

L'activité va-t-elle accélérer ou ralentir ? Quelques outils pour répondre à cette question

Fanny Mikol
Matthieu Cornec

Division Enquêtes de conjoncture

*L*a prévision de la croissance du PIB est l'un des principaux défis auxquels sont confrontés les conjoncturistes. Une riche littérature s'est accumulée à ce sujet depuis plusieurs décennies, et offre une grande diversité de techniques pour y répondre.

En revanche, peu d'études se sont penchées sur la prévision du « profil » de la croissance, c'est-à-dire, se sont intéressées à simplement prédire si le taux de croissance du trimestre de prévision sera supérieur ou inférieur à celui du dernier trimestre connu. Or, au-delà du chiffre exact, c'est sans doute cette information qui a de la valeur pour les utilisateurs des prévisions conjoncturelles. Et d'ailleurs, l'analyse conjoncturelle s'exprime bien souvent en termes d'accélération et de décélération du PIB, sans nécessairement faire référence à un niveau donné de croissance.

Ce dossier examine dès lors différentes méthodes permettant de prévoir le profil de la croissance du PIB. Basées sur les enquêtes de conjoncture, les simulations réalisées en temps réel montrent qu'il est de fait possible d'apporter une information pertinente pour la prévision de ce profil.

Cette analyse permet également de construire un indicateur de profil qui mesure, pour le trimestre à prévoir, les chances d'accélération (ou de décélération) du PIB. Construit à partir des soldes d'opinions des enquêtes de conjoncture, cet indicateur constitue un outil utile au conjoncturiste, en complément des indicateurs synthétiques (climat des affaires, indicateur de retournement...) déjà existants.

Cet indicateur se révèle très performant pour prévoir les à-coups de l'activité observés depuis la sortie de récession. Appliqué au quatrième trimestre de l'année 2011, cet indicateur indique un risque élevé de ralentissement de l'activité. ■

L'activité va-t-elle accélérer ou ralentir ? Quelques outils pour répondre à cette question

Répondre à une question simple : l'activité va-t-elle accélérer ou ralentir ?

L'analyse conjoncturelle s'articule beaucoup autour des idées d'accélération et de décélération des grandeurs modélisées...

Dans le discours du conjoncturiste, le scénario de prévision est très souvent décrit en termes d'accélération ou de décélération : « la production accélère », « l'activité ralentit », etc. D'une part, l'appréciation qualitative permet d'illustrer le sens des impulsions économiques à la croissance, sans se focaliser sur le chiffre de la croissance lui-même : elle permet ainsi, en quelque sorte, de s'abstraire des erreurs de prévision qui peuvent survenir dans la prévision de ce chiffre. D'autre part, prévoir le profil de croissance de certaines grandeurs macroéconomiques peut s'avérer essentiel dans certains domaines : Sinclair et al. (2010) estiment par exemple que ce serait particulièrement le cas pour les prévisions de PIB et d'inflation, notamment dans le but d'optimiser des décisions de politique monétaire.

... sans que cette question ne fasse l'objet d'une analyse directe

Pourtant, cette question est souvent éclipsée par celle de la prévision du chiffre de la croissance, comme en témoigne l'abondante littérature à ce sujet, sur données françaises (Dubois et Michaux, 2006, Erkel-Rousse et Minodier, 2009, Bessec, 2010) comme étrangères (Moore, 1983, Wallis, 1989, Sédillot et Pain, 2003).

Très peu d'études économiques s'intéressent ainsi à la prédiction du profil de la croissance. Cette méthodologie a plutôt été mobilisée, pour certaines grandeurs macroéconomiques, dans la finance (Merton, 1981, Henriksson et Merton, 1981). En macroéconomie, elle n'est apparue que tardivement, et a été surtout utilisée pour fournir un critère de validation *ex-post* de modèles de prévision de croissance (Stekler, 1994, Sédillot et Pain, 2003, Bessec, 2010).

Construire une prévision directe sur l'« accélération » ou la « décélération » du PIB

En dehors du champ de l'analyse économique, de nombreux domaines recourent à des méthodes qui consistent à classer des observations en deux groupes (ou plus), puis à utiliser ces classifications pour faire des prévisions. En médecine notamment, des patients peuvent être classés en fonction de leurs caractéristiques (attributs individuels et/ou mesures cliniques) entre un groupe « à risque » et un groupe « sain ». Une fois cette classification réalisée, elle peut servir à attribuer *a priori* à un patient son groupe d'appartenance, en observant uniquement ses caractéristiques, et donc à en déduire un traitement adéquat.

C'est cette logique qui est adoptée ici dans la prévision du profil de la croissance : en particulier, on peut classer le profil de la croissance, encore inconnu, entre les deux catégories « accélération du PIB » et « décélération du PIB ». En reproduisant séquentiellement le cadre dans lequel, chaque trimestre, sont réalisées les prévisions du conjoncturiste, et en s'appuyant notamment sur un certain nombre d'indicateurs économiques - en l'occurrence ici, les enquêtes de conjoncture - des méthodes permettent de prévoir le profil de la croissance du PIB pour le trimestre suivant.

