notamment Pesaran (1984, 1987), Smith et McAleer (1995), Driver et Urga (2004).

Dans l'administration économique, la plupart des indicateurs utilisés pour la prévision à court terme sont plutôt des soldes d'opinion spécifiques choisis pour leur caractère avancé ou des indicateurs composites combinant plusieurs soldes. Néanmoins certains conjoncturistes (tel Hild, 2003) recommandent l'utilisation d'indicateurs fondés sur des pourcentages de réponses positives, intermédiaires et négatives plutôt que sur des soldes d'opinion. Hild (2007) développe cette approche en proposant un indicateur synthétique tiré de la combinaison d'un certain nombre d'indicateurs élémentaires prenant en compte la dynamique des réponses individuelles à l'enquête *Industrie* de l'Insee. Ces indicateurs élémentaires sont élaborés à partir des pourcentages d'entrepreneurs modifiant leur réponse à telle ou telle question (par exemple, en passant de la réponse intermédiaire à la réponse positive ou encore de la réponse intermédiaire à la réponse négative) d'une enquête à la suivante. L'idée sous-jacente est que les évolutions des réponses individuelles sont susceptibles de constituer des signes avant-coureurs d'un changement de tendance de l'activité.

Les indicateurs agrégés sont souvent critiqués au motif qu'ils n'exploitent pas l'hétérogénéité des comportements de réponses des enquêtés. En effet, les réponses individuelles y sont pondérées en fonction de critères de représentativité (au regard de la taille, de la branche d'activité, etc.) ne tenant pas compte de la plus ou moins grande fidélité avec laquelle les chroniques des réponses individuelles reproduisent les évolutions de la variable macroéconomique d'intérêt sous-jacente à la question posée. Certains auteurs, tels Kaiser et Spitz (2000) ou, plus encore, MSW (2002, 2004, 2005) préfèrent les indicateurs dits *désagrégés* au sens où ils intègrent cette hétérogénéité. Les applications de MSW à des données britanniques et allemandes suggèrent que leurs indicateurs désagrégés permettraient d'obtenir des estimations précoces plus précises du taux de croissance de la production industrielle que trois indicateurs agrégés usuels. Ceci justifie la poursuite de l'évaluation de leurs indicateurs sur des données d'autres pays.

Les applications de MSW et leurs principaux résultats

MSW (2002, 2004-1) appliquent leur méthodologie (cf. encadré 1) à des données trimestrielles de l'enquête de la CBI tirées

des réponses à la question portant sur l'évolution passée de la production. Leur échantillon couvre les réponses de 5002 entrepreneurs sur la période allant du troisième trimestre de 1988 au troisième trimestre de 1997 (soit 37 trimestres). Les entreprises ayant répondu à moins de 20 enquêtes sont exclues, le calcul des composantes individuelles des indicateurs de MSW étant conditionné par la disponibilité d'un nombre minimum d'observations par entreprise. Cette sélection se traduit par une chute importante de la taille du panel. Il ne reste les réponses que de 643 entrepreneurs après éviction des entreprises insuffisamment « bonnes répondantes » (7). MSW (2002, 2004-1) concluent qu'en termes de prévision à court terme du taux de croissance de la production manufacturière, leurs indicateurs aboutissent à de meilleurs résultats que les indicateurs agrégés proposés par Carlson et Parkin (1975) (approche probabiliste), Pesaran (1984) (approche par la régression) et Cunningham *et al.* (1998) (approche par la régression inversée). Ils étendent ensuite leur étude comparative de ces trois indicateurs agrégés et de leurs deux indicateurs désagrégés dans deux directions.

D'une part, MSW (2004-2) appliquent leur méthodologie à des données d'enquêtes *Industrie* harmonisées allemandes, portugaises, suédoises et, à nouveau, britanniques, de périodicité mensuelle dans le cas des trois premiers pays, trimestrielle pour le dernier. Ils ne précisent pas les tailles des panels retenus après sélection des entreprises bonnes répondantes, mais ils en indiquent la taille initiale, qui s'élève, respectivement, pour l'Allemagne, le Portugal, le Royaume-Uni et la Suède, à 9703, 1528, 5519 et 1620 réponses d'entreprises en tout, représentant une moyenne de 3843, 832, 1142 ou 784 réponses par mois ou trimestre, selon le pays. Les critères de sélection des entreprises varient selon les pays. Pour l'Allemagne, par exemple, le critère est plus exigeant que pour le Royaume-Uni : sont conservées les entreprises ayant répondu au moins 96 fois sur la période. MSW (2004-2) concluent que leurs indicateurs désagrégés sont plus performants que les indicateurs agrégés sur les données allemandes et britanniques et le résultat s'avère opposé sur les données portugaises et suédoises.

7. L'expression entreprises « insuffisamment bonnes répondantes » renvoie à tout entrepreneur ayant donné moins de vingt réponses à l'enquête sur la période étudiée, soit parce qu'il ne répond pas régulièrement à l'enquête, soit parce qu'il n'a pas reçu le questionnaire sur toute la période (cas d'entreprises ayant disparu ou, au contraire, ayant été créées ou, plus généralement, intégrées à l'échantillon des enquêtées en cours de période).

LES INDICATEURS DE MITCHELL, SMITH ET WEALE (2002, 2004)

Les indicateurs de MSW (2002, 2004) sont conçus pour assurer que les réponses des entreprises enquêtées contribuent d'autant plus à la variabilité des indicateurs qu'elles relatent des fluctuations d'activité plus en phase avec celles de la production manufacturière dans son ensemble. Cet encadré reprend les notations de MSW.

Notons x_t le taux de croissance trimestriel de la production manufacturière des comptes nationaux trimestriels. Comme les résultats des enquêtes de conjoncture sont publiés avant les comptes nationaux, il est possible d'en inférer une estimation quantitative précoce de x_t . Considérons une enquête de conjoncture relative à la période t (trimestre ou mois), dans le cadre de laquelle un échantillon de N_t entrepreneurs de l'industrie manufacturière expriment leur opinion sur l'évolution de leur production passée (hausse, stable ou baisse) pendant la période de référence t , t indiquant donc la période de référence de l'agrégat x et de la question considérée, non la période d'interrogation de l'enquête.

Soit $j_{i,t}$ ($j_{i,t} \in \{-1, 0, +1\}$) la réponse de l'entrepreneur i à cette question pour la période t , où -1 , 0 et $+1$ correspondent respectivement aux réponses « baisse », « stable » et « hausse ». Au niveau microéconomique, la variable d'intérêt dans l'optique de l'obtention d'une prévision précoce de x_t est $E(x_t / j_{i,t}, i)$, l'espérance mathématique de x_t conditionnellement à la réponse $j_{i,t}$ donnée par l'entrepreneur i à la question considérée pour la période t . L'estimation de cette espérance conditionnelle conduit à une évaluation de la variable macroéconomique x_t sur la base de la réponse individuelle de l'entrepreneur i .

Les indicateurs introduits par MSW (2002, 2004) au niveau macroéconomique sont définis comme les moyennes arithmétiques simples, calculées sur l'échantillon des entrepreneurs « bons répondants » (cf. note 7 *supra*), des estimations $\hat{E}(x_t / j_{i,t}, i)$ des espérances mathématiques de x_t conditionnellement à la réalisation des réponses individuelles données pour la période t :

$$IND_t = \frac{1}{N_t} \sum_{1 \leq i \leq N_t} \hat{E}(x_t / j_{i,t}, i) \quad (1)$$

L'estimation des espérances conditionnelles individuelles nécessite celle des densités conditionnelles correspondantes $f(x_t / j, i)$, où l'indice j représente la variable aléatoire dont la réalisation est la réponse $j_{i,t}$ donnée par l'entrepreneur i pour la période t . Les auteurs proposent deux modes d'estimation possibles de ces densités conditionnelles individuelles. Le premier relève d'une approche non paramétrique très simple, le second d'une approche paramétrique un peu plus complexe à mettre en œuvre. Chaque approche

conduit à la définition d'un indicateur de type « désagrégé ».

L'indicateur non paramétrique de MSW

Les densités théoriques $f(x_t / j, i)$ sont approchées par les densités conditionnelles empiriques des réponses individuelles. Pour cela, il faut disposer :

- des variables indicatrices $(y_{i,t}^j)_{j=-1,0,+1}$ relatives à la question considérée, avec $y_{i,t}^j = 1$ si l'entrepreneur i a donné la réponse j pour la période t et $y_{i,t}^j = 0$ sinon ;
- du nombre T_i^j de fois où l'entrepreneur i a donné la même réponse j sur la période d'estimation globale $1, \dots, T$: $T_i^j = \sum_{\tau \in \{1, \dots, T\}} y_{i,\tau}^j$;
- du nombre total de réponses T_i données par l'entrepreneur i durant la période d'estimation $1, \dots, T$: $T_i = \sum_{j \in \{-1, 0, +1\}} T_i^j$.

L'espérance conditionnelle $E(x_t / j_{i,t}, i)$ est estimée par la moyenne simple des taux de croissance (x_τ) sur la sous-période où l'entreprise i a donné la même réponse $j_{i,t}$ qu'en t :

$$\hat{E}(x_t / j_{i,t}, i) = \frac{1}{T_i^j} \sum_{\tau=1}^T x_\tau y_{i,\tau}^j$$

L'indicateur non paramétrique de MSW, noté $MSWNP$, est défini comme la moyenne arithmétique simple des espérances conditionnelles ainsi estimées :

$$MSWNP_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \hat{E}(x_t / j_{i,t}, i) \quad (2)$$

Une illustration graphique permettant de visualiser cet indicateur est donnée en annexe 1.

L'indicateur paramétrique de MSW

Les densités conditionnelles individuelles $f(x_t / j, i)$ s'expriment (formule de Bayes) :

$$f(x_t / j, i) = \frac{P(j / x_t, i) f(x_t)}{P(j / i)} \quad (3)$$

MSW estiment la densité $f(x_t)$ de la loi de x_t dans l'ensemble des lois normales et de Pearson, qui approchent assez bien la distribution de x_t . Il s'avère que les performances prédictives des indicateurs de MSW ne dépendent pas significativement du choix entre une distribution normale ou une loi de Pearson. Soit $\hat{f}(x_t)$ l'estimateur de $f(x_t)$ retenu. Notons que, pour l'application sur données françaises, suivant les conclusions →

D'autre part, MSW (2004-2 et 2005) raffinent le calcul de l'indicateur paramétrique pour l'appliquer à la question prospective : ils traitent *ex ante* le problème d'endogénéité qui peut apparaître quand on estime le modèle (6) de l'encadré 1 en utilisant les réponses données à la question portant sur les perspectives personnelles de production ($y_{i,t}$ dans la partie gauche de l'équation, t se rapportant au mois ou au trimestre courant) et le taux de croissance de la production manufacturière observé au mois ou au trimestre suivant dans sa partie droite. Les résultats de MSW (2005) sont en faveur de l'indicateur paramétrique modifié sur les don-

nées britanniques. Dans les applications sur les données des autres pays, l'indicateur paramétrique modifié est peu testé, MSW (2004-2) se centrant sur l'indicateur non paramétrique lorsqu'ils traitent la question prospective.

Une application aux résultats de l'enquête *Industrie* de l'Insee : méthodologie

Notre application s'appuie sur des données françaises de l'enquête *Industrie* de l'Insee

Encadré 1 (suite)

du test de normalité de Shapiro-Wilk, une fonction de densité normale est choisie pour cette densité.

Dès qu'un estimateur $\hat{P}(j/x_t, i)$ de $P(j/x_t, i)$ est disponible, il est facile d'en inférer un estimateur du dénominateur de la formule de Bayes (3) en calculant l'intégrale $\int \hat{P}(j/x_t, i) \hat{f}(x_t) dx_t$, puis un estimateur $\hat{f}(x_t/j, i)$ de la densité conditionnelle $f(x_t/j, i)$ par la formule de Bayes (3). Un estimateur de l'espérance conditionnelle individuelle $E(x_t/j, i)$ en découle immédiatement :

$$\hat{E}(x_t/j, i) = \int x_t \hat{f}(x_t/j, i) dx_t \quad (4)$$

L'indicateur paramétrique de MSW, noté *MSWP*, est défini comme la moyenne arithmétique simple des espérances conditionnelles ainsi estimées :

$$MSWP_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \hat{E}(x_t/j_{i,t}, i) \quad (5)$$

Il reste à calculer les $\hat{P}(j/x_t, i)$, dont découle (5). Ces estimations résultent de celles de modèles à choix discrets ordonnés. MSW (2002, 2004) supposent que le taux de croissance trimestriel de la production de l'entreprise i à la période t , noté $y_{i,t}$, dépend du taux de croissance macroéconomique x_t selon le modèle conditionnel :

$$y_{i,t} = \alpha_i + \beta_j x_t + \varepsilon_{i,t} \quad (6)$$

où α_i et β_j sont des coefficients invariants dans le temps et spécifiques à l'entreprise i . x_t est supposé stationnaire. En outre, en dépit de la relation quasi comptable entre x_t et les $(y_{i,t})_{i=1, \dots, N_t}$, les tests de l'exogénéité faible de x_t dans (6) menés par MSW aboutissent à son acceptation pour la quasi-totalité des entreprises. Les $(\varepsilon_{i,t})_{t=1, \dots, T}$ désignent des composantes spécifiques (erreurs) supposées indépendamment et identiquement distribuées (*i.i.d.*) dans le temps et de fonction de répartition logistique commune F : $F(z) = 1/(1 + e^{-z})$, $-\infty < z < +\infty$.

