NOWCASTING DE L'INFLATION

Feuille de route d'un projet de recherche quantitative

Equipe BlaCS-Scholes

1 INTRODUCTION

L'objectif de ce projet est de construire un système capable d'estimer, en continu et avant leur publication officielle, le niveau de l'inflation américaine (CPI et PCE), puis de traduire cette estimation en une position de marché cohérente. Ce document en fixe la feuille de route : il part des fondements macroéconomiques nécessaires pour comprendre *pourquoi* l'inflation évolue comme elle le fait, avant d'aborder le cadre mathématique permettant de l'estimer en temps réel, puis la manière de transformer cette estimation en signal exploitable sur les marchés de taux.

Le parti pris de ce document est pédagogique : chaque notion économique un peu avancée est introduite avant d'être utilisée, de sorte qu'un lecteur ayant une solide formation quantitative mais peu d'exposition à la macroéconomie puisse suivre l'ensemble du raisonnement sans prérequis caché.

2 FONDEMENTS MACROÉCONOMIQUES

2.1 Ce que mesure réellement l'inflation

L'inflation n'est pas un nombre unique mais le résultat d'un choix de construction statistique. Le *Consumer Price Index* (CPI) et le *Personal Consumption Expenditures index* (PCE) — ce dernier étant la mesure privilégiée par la Réserve fédérale — sont tous deux des moyennes pondérées de milliers de prix individuels, mais ils diffèrent par leur panier de biens, leur méthode de pondération et leur traitement de la substitution entre biens. Il est également essentiel de distinguer :

— l'inflation *headline*, qui inclut l'ensemble des biens et services, y compris l'énergie et l'alimentation, réputées volatiles ;

— l'inflation *core*, qui les exclut pour isoler une tendance de fond ;

— les variations mensuelles (*month-over-month*), pertinentes pour le nowcasting à court terme, et les variations en glissement annuel (*year-over-year*), plus lisibles mais qui intègrent des effets de base parfois trompeurs.

Un projet de nowcasting sérieux doit choisir explicitement sur quel objet il travaille : prévoir la variation mensuelle du CPI headline n'a ni les mêmes déterminants, ni le même horizon utile pour le trading, que prévoir l'inflation *core* en glissement annuel.

2.2 La courbe de Phillips : une grammaire des causes

Le cadre théorique standard pour comprendre la formation de l'inflation est la courbe de Phillips néo-keynésienne, dans sa version dite « hybride » :

$$\pi_t = \gamma_f \, \mathbb{E}_t[\pi_{t+1}] + \gamma_b \, \pi_{t-1} + \kappa \, mc_t + \varepsilon_t \tag{1}$$

Chaque terme correspond à un mécanisme économique distinct :

— $\mathbb{E}_t[\pi_{t+1}]$: les entreprises fixent leurs prix en anticipant les coûts futurs, pas seulement les coûts présents. Si tout le monde anticipe une inflation élevée, chacun ajuste ses prix en conséquence — l'inflation devient en partie auto-réalisatrice ;

— π_{t-1} : une composante purement inertielle, liée à des rigidités contractuelles (contrats, indexations, habitudes de fixation de prix) ;

— mc_t : le coût marginal réel de production, dont le meilleur indicateur empirique est le coût salarial unitaire (croissance des salaires ajustée de la productivité) ;

— ε_t : les chocs d'offre exogènes — énergie, matières premières, ruptures d'approvisionnement — qui n'ont rien à voir avec la demande intérieure.

L'implication pratique est directe : un modèle qui tente de prévoir l'inflation agrégée sans distinguer ces canaux mélange des dynamiques à fréquences et à horizons différents, ce qui dégrade à la fois la précision et l'interprétabilité économique du résultat.

2.3 Les canaux de transmission, et leurs difficultés propres

Trois notions supplémentaires méritent d'être précisées, car elles sont chacune source de complexité réelle et non de simple bruit statistique.

L'écart de production (*output gap*). Il mesure la tension entre la production effective et la production potentielle de l'économie — une notion de « capacité maximale sans tension inflationniste ». Le problème est que cette capacité potentielle n'est pas observable : elle doit être estimée, et les méthodes usuelles (filtres statistiques comme le filtre de Hodrick-Prescott) souffrent d'un biais de bord bien connu, qui rend l'estimation en temps réel nettement moins fiable que l'estimation rétrospective. C'est une limite structurelle qu'il faut assumer plutôt que masquer.