L'originalité de ce travail réside également dans la variété des méthodes utilisées. En effet, au-delà des modèles économétriques usuels de prévisions en niveau, un large éventail de méthodes de classification peut également être implémenté, dont nous ne détaillerons ici qu'un exemple. En définitive, un indicateur de profil est présenté, qui résume en termes probabiliste le « risque » d'une future décélération (ou accélération) du PIB.

L'activité va-t-elle accélérer ou ralentir ? Quelques outils pour répondre à cette question

Prévoir le profil de la croissance d'un trimestre sur l'autre : données et méthode

*La variable d'intérêt :
les premiers résultats du PIB...*

Dans les conditions de réalisation de la note de conjoncture, l'objectif est de prévoir le profil du taux de croissance du PIB entre le trimestre précédent (où la croissance est connue) et le trimestre courant (où elle est encore inconnue). La croissance du PIB du trimestre précédent, au moment où la note de conjoncture est publiée, correspond à la première parution des résultats relatifs au trimestre précédent. Ces résultats paraissent 45 jours après la fin du trimestre. C'est sur cette publication que s'appuie l'exercice de prévision conjoncturelle à l'Insee, dont l'objectif est de prévoir le trimestre suivant, et un ou deux trimestres au-delà.

*... et plus précisément
leur profil*

Ainsi, notant y_t , le taux de croissance au trimestre t publié dans les premiers résultats du trimestre t , c'est-à-dire à $t + 45$ jours, le problème se ramène alors à la prévision du profil de la série des (y_t) : est-ce que (y_t) sera supérieur ou inférieur à (y_{t-1}) . Par exemple, les premiers résultats relatifs au troisième trimestre 2011, publiés mi-novembre 2011, donnent une croissance du PIB de 0,4 % : $y_{2011T3} = 0,4\%$. L'enjeu est alors de prédire si la croissance du PIB au quatrième trimestre, qui sera publiée à l'occasion des premiers résultats des comptes trimestriels mi-février 2012, sera supérieure ou inférieure à 0,4 % ($y_{2011T4} \geq y_{2011T3}$ ou $y_{2011T4} < y_{2011T3}$).

Pour chaque trimestre depuis 1997, on définit ainsi la variable d'intérêt, à savoir le sens de variation de la croissance entre $t-1$ et t par :

$$\varepsilon_t = 1\{y_t \geq y_{t-1}\}$$

Avec cette notation, il y a accélération au trimestre t (resp. décélération) si $\varepsilon_t = 1$ (resp. $\varepsilon_t = 0$). La série des profils sur le passé, pour la période [2000T1 ; 2011T3], est illustrée dans le tableau 1⁽¹⁾.

On y lit par exemple que $\varepsilon_t = 0$ pour le quatrième trimestre 2010, ce qui signifie que le premier résultat de la croissance du PIB au quatrième trimestre 2010 (+0,4 %) a été inférieur à celui enregistré au troisième trimestre 2010 (+0,5 %). Dans la suite du dossier, on cherche à prévoir si pour le trimestre en cours, la variable ε_t vaut 1 ou 0.

(1) On constate en pratique que les valeurs des profils aux dates considérées sont très stables : elles sont très peu modifiées au fil du temps, entre le premier résultat et les révisions successives des comptes trimestriels.

Tableau 1

Profil d'accélération et de décélération du PIB

| | T1 | T2 | T3 | T4 |
|------|----|----|----|----|
| 2000 | 0 | 0 | 0 | 1 |
| 2001 | 0 | 0 | 1 | 0 |
| 2002 | 1 | 1 | 0 | 1 |
| 2003 | 1 | 0 | 1 | 1 |
| 2004 | 1 | 1 | 0 | 1 |
| 2005 | 0 | 0 | 1 | 0 |
| 2006 | 1 | 1 | 0 | 1 |
| 2007 | 0 | 0 | 1 | 0 |
| 2008 | 1 | 0 | 1 | 0 |
| 2009 | 1 | 1 | 1 | 1 |
| 2010 | 0 | 1 | 0 | 0 |
| 2011 | 1 | 0 | 1 | |

L'activité va-t-elle accélérer ou ralentir ? Quelques outils pour répondre à cette question

L'information conjoncturelle pour faire cette prévision : les enquêtes de conjoncture

Pour réaliser cet exercice, les enquêtes de conjoncture auprès des entreprises réalisées par l'Insee dans les principaux secteurs de l'économie présentent de nombreux avantages. Tout d'abord, elles constituent une information fiable sur les tendances récentes de l'activité et sur les anticipations des chefs d'entreprises. Ensuite, elles sont très rapidement disponibles, et avec une fréquence mensuelle. En cela, elles correspondent aux données les plus récentes permettant d'établir un diagnostic conjoncturel. Enfin, elles ne sont soumises qu'à des révisions minimales.