$y_{i,t}$ est inobservée mais l'enquête de conjoncture apporte des informations sur son sens de variation. MSW utilisent des modèles à choix discrets ordonnés fondés sur la régression latente (6), en supposant l'existence de seuils μ_j^{-1} et μ_j^{+1} tels que :

$$y_{i,t} < \mu_j^{-1} \Leftrightarrow y_{i,t}^{-1} = 1$$

$$\mu_j^{-1} \leq y_{i,t} < \mu_j^{+1} \Leftrightarrow y_{i,t}^0 = 1$$

$$y_{i,t} \geq \mu_j^{+1} \Leftrightarrow y_{i,t}^{+1} = 1$$

Ils en déduisent la probabilité $P_{j,i,t} \equiv P(j/x_t, i)$ d'observer la réponse $y_{i,t}^j = 1$ pour la modalité j et la période t , conditionnellement à la valeur de x_t et à l'entrepreneur i :

$$\begin{cases} P_{-1,i,t} = F(\mu_i^{-1} - \alpha_i - \beta_j x_t) \\ P_{0,i,t} = F(\mu_i^{+1} - \alpha_i - \beta_j x_t) - F(\mu_i^{-1} - \alpha_i - \beta_j x_t) \\ P_{+1,i,t} = 1 - F(\mu_i^{+1} - \alpha_i - \beta_j x_t) \end{cases} \quad (7)$$

Les $(\varepsilon_{i,t})_{t=1, \dots, T}$ étant *i.i.d.*, la vraisemblance relative à l'entreprise i est définie par :

$$\mathcal{L}_i = \prod_t \prod_{j \in \{-1, 0, 1\}} (P_{j,i,t})^{y_{i,t}^j} \quad (8)$$

La maximisation de la vraisemblance (8) permet d'obtenir des estimateurs convergents quand $T \rightarrow +\infty$ des coefficients β_j et $(\mu_j^j - \alpha_i)$ (les modèles à choix discrets sont identifiés à un paramètre d'échelle près ; l'estimation de β_j , α_i et μ_j^j nécessiterait de contraindre, par exemple, le premier paramètre de seuil μ_j^{-1} à zéro). En découlent les probabilités conditionnelles estimées $\hat{P}_{j,i,t}$, par application de (7) en remplaçant les coefficients théoriques par leurs estimations. L'indicateur *MSWP* en est déduit par application successive des formules (3), (4) et (5), en remplaçant les probabilités théoriques $P_{j,i,t}$ par les probabilités estimées $\hat{P}_{j,i,t}$.

portant sur les branches manufacturières (réponses aux questions rétrospective et prospective sur la production des entreprises par produit) et sur des séries de taux de croissance de la production par grande branche manufacturière issues des comptes nationaux trimestriels. Pour chaque question (rétrospective, prospective) et chaque méthodologie (non paramétrique, paramétrique), on estime les indicateurs de MSW sur l'ensemble du panel des réponses d'une part et sur quatre sous-panels relatifs aux grandes branches manufacturières (biens de consommation, biens d'équipement, automobile et biens intermédiaires) d'autre part. Pour chaque sous-panel, on applique le même critère de sélection des unités répondantes que sur le panel total (cf. encadré 2). Lors des estimations sur les sous-panels, le taux de croissance de la production x_t considéré se

rapporte à la branche (cf. encadré 1). On tire des estimations sur les sous-panels un indicateur relatif à l'ensemble des produits manufacturés en procédant en deux étapes. On régresse tout d'abord chaque taux de croissance de la production par branche sur l'indicateur de MSW correspondant et une constante dans un but de normalisation. On effectue ensuite la moyenne des indicateurs par branche ainsi normalisés (taux de croissance par branche ajustés) pondérés par la part des branches dans la production manufacturière retardée d'un trimestre. Les indicateurs résultants sont qualifiés d'indicateurs « sectoriels » de MSW, par référence à leur obtention à partir d'agrégations d'estimations réalisées à un niveau plus fin que les indicateurs de MSW dits « macroéconomiques », estimés directement à partir du panel total.

Encadré 2

LES SOURCES

L'application présentée s'appuie sur des données françaises de l'enquête *Industrie* de l'Insee portant sur les branches manufacturières. Cette enquête harmonisée au niveau européen est réalisée tous les mois auprès d'environ 4 000 entreprises. Les résultats sont disponibles environ 5 jours ouvrables avant la fin du mois de réalisation de chaque campagne d'enquête. On utilise les réponses aux questions rétrospective et prospective relatives à la production des entreprises par produit. Ces deux questions sont formulées plus précisément de la façon suivante dans l'enquête :

Question rétrospective :

Votre production : évolution au cours des trois derniers mois ↗ → ↘

Question prospective :

Votre production : évolution probable au cours des trois prochains mois ↗ → ↘

Le taux de croissance de la production manufacturière en volume est tiré des comptes nationaux trimestriels de l'Insee (résultats détaillés des comptes trimestriels du quatrième trimestre de 2004 exprimés en base 1995). Son chiffrage à un trimestre donné est publié pour la première fois environ 50 jours après la fin du trimestre, puis il est révisé environ 90 jours après la fin du trimestre.

L'analyse couvre la période allant du premier trimestre de 1990 au quatrième trimestre de 2004. La périodicité trimestrielle est adoptée pour la raison suivante. Les prévisions de l'Insee publiées quatre fois par an dans ses *Notes et Points de conjoncture* consistent à extrapoler les principaux agrégats macroéconomiques des comptes nationaux trimestriels sur la base

de séries de même périodicité tirées notamment de résultats d'enquêtes de conjoncture. Ainsi, on choisit de tester les performances des indicateurs de Mitchell, Smith et Weale pour la prévision du taux de croissance de la production manufacturière des comptes trimestriels, variable d'intérêt privilégiée dans les campagnes de prévision. Une autre possibilité eût été de choisir l'indice de production industrielle (l'IPI) comme agrégat à prévoir et de travailler sur données mensuelles ; mais ce n'est pas ce qui est fait lors des préparations des *Notes et Points de conjoncture* de l'Insee.

On retient les résultats des enquêtes réalisées durant les mois de janvier, avril, juillet et octobre, dont les réponses aux questions sur la production passée et sur les perspectives personnelles de production se réfèrent à des trimestres pleins du fait des périodes de référence des questions (les trois derniers mois pour la question rétrospective et les trois prochains mois pour la question prospective), sachant que la plupart des réponses aux enquêtes sont retournées en début de mois. Ainsi, par exemple, les réponses données à l'enquête réalisée en janvier à la question rétrospective (respectivement prospective) se réfèrent au quatrième trimestre de l'année précédente (respectivement au premier trimestre de l'année courante). Ces réponses peuvent donc être directement rapprochées de l'agrégat macroéconomique à prévoir.

On applique la méthodologie de Mitchell, Smith et Weale (2002, 2004-1) à ces deux questions. L'indice de temps relatif aux réponses des entrepreneurs à la question rétrospective est décalé d'un trimestre par rapport à l'indice de temps relatif aux réponses à la question prospective. Ainsi, les réponses données à une enquête réalisée en janvier de l'année a (par exemple) à la question rétrospective (respectivement

→

Le calcul des indicateurs sectoriels de MSW est motivé par les arguments suivants. Les indicateurs macroéconomiques peuvent être biaisés si les unités dont on retient les réponses ne sont pas uniformément distribuées d'une branche à l'autre, particulièrement si les prévisions macroéconomiques tirées des réponses individuelles sont influencées par la conjoncture de leur branche d'activité. Les indicateurs sectoriels de MSW sont obtenus à partir de modèles à choix discrets liant les réponses élémentaires à l'enquête à des taux de croissance de la production calculés à un niveau moins agrégé. Le résultat attendu est une amélioration de l'ajustement, par correction d'au moins une part de cet éventuel biais. En outre, l'application de la méthodologie de MSW au niveau des grandes

branches de l'industrie manufacturière limite *a priori* les inconvénients du caractère statique du modèle (6) de l'encadré 1, qui exclut la prise en compte de phénomènes de diffusion. En effet, contrairement à l'application sur le panel total, celle au niveau des sous-panels permet de lever l'hypothèse peu réaliste que les conjonctures des branches des biens intermédiaires et des biens d'équipement, par exemple, sont coïncidentes.

En ce qui concerne les indicateurs agrégés, on privilégie les soldes d'opinion publiés par l'Insee (soldes pondérés), qui sont les indicateurs de référence usuels des conjoncturistes. Ainsi, le solde d'opinion portant sur une question donnée (rétrospective ou prospective) est défini comme la différence entre les pourcenta-

Encadré 2 (suite)

prospective) sont rapprochées du taux de croissance de la production manufacturière au quatrième trimestre de l'année $a - 1$ (respectivement au premier trimestre de l'année a). Ce décalage permet de conserver la convention utilisée dans l'encadré 1 pour l'indice de temps t : dès lors, celui-ci indique le trimestre sur lequel porte l'agrégat des comptes nationaux, quelle que soit la question.

L'enquête *Industrie* de l'Insee pose les questions portant sur la production passée et prévue au niveau non pas de l'entreprise mais de ses produits (un enquêteur pouvant déclarer au plus quatre produits), l'objectif étant de mesurer au plus près l'évolution de l'activité industrielle. Dans l'application sur données françaises, les unités élémentaires (représentées par l'indice i dans l'encadré 1) se réfèrent donc non à une entreprise mais à un produit d'une entreprise. Sur la période considérée, le nombre total de réponses à l'enquête *Industrie* s'élève à 185 204, recouvrant 9 918 unités élémentaires (croisement entreprises \times produits). Ainsi, une unité élémentaire a donné en moyenne 19 réponses sur la période. Ces réponses élémentaires proviennent de 6 955 entrepreneurs ; il est déclaré en moyenne 1,4 produit par entreprise. À l'instar de MSW (2002, 2004-1), pour chaque question, on exclut les unités élémentaires ayant répondu à moins de 20 enquêtes sur la période traitée ; il reste alors 131 955 réponses, représentant 3 627 unités élémentaires. Le nombre moyen de réponses par unité élémentaire retenue (36) est nettement supérieur à celui calculé sur le panel initial (19). Les réponses de 2 803 entreprises demeurent. Le nombre de produits déclarés en moyenne par entreprise (1,3) est similaire à celui calculé sur le panel initial. Les chiffres cités portent sur la question rétrospective. Les ordres de grandeur sont les mêmes pour la question prospective.

Le panel d'étude français diffère donc de ceux utilisés par MSW (2002, 2004, 2005) par sa plus grande taille. C'est certain dans le cas des panels britannique,

portugais et suédois, *a priori* un peu moins dans le cas du panel allemand. En effet, avant sélection des unités « bonnes répondantes » (cf. note 7 *supra*), le panel français contient un peu moins d'observations que le panel allemand en moyenne par campagne d'enquête (3 087 réponses contre 3 843), en dépit de sa taille totale sensiblement supérieure du fait de sa couverture temporelle plus longue. Cependant, le critère de sélection appliqué ensuite au panel allemand étant très exigeant (élimination des unités ayant répondu moins de 96 fois, contre 20 dans le cas français), si on se restreint aux unités effectivement retenues, le panel français est probablement sensiblement plus fourni que le panel allemand. Ces écarts de taille entre les panels pourraient induire des divergences de résultats d'un pays à l'autre. Intuitivement, en effet, on s'attend plutôt à ce que l'indicateur paramétrique de MSW se comporte mieux lorsqu'il est estimé sur des données nombreuses que sur des petits échantillons et que l'indicateur non paramétrique de MSW, plus parcimonieux, soit, quant à lui, moins pénalisé par de petits échantillons.

