

Les enquêtes de conjoncture Insee et leur utilisation pour l'analyse conjoncturelle

Réunion plénière OPC

O. Biau (DR Insee Paca)

25 octobre 2012

Objectifs de l'exposé

- Présenter rapidement les enquêtes de conjoncture Insee
- En décrire les principales utilisations pour l'analyse de la conjoncture : construction d'indicateurs synthétiques, prévisions à court terme.

Agenda

- Vue d'ensemble
- Indicateurs synthétiques
- Prévisions à court terme

- **Vue d'ensemble**
- Indicateurs synthétiques
- Prévisions à court terme

La conjoncture à l'Insee

- Le Dépt. Conjoncture fait partie de la Direction des études et synthèses économiques (DESE) :
 - Département des comptes nationaux :
Comptes nationaux annuels et trimestriels.
 - Départements des études économiques d'ensemble :
 - Division croissance et politiques macroéconomiques.
 - ...
 - Département de la conjoncture :
 - Division des enquêtes de conjoncture.
 - Division de la synthèse conjoncturelle.
- Le réseau des DR, dans les Services études et diffusion

Les enquêtes de conjoncture

- Objectif : obtenir des données économiques relatives au passé récent, au présent et à l'avenir proche,
 - rapidement,
 - simplement,
 - pour un coût raisonnable.
- En mesurant l'opinion des entrepreneurs concernant leur activité, leurs effectifs... et le « moral » des ménages

Principales caractéristiques des enquêtes de conjoncture

- Enquêtes d'opinion (qualitatives), périodicité infra-annuelle (mensuelle ou trimestrielle).
- Petits échantillons : de 1000 à 5000 entreprises.
- Questionnaires simples et légers, questions en faible nombre et auxquelles on peut répondre vite.
- Questions portant sur le passé récent, la situation actuelle et les anticipations.
- Questions généralement à choix multiple, à 3 modalités : « en hausse », « stable » ou « en baisse » :
 - Evolution observée de votre production au cours des 3 derniers mois : ↗ ⇒ ↘
 - Evolution probable de votre production au cours des 3 prochains mois : ↗ ⇒ ↘

Principales caractéristiques des enquêtes de conjoncture

- Collecte, exploitation et diffusion rapides des résultats (à la fin du mois d'enquête).
- Publication des résultats sous la forme de soldes d'opinion :

$$s = \%(\nearrow) - \%(\searrow) = \sum_{h=1}^H \frac{V_h}{V} \frac{C_{+h} - C_{-h}}{C_h}$$

- Publication des résultats en fin de mois, avant les indicateurs quantitatifs.
- Résultats très peu révisés.
- Les résultats des enquêtes Insee sont envoyés à la Commission européenne, qui les (re)diffuse et calcule les agrégats UE + EA (Système harmonisé européen des enquêtes de conjoncture).

Les enquêtes de conjoncture :

« une statistique sans chiffres »

- L'information collectée est qualitative
- Avantage des soldes d'opinion : lecture simple de l'évolution des réponses au cours du temps
- Inconvénient : perte d'information
- Proposé par Theil (1952), revisité par Fansten (1976), puis Biau & Ferrari (2006)...
- La justification théorique du solde nécessite des hypothèses très fortes sur la distribution des réponses individuelles des entrepreneurs.

Points forts / points faibles

- (+) Indications en général de très bonne qualité sur le passé récent.
- (+) Vue assez complète et cohérente d'un secteur d'activité.
- (+) Éclairage sur des domaines peu couverts ou très tardivement par les statistiques quantitatives.

- (-) La nature qualitative des réponses et la subjectivité qui s'y attache rendent l'interprétation délicate.
- (-) L'intérêt des indications fournies diminue dès que les statistiques quantitatives sont disponibles.

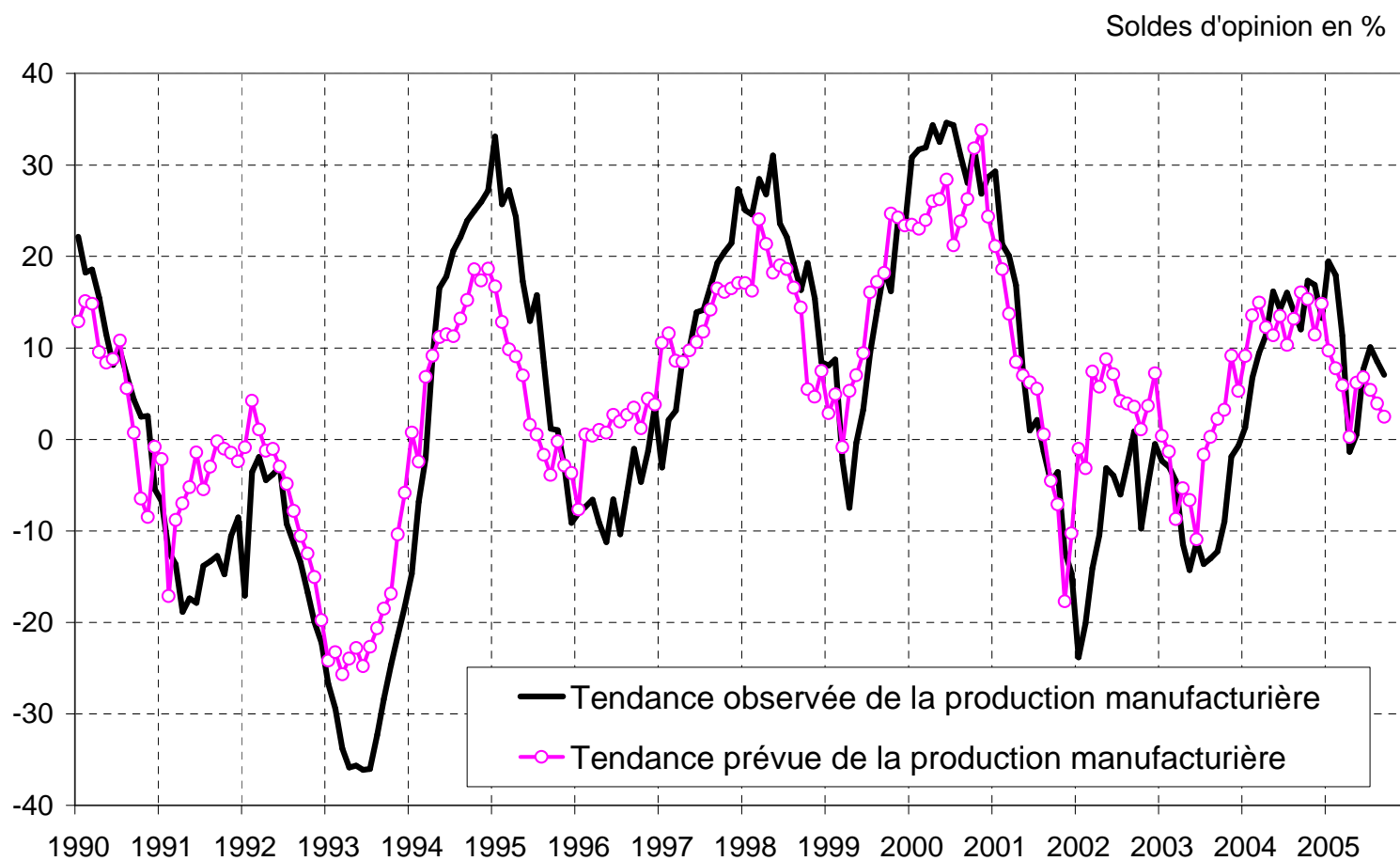
Les enquêtes de conjoncture nationales de l'Insee

- Un dizaine d'enquêtes régulières.
- Environ 80 publications par an.
- Secteurs enquêtés : industrie, construction, tertiaire à l'exclusion des services financiers.
- **Dans l'industrie** (4000 ent. enquêtées) :
 - Activité dans l'industrie (UE) : mensuelle.
 - Investissement (UE) : trimestrielle.
 - Situation de trésorerie : semestrielle.