Les réponses des chefs d'entreprise sont, en général, de nature qualitative : les questions comportent, par exemple, pour l'activité prévue, trois modalités de réponses : « en hausse », « stable », et « en baisse ». Dans ce dossier, deux types d'informations issues des enquêtes de conjoncture sont mobilisées. D'une part, les soldes d'opinion qui permettent de quantifier les réponses : il s'agit, pour une question donnée, de la différence entre le pourcentage de réponses positives et le pourcentage de réponses négatives. D'autre part, l'indicateur composite publié tous les mois par l'Insee et synthétisant l'information de l'ensemble des secteurs, le « climat des affaires » en France.

Une prévision dans un cadre séquentiel

Au moment de l'élaboration de la prévision pour le trimestre t , on dispose des données des premiers résultats de croissance du PIB - et donc aussi des profils de croissance - jusqu'au trimestre $t-1$:

$$\boldsymbol{\varepsilon}_{t-1} = (\varepsilon_1, \dots, \varepsilon_{t-1})$$

Les données des enquêtes de conjoncture sont déjà disponibles pour les deux premiers mois du trimestre : on note x_t ces informations relatives au deuxième mois du trimestre t . Le conjoncturiste dispose également des données passées des enquêtes de conjoncture x_1, \dots, x_{t-1} et il observe donc en définitive :

$$x_t = (x_1, \dots, x_{t-1}, x_t)$$

Le conjoncturiste doit donc prévoir la prochaine occurrence ε_t de la séquence des profils $(\varepsilon_1, \dots, \varepsilon_{t-1})$, connaissant également l'information conjoncturelle (x_1, \dots, x_t) .

De multiples stratégies de prévisions sont envisageables, avec des performances diverses

Pour prévoir le profil de croissance, de nombreuses stratégies sont envisageables. Pour les comparer, on utilise un critère quantifiant leur efficacité « en temps réel », c'est-à-dire dans les conditions dans lesquelles est mené l'exercice de prévision (cf. annexe 1).

Des stratégies plus ou moins complexes

Le premier ensemble de stratégies est constitué de stratégies dites « naïves » parce qu'elles n'exploitent pas l'information disponible dans les enquêtes de conjonctures (cf. annexe 2). La plus simple consiste à tirer au hasard la valeur de ε_t (égale à 1 ou 0). D'autres, légèrement plus complexes, consistent, par exemple, à prévoir systématiquement le sens de variation du taux de croissance comme l'opposé de sa valeur du trimestre précédent.

Une deuxième famille de stratégies s'appuie sur les méthodes quantitatives usuelles de prévision de la croissance. Ainsi « l'étalonnage seuillé » (cf. annexe 2) consiste à utiliser les modèles habituels de prévision du taux de croissance de l'activité à partir des enquêtes de conjoncture puis d'en déduire une prévision de ε_t si le modèle indique une accélération ou un ralentissement. Cette approche est, de fait, jusqu'à présent l'une des seules utilisées par les conjoncturistes dans la littérature puisque, comme on l'a rappelé dans l'introduction, la question de l'accélération ou de la décélération est rarement examinée pour elle-même.

L'activité va-t-elle accélérer ou ralentir ? Quelques outils pour répondre à cette question

Or, des stratégies plus complexes, s'appuyant sur des méthodes de classification et développées notamment en analyse financière ou dans le domaine médical, peuvent donner une estimation directe de ε_t (égale à 1 ou 0). Elles sont notamment présentées de manière détaillée dans Cornec et Mikol (2011) et, plus succinctement, résumées dans l'annexe 2. Parmi ces méthodes directes, le modèle « probit », présenté en encadré 1, est usuel en économie.

Les méthodes directes sont les plus performantes

Les stratégies naïves ont, logiquement, des performances peu élevées, avec des taux d'erreur de l'ordre de 35 %, c'est-à-dire que la prévision est inexacte dans 35 % des cas. Celles s'appuyant sur les outils habituels de prévision (« étalonnages ») ont des performances supérieures, avec un taux d'erreur de l'ordre de 18 % sur la période 1997T1-2011T4. Les enquêtes de conjoncture apportent ainsi une information très pertinente au conjoncturiste pour prévoir le profil de la croissance du PIB d'un trimestre à l'autre. Enfin, les stratégies directes de classification permettent, elles, d'obtenir des taux d'erreur encore plus faibles, de l'ordre de 12 % (cf. tableau 2).

Tableau 2

Performances des différentes stratégies de prévision

| Stratégies | Taux d'erreur |
|-------------------|--------------------|
| Stratégies naïves | De l'ordre de 35 % |
| Outils habituels | De l'ordre de 18 % |
| Méthodes directes | De l'ordre de 12 % |

Un indicateur de profil trimestriel très performant sur la période récente

Construire un indicateur qui vaut 1 (respectivement -1) quand la probabilité d'accélération (respectivement de décélération) est très forte...

Au-delà de la performance moyenne de la stratégie retenue, il peut être intéressant, pour un trimestre donné, d'estimer les chances de se tromper dans sa prévision de profil, conditionnellement aux variables conjoncturelles observées. En effet, selon le contexte conjoncturel dans lequel on se trouve un trimestre donné, il peut être plus ou moins facile de prévoir le profil⁽²⁾.