La plus grande taille du panel français est due au nombre supérieur d'entrepreneurs interrogés, mais aussi au fait que les réponses des enquêtés portent sur chacun de leurs principaux produits plutôt que sur l'activité de leur entreprise en général. Cette caractéristique de l'enquête de l'Insee permet de disposer d'un panel plus fourni pour l'étude empirique que MSW, ce qui constitue un point positif en termes de robustesse des résultats futurs, dans la mesure où les réponses données par un entrepreneur pour ses différents produits ne sont pas automatiquement très corrélées entre elles dans le temps. En d'autres termes, ces réponses multiples par entrepreneur sont porteuses d'une information conjoncturelle significative. Ce constat justifie l'interrogation des entrepreneurs au niveau de leurs produits plutôt que de leur entreprise dans son ensemble.

ges pondérés de réponses positives et négatives, ce qui, en utilisant les mêmes notations que plus haut, s'écrit :

$$Solde_t = \sum_{1 \leq i \leq N_t} \omega_i (y_{it}^{+1} - y_{it}^{-1}),$$

$$\text{avec : } \sum_{1 \leq i \leq N_t} \omega_i = 1$$

$(\omega_i)_{i=1, \dots, N_t}$ sont des pondérations qui reflètent la taille de l'entreprise et le poids de sa branche d'activité dans l'économie (pondérations en deux étapes, dites doubles) (8).

L'analyse consiste donc à comparer les performances prédictives de dix indicateurs (cinq par question), soit, pour chaque question : le solde d'opinion, les deux indicateurs macroéconomiques de MSW (paramétrique ; non paramétrique), les deux indicateurs sectoriels de MSW (paramétrique ; non paramétrique).

La nécessité d'une analyse « hors échantillon »

Sur leur période d'estimation, les indicateurs de MSW épousent mieux le profil du taux de croissance trimestriel de la production manufacturière que les soldes d'opinion (cf. tableau 1 et graphique II) (9). Ceci résulte cependant de ce que les soldes, contrairement aux indicateurs de MSW, sont calculés indépendamment de l'agrégat comptable. La meilleure adéquation des indicateurs de MSW à l'agrégat d'intérêt sur la période d'estimation apparaît donc largement tautologique. Ce n'est pas cela qui présente de l'intérêt et permet d'évaluer les différents indicateurs les uns par rapport aux autres.

Tableau 1
Corrélations entre les indicateurs étudiés et le taux de croissance de la production manufacturière sur la période d'estimation

Indicateurs	Question rétrospective	Question prospective
MSWNP macroéconomique	0,86	0,86
MSWP macroéconomique	0,85	0,87
MSWNP sectoriel	0,86	0,88
MSWP sectoriel	0,84	0,86
Solde d'opinion	0,71	0,64

Source : enquête Industrie de l'Insee et comptes nationaux trimestriels. Calculs des auteurs.

Pour être véritablement éclairante, la comparaison des dix indicateurs doit être menée en dehors de leur période d'estimation. Il s'agit alors d'étudier les performances relatives de chaque indicateur au regard de l'objectif retenu de prévision à un trimestre du taux de croissance x_t de la production manufacturière. Pour ce faire, on compare les séries d'erreurs de prévision simulées à partir de deux familles de modèles (la première, « statique », la seconde, « dynamique ») rapprochant la variable expliquée x_t d'une ou plusieurs variable(s) explicative(s) construites sur la base d'un type d'indicateurs ou d'un autre.

Première étape : régressions statiques

Dans une première phase, on estime les huit indicateurs de MSW relatifs aux questions sur les productions passée et prévue en mobilisant la chronique des observations (x_t) (10) restreinte à la période complète $[1, T]$ à l'exclusion du trimestre t , t compris entre 1 et T . On calcule les chroniques des huit indicateurs de MSW étudiés sur la totalité de la période $[1, T]$, à partir des données d'enquête disponibles sur cette période complète et des résultats d'estimation sur la sous-période $[1, T]$ à l'exclusion du trimestre t . On calcule aussi les soldes d'opinion sur cette même sous-période. Aucun de ces calculs ne mobilise l'observation x_t des comptes. Dans une deuxième phase, on régresse par les moindres carrés ordinaires (MCO) le taux de croissance trimestriel de la production manufacturière x_t sur une constante et un indicateur (Ind) parmi les dix considérés ($x_t = a Ind_t + b + u_t$) sur la même sous-période $[1, T]$ à l'exclusion du trimestre t , où τ représente l'indice de temps, a et b deux coefficients à estimer et u_t un terme d'erreur. On utilise ensuite chacune des dix équations estimées pour prévoir le taux de croissance de la production manufacturière x_t au trimestre t . On confronte enfin la prévision

8. Les raisons ayant guidé notre choix des pondérations utilisées pour chaque type d'indicateurs sont expliquées et discutées en détail dans Biau, Erkel-Rousse et Ferrari (2005).

9. On présente en annexe 2 des éléments descriptifs sur les distributions des coefficients estimés des paramètres (β_i) du modèle (6), qui entrent dans l'estimation de l'indicateur paramétrique macroéconomique de MSW.

10. Dans cette première phase, x_t représente le taux de croissance de la production manufacturière totale ou d'une branche, selon l'indicateur de MSW considéré (indicateur macroéconomique ou composante d'un indicateur sectoriel). Dans les phases suivantes, x_t se réfère exclusivement au niveau manufacturier total.

obtenue \hat{x}_t à sa valeur observée x_t en calculant l'erreur de prévision $e_t = x_t - \hat{x}_t$.

On itère ces opérations pour tout t compris entre 1 et T (premier et dernier trimestres pour lesquels on dispose de données) (11). On en tire dix séries d'erreurs de prévision, dont on compare les moyennes arithmétiques, les écarts-types et les erreurs quadratiques moyennes. Les séries d'erreurs de prévision sont suffisamment longues (60 observations chacune) pour se prêter à la réalisation de tests d'égalité des performances prédictives selon la méthodologie de Harvey, Leybourne et Newbold (1997) dans des bonnes conditions de robustesse (12), cf. encadré 3.

L'approche choisie ici ne reflète pas le travail en temps réel du conjoncturiste, qui se

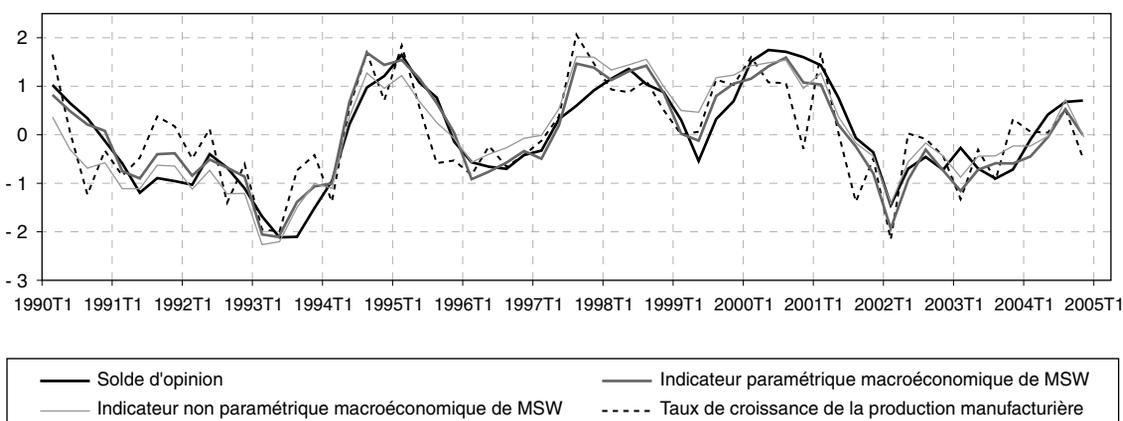
préoccupe exclusivement de prévoir la valeur de l'agrégat d'intérêt sur la période la plus récente. Cependant, pour artificielle qu'elle puisse paraître, cette approche présente des avantages certains pour l'analyse. D'une part, elle permet de travailler sur des périodes d'estimation de longueur identique et maximale à chaque itération (c'est-à-dire quel que soit le trimestre t prévu). D'autre part, elle maximise la longueur des séries d'erreurs simulées. La combinaison de ces deux propriétés renforce la robustesse des résultats.

11. L'origine de temps ($t = 1$) est contrainte par la disponibilité du détail des réponses individuelles à l'enquête Industrie de l'Insee, qui n'est aisée qu'à partir de janvier 1990.

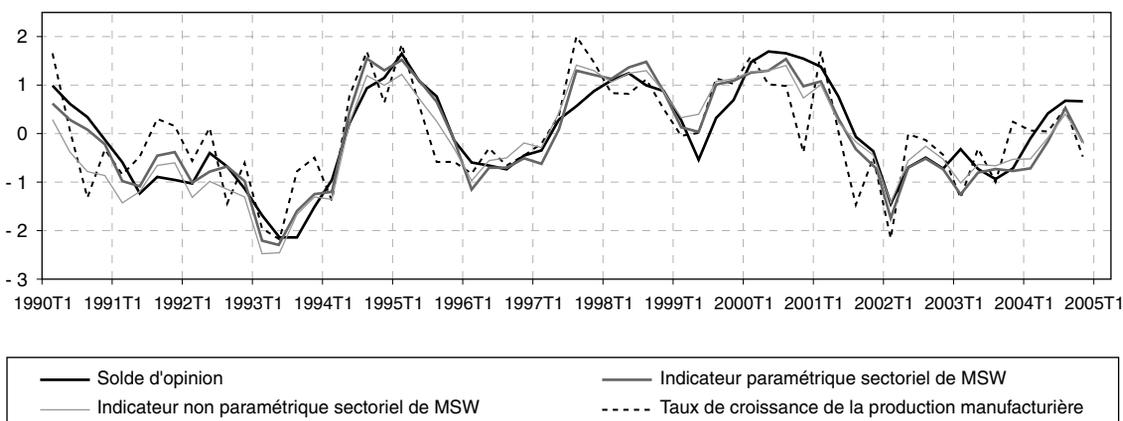
12. MSW (2002, 2004) évoquent la réalisation de tels tests, mais MSW (2005) mentionnent qu'il n'en ont pas effectué en raison de la taille insuffisante de leur échantillon simulé, suivant Ashley (2003).

Graphique II
Les indicateurs centrés réduits estimés sur l'ensemble de la période
Cas de la question portant sur la production passée

A – Indicateurs macroéconomiques de MSW



B – Indicateurs sectoriels de MSW



Lecture : les séries relatives aux indicateurs sont décalées d'un trimestre de sorte que le trimestre des comptes nationaux s'aligne avec la période de référence de la question rétrospective. Le lecteur peut consulter les graphiques des indicateurs portant sur la question prospective dans Biau et al. (2005).

Source : enquête Industrie de l'Insee et comptes nationaux trimestriels. Calculs des auteurs.

Seconde étape : régressions dynamiques

Les modèles de prévision utilisés dans l'administration économique combinent en général les effets dynamiques de plusieurs variables explicatives, notamment ceux des soldes d'opinion relatifs à la production passée et prévue – cf. par exemple Reynaud et Scherrer (1996) ou Dubois et Michaux (2006). On considère donc ici des modèles dont la spécification est plus proche de celles des modèles opérationnels, dynamiques. On estime tout d'abord les huit indicateurs de MSW sur la sous-période $1, \dots, t-1$, pour un trimestre t compris entre T_0 et T (13). Les valeurs de ces indicateurs sur la période totale $[1, T]$ sont déduites de ces estimations et de la totalité des données d'enquête, le point important étant que les observations (au niveau macroéconomique ou d'une grande branche) $(x_\tau)_{\tau \in \{1, \dots, t-1\}}$ ne sont pas prises en compte dans le calcul des

indicateurs sur la sous-période $[t, T]$. On calcule également les deux soldes d'opinion, dont la dérivation est indépendante des valeurs de x . On note Ind_PA et Ind_PR les indicateurs, d'un type donné, relatifs à la production passée et prévue qui résultent de ces calculs. Dans un second temps, on estime par les MCO cinq modèles dynamiques (un par type d'indicateurs) formulés comme suit, toujours sur la sous-période $1, \dots, t-1$:

$$x_\tau = a_1 Ind_PA_{\tau-1} + a_2 Ind_PA_{\tau-2} + b_0 Ind_PR_\tau + b_1 Ind_PR_{\tau-1} + c + u_\tau$$

13. T_0 est le premier trimestre de 1999. Commencer les simulations en 1999 permet de disposer d'un nombre minimum acceptable d'observations à la fois sur la période antérieure (de 1990 T1 à 1998 T4), pour effectuer des estimations relativement fiables, et sur la période simulée, pour que les résultats des simulations portent sur un nombre suffisant de points.

Encadré 3

PRINCIPE DU TEST DE HARVEY, LEYBOURNE ET NEWBOLD (1997) D'ÉGALITÉ DES PERFORMANCES PRÉDICTIVES

Considérons deux séries d'erreurs de prévisions à l'horizon h , $\{e_{1t}\}_{t=1}^T$ et $\{e_{2t}\}_{t=1}^T$ et une transformation de ces erreurs : $g(e_{it}), \forall i = 1, 2$. L'hypothèse nulle d'égalité des performances prédictives s'écrit $E[g(e_{1t})] = E[g(e_{2t})]$ ou bien $E[d_t] = 0$, où $d_t = g(e_{1t}) - g(e_{2t})$ désigne la fonction d'écart entre les erreurs. Si on cherche à évaluer les performances prédictives à partir de l'erreur quadratique moyenne (MSE) alors $g(e_{it}) = e_{it}^2$.

