

# PRÉVOIR LES CHANGEMENTS DE DIRECTION DU TAUX DE CROISSANCE DU PIB, AVEC UNE APPLICATION AUX ENQUÊTES DE CONJONCTURE FRANÇAISES

Matthieu Cornec

Fanny Mikol

matthieu.cornec@insee.fr

fanny.mikol@sante.gouv.fr

INSEE, DRESS

26 Janvier 2012

# 1 INTRODUCTION

## 2 NOTATIONS, DÉFINITIONS

## 3 STRATÉGIES BENCHMARK

## 4 PRÉDICTION SUR LES FUTURES ERREURS ?

## 5 INDICATEUR DE PROFIL

1 INTRODUCTION

2 NOTATIONS, DÉFINITIONS

3 STRATÉGIES BENCHMARK

4 PRÉDICTION SUR LES FUTURES ERREURS ?

5 INDICATEUR DE PROFIL

1 INTRODUCTION

2 NOTATIONS, DÉFINITIONS

3 STRATÉGIES BENCHMARK

4 PRÉDICTION SUR LES FUTURES ERREURS ?

5 INDICATEUR DE PROFIL

- 1 INTRODUCTION
- 2 NOTATIONS, DÉFINITIONS
- 3 STRATÉGIES BENCHMARK
- 4 PRÉDICTION SUR LES FUTURES ERREURS ?
- 5 INDICATEUR DE PROFIL

- 1 INTRODUCTION
- 2 NOTATIONS, DÉFINITIONS
- 3 STRATÉGIES BENCHMARK
- 4 PRÉDICTION SUR LES FUTURES ERREURS ?
- 5 INDICATEUR DE PROFIL

- Dans la Note de Conjoncture, la prévision du taux de croissance du PIB est **quantitative** : +0,5%, +0,2%, ....
- *mais* le discours économique est principalement **qualitatif** : *la production accélère, ou bien l'activité décélère ou le climat s'est stabilisé.*

## EXEMPLE

*La vue d'ensemble de la Note de juin 2010 contenait 7 fois le terme **augmenter**, et plus de 20 fois le terme **progresser***

- Dans la Note de Conjoncture, la prévision du taux de croissance du PIB est **quantitative** : +0,5%, +0,2%, ....
- *mais* le discours économique est principalement **qualitatif** : *la production accélère, ou bien l'activité décélère ou le climat s'est stabilisé.*

## EXEMPLE

*La vue d'ensemble de la Note de juin 2010 contenait 7 fois le terme **augmenter** , et plus de 20 fois le terme **progresser***



- Dans la Note de Conjoncture, la prévision du taux de croissance du PIB est **quantitative** : +0,5%, +0,2%, ....
- *mais* le discours économique est principalement **qualitatif** : *la production accélère, ou bien l'activité décélère ou le climat s'est stabilisé.*

## EXEMPLE

*La vue d'ensemble de la Note de juin 2010 contenait 7 fois le terme **augmenter** , et plus de 20 fois le terme **progresser***

## DÉFINITION

Une *accélération* (resp. *décélération*) du PIB correspond à un taux de croissance observé supérieur (resp. inférieur) au taux de croissance observé du trimestre précédent.

## EXEMPLE

Au T2 2010, le taux de croissance 0,7% et au T3 2010 0,5% : il y a eu *décélération* au T3 2010.

# DÉFINITION PROFIL

## DÉFINITION

Une *accélération* (resp. *décélération*) du PIB correspond à un taux de croissance observé supérieur (resp. inférieur) au taux de croissance observé du trimestre précédent.

## EXEMPLE

Au T2 2010, le taux de croissance 0,7% et au T3 2010 0,5% : il y a eu *décélération* au T3 2010.

- problématique **quantitative** classique : peut-on prévoir le niveau du taux de croissance pour le trimestre coïncident ?
- nouvelle problématique **qualitative** : pouvons-nous prévoir le profil du PIB pour le trimestre coïncident ? accélération ou décélération ?
- objectifs additionnels : définir un cadre d'analyse **séquentiel**, **non asymptotique** permettant de **comparer** les stratégies.