En utilisant les caractéristiques statistiques de l'estimation de ε_t , on va alors construire un indicateur prenant des valeurs comprises entre -1 et 1 directement exploitable pour la prévision du profil de la croissance (cf. encadré 2) : plus l'indicateur est proche de 1, plus la probabilité d'observer une accélération est forte. Symétriquement plus l'indicateur est proche de -1, plus la probabilité d'observer un ralentissement est forte.

Parmi les méthodes les plus performantes, c'est-à-dire les méthodes de classification, le modèle probit possède des propriétés permettant de construire cet indicateur de façon simple. C'est donc celle-ci qui sera utilisée ici (cf. encadré 1).

... avec une zone d'incertitude quand l'indicateur est proche de 0

Quand l'indicateur est proche de 0, son signe continue d'indiquer une prévision quant au profil de l'activité : accélération si le signe de l'indicateur est positif, ralentissement sinon. Mais, aux alentours de 0, le risque d'erreur est plus important. On définit alors, par convention, une zone d'incertitude quand l'indicateur prend des valeurs entre -0,5 et +0,5.

Quand l'indicateur est hors de la zone d'incertitude, le taux d'erreur est très faible

Sur la période 1997T1-2011T1, cet indicateur de profil se situe en dehors de la zone d'incertitude environ 61 % du temps (cf. graphique). À l'intérieur de cette zone de confiance, le taux moyen d'erreurs de prévision est inférieur à 4 % : sur toute la période 1997-2011, seuls deux trimestres de prévision sont incorrects. Les prévisions de profil apparaissent donc fiables lorsque l'indicateur de profil associé se situe dans la zone de confiance.

(2) cf. Cornec (2011) pour une application de cette théorie à la prévision de la croissance du PIB en niveau.

L'activité va-t-elle accélérer ou ralentir ? Quelques outils pour répondre à cette question

Encadré 1 - Prédiction par un modèle probit

Cet encadré détaille le modèle probit, usuel en économie, qui est une des méthodes de classification évoquées dans l'encadré 4.

Le sens de variation de l'activité (ε_t) peut être directement estimé à partir d'un modèle de classification probit, faisant intervenir la croissance du PIB comme variable endogène, et les variables issues des enquêtes de conjoncture comme exogènes (x_t). Avec cette modélisation, la probabilité d'observer un profil d'accélération du PIB s'écrit comme suit :

$$P(\varepsilon_t = 1 | X_t = x_t) = F(\beta x_t)$$

où F est la fonction de répartition de la loi normale. Les coefficients β sont estimés par maximum de vraisemblance. La série des profils (ε_t) est ainsi prévue de la manière suivante :

$$\hat{\varepsilon}_t = \Phi_t^{probit}(x_t, \varepsilon_{t-1}) = 1 \{ F_t(\hat{\beta} x_t) \geq 0,5 \}$$

Deux spécifications ont été retenues pour le choix des variables x_t issues des enquêtes de conjoncture. Il s'agit des deux mêmes sous-ensembles que ceux retenus dans le cadre de l'étalonnage seuillé :

- l'un basé sur l'indicateur synthétique France

$$\hat{\beta} x_t = 1,99 - 4,54 y_{t-1} + 0,19 F_t + 0,27 \Delta F_t |\Delta F_t|$$

- et l'autre basé sur les soldes d'opinion de l'enquête de conjoncture dans l'industrie (cf. supra) :

$$\begin{aligned} \hat{\beta} x_t = & 1,95 - 4,79 y_{t-1} - 0,37 y_{t-4} + 0,06 \text{manuf.} \text{apa}_{t,2} \\ & + 0,06 (\text{manuf.} \text{apa}_{t,2} - \text{manuf.} \text{apa}_{t,3}) \\ & + 0,04 (\text{manuf.} \text{apre}_{t,2} - \text{manuf.} \text{apre}_{t,3}) \\ & + 0,04 (\text{manuf.} \text{apre}_{t,3} - \text{manuf.} \text{apre}_{t-1,3}) \end{aligned}$$

Le taux moyen d'erreurs de prévision depuis 1997 est d'environ 14 % pour la première spécification (resp. 16 % pour la seconde). Ces résultats apparaissent légèrement meilleurs que ceux obtenus avec la stratégie de l'étalonnage seuillé. Ceci montre qu'une méthode de classification peut être sensiblement plus performante qu'une méthode quantitative pour la prédiction du sens de variation de la croissance. D'autres méthodes de classification que le modèle probit confirment également ces bons résultats (cf. annexe 2). ■

Encadré 2 - Estimation d'un indicateur de profil à l'aide d'un modèle probit

Construction d'un indicateur de profil : principe général

À partir de ce modèle probit, on construit un indicateur de profil : pour ce faire, on associe à la prévision de profil une « vraisemblance » conditionnelle, qui dépend des données d'enquête de conjoncture observées au même moment. L'objectif de l'indicateur de profil est de déterminer, chaque trimestre, la chance de succès (ou d'échec) associée à une prévision de profil, conditionnellement aux valeurs observées des enquêtes de conjoncture. Plus précisément, il permet de fournir la probabilité, pour un trimestre donné, d'être dans un état d'accélération du PIB ($\varepsilon_t = 1$) ou de décélération du PIB ($\varepsilon_t = 0$). Pour des raisons de symétrie, cet indicateur de profil est compris entre -1 (décélération) et +1 (accélération). Formellement, l'indicateur de profil se définit comme :

$$I_t = 2 \left[\hat{P}(\varepsilon_t = +1 | x_t) - 1/2 \right]$$

Le terme probabiliste dans la définition de l'indicateur s'interprète comme la probabilité estimée d'être dans un état d'accélération du PIB conditionnellement aux valeurs x_t des enquêtes de conjoncture au trimestre t .