On montre que : $\sqrt{T}(\bar{d} - \mu) \xrightarrow{L} N(0, 2\pi f_d(0))$

avec : $\bar{d} = \frac{1}{T} \sum_{t=1}^T [g(e_{1t}) - g(e_{2t})]$, $f_d(0) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{+\infty} \gamma_d(\tau)$,

la densité spectrale de la fonction d'écart à la fréquence zéro, $\gamma_d(\tau) = E[(d_t - \mu)(d_{t-\tau} - \mu)]$ l'autocovariance de la fonction d'écart et μ la moyenne théorique de celle-ci.

La statistique proposée par Diebold et Mariano (1995) pour tester l'hypothèse nulle d'égalité des performances prédictives est égale à :

$$S = \frac{\bar{d}}{\sqrt{\frac{2\pi f_d(0)}{T}}}$$

où $f_d(0)$ est estimée à partir des $h-1$ premiers termes de l'estimateur de Newey – West :

$$2\pi \hat{f}_d(0) = \hat{\gamma}_d(0) + 2 \sum_{\tau=1}^{h-1} \omega(\tau, m) \hat{\gamma}_d(\tau),$$

avec : $\hat{\gamma}_d(\tau) = \frac{1}{T} \sum_{t=|\tau|+1}^T (d_t - \bar{d})(d_{t-\tau} - \bar{d})$

et où $\omega(\tau, m) = 1 - \frac{|\tau|}{m+1}$ désigne la fenêtre de lissage, où m est le point de troncation de la fenêtre de Bartlett.

Harvey, Leybourne et Newbold (1997) montrent que le test proposé par Diebold et Mariano est peu puissant pour des échantillons de petites et moyennes tailles. Ils proposent de corriger la statistique de test en la multipliant par \sqrt{C} , où $C = \frac{T+1-2h+(1/T)h(h-1)}{T}$.

Pour tenir compte d'éventuelles queues de distributions épaisses, ils suggèrent en outre de comparer la statistique $\sqrt{C} \cdot S$ non pas aux valeurs critiques d'une loi normale mais à celles issues d'une loi de Student à $(T-1)$ degrés de liberté.

Dans le cas particulier (considéré dans la présente étude) où $h=1$, l'estimateur de Newey-West de la fonction de densité spectrale à la fréquence zéro

se réduit à $\hat{\gamma}_d(0) = \frac{1}{T} \sum_{t=1}^T (d_t - \bar{d})^2$ et la statistique de

Harvey *et al.* à $\sqrt{T-1} \frac{\bar{d}}{\sqrt{\hat{\gamma}_d(0)}}$.

Les résultats des tests de Harvey *et al.* effectués ici sont présentés en annexe 3.

où l'indice τ représente le trimestre courant (au sens de la convention expliquée dans l'encadré 1) et x_τ représente le taux de croissance de la production manufacturière (14). De chaque modèle estimé sur la période $1, \dots, t-1$ et des valeurs des indicateurs au trimestre t , on tire une prévision \tilde{x}_t de x_t pour le trimestre t (15), puis on en déduit la valeur de l'erreur de prévision $e_t = x_t - \tilde{x}_t$ au trimestre t .

On itère ces opérations pour tout trimestre t compris entre T_0 et T . On dispose *in fine* d'une série d'erreurs de prévision à l'horizon d'un trimestre sur la sous-période allant du deuxième trimestre de 1999 au quatrième trimestre de 2004 pour chaque type d'indicateurs. La comparaison de ces cinq séries d'erreurs donne des éléments d'information sur la manière dont les cinq types d'indicateurs pourraient se comporter dans le cadre des exercices de prévision effectués par les conjoncturistes.

Dans cette seconde étape, l'approche retenue est, cette fois, effectivement plus proche de la pratique conjoncturelle que dans la première étape : à chaque itération, on prévoit la valeur de l'agrégat d'intérêt pour le trimestre correspondant à la première observation postérieure à la période d'estimation. L'exercice est donc moins artificiel que celui mené dans la première étape de l'analyse hors échantillon. Cependant, on y combine dès lors des périodes d'estimation de longueurs plus courtes et hétérogènes d'une itération à l'autre. En outre, les séries d'erreurs simulées comptent moins d'observations que dans la première étape (23 au lieu de 60 chacune). Par conséquent, les tests systématiques d'égalité des performances prédictives d'Harvey *et al.* (cf. encadré 3) réalisés sur

leur base risquent de conduire à des conclusions plus fragiles. Les deux étapes de l'analyse hors échantillon se complètent donc, chacune présentant des avantages et des inconvénients différents.

Les résultats obtenus sur données françaises : le cas des modèles statiques

Les prévisions à un trimestre tirées de l'utilisation d'indicateurs portant sur la question rétrospective sont associées à des erreurs quadratiques moyennes des erreurs de prévision numériquement un peu inférieures à celles tirées de l'utilisation d'indicateurs fondés sur la question prospective (cf. tableau 2). Ceci est conforme à l'intuition puisque les réponses à la question rétrospective reposent sur une information plus récente (d'un trimestre) que les réponses à la question prospective. Cependant, les résultats frappent surtout par la disparition de l'avantage relatif des indicateurs de MSW sur les soldes. En effet, d'après une série de tests de Harvey *et al.* (1997), les performances prédictives de trois des quatre modèles utilisant un indicateur désagrégé sont non significativement différentes de celles du modèle utilisant le solde d'opinion correspondant ; quant au qua-

14. Les résultats qualitatifs sont inchangés quand on utilise des modèles à deux retards.

15. À la fin juillet, par exemple, on dispose des résultats d'enquête jusqu'à juillet et des comptes trimestriels jusqu'au deuxième trimestre. On prévoit le troisième trimestre des comptes (x_t) à partir des modèles en mobilisant l'information la plus récente, tirée de la dernière enquête (résumée dans Ind_PA_{t-1} et Ind_PR_t) et de l'enquête précédente (résumée dans Ind_PA_{t-2} et Ind_PR_{t-1}).

Tableau 2
Erreurs de prévision à un trimestre sur le taux de croissance de la production manufacturière (régressions simples)

Question	Indicateurs	Erreur moyenne (en point)	Écart-type des erreurs (en point)	Erreur quadratique moyenne
Production passée	MSWNP macroéconomique	0,025	1,058	1,101
	MSWP macroéconomique	0,017	0,965	0,915
	MSWNP sectoriel	0,019	0,942	0,993
	MSWP sectoriel	0,014	0,932	0,854
	Solde d'opinion	- 0,001	0,955	0,897
Production prévue	MSWNP macroéconomique	0,038	1,110	1,213
	MSWP macroéconomique	0,036	1,041	1,068
	MSWNP sectoriel	0,036	1,057	1,100
	MSWP sectoriel	0,035	1,018	1,021
	Solde d'opinion	- 0,003	1,030	1,043

Lecture : les statistiques sont calculées à partir des 60 observations simulées lors de la première étape de l'analyse hors échantillon. Sources : enquête Industrie de l'Insee et comptes nationaux trimestriels. Calculs des auteurs.

trième modèle, fondé sur l'indicateur non paramétrique macroéconomique de MSW, il apparaît légèrement moins performant que le modèle avec solde (au seuil de 10 %), du moins pour la question rétrospective (cf. annexe 3, tableaux A et B). D'autres tests de même type concluent que le modèle mobilisant l'indicateur paramétrique « sectoriel » de MSW se comporte un peu mieux (au seuil de 5 ou 10 %, selon les cas) que les modèles utilisant un des autres indicateurs de MSW.

Ces résultats suggèrent que les indicateurs de MSW n'anticipent pas mieux la production manufacturière que le solde d'opinion à l'horizon d'un trimestre. Les soldes d'opinion, indépendants de l'agrégat à prévoir, ne seraient finalement pas moins fiables pour anticiper cet agrégat.

La différence entre les résultats obtenus dans l'analyse en échantillon et dans cette première analyse hors échantillon peut sembler à première vue peu intuitive. En effet, on aurait pu s'attendre à ce qu'une étude à base de régressions simples aboutisse aux mêmes conclusions que l'observation de corrélations. Ce n'est pas le cas parce que les valeurs des indicateurs de MSW au trimestre t peuvent différer sensiblement selon qu'elles résultent de l'estimation sur l'ensemble de la période $[1, T]$ ou sur la sous-période $[1, T]$ à l'exclusion du trimestre t . Ceci est illustré sur un exemple par le graphique III.

Enfin, les erreurs de prévision relativement élevées obtenues quel que soit le modèle utilisé méritent un bref commentaire. Les prévisions sont en moyenne non biaisées (erreurs moyennes proches de zéro, cf. tableau 2), particulièrement dans le cas des modèles utilisant un solde d'opinion, mais leurs écarts-types sont assez élevés au regard de l'ordre de grandeur de l'agrégat d'intérêt (qui est exprimé en pourcentages). Ceci reflète le fait que les indicateurs tirés des enquêtes de conjoncture sont notablement plus lisses que les agrégats macroéconomiques des comptes trimestriels. On reviendra plus loin sur ce point.

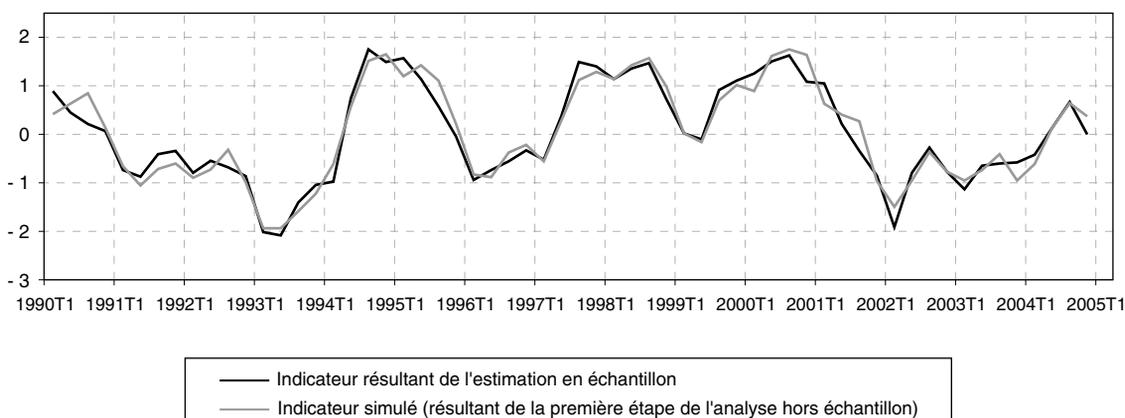
Les résultats obtenus sur données françaises : le cas des modèles dynamiques

Les résultats des simulations tirées de la seconde étape de l'analyse hors échantillon suggèrent que le modèle mobilisant les soldes d'opinion conduit à des prévisions à un trimestre du taux de croissance de la production manufacturière plus précises que les autres modèles (cf. tableau 3 et graphique IV), ce que confirment les tests de Harvey *et al.* (cf. tableau C, annexe 3).

Une explication possible des meilleures performances des soldes d'opinion dans cette seconde

Graphique III

L'indicateur paramétrique macroéconomique de MSW portant sur la production passée : valeurs estimées sur l'ensemble de la période et simulées à partir de régressions simples sur la sous-période $[1, T]$ à l'exclusion du trimestre t



Lecture : la valeur de l'indicateur de MSW simulée au trimestre t est obtenue en calculant l'indicateur paramétrique macroéconomique de MSW au trimestre t à partir de l'estimation effectuée sur la sous-période $[1, T]$ à l'exclusion du trimestre t . Sources : enquête Industrie de l'Insee et comptes nationaux trimestriels. Calculs des auteurs.

étape de l'analyse hors échantillon pourrait provenir du nombre limité d'observations sur lesquelles est fondée l'analyse dynamique. D'un côté, des périodes d'estimation courtes pourraient favoriser les indicateurs les plus simples et les plus parcimonieux, c'est-à-dire les soldes d'opinion. D'un autre côté, la longueur réduite de la période de simulation (23 observations, au lieu de 60 dans la première étape de l'analyse hors échantillon) risque de fragiliser les résultats, surtout si la période de simulation est assez spécifique. On ne peut pas non plus exclure que les modèles dynamiques discriminent mieux les différents types d'indicateurs que les régres-

sions statiques. Cependant, les simulations du taux de croissance de la production manufacturière tirées des modèles dynamiques conduisent à des prévisions moins précises que celles obtenues à partir des régressions simples, en dépit du caractère *a priori* plus convaincant des modèles dynamiques. En effet, les erreurs quadratiques moyennes des erreurs de prévision dans le tableau 3 sont un peu plus élevées que celles du tableau 2, en raison de l'existence de biais négatifs sur la période d'environ -0,3 point, c'est-à-dire du même ordre de grandeur que l'agrégat à prévoir. Là encore, ce résultat peut avoir plusieurs origines.