- problématique **quantitative** classique : peut-on prévoir le niveau du taux de croissance pour le trimestre coïncident ?
- nouvelle problématique **qualitative** : pouvons-nous prévoir le profil du PIB pour le trimestre coïncident ? accélération ou décélération ?
- objectifs additionnels : définir un cadre d'analyse **séquentiel**, **non asymptotique** permettant de **comparer** les stratégies.

- problématique **quantitative** classique : peut-on prévoir le niveau du taux de croissance pour le trimestre coïncident ?
- nouvelle problématique **qualitative** : pouvons-nous prévoir le profil du PIB pour le trimestre coïncident ? accélération ou décélération ?
- objectifs additionnels : définir un cadre d'analyse **séquentiel**, **non asymptotique** permettant de **comparer** les stratégies.

# Plan

- 1 INTRODUCTION
- 2 NOTATIONS, DÉFINITIONS
- 3 STRATÉGIES BENCHMARK
- 4 PRÉDICTION SUR LES FUTURES ERREURS ?
- 5 INDICATEUR DE PROFIL

# DÉFINITION PROFIL

## DÉFINITION

Chaque trimestre  $q$ , nous définissons le sens de la conjoncture par

$$\varepsilon_q := 1 \{y_q^{PR} > y_{q-1}^{PR}\}$$

avec  $y_q^{PR}$  le premier résultat du taux de croissance du PIB *relatif au trimestre*  $q$  mais *connu au trimestre*  $q + 1$ .

## REMARQUE

Nous avons donc la convention suivante :

$$\begin{cases} \varepsilon_q = 1 & \text{accélération} \\ \varepsilon_q = 0 & \text{décélération} \end{cases}$$



# DÉFINITION PROFIL

## DÉFINITION

Chaque trimestre  $q$ , nous définissons le sens de la conjoncture par

$$\varepsilon_q := 1 \{y_q^{PR} > y_{q-1}^{PR}\}$$

avec  $y_q^{PR}$  le premier résultat du taux de croissance du PIB *relatif au trimestre*  $q$  mais *connu au trimestre*  $q + 1$ .

## REMARQUE

Nous avons donc la convention suivante :

$$\begin{cases} \varepsilon_q = 1 & \text{accélération} \\ \varepsilon_q = 0 & \text{décélération} \end{cases}$$

# SÉRIES DES PROFILS

	Qtr1	Qtr2	Qtr3	Qtr4
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	

# DÉFINITION STRATÉGIE DE PRÉVISION DE PROFIL

- A chaque trimestre  $q$ , le conjoncturiste prend connaissance des données économiques jusqu'au trimestre  $q$  :  $x_{\bar{q}}$  et des profils passés  $\varepsilon_{q-1}$ .
- Le **but** est à chaque trimestre  $q$  de deviner le nouveau sens de l'activité connaissant le passé.

## DÉFINITION

Une *stratégie de prévision* est définie comme une famille de prédicteurs :

$$\hat{\varepsilon}_q := \phi_q(x_{\bar{q}}, \varepsilon_{q-1})$$

## REMARQUE

Dans cette étude,  $x_q$  représentera les principaux soldes d'opinion des enquêtes de conjoncture.

# DÉFINITION STRATÉGIE DE PRÉVISION DE PROFIL

- A chaque trimestre  $q$ , le conjoncturiste prend connaissance des données économiques jusqu'au trimestre  $q$  :  $x_{\bar{q}}$  et des profils passés  $\varepsilon_{q-1}$ .
- Le **but** est à chaque trimestre  $q$  de deviner le nouveau sens de l'activité connaissant le passé.

## DÉFINITION

Une *stratégie de prévision* est définie comme une famille de prédicteurs :

$$\hat{\varepsilon}_q := \phi_q(x_{\bar{q}}, \varepsilon_{q-1})$$

## REMARQUE

Dans cette étude,  $x_q$  représentera les principaux soldes d'opinion des enquêtes de conjoncture.