Les enquêtes de conjoncture nationales de l'Insee

- **Dans le BTP** (1000 à 4000 ent. par enquête) :
 - Bâtiment (UE) (avec la FFB) : mensuelle.
 - Travaux publics (UE) (avec la FNTP) : trimestrielle.
 - Construction immobilière : trimestrielle.
 - Artisanat du bâtiment (10 salariés ou moins) : trimestrielle.
- **Dans le tertiaire** (4000 à 5000 ent. par enquête) :
 - Services (UE) : mensuelle depuis juin 2000.
 - Commerce de détail (UE) : mensuelle.
 - Commerce de gros : bimestrielle.

Exemple : soldes de l'enquête sur l'activité et les perspectives dans l'industrie

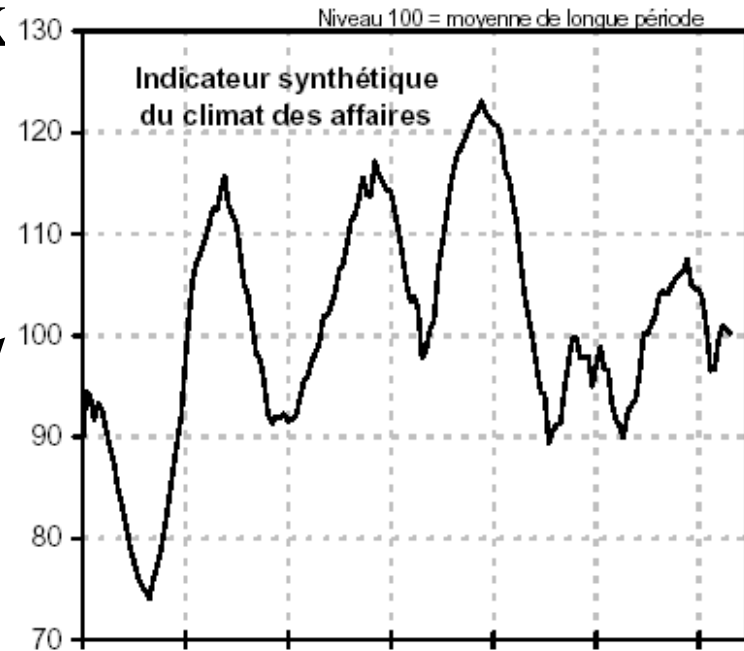
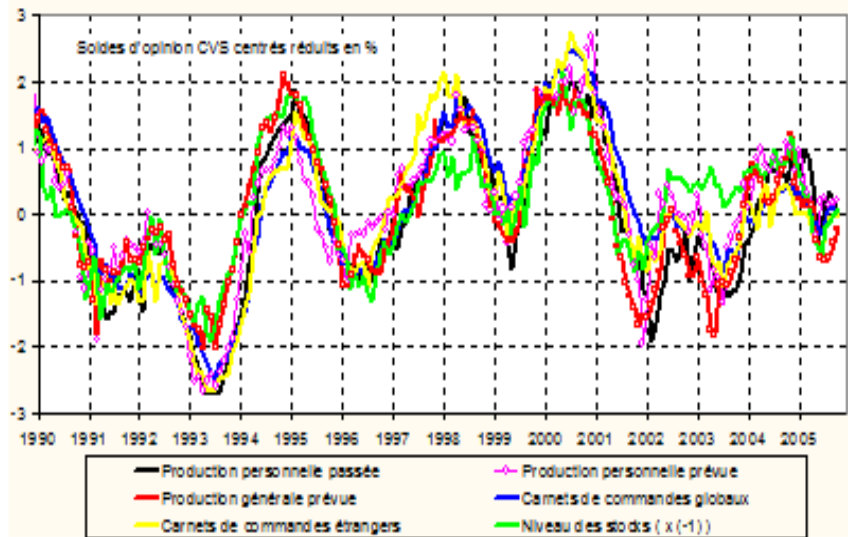


Interprétation des résultats

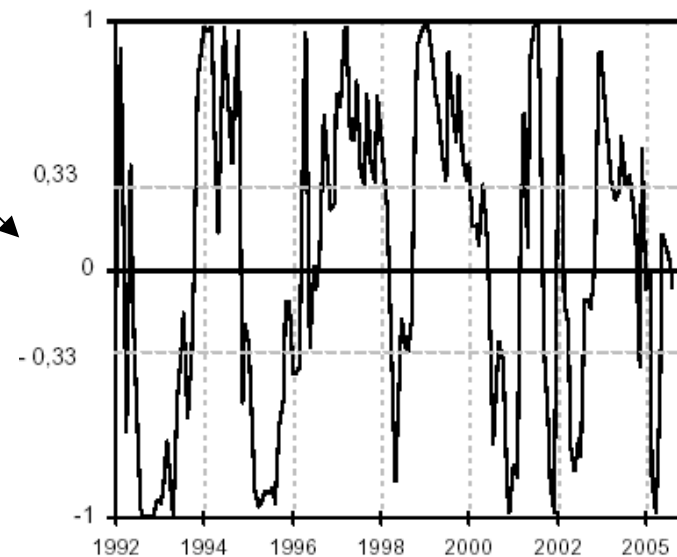
- Une première approche : centrer et réduire \Rightarrow comparer à la moyenne de long terme.
- Le signe du solde n'est pas interprétable, la variation du solde l'est.
- Comparer aux fluctuations passées.
- Comparer à l'écart-type calculé sur le passé de la série considérée.
- Comparer à l'écart-type lié à l'échantillonnage (interne à l'Insee).