Les anticipations et leur ancrage. On distingue les anticipations adaptatives (extrapoler le passé récent), rationnelles (utiliser tout le modèle disponible) et « ancrées » (rester proches d'une cible annoncée par la banque centrale tant que sa crédibilité n'est pas remise en cause). Le désancrage des anticipations — lorsque les agents cessent de croire à la cible d'inflation — est précisément le mécanisme qui a transformé des chocs pétroliers ponctuels en inflation persistante dans les années 1970.

Les chocs de coûts. Contrairement aux deux canaux précédents, ils sont exogènes à la dynamique interne de l'économie et se transmettent avec des délais et des élasticités qui varient fortement selon la composante de l'indice concernée — ce qui justifie l'approche par décomposition présentée dans la section suivante.

3 Une architecture par composantes structurelles

3.1 Pourquoi décomposer plutôt que modéliser l'agrégat

Le CPI et le PCE sont des sommes pondérées de composantes publiées séparément, avec des poids connus à l'avance (mis à jour annuellement, donc sans incertitude au moment de la prévision) :

$$\pi_t^{\text{CPI}} = \sum_i w_i \, \pi_{i,t} \tag{2}$$

Puisque chaque composante i obéit à un mécanisme économique différent, la démarche la plus rigoureuse consiste à construire un modèle dédié par composante — chacun avec ses propres indicateurs avancés — puis à ré-agréger via les poids officiels. Quatre blocs concentrent l'essentiel de la dynamique.

3.2 Le logement : le bloc dominant et son mécanisme de retard

Le logement (*shelter*) représente à lui seul environ un tiers de l'indice headline et plus de 40 % de l'indice core — c'est de loin le bloc le plus important, et paradoxalement le plus prévisible, pour une raison structurelle rarement enseignée en dehors des cercles spécialisés.

Le CPI ne mesure pas les loyers de marché (ceux payés par un nouveau locataire), mais une moyenne de tous les loyers en cours, y compris ceux des locataires en place. Trois mécanismes créent un décalage mesurable et quasi mécanique entre le loyer de marché observé aujourd'hui et sa traduction dans l'indice :

1. la majorité des baux sont signés pour douze mois, de sorte qu'un loyer ne peut s'ajuster au marché qu'au moment du renouvellement ;
2. lors du renouvellement, les propriétaires lissent en général la hausse plutôt que de répercuter intégralement et immédiatement le prix de marché ;
3. la construction même de l'indice compare le loyer du mois courant à celui de six mois auparavant, ce qui ajoute un délai supplémentaire à la mesure.

Des travaux de recherche récents ont formalisé ce mécanisme et montré qu'il reproduit fidèlement la trajectoire observée du CPI logement à partir des seules données de loyers de marché sur la dernière décennie. La conséquence pour un projet de nowcasting est importante : connaître aujourd'hui les loyers de marché revient, avec un délai d'environ douze à dix-huit mois, à *déjà connaître* une bonne part de la trajectoire future du CPI logement. C'est l'un des rares cas en macroéconomie où une prévision se rapproche d'une quasi-certitude mécanique plutôt que d'un pari probabiliste — et donc l'un des rares cas où un edge de marché peut être défendu avec un haut degré de confiance.

3.3 L'énergie : une transmission rapide et volatile

À l'opposé du logement, la composante énergie se transmet des prix spot (pétrole, essence) à l'indice en quelques semaines, avec une élasticité relativement stable mais une volatilité intrinsèque élevée. Elle est facile à « nowcaster » au sens statistique, mais peu porteuse d'edge économique, car l'information est déjà largement intégrée par le marché au moment où elle est disponible.

3.4 *Les biens durables : la chaîne de gros vers le détail*

Les prix de biens comme les véhicules d'occasion suivent avec un décalage de quelques semaines les indices de prix de gros, reflétant une dynamique classique de chaîne d'approvisionnement et de gestion des stocks.

3.5 *Les services « core » hors logement : le canal salarial*

C'est la composante que les banques centrales surveillent le plus attentivement depuis les épisodes inflationnistes récents, car elle est directement liée au terme mc_t de la courbe de Phillips : sa dynamique est tirée par la croissance des salaires plutôt que par des chocs de matières premières, ce qui la rend à la fois plus persistante et plus difficile à faire baisser rapidement par une politique monétaire restrictive.