# DÉFINITION STRATÉGIE DE PRÉVISION DE PROFIL

- A chaque trimestre  $q$ , le conjoncturiste prend connaissance des données économiques jusqu'au trimestre  $q$  :  $x_{\bar{q}}$  et des profils passés  $\varepsilon_{\overline{q-1}}$ .
- Le **but** est à chaque trimestre  $q$  de deviner le nouveau sens de l'activité connaissant le passé.

## DÉFINITION

Une *stratégie de prévision* est définie comme une famille de prédicteurs :

$$\hat{\varepsilon}_q := \phi_q(x_{\bar{q}}, \varepsilon_{\overline{q-1}})$$

## REMARQUE

Dans cette étude,  $x_q$  représentera les principaux soldes d'opinion des enquêtes de conjoncture.

# DÉFINITION STRATÉGIE DE PRÉVISION DE PROFIL

- A chaque trimestre  $q$ , le conjoncturiste prend connaissance des données économiques jusqu'au trimestre  $q$  :  $x_{\bar{q}}$  et des profils passés  $\varepsilon_{\overline{q-1}}$ .
- Le **but** est à chaque trimestre  $q$  de deviner le nouveau sens de l'activité connaissant le passé.

## DÉFINITION

Une *stratégie de prévision* est définie comme une famille de prédicteurs :

$$\hat{\varepsilon}_q := \phi_q(x_{\bar{q}}, \varepsilon_{\overline{q-1}})$$

## REMARQUE

Dans cette étude,  $x_q$  représentera les principaux soldes d'opinion des *enquêtes de conjoncture*.

# DÉFINITION COÛT D'UNE STRATÉGIE

## DÉFINITION

Le *coût* moyen cumulé à la date  $q$  par la stratégie  $\Phi := (\phi_q)_q$  est défini par :

$$L_q(\Phi) := \frac{1}{q} \sum_{t=1}^q 1(\varepsilon_t \neq \hat{\varepsilon}_t).$$

Il s'agit de l'*erreur moyenne*.

## REMARQUE

*Erreurs de prévision étudiées depuis 1997 (pour comparer à l'Insee) ou depuis 1992.*

# DÉFINITION COÛT D'UNE STRATÉGIE

## DÉFINITION

Le *coût* moyen cumulé à la date  $q$  par la stratégie  $\Phi := (\phi_q)_q$  est défini par :

$$L_q(\Phi) := \frac{1}{q} \sum_{t=1}^q 1(\varepsilon_t \neq \hat{\varepsilon}_t).$$

Il s'agit de l'*erreur moyenne*.

## REMARQUE

Erreurs de prévision étudiées depuis 1997 (pour comparer à l'Insee) ou depuis 1992.



# Plan

- 1 INTRODUCTION
- 2 NOTATIONS, DÉFINITIONS
- 3 STRATÉGIES BENCHMARK**
- 4 PRÉDICTION SUR LES FUTURES ERREURS ?
- 5 INDICATEUR DE PROFIL

# STRATÉGIE AGENT NON-INFORMÉ

- Un agent non-informé (n'ayant pas connaissance des données) pourra prévoir au hasard en tirant dans une loi Bernouilli indépendante  $U_q$  de paramètre  $1/2$ .

$$\widehat{\varepsilon}_q := u_q.$$

- Son **erreur moyenne** depuis 1997 aurait été avec grande probabilité (95%) supérieure à

$$L_q(\Phi^B) \geq 35\%.$$

# STRATÉGIE AGENT NON-INFORMÉ

- Un agent non-informé (n'ayant pas connaissance des données) pourra prévoir au hasard en tirant dans une loi Bernouilli indépendante  $U_q$  de paramètre  $1/2$ .

$$\widehat{\varepsilon}_q := u_q.$$

- Son **erreur moyenne** depuis 1997 aurait été avec grande probabilité (95%) supérieure à

$$L_q(\Phi^B) \geq 35\%.$$

# ÉTALONNAGE SEUILLÉ

- Définissons un étalonnage utilisé pour prévoir le PIB en niveau :

$$\widehat{y}_q^{PR} = \widehat{\beta}_0 + \widehat{\beta}_1 y_{q-1}^{PR} + \widehat{\beta}_2 F_q + \widehat{\beta}_3 \Delta F_q |\Delta F_q|$$

- La prévision en **profil** est alors donnée par :

$$\widehat{\varepsilon}_q := 1 \left\{ \widehat{y}_q^{PR} \geq y_{q-1}^{PR} \right\}$$

- L'erreur moyenne de l'étalonnage seuillé depuis 1997 est de 18%.