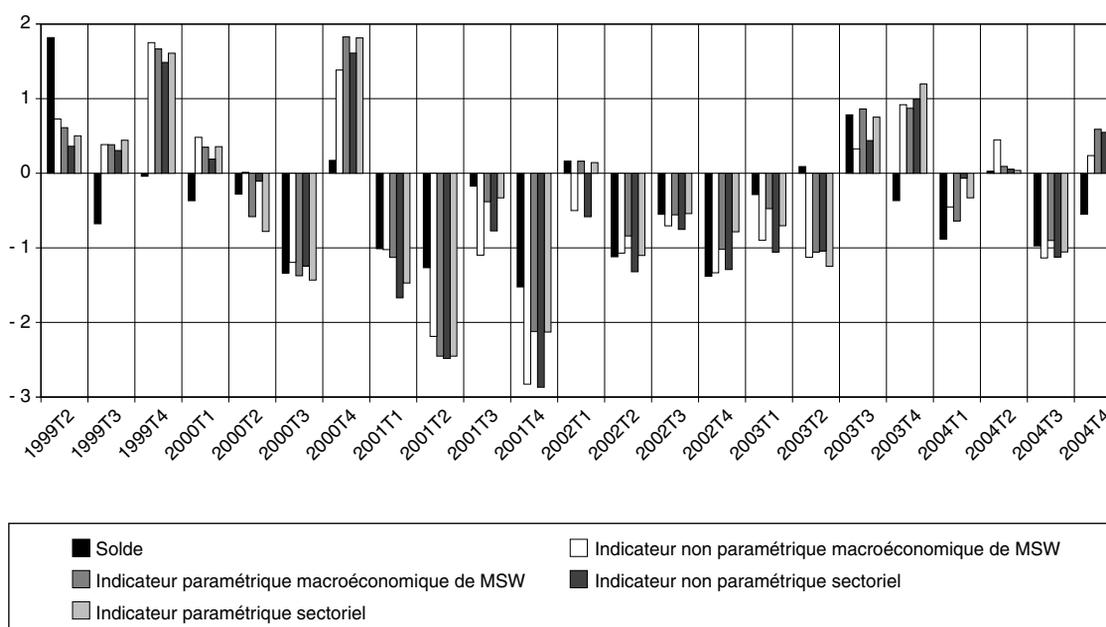
Tableau 3
Erreurs de prévision à un trimestre sur le taux de croissance de la production manufacturière tirées des modèles dynamiques

Indicateurs	Erreur moyenne (en point)	Écart-type des erreurs (en point)	Erreur quadratique moyenne
MSWNP macroéconomique	- 0,382	1,105	1,315
MSWP macroéconomique	- 0,263	1,073	1,172
MSWNP sectoriel	- 0,448	1,138	1,438
MSWP sectoriel	- 0,299	1,110	1,267
Soldes d'opinion	- 0,419	0,762	0,732

Lecture : les statistiques sont calculées à partir des 23 observations simulées lors de la seconde étape de l'analyse hors échantillon. La case grisée indique que l'erreur quadratique moyenne des erreurs de prévision associées au modèle avec soldes d'opinion est significativement inférieure à celle des erreurs de prévision associées aux autres modèles, selon plusieurs tests de Harvey et al. (1997). Pour plus de détails sur ces tests, cf. le tableau C de l'annexe 3.

Sources : enquête Industrie de l'Insee et comptes nationaux trimestriels. Calculs des auteurs.

Graphique IV
Erreurs de prévision à un trimestre tirées des modèles dynamiques



Lecture : pour les 23 trimestres simulés (1999 T2 à 2004 T4) sont présentées les erreurs de prévision e_t obtenues en effectuant la différence entre le taux de croissance de la production manufacturière observé et sa prévision tirée de chacun des cinq modèles dynamiques utilisés. Les valeurs absolues des erreurs de prévision tirées du modèle avec soldes d'opinion (bâtons noirs) sont en moyenne légèrement inférieures à celles des erreurs de prévision tirées des autres modèles.

Sources : enquête Industrie de l'Insee et comptes nationaux trimestriels. Calculs des auteurs.

Pour isoler en partie l'impact relatif des modèles utilisés, de la longueur de la période d'estimation et de la spécificité des périodes d'estimation et de simulation, on restreint l'étude des erreurs de prévision tirées des régressions simples à la période allant du deuxième trimestre de 1999 au quatrième trimestre de 2004, correspondant à celle de la seconde étape de l'analyse hors échantillon. Les résultats obtenus sont résumés dans le tableau 4, ainsi que dans les tableaux D et E de l'annexe 3.

Au seuil de 5 %, les tests de Harvey *et al.* confirment que, sur la sous-période 1999 T2 - 2004 T4 (cf. tableau 4) aussi bien que sur l'ensemble de la période (cf. tableau 2), les performances prédictives moyennes des indicateurs de MSW, lorsqu'ils sont mobilisés au sein de régressions simples, ne diffèrent pas significativement de celles du solde d'opinion correspondant. Sur la sous-période 1999 T2 - 2004 T4, le solde portant sur la question prospective apparaît légèrement plus performant que les indicateurs non paramétriques au seuil de 10 %. Néanmoins, dans l'ensemble, les performances prédictives relatives des cinq types d'indicateurs sont peu modifiées lorsqu'on se concentre sur la sous-période 1999 T2 - 2004 T4. L'effet « modèle » semble prédominer, la combinaison des indicateurs portant sur la production passée et prévue au sein de modèles *a priori* plus convaincants conduisant à un diagnostic plus clairement en faveur des soldes.

En outre, contrairement au tableau 2, le tableau 4 fait apparaître des biais de prévision négatifs fréquents, quoique moins notables que le tableau 3. Sur ce point, un léger impact des types de modèles (statique, dynamique) pourrait être détecté, mais qui ne paraît pas prédominer. Ici, le problème est probablement dû au fait que,

plus la période de simulation est courte, plus les résultats sont influencés par sa spécificité. La sous-période 1999 T2 - 2004 T4 englobe le début des années 2000 et l'année 2004 durant lesquelles les enquêtes de conjoncture françaises suggéraient des taux de croissance plus élevés que ceux publiés par les comptes trimestriels.

Au total, les occurrences quasi systématiques de biais de prévision négatifs dans les tableaux 3 et 4 ne signifient pas que les enquêtes de conjoncture ne peuvent pas être utilisées pour prévoir la croissance de la production manufacturière, ainsi que l'illustre l'absence de biais significatif dans le premier jeu de prévisions tirées d'une période de simulation plus longue (cf. tableau 2). Néanmoins, il arrive qu'à certaines périodes (comme au début des années 2000 ou en 2004), les évaluations de l'activité économique sur la base des enquêtes de conjoncture et des comptes nationaux trimestriels ne soient pas en ligne. Souvent, l'ampleur des écarts d'évaluation tend à décroître au fur et à mesure que les comptes sont révisés. Néanmoins, certains écarts peuvent persister, qui sont susceptibles de conduire à des erreurs de prévision élevées sur certaines périodes. En dehors de ces épisodes de divergence temporaire entre les sources, les principaux écarts à court terme entre les évaluations données par les enquêtes de conjoncture d'une part et les comptes trimestriels d'autre part proviennent du caractère plus lisse des données d'enquêtes. Tout ceci explique les écarts-types relativement élevés des erreurs de prévision tirées des modèles étudiés, quel que soit le modèle utilisé et la période de simulation. Il n'en reste pas moins que les enquêtes de conjoncture demeurent des outils très utiles pour la prévision à court terme de l'activité économique, au sens où elles permettent en général

Tableau 4

Erreurs de prévision à un trimestre sur le taux de croissance de la production manufacturière, selon les indicateurs – régressions simples, résultats sur la sous-période 1999-T2 à 2004-T4

Question	Indicateurs	Erreur moyenne (en point)	Écart-type des erreurs (en point)	Erreur quadratique moyenne
Production passée	MSWNP macroéconomique	- 0,271	1,074	1,177
	MSWP macroéconomique	- 0,023	1,023	1,001
	MSWNP sectoriel	- 0,302	1,031	1,107
	MSWP sectoriel	- 0,108	0,985	0,939
	Solde d'opinion	- 0,174	1,001	0,989
Production prévue	MSWNP macroéconomique	- 0,355	1,151	1,394
	MSWP macroéconomique	- 0,096	1,107	1,181
	MSWNP sectoriel	- 0,358	1,146	1,384
	MSWP sectoriel	- 0,167	1,097	1,179
	Solde d'opinion	- 0,278	0,985	1,005

Sources : enquête Industrie de l'Insee et comptes nationaux trimestriels. Calculs des auteurs.

d'évaluer relativement bien les tendances de ses inflexions, à défaut de l'ampleur de certaines de ses variations d'un trimestre à l'autre.

Bilan et perspectives

Les applications empiriques successives de MSW (2002, 2004) indiquent que leurs indicateurs désagrégés sont plus performants que trois indicateurs agrégés usuels pour la prévision conjoncturelle du taux de croissance de la production manufacturière sur des données allemandes et britanniques, mais non sur des données portugaises et suédoises. En ce qui concerne ces deux derniers pays, le résultat le plus frappant réside dans les valeurs absolues très élevées des erreurs de prévision quel que soit l'indicateur utilisé, sans doute en lien avec la très petite taille des panels sous-jacents ; dès lors, étudier des écarts de performances prédictives d'un indicateur à l'autre est d'un intérêt limité. Les commentaires qui suivent se centrent donc sur la comparaison des résultats tirés des panels français, allemands et britanniques.

L'application sur données françaises compare les performances prédictives à un trimestre de cinq types d'indicateurs : un indicateur agrégé (le solde d'opinion) et quatre indicateurs désagrégés tirés de la méthodologie de MSW (2002, 2004-1). Les soldes d'opinion apparaissent plutôt plus performants que les indicateurs désagrégés. En outre, contrairement à l'application de MSW sur données britanniques, l'étude sur données françaises fait ressortir des indicateurs paramétriques un peu meilleurs en prévision que les indicateurs non paramétriques. Enfin, elle suggère que les indicateurs obtenus en appliquant la méthodologie de MSW au niveau des grandes branches d'activité, avant agrégation plutôt que directement au niveau agrégé, seraient un peu plus performants.

Dans l'application française, les indicateurs doublement pondérés (soldes) se comportent un peu mieux que les indicateurs de MSW, qui sont non pondérés ou simplement pondérés. La correction par la pondération d'un biais de sélection lié à une distribution hétérogène des non réponses pourrait être en cause, dans une configuration où la représentativité par branche (et peut-être aussi par taille) des réponses importerait en raison d'un effet de structure (16). Dans une telle configuration, un système de pondérations reflétant la taille et la branche d'activité des entreprises pourrait permettre d'améliorer

les performances des indicateurs. C'est le cas des soldes d'opinion et, indirectement et dans une moindre mesure, des indicateurs sectoriels de MSW. Bien qu'on dispose de peu d'informations sur le comportement de réponse des entrepreneurs selon la branche d'activité et la taille de l'entreprise, du moins les résultats auxquels aboutit l'application française ne contredisent-ils pas la possibilité d'une telle configuration. Il n'en est pas de même chez MSW (2004-2) : étudiant l'impact de différents modes de pondération sur la construction de leurs indicateurs désagrégés, ceux-ci trouvent que les poids égaux surpassent tout autre jeu de pondérations, dont celles par la taille des entreprises. Ils expliquent ce résultat par le fait que les unités répondantes sont implicitement pondérées dans leurs indicateurs selon leur capacité à évaluer plus ou moins correctement la croissance passée de la production manufacturière globale, au sens où les répondants qui donnent peu d'informations sur l'agrégat d'intérêt contribuent marginalement à la variabilité des indicateurs de MSW. Ils suggèrent aussi, avec raison, que la réflexion sur les pondérations doit être poursuivie, point sur lequel on va revenir.

Au-delà de la question des pondérations, les résultats de l'application française sont directement comparables à ceux des applications de MSW sur données allemandes et britanniques du fait que l'approche par la régression sur données françaises conduit à un indicateur de Pesaran qui ne diffère pas significativement du solde d'opinion, que ce soit pour la question rétrospective ou la question prospective. En effet, si on estime le modèle : $x_t = \alpha U_t + \beta D_t + \chi + u_t$, où U et D désignent les proportions des unités élémentaires répondant respectivement en hausse et en baisse, u le terme d'erreur et α , β et χ des coefficients à estimer, l'hypothèse nulle de coefficients α et β d'amplitude égale et de signes opposés est acceptée sans ambiguïté pour les deux questions (17). Ainsi, l'opposition entre les résultats de MSW et les nôtres n'est pas attribuable à la nature différente des indicateurs agrégés retenus comme références, puisque le solde d'opinion de l'application française

16. Un effet de structure peut survenir si la concentration et la taille des entreprises diffèrent selon les branches et que les réponses des grandes entreprises apportent une information plus riche sur la conjoncture que les entreprises plus petites.