- Vue d'ensemble
- **Indicateurs synthétiques**
- Prévisions à court terme

Des soldes d'opinion aux indicateurs synthétiques

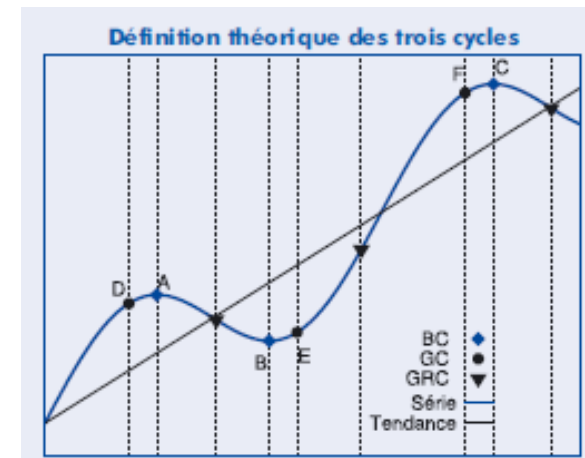


Indicateur de retournement



Les indicateurs synthétiques

- Objectifs :
 - favoriser la diffusion de l'enquête auprès du grand public
 - repérer facilement la position dans le cycle
- Comment définir le(s) cycle(s) économique(s) ?
 - Business cycle (niveau)
 - Growth cycle (écart à la tendance)
 - Growth rate cycle (points d'inflexions)



- Identifier les retournements : passage d'une phase haute à une phase basse ou inversement.

Le facteur commun, un indicateur « coïncident »

- Le facteur commun = résumé de l'information contenue dans les enquêtes (d'abord dans l'industrie, puis dans les services,... généralisé à l'ensemble des secteurs)
- C'est un indicateur du climat conjoncturel (« climat des affaires »).
- Méthode Insee :
extraction d'un facteur commun à partir de la technique de l'analyse factorielle (statique et dynamique).
- D'autres méthodes (plus simples) existent : DG ECFIN, IFO...

L'analyse factorielle statique :

- Principe : chaque solde est vu comme la somme de 2 termes :
 - un terme commun,
 - un terme propre à chaque série.
- Le modèle s'écrit :

$$y_{it} = \lambda_i F_t + u_{it}$$

$i = 1 \dots n$ (soldes d'opinion)

$t = 1 \dots T$ (observations)

- Hypothèses : facteurs communs et résidus non auto-corrélés et non corrélés entre eux...
- Cadre normalement adapté à des observations i.i.d.

Source : Doz, Lenglart [1999].

L'analyse factorielle dynamique

- Cadre adapté aux séries temporelles (variables auto-corrélées).
- Cadre adapté à des séries de fréquences différentes et présentant des changements de fréquence (enquête Services).
- Modélisation du facteur commun et des résidus par des processus ARMA.
- Ecriture du modèle à composantes inobservables sous la forme espace-état.
- Estimation des paramètres par la méthode du maximum de vraisemblance à l'aide du filtre de Kalman.

L'analyse factorielle dynamique

- Modèle à composantes inobservables :

$$\begin{cases} y_{it} = \lambda_i F_t + u_{it} \\ F_t = \varphi_1 F_{t-1} + \varphi_2 F_{t-2} + \varepsilon_t - \theta \varepsilon_{t-1} \\ u_{it} = \rho_i u_{it-1} + \varepsilon_{it} \end{cases}$$

- Représentation espace-état :

$$\begin{aligned} y_t &= Z \alpha_t \\ \alpha_t &= A \alpha_{t-1} + R \eta_t \end{aligned} \quad y_t = \begin{pmatrix} y_{1t} \\ \vdots \\ y_{It} \end{pmatrix} \quad \alpha_t = \begin{pmatrix} F_t \\ F_{t-1} \\ \varepsilon_t \\ u_{1t} \\ \vdots \\ u_{It} \end{pmatrix}$$

L'analyse factorielle dynamique

- Représentation espace-état (suite) :

$$y_t = Z \alpha_t$$

$$\alpha_t = A \alpha_{t-1} + R \eta_t$$

$$Z = \left(\begin{array}{ccc|c} \lambda_1 & 0 & 0 & \\ \vdots & \vdots & \vdots & Id(I) \\ \lambda_I & 0 & 0 & \end{array} \right)$$

$$A = \left(\begin{array}{ccc|ccc} \varphi_1 & \varphi_2 & -\theta & & & \\ 1 & 0 & 0 & & & 0 \\ 0 & 0 & 0 & & & \\ \hline & & & \rho_1 & & \\ & 0 & & & \ddots & \\ & & & & & \rho_I \end{array} \right)$$

$$R = \left(\begin{array}{c|c} 1 & 0 \\ 0 & 0 \\ 1 & 0 \\ \hline 0 & Id(I_t) \end{array} \right)$$

$$\eta_t = \begin{pmatrix} \varepsilon_t \\ \varepsilon_{1t} \\ \vdots \\ \varepsilon_{It} \end{pmatrix}$$

L'analyse factorielle dynamique

- La vraisemblance du modèle peut s'écrire :

$$f_{\theta}(Y^t) = \prod_{i=2}^t f_{\theta}(Y_i / Y^{i-1}) f_{\theta}(Y_1); Y^t = (y_1, \dots, y_t)$$

- Hypothèse : $\eta_t \sim \text{i.i.d. } N(0, Q)$

→ Le filtre de Kalman permet de calculer la vraisemblance par récurrence

- Notations : $\alpha_{t/t} = E(\alpha_t / Y^t)$; $\alpha_{t/t-1} = E(\alpha_t / Y^{t-1})$

- Mise à jour : $\alpha_{t/t} = \alpha_{t/t-1} + K_t (y_t - Z\alpha_{t/t-1})$

La représentation espace-état

Écriture et estimation du modèle

- Forme générale de la représentation espace-état :