À cet indicateur de profil peut être associée une « zone d'incertitude », qui correspond à une prévision de profil « risquée » (c'est-à-dire que le risque d'erreur est grand) et qui devra être considérée avec précaution. Par convention, cette zone d'incertitude correspond à l'intervalle $[-0,5 ; +0,5]$ des valeurs prises par l'indicateur de profil. De façon complémentaire, la « zone de confiance » correspond aux valeurs de l'indicateur de profil comprises dans les intervalles $[-1 ; -0,5[$ et $] +0,5 ; +1]$.

Construction d'un indicateur de profil : application au modèle probit

Dans le cas du modèle probit, les hypothèses associées à cette stratégie font en effet apparaître la probabilité conditionnelle, pour chaque trimestre, d'être dans un état d'accélération du PIB (cf. supra) :

$$P(\varepsilon_t = 1 | X_t = x_t) = F(\beta x_t)$$

En reprenant les notations précédentes, il est donc aisé de définir l'indicateur de profil comme :

$$I_t = 2 \left[F(\hat{\beta} x_t) - 1/2 \right]$$

■

L'activité va-t-elle accélérer ou ralentir ? Quelques outils pour répondre à cette question

Un indicateur qui fournit très souvent un signal non équivoque...

L'indicateur de profil est particulièrement performant pour détecter les variations du sens de l'activité dans les périodes de turbulences (cf. tableau 3). Ainsi, sur les 14 trimestres observés depuis le premier trimestre 2008, l'indicateur de profil est dans la zone de confiance pour 12 trimestres. Sur ces 12 trimestres, il signale systématiquement le sens de variation correct : accélération ou décélération du PIB. Pour les deux trimestres où l'indicateur est dans la zone d'incertitude (2009T3 et 2010T4), de fait l'accélération ou le ralentissement de l'activité observé au moment des premiers résultats a été particulièrement faible.

... et utile pour prévoir les à-coups de l'activité depuis 2008, dans la récession et en sortie de récession

Au plus fort de la crise fin 2008 et début 2009, alors que les prévisions de croissance de PIB étaient particulièrement difficiles à réaliser, notre indicateur de profil se maintient ainsi nettement dans la zone de confiance : les valeurs avoisinent -1 au quatrième trimestre 2008 quand le recul de l'activité s'accroît et prennent une valeur +1 au premier trimestre 2009 quand la récession perd en intensité puis quand l'activité progresse à nouveau au deuxième trimestre 2009.

De même, depuis la sortie de récession, une des thématiques importantes est celle des « à-coups », c'est-à-dire que la reprise n'est pas lisse mais suit un profil nettement heurté, lié par exemple, dans le cas de la France, aux effets de la prime à la casse. L'indicateur de profil prévoit ces à-coups avec fiabilité.

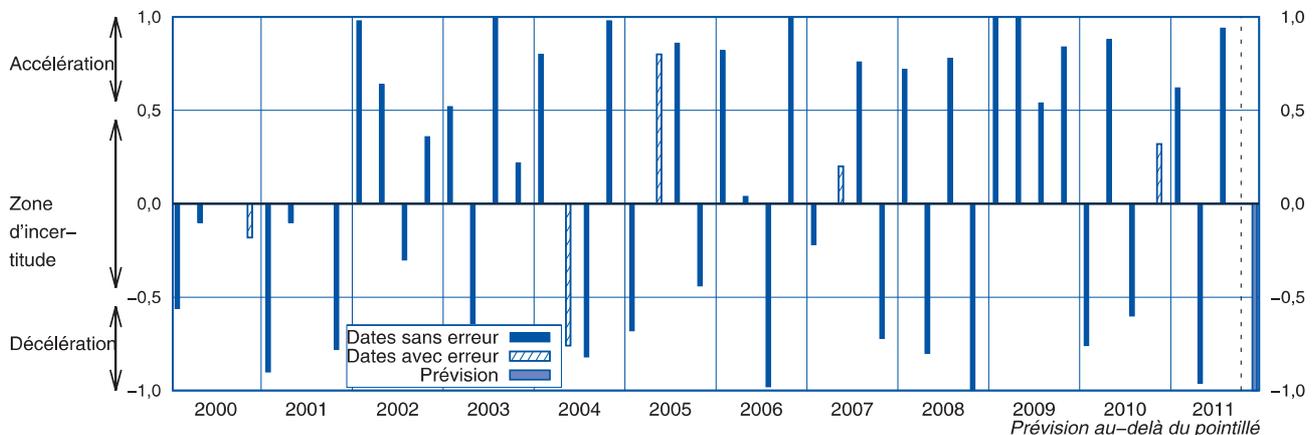
Pour le quatrième trimestre, la probabilité d'un ralentissement apparaît forte

Qu'en est-il de la prévision de cette note de conjoncture? Pour la prévision du sens de variation de l'activité au quatrième trimestre 2011, l'indicateur de profil se situe nettement dans la zone de confiance et prend une valeur -1. Il signale donc un risque élevé de ralentissement de l'activité.