4 Cadre mathématique du nowcasting

4.1 *Le problème du bord irrégulier des données*

Les indicateurs utiles arrivent à des fréquences et à des dates de publication différentes : certains prix sont quotidiens, les loyers sont mensuels avec retard de publication, les salaires trimestriels. À tout instant, le panel de données disponible est donc incomplet et asynchrone — un problème que la littérature désigne sous le nom de *jagged edge*. La méthode retenue doit pouvoir produire une estimation à jour à chaque nouvelle publication, sans attendre que l'ensemble du panel mensuel soit complet.

4.2 *Modèles à facteurs dynamiques en représentation espace-état*

L'outil standard consiste à supposer qu'un petit nombre de facteurs latents communs f_t explique l'essentiel des co-mouvements observés dans un grand panel de séries y_t :

$$f_t = A f_{t-1} + B\varepsilon_t \qquad \text{(équation d'état)} \tag{3}$$

$$y_t = \Lambda f_t + \xi_t \qquad \text{(équation de mesure)} \tag{4}$$

L'équation d'état décrit la dynamique propre du facteur inobservé ; l'équation de mesure relie ce facteur aux indicateurs effectivement publiés, chacun étant une combinaison linéaire bruitée du facteur commun. Cette représentation a l'avantage de traiter naturellement les données manquantes ou asynchrones : il suffit d'omettre, à une date donnée, les lignes de l'équation de mesure correspondant aux séries non encore publiées.

4.3 *Le filtre de Kalman comme mise à jour bayésienne séquentielle*

Le filtre de Kalman est le mécanisme qui exploite cette représentation : à chaque instant, il combine une croyance préalable sur l'état courant (issue de la dynamique du modèle) avec une nouvelle observation, pondérée selon la fiabilité relative de chacune, pour produire une estimation mise à jour ainsi que son incertitude associée. C'est très exactement ce qui permet de publier un nowcast *quotidien* : chaque nouvelle donnée, aussi partielle soit-elle, vient rafraîchir l'estimation sans qu'il soit nécessaire d'attendre la publication complète du mois.

5 Anticipations de marché et prime de risque

5.1 Le taux d'inflation implicite n'est pas l'anticipation d'inflation

Le marché des obligations indexées sur l'inflation (TIPS aux États-Unis) permet de calculer un « taux d'inflation implicite » (*breakeven*), simplement défini comme l'écart entre le rendement d'une obligation nominale et celui d'une obligation indexée de même maturité. L'interprétation naïve — ce taux serait l'anticipation d'inflation du marché — est incorrecte. En réalité :

$$\text{Breakeven}_t = \mathbb{E}_t[\pi] + \text{prime de risque d'inflation}_t + \text{prime de liquidité}_t \tag{5}$$

La prime de risque existe parce que l'inflation future est incertaine et que les investisseurs exigent une compensation pour porter ce risque ; la prime de liquidité existe parce que le marché des obligations indexées est structurellement moins liquide que celui des obligations nominales. Les deux primes varient dans le temps, parfois fortement, indépendamment de toute évolution des anticipations réelles.

5.2 Isoler l'anticipation « pure »

Les modèles de structure par terme sans arbitrage (dont la classe la plus connue s'appuie sur les travaux d'Adrian, Crump et Moench) permettent de décomposer empiriquement le breakeven observé en ses trois composantes, en exploitant les contraintes de cohérence entre maturités. C'est un exercice économétrique non trivial, mais il est décisif pour la suite du projet.

5.3 Pourquoi cette distinction est le cœur de l'edge

Si le nowcast produit par le système diffère du breakeven observé, cet écart peut venir soit d'un désaccord réel sur l'inflation future, soit d'une simple variation de prime de risque ou de liquidité. Seul le premier cas constitue un signal exploitable ; confondre les deux revient à trader du bruit. La capacité à isoler la composante « anticipation pure » du breakeven est ce qui distingue un edge économique authentique d'une corrélation statistique fragile.

6 De la prévision à la stratégie de marché

La nature de l'edge recherché doit être précisée : il ne s'agit pas de « prévoir mieux » au sens statistique — le marché intègre déjà une grande partie de l'information publique — mais d'exploiter des faits mécaniques que le marché a tendance à sous-pondérer, au premier rang desquels la convergence quasi certaine du logement décrite en section 3.2. Les instruments naturels sont les breakevens, les swaps d'inflation et les contrats à terme de taux courts. Le dimensionnement des positions doit refléter la nature de la confiance : un signal mécanique et quasi déterministe (logement) justifie une taille de position plus importante qu'un signal fondé sur un choc incertain (énergie, chaîne d'approvisionnement). Enfin, la capacité du projet est bornée par la liquidité réelle du marché des instruments indexés sur l'inflation, sensiblement plus étroit que celui des taux nominaux.