- Définissons un étalonnage utilisé pour prévoir le PIB en niveau :

$$\widehat{y}_q^{PR} = \widehat{\beta}_0 + \widehat{\beta}_1 y_{q-1}^{PR} + \widehat{\beta}_2 F_q + \widehat{\beta}_3 \Delta F_q | \Delta F_q |$$

- La prévision en **profil** est alors donnée par :

$$\widehat{\varepsilon}_q := 1 \left\{ \widehat{y}_q^{PR} \geq y_{q-1}^{PR} \right\}$$

- L'erreur moyenne de l'étalonnage seuillé depuis 1997 est de 18%.

- Définissons un étalonnage utilisé pour prévoir le PIB en niveau :

$$\widehat{y}_q^{PR} = \widehat{\beta}_0 + \widehat{\beta}_1 y_{q-1}^{PR} + \widehat{\beta}_2 F_q + \widehat{\beta}_3 \Delta F_q | \Delta F_q |$$

- La prévision en **profil** est alors donnée par :

$$\widehat{\varepsilon}_q := 1 \left\{ \widehat{y}_q^{PR} \geq y_{q-1}^{PR} \right\}$$

- L'erreur moyenne de l'étalonnage seuillé depuis 1997 est de 18%.

# CLASSIFICATION VERSUS RÉGRESSION

- Formuler le problème des profils comme un problème de **classification** permet de tester un large éventail de méthodes.
- Ces méthodes sont spécifiques au problème de classification : LDA, QDA, Probit, SVM, Arbre de classification,...
- Ces méthodes peuvent en théorie surpasser la régression.

# CLASSIFICATION VERSUS RÉGRESSION

- Formuler le problème des profils comme un problème de **classification** permet de tester un large éventail de méthodes.
- Ces méthodes sont spécifiques au problème de classification : **LDA, QDA, Probit, SVM, Arbre de classification,...**
- Ces méthodes peuvent en théorie surpasser la régression.



# CLASSIFICATION VERSUS RÉGRESSION

- Formuler le problème des profils comme un problème de **classification** permet de tester un large éventail de méthodes.
- Ces méthodes sont spécifiques au problème de classification : **LDA, QDA, Probit, SVM, Arbre de classification,...**
- Ces méthodes peuvent en théorie surpasser la régression.

# LINEAR DISCRIMINANT ANALYSIS

- Avec 2 classes, LDA est équivalente à une classification par régression linéaire
- Estimation le modèle  $\varepsilon_q = \beta_0 + \beta_1 y_{q-1}^{PR} + \beta_2 F_q + \beta_3 \Delta F_q |\Delta F_q| + \xi_q$ ,
- puis classement selon la position de  $\hat{\varepsilon}_q$  par rapport à 0,5.
- L'erreur moyenne LDA depuis 1997 est de 12%.

# LINEAR DISCRIMINANT ANALYSIS

- Avec 2 classes, LDA est équivalente à une classification par régression linéaire
- Estimation le modèle  $\varepsilon_q = \beta_0 + \beta_1 y_{q-1}^{PR} + \beta_2 F_q + \beta_3 \Delta F_q |\Delta F_q| + \xi_q$ ,
- puis classement selon la position de  $\hat{\varepsilon}_q$  par rapport à 0,5.
- L'erreur moyenne LDA depuis 1997 est de 12%.

# LINEAR DISCRIMINANT ANALYSIS

- Avec 2 classes, LDA est équivalente à une classification par régression linéaire
- Estimation le modèle  $\varepsilon_q = \beta_0 + \beta_1 y_{q-1}^{PR} + \beta_2 F_q + \beta_3 \Delta F_q |\Delta F_q| + \xi_q$ ,
- puis classement selon la position de  $\hat{\varepsilon}_q$  par rapport à 0,5.
- L'erreur moyenne LDA depuis 1997 est de 12%.