17. Par conséquent, l'indicateur de Pesaran, défini par : $PI_t = \hat{\alpha}U_t + \hat{\beta}D_t + \hat{\chi}$, où $\hat{\alpha}$, $\hat{\beta}$ et $\hat{\chi}$ sont les estimateurs des MCO de α , β et χ , ne diffère pas significativement de : $RPI_t = \hat{\alpha}(U_t - D_t) + \hat{\chi}$, où $\hat{\alpha}$ et $\hat{\chi}$ sont les estimateurs des moindres carrés de α et χ contraints par l'hypothèse nulle $H_0 : \alpha + \beta = 0$ et $(U_t - D_t)$ est le solde d'opinion.

est homogène à l'indicateur de Pesaran étudié par MSW.

Aucune explication aux divergences entre les résultats des trois applications n'apparaît non plus clairement du fait des différences de périodicité des séries et des périodes d'estimation. Les applications française et allemande sont fondées sur des données trimestrielles, l'application britannique sur des données mensuelles, sans conséquence visible sur les résultats. Les périodes d'estimation des trois applications sont 1990-2004 pour la France, 1991-2000 pour l'Allemagne et 1988-1999 pour le Royaume-Uni. La première partie des années 2000, durant laquelle les enquêtes de conjoncture et les comptes nationaux trimestriels ne donnaient pas tout à fait la même image du dynamisme de l'activité économique en France (cf. *supra*), est largement absente des autres panels. Ceci pourrait influencer les résultats, mais dans un sens qui n'est pas immédiat. De même, au sein de la période couverte, les cycles des affaires sont plus corrélés entre la France et l'Allemagne qu'avec le Royaume-Uni ; là aussi, aucun impact clair n'est visible sur les résultats.

En revanche, une partie des divergences entre les résultats des trois applications est probablement due à la plus grande taille du panel français. Intuitivement, un nombre élevé d'observations favorise sans doute l'application de techniques d'estimations plus complexes (indicateur paramétrique) tandis que, pour des panels de petite taille, des indicateurs simples et parcimonieux (indicateur non paramétrique) sont plus indiqués. Cette intuition est cohérente avec les meilleures performances des indicateurs paramétriques dans l'étude française, fondée sur le panel de plus grande taille, ainsi qu'avec le résultat opposé obtenu sur le panel britannique, plus petit, tandis que les performances des deux types d'indicateurs dans l'application sur données allemandes, fondée sur un panel de taille vraisemblablement intermédiaire, sont équivalentes. Toutefois, il faut reconnaître que ce raisonnement rencontre ses limites : sur le panel français, les soldes d'opinion, de calcul très simple, se comportent au moins aussi bien que les indicateurs paramétriques de MSW.

Plus fondamentalement, la taille des panels joue sur la robustesse des résultats. Ainsi, les résultats de l'application française jettent davantage le doute sur les performances des indicateurs de MSW que ceux des applications portugaise et suédoise du fait que la taille du panel français assure une plus grande robustesse des résultats.

Celle des résultats sur données britanniques pourrait être moins certaine en raison du nombre beaucoup plus réduit d'observations sous-jacentes.

Les avantages des soldes d'opinion

À la lumière de l'expérience empirique, au-delà de l'usage prédictif pour lequel ils ont été configurés et au regard duquel leurs performances relatives sont à ce stade non tranchées par les trois applications qui viennent d'être évoquées, les indicateurs de MSW présentent des inconvénients par rapport aux soldes d'opinion.

L'élaboration de l'indicateur paramétrique est lourde en temps de calcul, même pour le traitement d'une seule question. Le calcul du solde d'opinion permet une mise à disposition plus rapide de l'information conjoncturelle contenue dans l'intégralité de l'enquête, ce qui est un avantage notable pour un usage en production.

Le calcul des indicateurs de MSW nécessite de disposer de séries de réponses individuelles relativement longues, avec la conséquence pratique d'éliminer un nombre notable d'observations (celles issues d'entrepreneurs ne répondant pas suffisamment régulièrement à l'enquête ou à la question considérée). Ceci soulève un problème statistique potentiel, tant en termes de degré d'exhaustivité que de biais de sélection (18). Le calcul des indicateurs sectoriels de MSW limite ce biais, sans toutefois le supprimer avec certitude. Le solde d'opinion, quant à lui, est calculé sur un nombre d'observations beaucoup plus élevé, son élaboration ne nécessitant pas un aussi grand nombre de réponses par unité élémentaire. De ce fait, le calcul du solde n'induit pas l'élimination de réponses très probablement informatives sur la conjoncture.

Le solde d'opinion est soumis à des révisions limitées ou nulles, d'une enquête sur l'autre, selon la méthodologie de l'institut producteur de l'enquête (19). Les indicateurs de MSW peuvent, eux, subir des révisions plus notables sur une période plus longue, à la fois parce que toute réestimation suite à l'ajout d'observations supplémentaires touche l'intégralité des séries

18. Même si les vérifications empiriques de MSW (2004-2) ne font pas apparaître de signe tangible de l'existence de tels biais.
19. Pour l'enquête Industrie de l'Insee : les soldes bruts relatifs à l'enquête N - 1 deviennent définitifs à l'intégration de réponses tardives lors de la sortie de l'enquête N ; les soldes corrigés des variations saisonnières sont modifiés lors de chaque campagne annuelle de désaisonnalisation, mais très marginalement du fait de la longueur des séries.

et non seulement leurs derniers points et, surtout, du fait que les révisions régulières du taux de croissance de la production manufacturière dans les comptes trimestriels se répercutent automatiquement sur l'ensemble des séries des indicateurs de MSW.

La sensibilité des indicateurs de MSW au taux de croissance de la production manufacturière des comptes trimestriels, agrégat moins lisse que les soldes d'opinion, se traduit sur le profil de ces indicateurs, qui tend à être un peu plus heurté que celui du solde correspondant.

Toutes ces propriétés comparées des indicateurs de MSW et des soldes d'opinion pèsent en faveur de ces derniers. Toutefois, la présente étude ne prétend en aucun cas donner une conclusion définitive à la question de la meilleure combinaison possible des réponses individuelles aux enquêtes de conjoncture. De nombreuses pistes d'améliorations et de prolongement restent à explorer à cet égard.

Quelques pistes de prolongement

La question du système de pondérations des réponses individuelles à retenir au sein des indicateurs n'est véritablement résolue ni par MSW (2002, 2004), ni par la présente étude. Un des prolongements possibles de cette dernière serait d'analyser l'impact de systèmes de pondérations supplémentaires sur les performances relatives des indicateurs. Une piste serait de reproduire le schéma de doubles pondérations des soldes pour les indicateurs de MSW. Ceci serait justifié si la taille des entreprises est corrélée avec la qualité des indications sur la conjoncture macroéconomique contenues dans leurs réponses individuelles. Une amélioration des indicateurs de MSW viendrait à l'appui d'une telle hypothèse.

Cependant, une contribution tout récente de MSW (2006) règle la question des pondérations sur le plan théorique. La définition (1) de l'indicateur paramétrique (cf. encadré 1) y est remplacée par l'estimation directe, plus rigoureuse, de l'espérance de x_t conditionnellement à l'ensemble des réponses $(j_{i,t})_{i=1,\dots,I}$ données à l'enquête de conjoncture. La formule de Bayes est appliquée à ce nouveau conditionnement pour établir le lien nécessaire avec les modèles à choix discrets ordonnés de l'encadré 1. Il en résulte un nouvel indicateur pour lequel la question de la pondération des réponses ne se pose plus. Malheureusement, le test

sur données britanniques par MSW (2006) des propriétés hors échantillon de ce nouvel indicateur s'avère décevant. Ce dernier prévoit le taux de croissance de la production manufacturière sensiblement moins bien que l'indicateur paramétrique de MSW initial défini par (1) et se montre même moins performant que le solde d'opinion, nouvellement intégré par MSW (2006) dans leur liste d'indicateurs agrégés de référence. MSW interprètent cette performance décevante de leur nouvel indicateur comme une conséquence probable de la petite taille de leur panel et recommandent de procéder à un nouveau test lorsque des séries plus longues seront disponibles et qu'une analyse robuste pourra être menée. Au total, cette nouvelle contribution de MSW incite à tester les performances prédictives de leur nouvel indicateur sur le panel français, susceptible de conduire à des résultats plus fiables en raison de sa taille. Il reste que, parce que l'indicateur paramétrique initial de MSW (2002, 2004) demeure le plus performant dans MSW (2006), des pistes de prolongement dans la ligne directe de la présente étude demeurent envisageables, particulièrement pour ce qui concerne les indicateurs paramétriques.

Les indicateurs paramétriques portant sur la question prospective sont théoriquement perfectibles en appliquant la correction des problèmes potentiels d'endogénéité introduite par MSW (2005). Il serait souhaitable également de porter une plus grande attention aux questions d'endogénéité dans le cas de la question rétrospective, en dépit de l'absence de problème détecté dans les tests empiriques menés par MSW sur données britanniques.

Cependant, on envisage plutôt des prolongements d'inspiration différente. L'un serait de tenter d'intégrer directement des indicateurs de dispersion des réponses individuelles (combinés avec des indications de sens des inflexions de tendance) dans des modèles de prévision macroéconomique. L'intuition sous-jacente est que l'augmentation de la dispersion des perceptions individuelles peut refléter le degré d'incertitude dans l'économie, qui croît avant une inflexion ou un point de retournement (20). Par ailleurs, l'un des auteurs du présent article, en collaboration avec deux mathématiciens, a tout récemment exploré une piste innovante consistant à appliquer des méthodes non para-

20. Cette piste de prolongement peut être rapprochée des travaux de Zarnowitz et Lambros (1987), Lahiri, Teigland et Zaporowski (1988) et de beaucoup d'autres auteurs, même si le contexte n'est pas exactement le même, la plupart de ces travaux portant sur les anticipations d'inflation.

métriques du type « plus proches voisins » et « forêts d'arbres aléatoires » à la prévision conjoncturelle de la production manufacturière. Ces méthodes, sans rapport avec celle de MSW (2002, 2004, 2005 et 2006), tiennent toutefois compte également de l'hétérogénéité des comportements de réponses des entreprises. Une application à des données françaises donne des résultats prometteurs concernant la méthode des forêts d'arbres aléatoires (Biau, Biau et Rouvière, 2006), ce qui suggère d'approfondir la réflexion en ce sens. Enfin, pour un pays comme la France, pour lequel on dispose d'une source infra-annuelle quantitative sur la production des entreprises (les enquêtes de branches sources de l'indice de la produc-

tion industrielle IPI), il serait intéressant de confronter les réponses qualitatives données à l'enquête de conjoncture aux réponses chiffrées apportées à l'enquête de branche par une même unité de production. L'utilisation des IPI élémentaires tirés des enquêtes de branche et calculés au niveau de branches relativement fines permettrait de tester la méthode des indicateurs paramétriques sectoriels à un niveau plus désagrégé. Surtout, des modèles de prévision du taux de croissance de la production manufacturière potentiellement plus convaincants pourraient être élaborés par la combinaison des sources qualitatives et quantitatives de données individuelles d'entreprises. Ceci pourrait faire l'objet d'une étude ultérieure. □

BIBLIOGRAPHIE

Anderson O. (1952), « The Business Test of the IFO-Institute for Economic Research, Munich, and its Theoretical Model », *Review of the International Statistical Institute*, vol. 20, n° 1, pp. 1-17.

Ashley R. (2003), « Statistically Significant Forecasting Improvements: How much Out-of-Sample Data is Likely Necessary », *International Journal of Forecasting*, vol. 19, n° 2, pp. 229-239.

Balcombe K. (1996), « The Carlson-Parkin Method Applied to NZ Price Expectations Using QSBO Survey Data », *Economic Letters*, vol. 51, n° 1, pp. 51-57.

Batchelor R.A. (1981), « Aggregate Expectations Under the Stable Laws », *Journal of Econometrics*, vol. 16, n° 2, pp. 199-210.

Batchelor R.A. et Orr A.B. (1987), « Inflation Expectations Revisited », *Economica*, vol. 55, n° 219, pp. 317-331.

Bengoechea P. et Pérez-Quirós G. (2004), « A Useful Tool to Identify Recessions in the Euro Area », *Documentos de trabajo*, Banco de España, n° 0419, et *European Economy*, n° 215.

Berk J. (1999), « Measuring Inflation Expectations: A Survey Data Approach », *Applied Economics*, vol. 31, n° 11, pp. 1467-1480.