7 Validation scientifique et gestion du risque de modèle

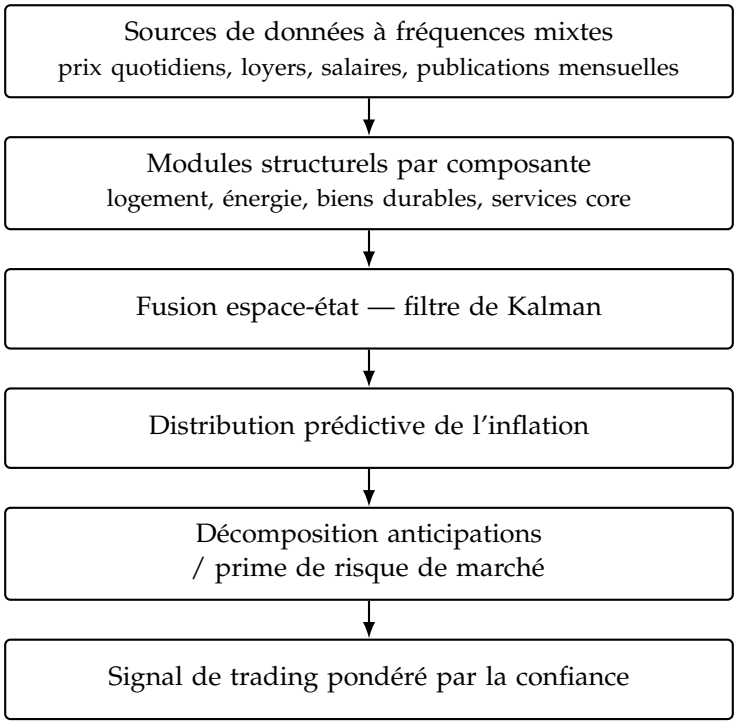
7.1 Le piège des données révisées

Les chiffres officiels de CPI et de PCE sont révisés après leur première publication. Un backtest qui utiliserait les données *actuelles* pour évaluer un modèle censé produire une prévision *en temps réel* introduit un biais de anticipation (*look-ahead bias*) qui invalide silencieusement la plupart des résultats obtenus de cette manière. La validation doit impérativement s'appuyer sur des bases de données conservant l'historique des publications telles qu'elles étaient connues à chaque date passée (« vintages »).

7.2 Comparer rigoureusement des prévisions

Comparer deux modèles sur la seule base d'une erreur quadratique moyenne ne suffit pas à établir qu'une différence de performance est statistiquement significative. Les tests dédiés à la comparaison de prévisions (Diebold-Mariano, et sa généralisation par Giacomini et White pour des modèles ré-estimés dans le temps) permettent de trancher formellement si un modèle domine réellement un autre, ou si l'écart observé n'est que du bruit d'échantillonnage.

8 Feuille de route



Phase	Objectif	Jalon de validation
0	Infrastructure de données	Panel reconstitué en « vintages » réelles, sans aucune donnée révisée après coup
1	Modèle de référence agrégé	Modèle à facteurs dynamiques sur l'indice headline, comparé aux nowcasts publics existants
2	Modules structurels	Un modèle dédié par composante, validé indicateur par indicateur avant intégration
3	Décomposition attentes / prime	Séparation empirique du breakeven en anticipation pure et prime de risque
4	Fusion bayésienne	Distribution prédictive complète, pondération dynamique des modules selon leur performance récente
5	Validation rigoureuse	Tests de comparaison de prévisions sur historique reconstitué en temps réel
6	Stratégie de marché	Construction et simulation d'une position sur instruments d'inflation, coûts de transaction inclus

9 Conclusion

La difficulté de ce projet n'est pas de produire un chiffre plus précis que les acteurs déjà établis, mais de comprendre suffisamment finement les mécanismes économiques sous-jacents à chaque composante de l'inflation pour identifier où se trouve une information réellement nouvelle — et de ne trader que cette information-là, en écartant tout ce qui relève de la prime de risque ou du bruit statistique. C'est cette discipline, plus que la sophistication du modèle, qui distingue un projet académique d'un projet transposable en gestion réelle.