## DÉFINITION

Nous définissons le score d'une stratégie  $\phi_q$  par

$$p_q := \mathbb{P}(L_q(\phi_q^{\text{random}}) \leq L_q(\phi_q)).$$

Par convention, une stratégie est significative au seuil  $\alpha$  si  $p_q \leq \alpha$ .

## REMARQUE

Si  $p_q \leq \alpha$ , cela signifie qu'un agent non informé a moins de  $\alpha\%$  de chance d'avoir une erreur moyenne plus petite que l'erreur moyenne  $L_q(\phi_q)$  de la stratégie  $\phi_q$ .

# TABLEAU RÉCAPITULATIF

Stratégie	Erreur depuis 1997	score
Hasard	0,50	-
Dernière valeur (opposée)	0,36	$1,8 \times 10^{-2}$
Moyenne de long terme (opposée)	0,39	$5,0 \times 10^{-2}$
Markov	0,36	$1,8 \times 10^{-2}$
Régression linéaire (facteur France...)	0,18	$8,4 \times 10^{-7}$
Régression linéaire (variables industrie)	0,18	$8,4 \times 10^{-7}$
Probit (facteur France...)	0,14	$3,6 \times 10^{-8}$
Probit (variables industrie)	0,16	$1,8 \times 10^{-7}$
LDA (facteur France...)	0,12	$6,5 \times 10^{-9}$
QDA (facteur France...)	0,14	$3,6 \times 10^{-8}$
RPART (facteur France...)	0,25	$9,1 \times 10^{-5}$
SVM ((facteur France...)	0,16	$1,8 \times 10^{-7}$

# Plan

- 1 INTRODUCTION
- 2 NOTATIONS, DÉFINITIONS
- 3 STRATÉGIES BENCHMARK
- 4 PRÉDICTION SUR LES FUTURES ERREURS ?
- 5 INDICATEUR DE PROFIL

# PRÉVISION DU TAUX D'ERREUR

- Taux d'erreur de la prévision LDA sur  $Q = 8$  prochaines prévisions (2 ans) :  $[0; 0, 5]$  (intervalle de confiance à 90%).
- Taux d'erreur de la prévision LDA sur  $Q = 60$  (15 ans) prochaines prévisions :  $[0; 0, 33]$  (intervalle de confiance à 90%).
- pour agent non informé, intervalle de confiance à 90% pour le taux d'erreur de la prévision sur  $Q = 8$  (2 ans) :  $[0, 2; 0, 8]$ .

## REMARQUE

*Ces intervalles de confiance prennent en compte à la fois l'erreur de prévision mais aussi l'erreur d'estimation sur l'échantillon historique.*



# PRÉVISION DU TAUX D'ERREUR

- Taux d'erreur de la prévision LDA sur  $Q = 8$  prochaines prévisions (2 ans) :  $[0; 0, 5]$  (intervalle de confiance à 90% ).
- Taux d'erreur de la prévision LDA sur  $Q = 60$  (15 ans) prochaines prévisions :  $[0; 0, 33]$  (intervalle de confiance à 90%).
- pour agent non informé, intervalle de confiance à 90% pour le taux d'erreur de la prévision sur  $Q = 8$  (2 ans) :  $[0, 2; 0, 8]$ .

## REMARQUE

*Ces intervalles de confiance prennent en compte à la fois l'erreur de prévision mais aussi l'erreur d'estimation sur l'échantillon historique.*

# PRÉVISION DU TAUX D'ERREUR

- Taux d'erreur de la prévision LDA sur  $Q = 8$  prochaines prévisions (2 ans) :  $[0; 0, 5]$  (intervalle de confiance à 90%).
- Taux d'erreur de la prévision LDA sur  $Q = 60$  (15 ans) prochaines prévisions :  $[0; 0, 33]$  (intervalle de confiance à 90%).
- pour agent non informé, intervalle de confiance à 90% pour le taux d'erreur de la prévision sur  $Q = 8$  (2 ans) :  $[0, 2; 0, 8]$ .