$$y_t = Z_t \alpha_t + d_t + e_t \quad (M)$$

$$\alpha_t = A_t \alpha_{t-1} + c_t + R_t \eta_t \quad (T)$$

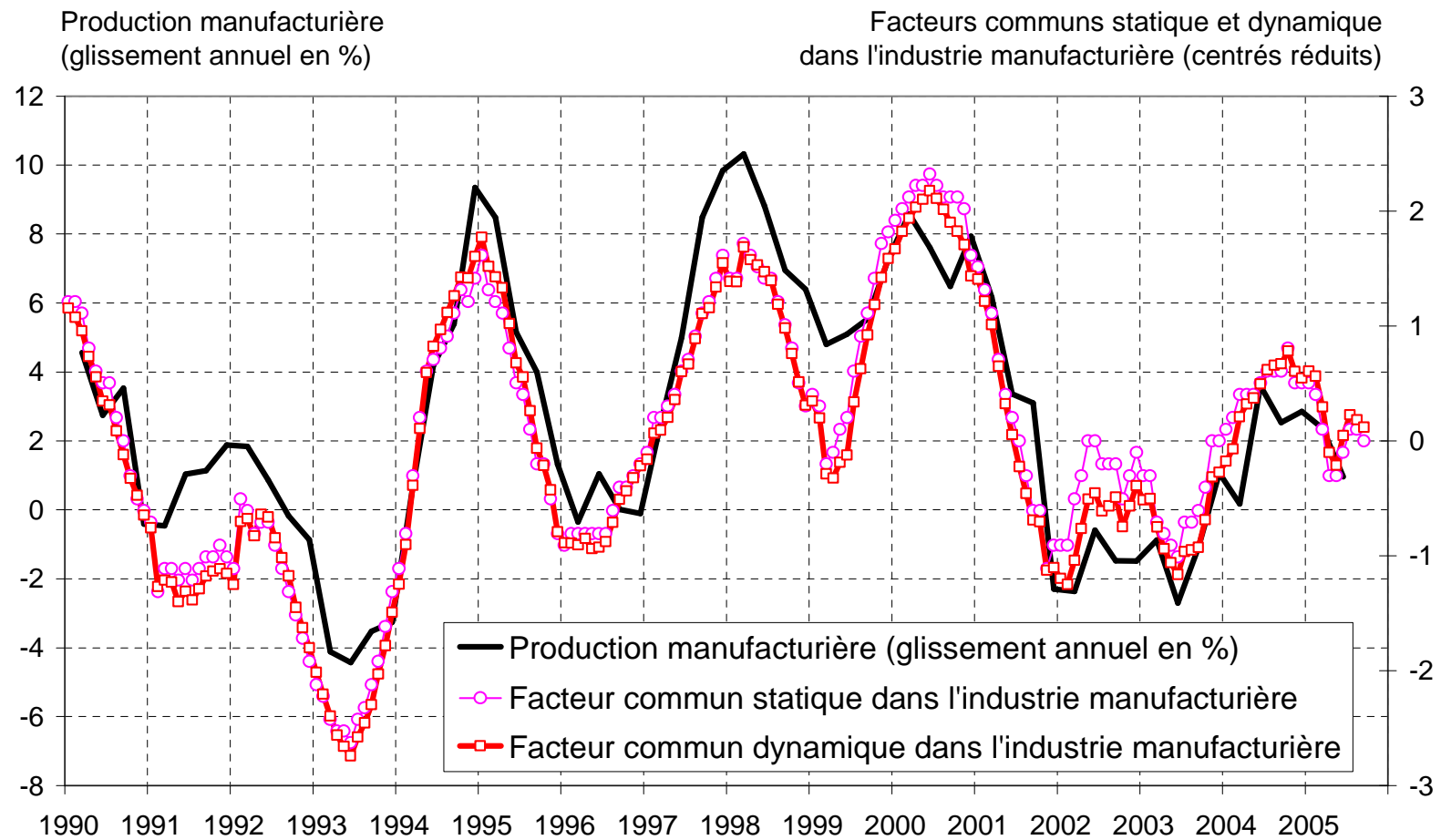
(*M*) : équation de mesure

(*T*) : équation de transition

hypothèse : (e_t) et (η_t) bruits blancs indépendants

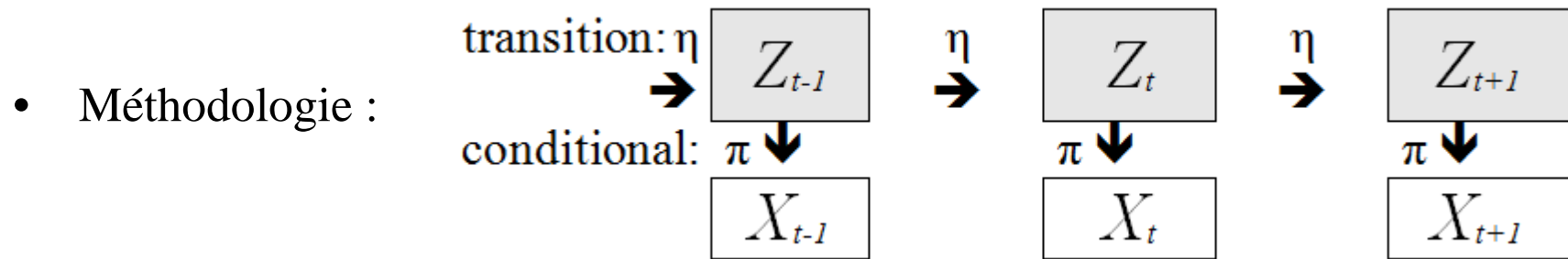
- Estimation par la méthode du maximum de vraisemblance à l'aide du filtre de Kalman.
(hypothèse : vecteurs gaussiens)

Comparaison des indicateurs Industrie statique et dynamique



L'indicateur de retournement, un indicateur « avancé »

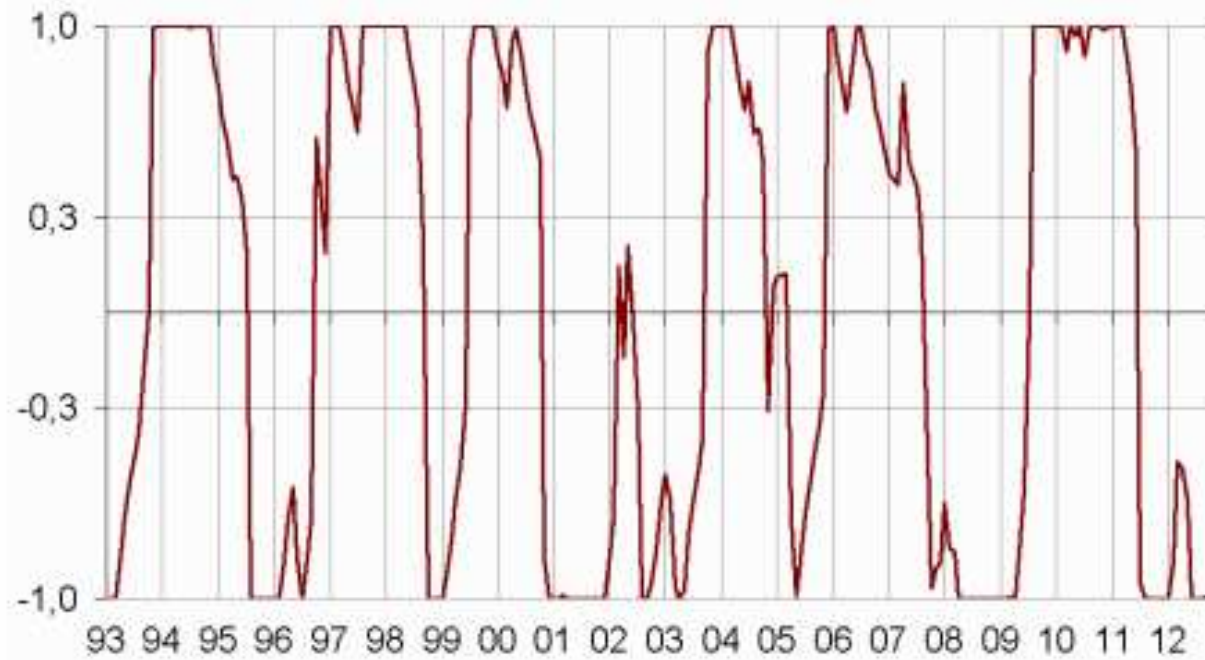
de la probabilité d'un retournement de tendance dans le futur proche.