Biau G., Biau O. et Rouvière L. (2006), « Non Parametric Forecasting of the Manufactured

Output with Firm-Level Survey Data », *document de travail*, Insee, n° G2006/06.

Biau O., Erkel-Rousse H. et Ferrari N. (2005), « Réponses individuelles aux enquêtes de conjoncture et prévision de la production manufacturière », *document de travail*, Insee, n° G2005/12.

Biau O. et Ferrari N. (2006), « Théorie de l'opinion : Faut-il pondérer les réponses individuelles ? », *document de travail*, Insee, n° G2006/12.

Bouton F. et Erkel-Rousse H. (2003), « Conjonctures sectorielles et prévision à court terme de l'activité : l'apport de l'enquête de conjoncture dans les services », *Économie et Statistique*, numéro spécial *Analyse conjoncturelle : entre statistique et économie*, n° 359-360 – 2002, publié en avril 2003, pp. 35-68.

Britton E., Cutler J. et Wardlow A. (1999), « The Bank's Use of Survey Data », *Bank of England Quarterly Bulletin*, vol. 3, n° 2, pp. 177-182.

Carlson J.A. (1975), « Are Price Expectations Normally Distributed? », *Journal of the American Statistical Association*, vol. 70, n° 352, pp. 749-754.

Carlson J.A. et Parkin M. (1975), « Inflation Expectations », *Economica*, vol. 42, n° 166, pp. 123-138.

Commission européenne (2006), *The Joint Harmonised EU Program of Business and Consumer Surveys, User Guide*. Téléchargeable

sur le site : http://europa.eu.int/comm/economy_finance/indicators/business_consumer_surveys/userguide_en.pdf.

Cornec M. et Deperraz T. (2007), « Un nouvel indicateur synthétique mensuel résumant le climat des affaires dans les services en France », *Économie et Statistique*, n° 395-396, ce numéro.

Cunningham A.W.F., Smith R.J. et Weale M.R. (1998), *Measurement Errors and Data Estimation: The Quantification of Survey Data*, Applied Economics and Public Policy, Cambridge University Press, Cambridge.

Dasgupta S. et Lahiri K. (1992), « A Comparative Study of Alternative Methods of Quantifying Qualitative Survey Responses Using NAPM Data », *Journal of Business and Economic Statistics*, vol. 10, n° 4, pp. 391-400.

D'Elia E. (2005), « Using the Results of Qualitative Surveys in Quantitative Analysis », *Documenti di Lavoro*, ISAE, n° 56.

Diebold F.X. et Mariano R.S. (1995), « Comparing Predictive Accuracy », *Journal of Business and Economic Statistics*, vol. 13, n° 3, pp. 253-265.

Doz C. et Lenglart F. (1999), « Analyse factorielle dynamique : test du nombre de facteurs, estimation et application à l'enquête de conjoncture dans l'industrie », *Annales d'Économie et de Statistique*, n° 54, pp. 91-127.

Driver C. et Urga G. (2004), « Transforming Qualitative Survey Data, Performance Comparisons for the UK », *Oxford Bulletin of Economics and Statistics*, vol. 66, n° 1-iv, pp. 71-89.

Dubois É. et Michaux E. (2006), « Étalonnages à l'aide d'enquêtes de conjoncture : de nouveaux résultats », *Économie et Prévision*, n° 172, 2006-1, pp. 11-28.

Fansten M. (1976), « Introduction à une théorie mathématique de l'opinion », *Annales de l'Insee*, n° 21, pp. 3-55.

Fishe R.P. et Lahiri K. (1981), « On the Estimation of Inflationary Expectations from Qualitative Responses », *Journal of Econometrics*, vol. 16, n° 1, pp. 89-102.

Forni M., Hallin M., Lippi M. et Reichlin L. (2001), « Coincident and Leading Indicators for the Euro Area », *Economic Journal*, vol. 111, n° 471, pp. 62-85.

Gregoir S. et Lenglart F. (1998), « Un nouvel indicateur pour saisir les retournements de conjoncture », *Économie et Statistique*, n° 314, pp. 39-60.

Gregoir S. et Lenglart F. (2000), « Measuring the Probability of a Business Cycle Turning Point by Using a Multivariate Qualitative Hidden Markov Model », *Journal of Forecasting*, vol. 19, n° 2, pp. 81-102.

Hamilton J.D. (1989), « A New Approach to the Economic Analysis of Non Stationary Time Series and the Business Cycle », *Econometrica*, vol. 57, n° 2, pp. 357-384.

Hamilton J.D. (1990), « Analysis of Time Series Subject to Changes in Regimes », *Journal of Econometrics*, vol. 45, n° 1-2, pp. 39-70.

Harvey D.I., Leybourne S.J. et Newbold P. (1997), « Testing the Equality of Prediction Mean Square Errors », *International Journal of Forecasting*, vol. 13, n° 2, pp. 281-291.

Hild F. (2003), « Une lecture enrichie des réponses aux enquêtes de conjoncture », *Économie et Statistique*, numéro spécial *Analyse conjoncturelle : entre statistique et économie*, n° 359-360-2002, publié en avril 2003, pp. 13-34.

Hild F. (2007), « Un nouvel indicateur synthétique prenant en compte la dynamique des réponses individuelles à l'enquête *Industrie* », *Économie et Statistique*, n° 395-396, ce numéro.

Kaiser U. et Spitz A. (2000), « Quantification of Qualitative Data Using Ordered Probit Models with an Application to a Business Survey in the German Service Sector », 25^e conférence du Ciret, Paris, et *ZEW Discussion Paper*, n° 00-58.

Lahiri K. et Moore G.H. eds. (1991), *Leading Economic Indicators: New Approaches and Forecasting Records*, Cambridge University Press, Cambridge, 464 p.

Lahiri K., Teigland C. et Zaporowski M. (1988), « Interest Rates and the Subjective Probability Distribution of Inflation Forecasts », *Journal of Money, Credit, and Banking*, vol. 20, n° 2, pp. 233-248.

Lahiri K. et Wang J.G. (1994), « Predicting Turning Points with Leading Index in a Markov Switching Model », *Journal of Forecasting*, vol. 13, pp. 245-263.

- Lankes F. et Wolters J. (1988)**, « Das IFO-Geschäftsklima und seine Komponenten als vorlaufende Indikatoren, Empirische Ergebnisse für die Verbrauchs- und Investitionsgüterindustrie für 1975-1986 », *FU Berlin, Fachbereich Wirtschaftswissenschaft, Diskussionsarbeit*, n° 5.
- Lee K.C. (1994)**, « Formation of Price and Cost Inflation Expectations in British Manufacturing Industries: A Multisectoral Analysis », *Economic Journal*, vol. 104, n° 423, pp. 372-385.
- Lenglart F. et Toutlemonde F. (2003)**, « Mieux appréhender le climat conjoncturel de la zone euro », *Économie et Statistique*, numéro spécial Analyse conjoncturelle : entre statistique et économie, n° 359-360 – 2002, publié en avril 2003, pp. 69-81.
- Mitchell J., Smith R.J. et Weale M.R. (2002)**, « Aggregate versus Disaggregate Survey-Based Indicators of Economic Activity », *mimeo*, National Institute of Economic and Social Research.
- Mitchell J., Smith R.J. et Weale M.R. (2004-1)**, « Aggregate Versus Disaggregate Survey-Based Indicators of Economic Activity », *National Institute of Economic and Social Research Discussion Paper*, n° 194.
- Mitchell J., Smith R.J. et Weale M.R. (2004-2)**, « The Impact of Survey Aggregation Methods on the Quality of Business Survey Indicators », *Rapport pour la Commission européenne, ECFIN/2003/A3-04*.
- Mitchell J., Smith R.J. et Weale M.R. (2005)**, « Forecasting Manufacturing Output Growth Using Firm-Level Survey Data », *The Manchester School*, vol. 73, n° 4, pp. 479-499.
- Mitchell J., Smith R.J. et Weale M.R. (2006)**, « A Bayesian Indicator of Manufacturing Output from Qualitative Business Panel Survey Data », *mimeo*, National Institute of Economic and Social Research.
- Nardo M. (2003)**, « The Quantification of Qualitative Survey Data: A Critical Assessment », *Journal of Economic Surveys*, vol. 17, n° 5, pp. 645-668.
- Pesaran M.H. (1984)**, *Expectations Formation and Macroeconomic Modelling*, Contemporary Macroeconomic Modelling, Blackwell, Oxford.
- Pesaran M.H. (1987)**, *The Limits to Rational Expectations*, Basil Blackwell, Oxford.
- Reynaud M. et Scherrer S. (1996)**, « Une modélisation VAR de l'enquête de conjoncture de l'Insee dans l'industrie », *document de travail*, Direction de la Prévision, n° 96-12.
- Sargent T.J. et Sims C.A. (1977)**, *Business Cycle Modeling without Pretending to Have too much a-priori Economic Theory*, New Methods in Business Cycle Research, Federal Reserve Bank of Minneapolis.
- Smith J. et McAleer M. (1995)**, « Alternative Procedures for Converting Qualitative Response Data to Quantitative Expectations: An Application to Australian Manufacturing », *Journal of Applied Econometrics*, vol. 10, n°2, pp. 165-185.
- Stock J.H. et Watson M.W. (2002)**, « Macroeconomic Forecasting Using Diffusion Indexes », *Journal of Business and Economic Statistics*, vol. 20, n° 2, pp. 147-162.
- Stone J.R.N. (1947)**, « On the Interdependence of Blocks of Transactions », *Supplement of the Journal of the Royal Statistical Society*, vol. 9, n° 1, pp. 1-32.
- Theil H. (1952)**, « On the Time Shape of Economic Microvariables and the Munich Business Test », *Revue de l'Institut International de Statistique*, vol. 20, n° 2, pp. 105-120.
- Watchel P. (1977)**, *Survey Measures of Expected Inflation and Their Potential Usefulness, in Analysis of Inflation*, Ballinger, Cambridge.
- Zarnowitz V. et Lambros L.A. (1987)**, « Consensus and Uncertainty in Economic Prediction », *Journal of Political Economy*, vol. 95, n° 3, pp. 591-621.

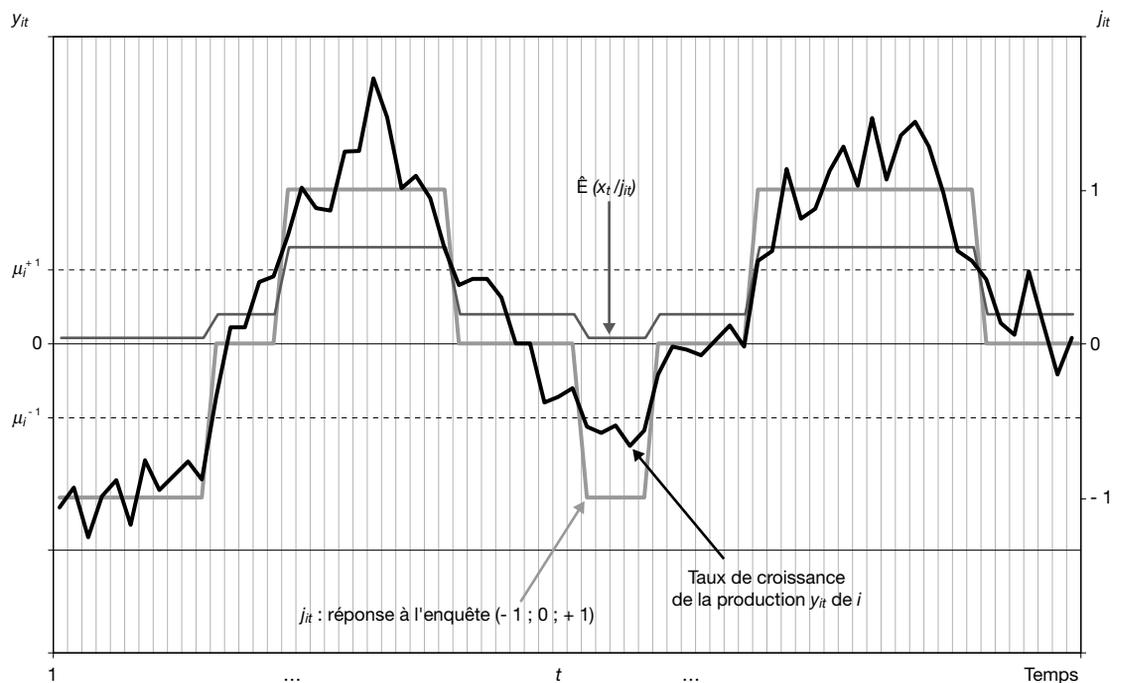
UNE ILLUSTRATION DE L'INDICATEUR NON PARAMÉTRIQUE DE MITCHELL, SMITH ET WEALE

La chronique fictive du taux de croissance de la production d'une unité répondante i ainsi que la série (également fictive) de ses réponses à la question portant sur sa production dans l'enquête *Industrie* de l'Insee sont représentées sur le graphique ci-dessous. Cette illustration n'est pas suffisamment précise pour qu'on puisse distinguer production passée ou prévue, ce n'en est pas l'objet. En phase de production croissante (respectivement décroissante), on s'attend à ce que la réponse de l'unité de production soit « en hausse », soit $j_{i,t} = +1$ (respectivement « en baisse », soit $j_{i,t} = -1$) avec, entre deux phases de sens opposés, des moments de stabilisation durant lesquels l'unité de production répond « stable », soit $j_{i,t} = 0$. Hild (2007) montre en effet que la plupart des entrepreneurs interrogés à l'enquête *Industrie* de l'Insee ne passent pas instantanément d'une réponse extrême à l'autre, mais répondent « stable » durant une période intermédiaire.