## REMARQUE

*Ces intervalles de confiance prennent en compte à la fois l'erreur de prévision mais aussi l'erreur d'estimation sur l'échantillon historique.*

# PRÉVISION DU TAUX D'ERREUR

- Taux d'erreur de la prévision LDA sur  $Q = 8$  prochaines prévisions (2 ans) :  $[0; 0, 5]$  (intervalle de confiance à 90%).
- Taux d'erreur de la prévision LDA sur  $Q = 60$  (15 ans) prochaines prévisions :  $[0; 0, 33]$  (intervalle de confiance à 90%).
- pour agent non informé, intervalle de confiance à 90% pour le taux d'erreur de la prévision sur  $Q = 8$  (2 ans) :  $[0, 2; 0, 8]$ .

## REMARQUE

*Ces intervalles de confiance prennent en compte à la fois l'erreur de prévision mais aussi l'erreur d'estimation sur l'échantillon historique.*

- en moyenne, le taux d'erreur des meilleures méthodes se situent entre 12% et 20% : **approche inconditionnelle**.
- **Question** : pour un trimestre donné  $q$ , est-il possible d'associer à notre prévision en profil un indicateur de risque **conditionnellement à l'information disponible** (pour nous, les enquêtes de conjoncture) ?
- **Intérêt** : possibilité de donner un scénario sur LE prochain trimestre et non les  $Q = 8$  prochains trimestres.

- en moyenne, le taux d'erreur des meilleures méthodes se situent entre 12% et 20% : **approche inconditionnelle**.
- **Question** : pour un trimestre donné  $q$ , est-il possible d'associer à notre prévision en profil un indicateur de risque **conditionnellement à l'information disponible** (pour nous, les enquêtes de conjoncture) ?
- **Intérêt** : possibilité de donner un scénario sur LE prochain trimestre et non les  $Q = 8$  prochains trimestres.

- en moyenne, le taux d'erreur des meilleures méthodes se situent entre 12% et 20% : **approche inconditionnelle**.
- **Question** : pour un trimestre donné  $q$ , est-il possible d'associer à notre prévision en profil un indicateur de risque **conditionnellement à l'information disponible** (pour nous, les enquêtes de conjoncture) ?
- **Intérêt** : possibilité de donner un scénario sur LE prochain trimestre et non les  $Q = 8$  prochains trimestres.

# Plan

- 1 INTRODUCTION
- 2 NOTATIONS, DÉFINITIONS
- 3 STRATÉGIES BENCHMARK
- 4 PRÉDICTION SUR LES FUTURES ERREURS ?
- 5 INDICATEUR DE PROFIL

Nous construisons un indicateur de profil  $I_q$  en **temps réel**, à l'aide des **enquêtes** tel que :

- ces valeurs sont comprises entre  $-1$  et  $+1$ .
- Proche de  $+1$  (respectivement  $-1$ ), il indique de fortes chances d'**accélération** (respectivement de **décélération**).
- Entre  $-0,5$  et  $+0,5$ , il souligne une **zone d'incertitude**.



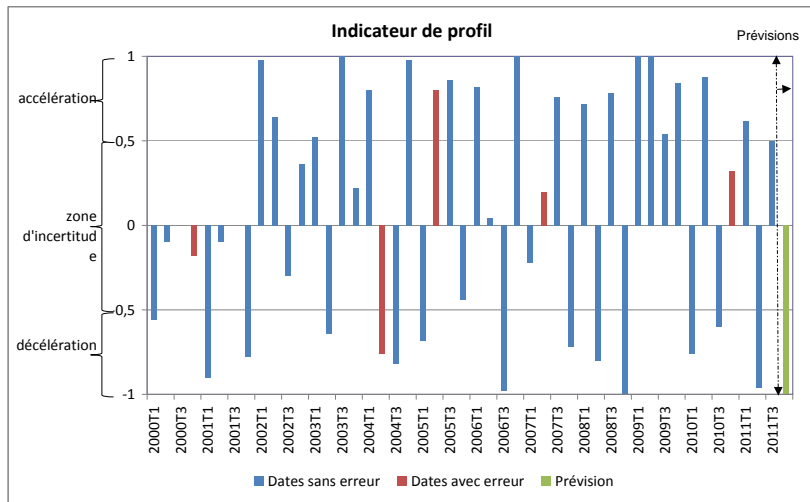
Nous construisons un indicateur de profil  $I_q$  en **temps réel**, à l'aide des **enquêtes** tel que :