Tableau 3

Indicateur de profil

| en % des trimestres observés depuis 1997T1 | L'indicateur prévoit une accélération et... | L'indicateur prévoit une décélération et... |
|--|---|---|
| ... on observe une accélération | 49 % | 1 % |
| ... on observe une décélération | 14 % | 36 % |



Source : calculs Insee

L'activité va-t-elle accélérer ou ralentir ? Quelques outils pour répondre à cette question

Conclusion

Les enquêtes de conjoncture apportent une information très pertinente pour prévoir le profil de la croissance du PIB d'un trimestre à l'autre. Ce résultat est de premier abord intuitif, dans la mesure où il est déjà démontré de longue date que les enquêtes de conjoncture figurent parmi les meilleurs indicateurs de prévision de la croissance du PIB en niveau. Cependant leur utilisation dans la prévision de profil de l'activité, qui correspond donc à une prévision qualitative, est originale et pertinente pour au moins deux raisons.

D'une part, elle permet d'élargir l'éventail des méthodes possibles pour la prévision. En particulier, les méthodes de classification donnent des résultats très satisfaisants, voire meilleurs que celles basées sur des étalonnages quantitatifs classiques. Cette approche rend en outre possible la comparaison de stratégies utilisant l'information conjoncturelle à des stratégies naïves d'agents non informés. De fait, les premières stratégies surpassent très nettement les secondes.

D'autre part, l'indicateur de profil permet, chaque trimestre, d'associer à la prévision de profil une « vraisemblance » conditionnelle aux enquêtes de conjoncture. Ceci offre la possibilité de définir une zone de confiance au sein de laquelle la prévision d'une accélération ou d'une décélération peut être considérée comme relativement fiable. Cet indicateur constitue, de ce fait, un outil utile et performant, complémentaire des indicateurs déjà existants pour l'analyse conjoncturelle. ■

Annexe 1 - Critères de jugement des stratégies de prévision du profil de la croissance

Simulations hors échantillon

Pour évaluer une stratégie de prévision, il est nécessaire tout d'abord de la simuler « en temps réel » (ou « hors échantillon ») sur le passé. Pour une simulation « hors échantillon » de 1997T1 à 2011T3 par exemple, cela signifie que l'on estime d'abord le modèle sur la période 1989T1-1996T4, et que l'on en déduit une prévision du profil de la croissance pour 1997T1. Puis, on ajoute un point à l'estimation (ce qui donne donc la période 1989T1-1997T1) et on prévoit le profil de 1997T2 ; et ainsi de suite jusqu'à la prévision de 2011T4.

Une telle simulation dite « en temps réel » n'est valable que si les variables utilisées correspondent effectivement, pour chaque trimestre du passé, à celles qui étaient effectivement disponibles à ce moment-là (on parle de « données en temps réel »). Or par construction, la série des premiers résultats du PIB et donc aussi celle des profils de la croissance constituent bien des données en temps réel (cf. *supra*). Les soldes d'opinion des enquêtes de conjoncture, ensuite, obéissent à du « pseudo-temps réel » dans la mesure où ils ne sont que révisés qu'une seule fois (le mois suivant leur première publication) et de façon très marginale.

Fonction de perte

L'étape suivante consiste à évaluer la capacité prédictive d'une stratégie de prévision. À cet égard, il est nécessaire de définir une « fonction de perte ». Les prévisions simulées précédemment sont comparées avec les premiers résultats de la croissance du PIB effectivement publiés sur la même période. La fonction de perte attribue

ainsi une certaine pénalité aux erreurs de prévisions ; ici, elle est simplement définie de la façon suivante :

$$L_t(\phi_t) := \frac{1}{t} \sum_{i=1}^t 1\{\epsilon_i \neq \hat{\epsilon}_i\}$$

Il s'agit en fait d'un « coût cumulé » correspondant simplement au taux moyen des erreurs de prévisions faites « en temps réel » sur la période $[1, t]$:

$$L_t(\phi_t) = \bar{\eta}_t$$

avec :

$$\eta_t := 1\{\epsilon_t \neq \hat{\epsilon}_t\}$$

Avec cette fonction, une accélération du PIB mal prévue « coûte » autant qu'une décélération mal prévue. On aurait pu envisager une fonction qui, par exemple, pénalise davantage les moments où une décélération est mal anticipée (périodes de crise notamment). La fonction utilisée ici, a cependant le mérite d'être relativement simple.

À toute stratégie de prévision, on peut associer sa fonction de perte sur la période $[1, t]$ (en l'occurrence, $[1997T1, 2011T3]$ pour l'exercice présent). La fonction de perte prend naturellement ses valeurs dans $[0, 1]$: une valeur de 0 (resp. 1) signifie que la stratégie en question a parfaitement prévu le signe des évolutions de croissance sur la période (resp. s'est trompée sur toute la période).