Le terme $\hat{E}(x_t / j_{i,t}, i) = \frac{1}{T^{j_{i,t}}} \sum_{\tau=1}^T x_{\tau} y_{i,\tau}^{j_{i,t}}$ relatif à l'unité

de production i dans la somme définissant l'indicateur non paramétrique de Mitchell, Smith et Weale a un profil en paliers (courbe en trait fin, comportant trois paliers possibles, prenant les valeurs des moyennes arithmétiques des (x_t) prises sur les sous-périodes où l'unité de production a répondu, respectivement, « en hausse », « stable » et « en baisse »). La sommation sur de nombreuses unités de production conduit à un indicateur non paramétrique de Mitchell, Smith et Weale dont le profil ne fait plus apparaître les paliers des termes élémentaires $\hat{E}(x_t / j_{i,t}, i)$, du fait que les réponses individuelles à l'enquête *Industrie* ne sont pas parfaitement synchronisées.

L'indicateur non paramétrique de MSW : une illustration fictive



Source : illustration inspirée de la discussion de F. Toutlemonde au séminaire du Département des Études Économiques d'Ensemble du 29 juin 2005, avec sa permission.

**LES DISTRIBUTIONS DES COEFFICIENTS ESTIMÉS ($\hat{\beta}_i$) DU MODÈLE (6) :
CAS DES INDICATEURS MACROÉCONOMIQUES DE MSW**

Le modèle (6) introduit à l'encadré 1 lie le taux de croissance de la production de l'unité de production i à la période t $y_{i,t}$ (variable latente) au taux de croissance de la production manufacturière x_t : $y_{i,t} = \alpha_i + \beta_i x_t + \varepsilon_{i,t}$ (un modèle par unité répondante sélectionnée). La distribution des ($\hat{\beta}_i$), estimateurs des MCO des coefficients (β_i) dans ce modèle donne des informations sur la manière dont l'activité productive s'articule entre les niveaux micro- et macroéconomiques. Ces coefficients estimés entrent en outre dans l'estimation des indicateurs paramétriques de MSW.

Comme dans les applications de Mitchell, Smith et Weale, le nombre d'unités à production contra-cyclique ou acy-

clique (associées à un coefficient estimé $\hat{\beta}_i$ négatif ou proche de zéro) est non négligeable (cf. tableaux A et B). À l'instar de Mitchell, Smith et Weale, on a conservé ces unités dans le calcul des indicateurs. Celles à production contra-cyclique peuvent en effet véhiculer une information conjoncturelle significative. Quant à la contribution des réponses des unités i à production acyclique à l'évolution des indicateurs paramétriques de Mitchell, Smith et Weale, elle est faible par construction, puisque leur composante $\hat{E}(x_t / j_{i,t}, t)$ conserve une valeur voisine de la moyenne historique des taux de croissance de la production manufacturière sur la période $[1, T]$, pour tout trimestre t , quelle que soit leur réponse $j_{i,t}$.

Tableau A
Production passée (quantiles)

1 %	5 %	10 %	25 %	50 % (médiane)	75 %	90 %	95 %	99 %
- 0,63	- 0,31	- 0,18	0,04	0,28	0,52	0,75	0,91	1,32

Lecture : 673 coefficients estimés sont négatifs sur 3076. La moyenne de la distribution est 0,28. Son écart-type vaut 0,39.

Tableau B
Production prévue (quantiles)

1 %	5 %	10 %	25 %	50 % (médiane)	75 %	90 %	95 %	99 %
- 0,65	- 0,38	- 0,24	- 0,04	0,18	0,40	0,63	0,79	1,13

Lecture : 753 coefficients estimés sont négatifs sur 2589. La moyenne de la distribution est 0,19. Son écart-type vaut 0,36.

RÉSULTATS DES TESTS DE PERFORMANCE PRÉDICTIVE DE HARVEY, LEYBOURNE ET NEWBOLD (1997)

Tableau A
Régressions simples – question sur la production passée

Test du modèle avec solde contre le modèle avec...	Différence moyenne des erreurs de prévision \bar{d}	Statistique de test $\sqrt{C} \times S$	P-value
... MSWNP macroéconomique	- 0,205	- 1,364	0,089
... MSWP macroéconomique	- 0,019	- 0,206	0,419
... MSWNP sectoriel	- 0,960	- 0,750	0,228
... MSWP sectoriel	0,043	0,506	0,307

Lecture : les notations \bar{d} et $\sqrt{C} \times S$ sont définies en encadré 3. Dans la deuxième ligne de la deuxième colonne, la valeur de \bar{d} est égale à la différence des erreurs quadratiques moyennes des erreurs de prévision tirées du modèle utilisant le solde et du modèle mobilisant l'indicateur non paramétrique macroéconomique, soit, d'après le tableau 2 : $0,897 - 1,101 = - 0,205$ (chiffres arrondis à 3 chiffres après la virgule).

Une P-value supérieure à 10 % indique l'acceptation de l'égalité des performances prédictives des deux modèles dont sont tirées les erreurs de prévision au seuil de significativité de 10 % (case non grisée). Une P-value comprise entre 5 % et 10 % indique l'acceptation de l'égalité des performances prédictives des deux modèles dont sont tirées les erreurs de prévision au seuil de significativité de 5 %, mais un refus à celui de 10 % (case grisée clair).

Ces tests expriment donc que les modèles mobilisant un indicateur de MSW conduisent à des prévisions non significativement plus performantes que le modèle avec solde d'opinion. Ce dernier paraît même induire des prévisions légèrement plus performantes (au seuil de 10 %) que l'indicateur non paramétrique macroéconomique. Des tests complémentaires suggèrent que le modèle avec indicateur paramétrique sectoriel serait significativement plus performant que les modèles mobilisant les autres indicateurs de MSW au seuil de 10 %.

Tableau B
Régressions simples – question sur la production prévue

Test du modèle avec solde contre le modèle avec...	Différence moyenne des erreurs de prévision \bar{d}	Statistique de test $\sqrt{C} \times S$	P-value
... MSWNP macroéconomique	- 0,170	- 1,085	0,141
... MSWP macroéconomique	- 0,025	- 0,183	0,428
... MSWNP sectoriel	- 0,057	- 0,389	0,349
... MSWP sectoriel	0,023	0,174	0,431

Lecture : les notations \bar{d} et $\sqrt{C} \times S$ sont définies en encadré 3. Dans la deuxième ligne de la deuxième colonne, la valeur de \bar{d} est égale à la différence des erreurs quadratiques moyennes des erreurs de prévision tirées du modèle utilisant le solde et du modèle mobilisant l'indicateur non paramétrique macroéconomique, soit, d'après le tableau 2 : $1,043 - 1,213 = - 0,170$.

Une P-value supérieure à 10 % indique l'acceptation de l'égalité des performances prédictives des deux modèles dont sont tirées les erreurs de prévision au seuil de significativité de 10 % (case non grisée).

Les tests effectués expriment donc que les performances prédictives des modèles associés à l'un des indicateurs de MSW sont non significativement différentes de celles du modèle avec solde d'opinion.

Tableau C
Modèles dynamiques

Test du modèle avec soldes contre le modèle avec :	Différence moyenne des erreurs de prévision \bar{d}	Statistique de test $\sqrt{C} \times S$	P-value
... MSWNP macroéconomique	- 0,583	- 1,725	0,049
... MSWP macroéconomique	- 0,440	- 1,425	0,084
... MSWNP sectoriel	- 0,707	- 1,890	0,036
... MSWP sectoriel	- 0,535	- 1,695	0,052

Lecture : les notations \bar{d} et $\sqrt{C} \times S$ sont définies en encadré 3. Dans la deuxième ligne de la deuxième colonne, la valeur de \bar{d} est égale à la différence des erreurs quadratiques moyennes des erreurs de prévision tirées du modèle utilisant les soldes et du modèle mobilisant les indicateurs MSWNP macroéconomiques, soit, d'après le tableau 3 : $0,732 - 1,315 = - 0,583$.

Une P-value comprise entre 5 % et 10 % indique l'acceptation de l'égalité des performances prédictives des deux modèles dont sont tirées les erreurs de prévision au seuil de significativité de 5 %, mais un refus à celui de 10 % (case grisée clair). Une P-value inférieure à 5 % indique le refus de l'hypothèse d'égalité des performances prédictives des deux modèles dont sont tirées les erreurs de prévision au seuil de significativité de 5 % (case grisée foncée).

Les tests effectués expriment donc que le modèle avec soldes d'opinion conduit à des prévisions significativement plus performantes que ceux mobilisant des indicateurs de MSW, non paramétriques particulièrement.

Tableau D
Régressions simples – question sur la production passée
Résultats sur la sous-période 1999T2-2004T4

Test du modèle avec solde contre le modèle avec...	Différence moyenne des erreurs de prévision \bar{d}	Statistique de test $\sqrt{C} \times S$	P-value
... MSWNP macroéconomique	- 0,188	- 0,831	0,208
... MSWP macroéconomique	- 0,012	- 0,075	0,470
... MSWNP sectoriel	- 0,118	- 0,681	0,252
... MSWP sectoriel	0,050	0,340	0,369

Lecture : les notations \bar{d} et $\sqrt{C} \times S$ sont définies en encadré 3. Dans la deuxième ligne de la deuxième colonne, la valeur de \bar{d} est égale à la différence des erreurs quadratiques moyennes des erreurs de prévision tirées du modèle utilisant le solde et du modèle mobilisant l'indicateur non paramétrique macroéconomique, soit, d'après le tableau 4 : $0,989 - 1,177 = - 0,188$.

Une P-value supérieure à 10 % indique l'acceptation de l'égalité des performances prédictives des deux modèles dont sont tirées les erreurs de prévision au seuil de significativité de 10 % (case non grisée).

Ces tests expriment donc que les modèles mobilisant un indicateur de MSW conduisent à des prévisions non significativement plus performantes que le modèle avec solde d'opinion. Des tests complémentaires suggèrent que les modèles mobilisant un indicateur de MSW ont des performances prédictives équivalentes.

Tableau E
Régressions simples – question sur la production prévue
Résultats sur la sous-période 1999T2-2004T4

Test du modèle avec solde contre le modèle avec...	Différence moyenne des erreurs de prévision \bar{d}	Statistique de test $\sqrt{C} \times S$	P-value
... MSWNP macroéconomique	- 0,388	- 1,428	0,084
... MSWP macroéconomique	- 0,175	- 0,794	0,218
... MSWNP sectoriel	- 0,378	- 1,547	0,068
... MSWP sectoriel	- 0,174	- 0,800	0,216

Lecture : les notations \bar{d} et $\sqrt{C} \times S$ sont définies en encadré 3. Dans la deuxième ligne de la deuxième colonne, la valeur de \bar{d} est égale à la différence des erreurs quadratiques moyennes des erreurs de prévision tirées du modèle utilisant le solde et du modèle mobilisant l'indicateur non paramétrique macroéconomique, soit, d'après le tableau 4 : $1,005 - 1,394 = - 0,388$.

Une P-value supérieure à 10 % indique l'acceptation de l'égalité des performances prédictives des deux modèles dont sont tirées les erreurs de prévision au seuil de significativité de 10 % (case non grisée). Une P-value comprise entre 5 % et 10 % indique l'acceptation de l'égalité des performances prédictives des deux modèles dont sont tirées les erreurs de prévision au seuil de significativité de 5 %, mais un refus à celui de 10 % (case grisée clair).

Les tests effectués expriment donc que les performances prédictives des modèles associés à l'un des indicateurs de MSW sont non significativement meilleures que celles du modèle avec solde d'opinion au seuil de 5 %. Le modèle avec solde conduit à des prévisions légèrement meilleures (au seuil de 10 %) que les modèles mobilisant l'un ou l'autre des indicateurs non paramétriques de MSW. Des tests complémentaires suggèrent que les modèles mobilisant un indicateur de MSW ont des performances prédictives équivalentes.