- Z : régime de l'économie (inobservable), Z_t possède la dynamique d'une chaîne de Markov homogène d'ordre 1.
- X : séries « observées », information nouvelle (+) ou (-) extraite à chaque date t et pour chaque solde
- Filtre de Kitagawa-Hamilton au cœur de la méthode :
- La connaissance des (π) et des (η) , suffit à déduire et des probabilités conditionnelles :
$$P(Z=+1/\underline{X}) - P(Z=-1/\underline{X})$$
- Sources : Lengart [1997], Grégoir et Lengart [2000].

Indicateur de retournement France

Proche de +1 (resp. -1) : climat conjoncturel favorable (défavorable)



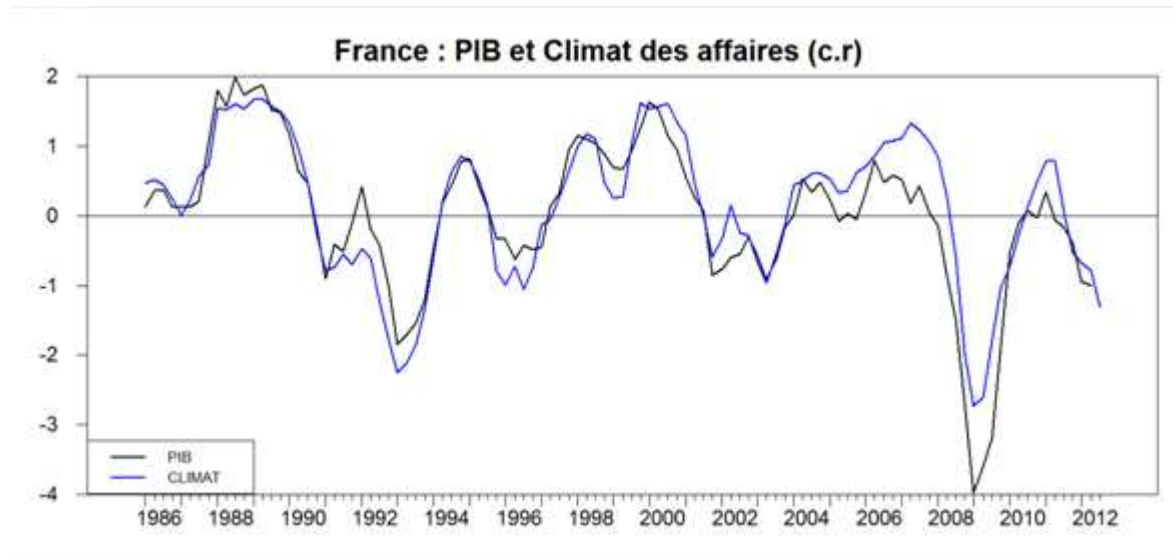
L'indicateur de retournement calculé au niveau France reste en zone défavorable. C'est également le cas de tous les indicateurs de retournement sectoriels sauf dans les services où l'indicateur se redresse tout en restant en zone d'incertitude.

- Vue d'ensemble
- Indicateurs synthétiques
- **Prévisions à court terme**

« Prediction is very difficult, especially if it is about the future »

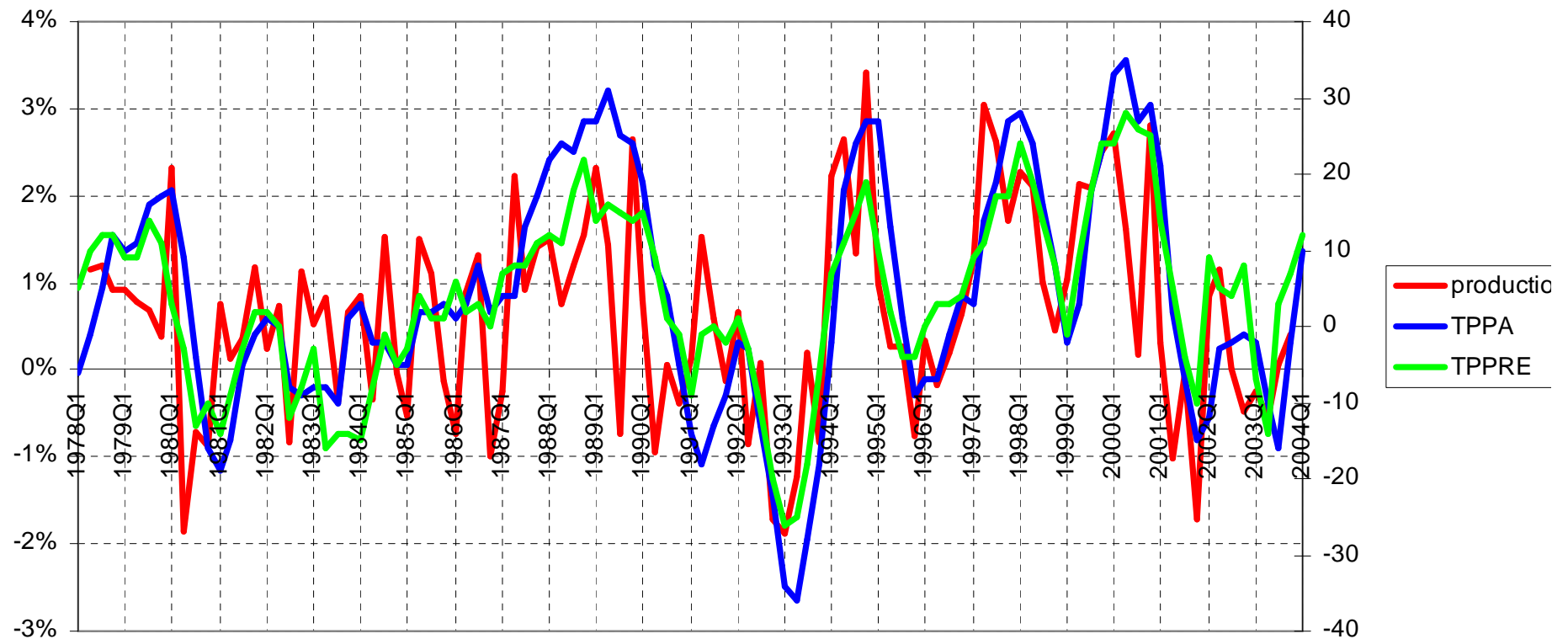
Niels Bohr

What else?