- ces valeurs sont comprises entre  $-1$  et  $+1$ .
- Proche de  $+1$  (respectivement  $-1$ ), il indique de fortes chances d'**accélération** (respectivement de **décélération**).
- Entre  $-0,5$  et  $+0,5$ , il souligne une **zone d'incertitude**.

Nous construisons un indicateur de profil  $I_q$  en **temps réel**, à l'aide des **enquêtes** tel que :

- ces valeurs sont comprises entre  $-1$  et  $+1$ .
- Proche de  $+1$  (respectivement  $-1$ ), il indique de fortes chances d'**accélération** (respectivement de **décélération**).
- Entre  $-0,5$  et  $+0,5$ , il souligne une **zone d'incertitude**.

# INDICATEUR DE PROFIL (GRAPHIQUE)



Sur l'échantillon historique (en out-sample) :

- l'indicateur se situe en dehors de la zone d'incertitude plus de 60% du temps.
- quand l'indicateur est en dehors de la zone d'incertitude, le taux de réussite est supérieur à 96%.

## REMARQUE

*Ne pas confondre avec l'indicateur de profil (problématique de court terme sur le trimestre coïncident) avec l'indicateur de retournement (moyen terme sur l'output gap).*

# RÉSULTATS NUMÉRIQUES

Sur l'échantillon historique (en out-sample) :

- l'indicateur se situe en dehors de la zone d'incertitude plus de 60% du temps.
- quand l'indicateur est en dehors de la zone d'incertitude, le taux de réussite est supérieur à 96%.

## REMARQUE

*Ne pas confondre avec l'indicateur de profil (problématique de court terme sur le trimestre coïncident) avec l'indicateur de retournement (moyen terme sur l'output gap).*

# RÉSULTATS NUMÉRIQUES

Sur l'échantillon historique (en out-sample) :

- l'indicateur se situe en dehors de la zone d'incertitude plus de 60% du temps.
- quand l'indicateur est en dehors de la zone d'incertitude, le taux de réussite est supérieur à 96%.

## REMARQUE

*Ne pas confondre avec l'indicateur de profil (problématique de **court terme** sur le trimestre coïncident) avec l'indicateur de retournement (**moyen terme** sur l'output gap).*

- Nouvelle problématique : **prévoir les profils.**
- Méthodes économétriques avec enquêtes avec taux de réussite entre 80% et 90%
- Taux de réussite moyen de LDA pour les 8 prochains trimestres : [50%, 100%] (intervalle à 90%), bien meilleur que celui d'agent non informé [20%, 80%]
- Construction d'un **indicateur de profil** permettant le scénario en profil pour un trimestre donné.

# CONCLUSIONS-SYNTHESE

- Nouvelle problématique : prévoir les profils.
- Méthodes économétriques avec enquêtes avec taux de réussite entre 80% et 90%
- Taux de réussite moyen de LDA pour les 8 prochains trimestres : [50%, 100%] (intervalle à 90%), bien meilleur que celui d'agent non informé [20%, 80%]
- Construction d'un indicateur de profil permettant le scénario en profil pour un trimestre donné.



# CONCLUSIONS-SYNTHESE

- Nouvelle problématique : prévoir les profils.
- Méthodes économétriques avec enquêtes avec taux de réussite entre 80% et 90%
- Taux de réussite moyen de LDA pour les 8 prochains trimestres : [50%, 100%] (intervalle à 90%), bien meilleur que celui d'agent non informé [20%, 80%]
- Construction d'un indicateur de profil permettant le scénario en profil pour un trimestre donné.