Les erreurs de prévision des différentes stratégies étudiées sont comparées depuis 1997, date à laquelle les variables conjoncturelles observées sont toutes disponibles depuis plusieurs années. ■

L'activité va-t-elle accélérer ou ralentir ? Quelques outils pour répondre à cette question

Annexe 2 - De multiples stratégies de prévision sont envisageables

1 - Exemples de stratégies de prévision "naïves"

Les stratégies naïves n'exploitent pas l'information disponible dans les enquêtes de conjoncture mais se contentent, au mieux, de répliquer une certaine régularité dans les sens de variation. Elles sont logiquement peu performantes et fournissent ainsi un « point de référence » pour juger la performance des méthodes plus sophistiquées.

Stratégie du hasard

Un agent non informé peut prévoir au hasard le prochain signe de l'activité. Sa prévision sera simplement tirée dans une loi de Bernoulli U_t de paramètre $1/2$:

$$\hat{\epsilon}_t = \phi_t^{\text{hasard}}(\epsilon_{t-1}) = u_t$$

Par définition son erreur moyenne depuis 1997 aurait été autour de 50 %. Plus précisément, l'application du théorème central limite conduit à prévoir qu'avec une forte probabilité (95 %), son erreur moyenne sur la période 1997T1-2011T1 aurait été supérieure à 35 %.

Stratégie du contrepied

Depuis 1997, la série des signes de l'activité (ϵ_t) a alterné 63 % du temps. Une stratégie serait donc de prendre l'exact contrepied du profil pour le trimestre t de celui observé en $t-1$:

$$\hat{\epsilon}_t = \phi_t^{\text{contrepied}}(\epsilon_{t-1}) = u_t$$

Sur la période 1997T1-2011T1, la moyenne des erreurs de la stratégie du contrepied est d'environ 36 %.

Stratégie de la moyenne de long terme

Cette stratégie consiste simplement à prolonger la série des profils pour le trimestre t par sa moyenne de long terme sur la période observée $[1, t-1]$ (arrondie à 0 ou 1) :

$$\hat{\epsilon}_t = \phi_t^{\text{moyenne}}(\epsilon_{t-1}) = 1 \left\{ \frac{1}{t-1} \sum_{i=1}^{t-1} \epsilon_i \geq 0,5 \right\}$$

La moyenne des erreurs de cette stratégie est d'environ 41 % depuis 1997. ■

2 - Étalonage « seuillé »

Cette stratégie consiste à déduire, chaque trimestre, le sens de variation de l'activité à partir d'une prévision de croissance du PIB en niveau. La prévision quantitative est obtenue à travers un modèle de régression usuel, faisant intervenir la croissance du PIB comme variable endogène, et les variables issues des enquêtes de conjoncture comme exogènes. Le profil de l'activité pour le trimestre t est alors simplement dérivé de la comparaison entre la prévision de croissance en t , notée y_t , et la croissance du trimestre précédent $t-1$ effectivement observée :

$$\hat{\epsilon}_t = \phi_t^{\text{reg}}(x_t, \epsilon_{t-1}) = 1 \{ \hat{y}_t \geq y_{t-1} \}$$

Les variables issues des enquêtes de conjoncture (soldes d'opinion, climat des affaires) étant potentiellement très nombreuses, il est nécessaire de se restreindre à un sous-ensemble réduit pour réaliser les étalonnages. Plusieurs spécifications ont ainsi été testées pour déterminer les prévisions quantitatives de croissance (y_t), et conduisent à des prévisions de profil de qualité comparable.

Deux spécifications peuvent notamment être mises en avant :

- La première est basée sur l'indicateur synthétique France en niveau sur le trimestre $t^{(1)}$, F_t , et en accélération « signée » entre

les trimestres $t-1$ et t , $\Delta F_t \times |\Delta F_t|$. Cette spécification fait également intervenir le niveau de croissance observée pour le trimestre précédent y_{t-1} :

$$\hat{y}_t = 0,60 - 0,40 y_{t-1} + 0,09 F_t + 0,09 \Delta F_t |\Delta F_t|$$

- La seconde spécification est basée sur une sélection de soldes d'opinion relatifs au secteur de l'industrie manufacturière : activité passée ($manuf.apa$) et activité prévue ($manuf.apre$). L'activité du secteur manufacturier contribue en effet fortement à la variabilité de la production marchande totale. Cette spécification mobilise également les niveaux antérieurs de croissance du PIB. Au final elle s'écrit donc comme suit⁽²⁾ :

$$\begin{aligned} \hat{y}_t = & 0,47 - 0,37 y_{t-1} - 0,10 y_{t-4} + 0,03.manuf.apa_{t,2} \\ & + 0,01.(manuf.apa_{t,2} - manuf.apa_{t,1}) \\ & + 0,02.(manuf.apre_{t,2} - manuf.apre_{t,1}) \\ & + 0,02.(manuf.apre_{t,1} - manuf.apre_{t-1,3}) \end{aligned}$$

Chacune de ces spécifications donne donc lieu à une stratégie de prévision du profil. Le taux moyen d'erreurs de prévision sur la période 1997T1-2011T3 est, en définitive, d'environ 18 % pour ces deux stratégies. ■

(1) En fait, il s'agit de la moyenne de l'indicateur pour les trois derniers mois observés au moment où l'on réalise la prévision, c'est-à-dire les deux premiers mois du trimestre et le troisième mois du trimestre précédent.