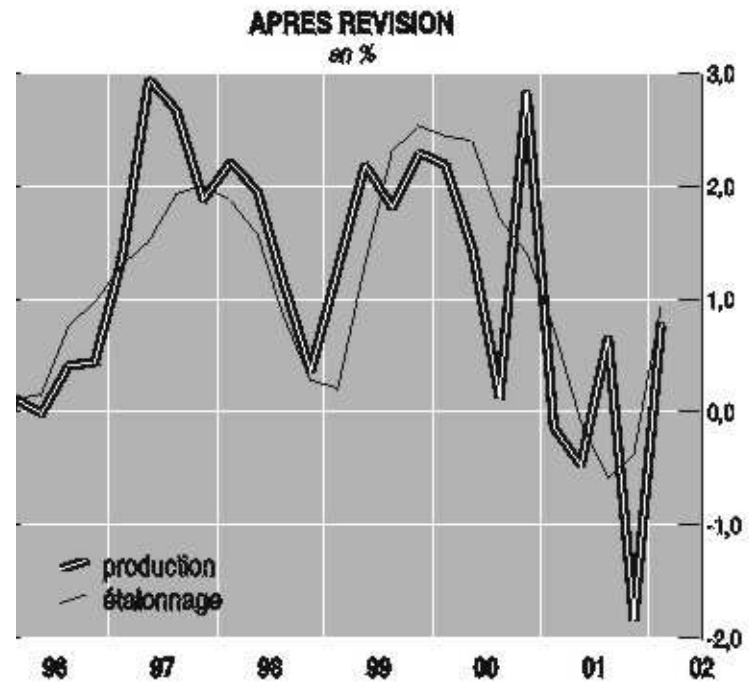
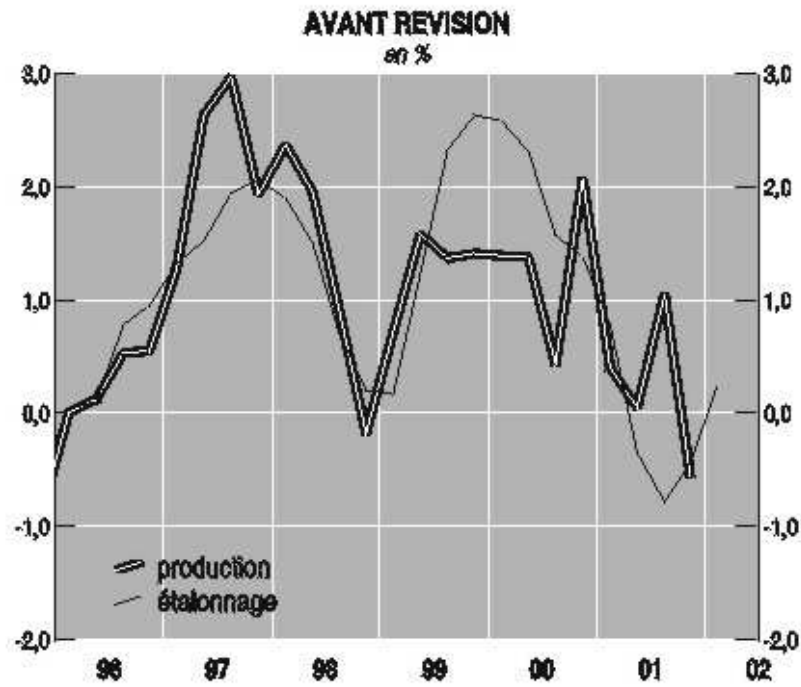
Des résultats d'enquête aux prévisions

Taux de croissance de la production manufacturière



Les soldes d'opinion sont très corrélés aux indicateurs quantitatifs et disponibles avant eux.

Question existentielle...



Utilisation des soldes d'opinion pour la prévision à court terme

- Prévisions à 3 ou 6 mois des taux de croissance trimestriels de grands agrégats macroéconomiques :
 - PIB,
 - production (et effectifs par secteur),
 - investissement,
 - importations,
 - stocks,
 - emplois,...
- Contribution au diagnostic conjoncturel publié tous les trimestres dans les notes et points de conjoncture.

Étalonnages univariés – modèle linéaire

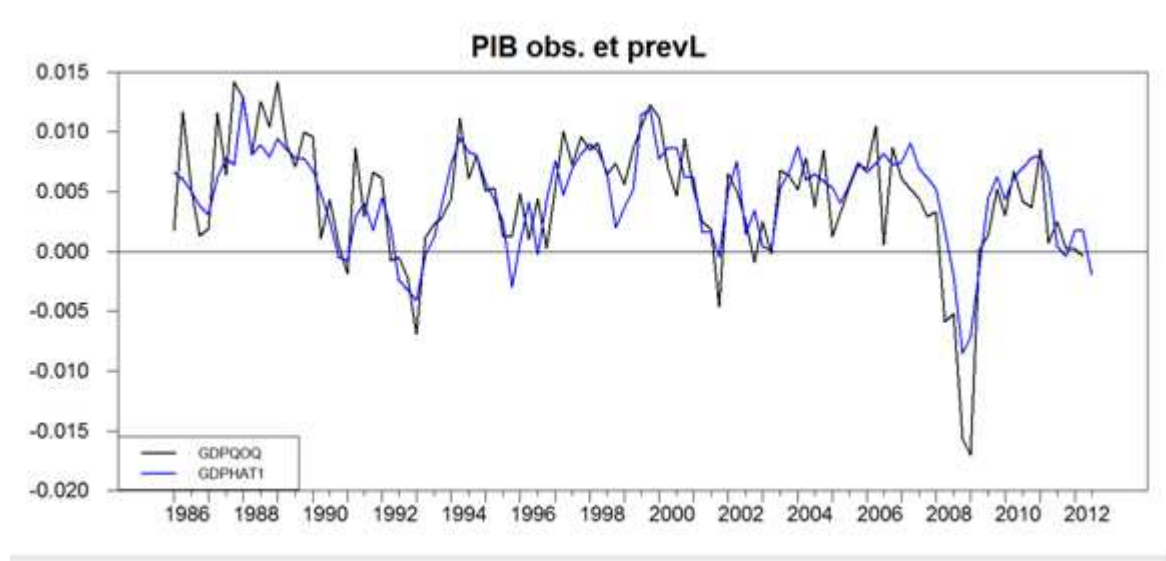
$$\Delta \log PIB_t = c + a_0 \cdot I_t + a_1 \cdot I_{t-1} + b \cdot \Delta \log PIB_{t-1}$$

- En pratique, plusieurs modifications peuvent intervenir...
 - Certains coefficients sont rejetés de manière très pragmatique et/ou d'après des tests statistiques (nullité des coefficients, stabilité dans le temps).
 - On ne garde pas toujours le taux de croissance retardé de la variable endogène (on prend alors uniquement les données d'enquêtes)
 - en particulier en cas de suspicion quant à la qualité du dernier point de la variable endogène (risque de forte révision),
 - mais, généralement, pouvoir prédictif moindre.
 - Il est possible d'enrichir la spécification en y incluant d'autres variables explicatives.
 - On peut faire apparaître la variable l'indicateur **en niveau** et **en variation**. Ainsi, ces deux spécifications sont équivalentes :

$$\Delta \log PIB_t = c + aI_t + bI_{t-1} = c + (a + b)I_t - b\Delta I_t$$

La deuxième a l'avantage de faire ressortir l'influence des variations du solde, facteur important notamment en période de retournement.