# CONCLUSIONS-SYNTHESE

- Nouvelle problématique : prévoir les profils.
- Méthodes économétriques avec enquêtes avec taux de réussite entre 80% et 90%
- Taux de réussite moyen de LDA pour les 8 prochains trimestres : [50%, 100%] (intervalle à 90%), bien meilleur que celui d'agent non informé [20%, 80%]
- Construction d'un indicateur de profil permettant le scénario en profil pour un trimestre donné.

# Plan

- 1 INTRODUCTION
- 2 NOTATIONS, DÉFINITIONS
- 3 STRATÉGIES BENCHMARK
- 4 PRÉDICTION SUR LES FUTURES ERREURS ?
- 5 INDICATEUR DE PROFIL

# TEST INDÉPENDANCE DES ERREURS

## PROPOSITION

Sous l'hypothèse markovienne d'ordre 1, un test asymptotique de niveau  $\alpha$  pour l'hypothèse

$$H_0 := \{\text{Indépendance des erreurs}\} = \{\text{Cov}(\varepsilon_t, \varepsilon_{t-1}) = 0\}$$

$$\left\{ \left| \frac{\sqrt{Q} \widehat{\text{Cov}}(\varepsilon_t, \varepsilon_{t-1})}{\sqrt{\hat{\lambda}' \hat{\Sigma} \hat{\lambda}}} \right| \geq q_{1-\alpha/2} \right\}$$

$$\text{avec } Z_i := \begin{pmatrix} \varepsilon_{i-1} \varepsilon_i \\ \varepsilon_i \\ \varepsilon_{i-1} \end{pmatrix}, \hat{\lambda} := (1, \hat{E}(Z_t^3), \hat{E}(Z_t^2))' \text{ et}$$

$$\hat{\Sigma} := \widehat{\text{Var}}(Z_1) + 2[\widehat{\text{cov}}(Z_1, Z_2) + \widehat{\text{cov}}(Z_1, Z_3)]$$

# INTERVALLE ASYMPTOTIQUE DE L'ERREUR DE PRÉVISION

## PROPOSITION

*Sous l'hypothèse d'indépendance des erreurs, nous obtenons un intervalle de confiance asymptotique pour le taux d'erreur moyen de prévision*

$$\underbrace{\frac{1}{Q} \sum_{i=1}^Q \varepsilon_i}_{\text{Moyenne des futures erreurs}} \in \hat{p}_1 \pm \hat{\sigma} q_{1-\alpha/4} \left[ \underbrace{\frac{1}{\sqrt{Q}}}_{\text{Erreur Prévision}} + \underbrace{\frac{1}{\sqrt{N}}}_{\text{Erreur Estimation}} \right]$$

*avec N le nombre de trimestres pour l'estimation, et Q pour la prévision.*

# INTERVALLE NON-ASYMPTOTIQUE DE L'ERREUR DE PRÉVISION

## PROPOSITION

*Sous l'hypothèse d'indépendance des erreurs, nous obtenons un intervalle de confiance non-asymptotique pour le taux d'erreur moyen de prévision avec grande probabilité au moins  $1 - \delta$  :*

$$\left| \frac{1}{Q} \sum_q \varepsilon_q - \hat{\rho}_1 \right| \leq \underbrace{\hat{\sigma}_N \sqrt{2 \ln(8/\delta)} \left( \frac{1}{\sqrt{Q}} + \frac{1}{\sqrt{N}} \right)}_{\text{Erreur asymptotique}} + \underbrace{\frac{7 \ln(8/\delta)}{3(N-1)} + \frac{\ln(8/\delta)}{3Q} + \frac{2 \ln(8/\delta)}{\sqrt{Q(N-1)}}}_{\text{Erreur approximation}}$$

avec  $N$  le nombre de trimestres pour l'estimation, et  $Q$  pour la prévision.

## DÉFINITION

*L'indicateur de profil est défini par :*

$$I_q := 2(\hat{\mathbb{P}}_q(\varepsilon_q = +1) - 1/2)$$

*avec  $\hat{\mathbb{P}}_q$  la probabilité conditionnelle aux enquêtes de conjoncture au trimestre  $q$ .*