(2) Les soldes d'opinion étant mensuels, les indices correspondants se réfèrent à un trimestre ainsi qu'au mois de ce trimestre. Par exemple $manuf.apa_{t,2}$ représente le solde relatif à l'activité passée pour le deuxième mois du trimestre t .

L'activité va-t-elle accélérer ou ralentir ? Quelques outils pour répondre à cette question

3 - Des méthodes de classification plus complexes pour prévoir le sens de l'activité

Ces méthodes apparaissent en général comme les plus performantes. Le lecteur trouvera une présentation complète de ces stratégies dans Cornec et Mikol (2011).

Les techniques d'apprentissage automatique (*machine learning* en anglais) sont de plus en plus fréquemment utilisées dans les domaines de l'analyse financière, de la reconnaissance d'objets (visages, schémas...), des diagnostics médicaux etc. Elles sont en revanche encore très peu employées dans le domaine de l'analyse économique. Ces techniques sont basées sur des méthodes algorithmiques et permettent, à partir de données empiriques (données « d'apprentissage »), d'analyser et de classer des systèmes complexes. Dans notre cas, l'analyse porte sur les données économiques observées du trimestre en cours, que l'on souhaite « classifier » en fonction du sens de variation de la croissance non encore observé. Les données d'apprentissage correspondent alors aux observations conjoncturelles passées : PIB, profils de la

croissance et enquêtes de conjoncture. Parmi les algorithmes les plus couramment utilisés dans le *machine learning*, on peut citer :

- l'analyse discriminante linéaire
- les arbres de décision
- les machines à vecteur de support.

Une documentation précise et complète de ces trois méthodes figure notamment dans Hastie, Tibshirani et Friedman (2009). Appliquées à notre problématique de prévision du sens de variation de l'activité, ces trois méthodes donnent des résultats satisfaisants (Cornec et Mikol, 2011). En particulier, la méthode d'analyse discriminante linéaire, utilisant comme variables d'apprentissage le climat des affaires France (en niveau moyen trimestriel et en variation) et le premier retard de croissance du PIB, offre la meilleure performance de prédiction, avec seulement 12 % d'erreurs de prévision répertoriées sur la période 1997T1-2011T1. ■

Bibliographie

J. Bardaji, C. Minodier, L. Clavel, et F. Tallet : « Deux nouveaux indicateurs pour aider au diagnostic conjoncturel en France », *Note de Conjoncture* (décembre), Insee, 2008.

M. Bessec : « Étalonnages du taux de croissance du PIB français sur la base des enquêtes de conjoncture », *Économie & prévision*, (2):77-99, 2010.

G. Biau, K. Bleakley, L. Györfi, et G. Ottucsák : « Nonparametric sequential prediction of time series », *Journal of Nonparametric Statistics*, 22(3):297-317, 2010.

M. Cornec : « Constructing a conditional GDP fan chart with an application to French business survey data », (2011 - à paraître) http://ec.europa.eu/economy_finance/db_indicators/surveys/documents/workshops/2010/ec_meeting/14_m_cornec_conditional_gdp_fan_chart.pdf

M. Cornec et F. Mikol « Nowcasting GDP directional change with an application to French business survey data », (2011 - à paraître)

É. Dubois and E. Michaux : « Étalonnages à l'aide d'enquêtes de conjoncture : de nouveaux résultats. », *Économie & prévision*, (1):11-28, 2006.

A. Gammerman, V. Vovk et G. Shafer : « Algorithmic learning in a random world », *Springer*, New York, 2005.

H. Erkel-Rousse and C. Minodier : « Do business tendency surveys in industry and services help in forecasting GDP growth? A real-time analysis on french data », *Documents de Travail de la Dese*, Insee, 2009.

T. Hastie, R. Tibshirani and J.H. Friedman : « The elements of statistical learning: data mining, inference, and prediction », *Springer Verlag*, 2009.

G.H. Moore : « Forecasting short-term economic change », 1983.

M.H. Pesaran and A. Timmermann : « A simple nonparametric test of predictive performance », *Journal of Business & Economic Statistics*, 10(4):461-465, 1992.

F. Sedillot and N. Pain : « Indicator models of real GDP growth in selected OECD countries », *OECD Economics Department Working Papers*, 2003.

T.M. Sinclair, HO Stekler, and L. Kitzinger : « Directional forecasts of GDP and inflation: a joint evaluation with an application to Federal Reserve predictions », *Applied Economics*, 42(18):2289-2297, 2010.

K.F. Wallis : « Macroeconomic forecasting: a survey », *The Economic Journal*, 99 (394):28-61, 1989.