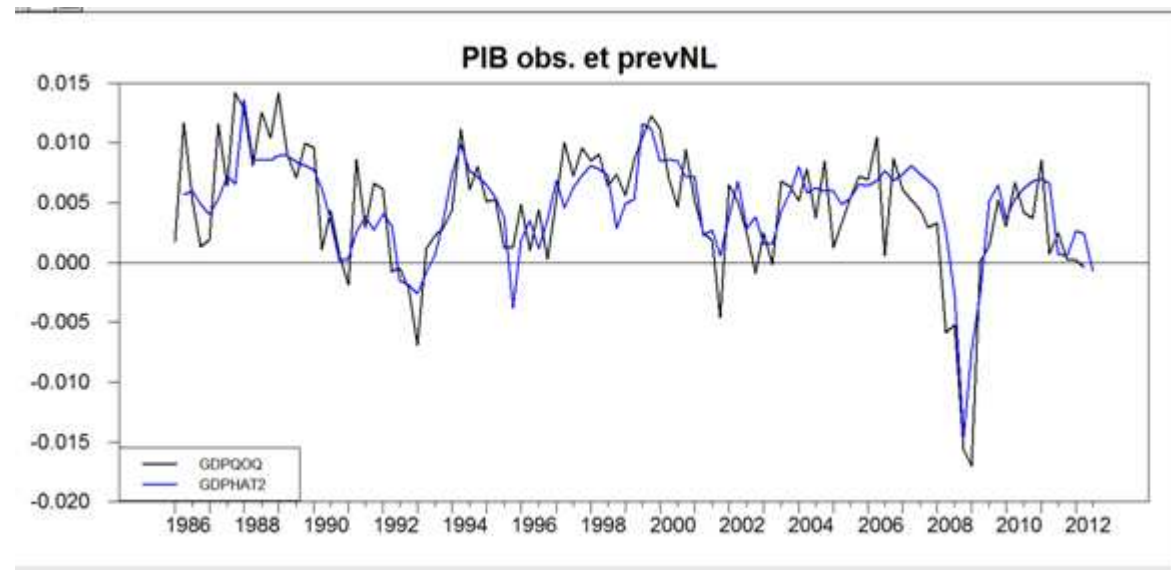
Modèle : $\log(\text{PIB}) / 1, S, \Delta S$



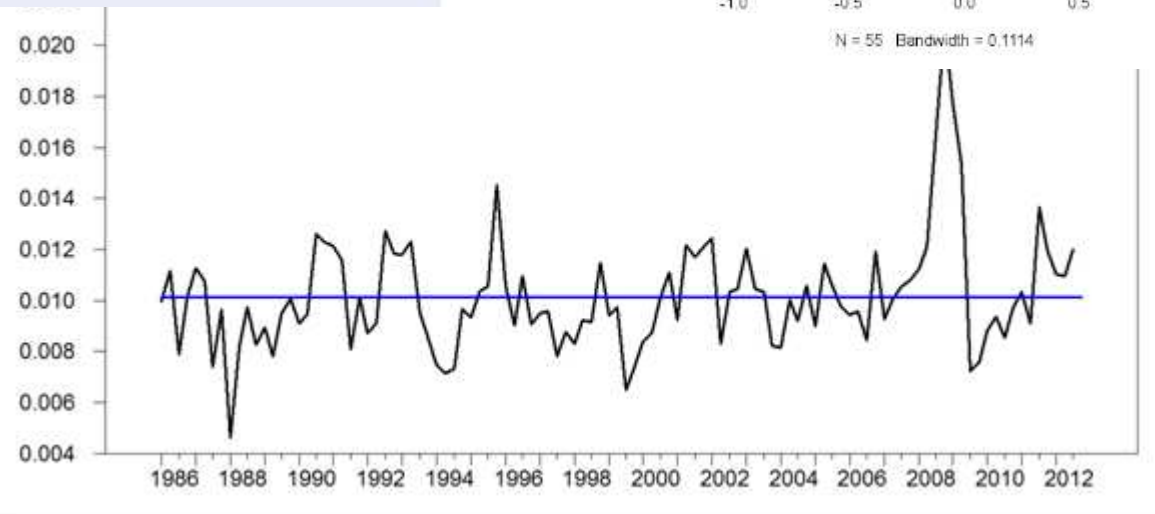
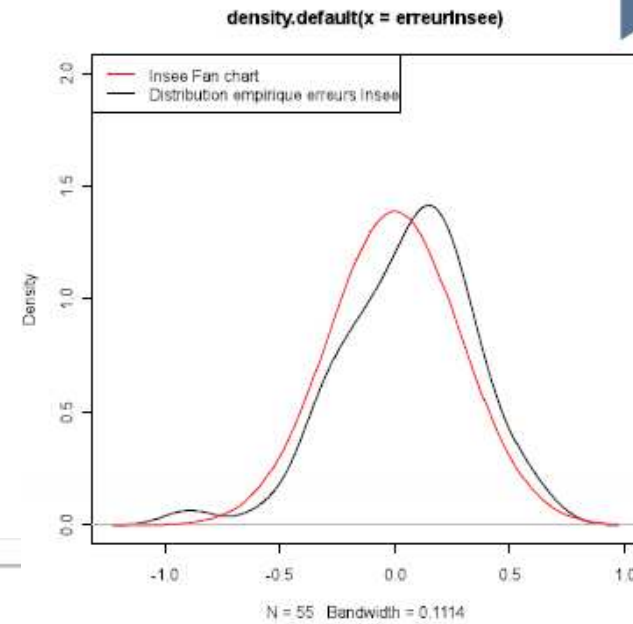
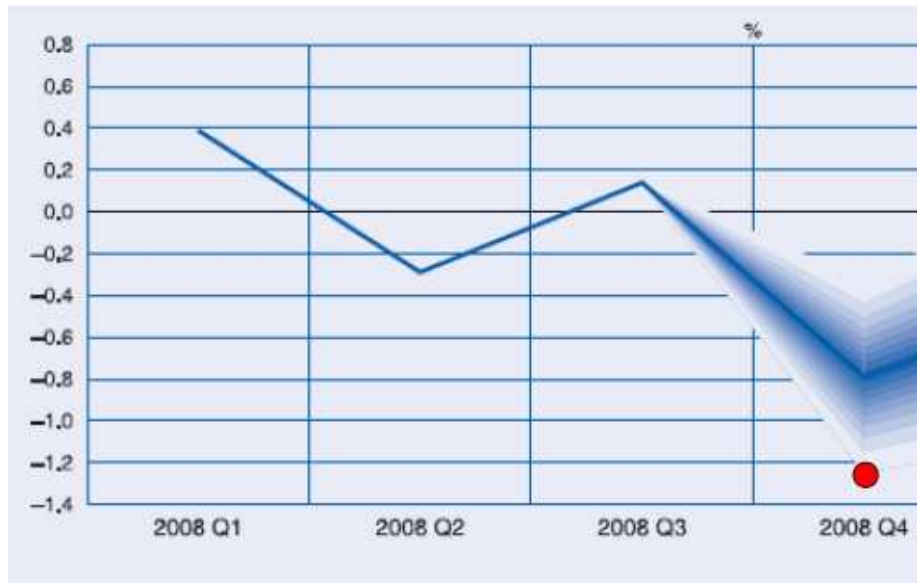
Les derniers développements...

- Modèles à facteurs dynamiques sur des grandes bases de données, prévisions en temps réel
- Avec la crise de 2008, problème de « non linéarité aux valeurs extrêmes » : recours à des spécifications non linéaires
- Utilisation des données individuelles :
 - techniques nouvelles (non paramétriques) pour faire mieux que les modèles linéaires avec des soldes
 - explorer de nouveaux domaines (perception de l'inflation)
- Utiliser de nouvelles techniques (régression quantile, non-paramétriques, TVP) pour estimer la difficulté de l'exercice de prévision, estimer si le PIB va accélérer / décélérer, le PIB potentiel.

Modèle : $\log(\text{PIB}) / 1, \log(\text{PIB})_{-1} S, \Delta S.|\Delta S|$

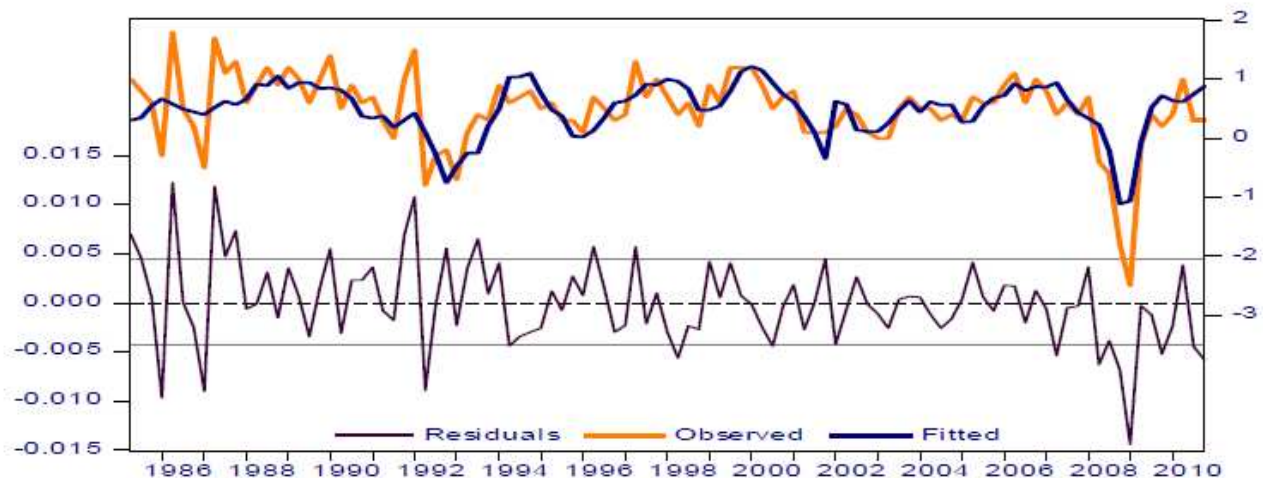


Forecasting Risk Index (M. Cornec 2010)

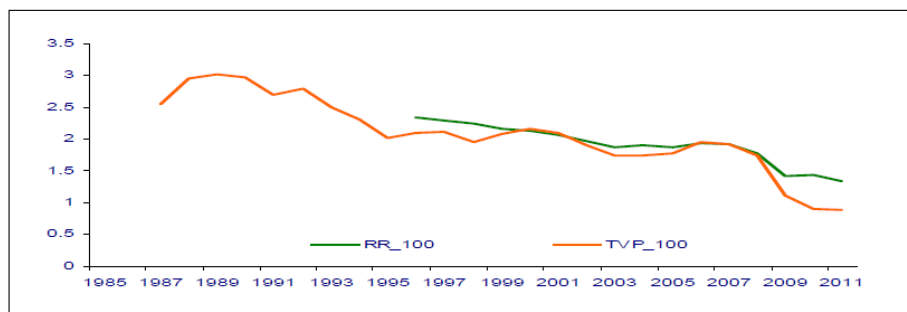


Is there a decoupling between soft and hard data? The relationship between GDP growth and the ESI

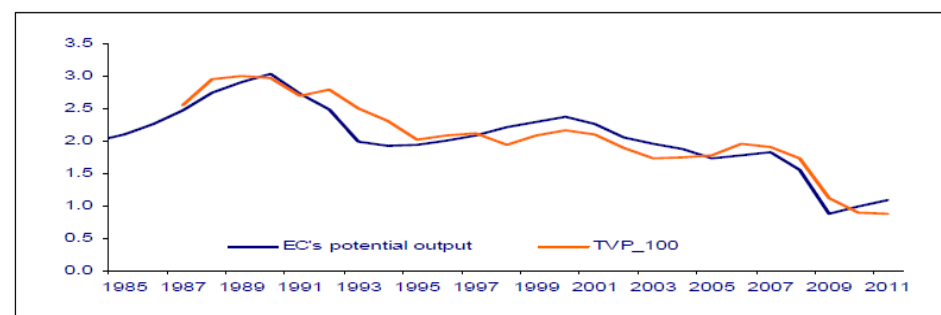
Olivier BIAU[§] - Angela D'ELIA^{**}



$$Y_t = b_0 + b_1 ESI_t + b_2 \Delta ESI_t + u_t$$



Graph 2: Projected annual GDP growth (%) corresponding to a level of 100 in the ESI, euro area (1985 to 2011). RR_100 and TVP_100: outcome from rolling regression and TVP model, respectively.



Graph 3: Annual potential GDP growth (%), euro area (1985 to 2011)
Source: EC's AMECO database and our computations

Références

- Doz C. et Lengart 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.
- Gregoir S. et Lengart F. (2000). "Measuring the probability of a business cycle turning point by using a multivariate qualitative hidden Markov model", *Journal of Forecasting*, vol. 19, pp. 81-102.
- Dubois E., Michaux E. (2006) : « Étalonnages à l'aide d'enquêtes de conjoncture : de nouveaux résultats », *Économie et Prévision*, n°172.
- Cornec, Deperraz (2006), « Un nouvel indicateur synthétique mensuel résumant le climat des affaires dans les services en France », *Économie et Statistique*, n°395-396, 2006.
- Bardaji, Minodier, Clavel Tallet (2008) « Deux nouveaux indicateurs pour aider au diagnostic conjoncturel en France ».
- Cornec M (2010), « Constructing a conditional GDP fan chart with an application to French business survey data ».
- Biau O, Ferrari N (2006), « Théorie de l'opinion Faut-il pondérer les réponses individuelles ? », document de travail Insee n° G 2006/12.
- Biau G, Biau O, Rouvière L (2007), « Nonparametric forecasting of the manufacturing output growth with firm-level survey data », document de travail Insee n°G 2006/06 et JBCMA n°200 7, 3-03.
- Biau O, Erkel-Rousse H, Ferrari N (2006), « Réponses individuelles aux enquêtes de conjoncture et prévision de la production manufacturière », document de travail Insee n°G 2005/12 et *Économie et Statistique*, n°395-396, 2006.
- Biau O, D'Elia A (2010), « Euro area GDP forecast using large survey dataset - A random forest approach ».
- Biau O, D'Elia A (2011), « Is there a decoupling between soft and hard data ? ».