



Master 1 DAC

UE DALAS

Et si nos décisions étaient guidées par les taux d'intérêt ?

Elio Bteich
Abderraouf Touzene

Mai 2025

Table des matières

1	Introduction générale	3
2	Problématique	3
3	Objectifs du rapport	3
4	Méthodologie générale	3
5	Données utilisées	4
5.1	Séries utilisées	4
5.2	Collecte de Données Complémentaires par Web Scraping	4
5.3	Prétraitement des données	5
6	Immobilier	6
6.1	Analyse exploratoire (EDA)	7
6.1.1	Visualisation des variables clés	7
6.1.2	Corrélations entre variables immobilières et taux d'intérêt	7
6.1.3	Taux d'intérêt et prix immobiliers : un regard différé	11
6.2	Modélisation prédictive : prévoir les prix immobiliers à 3 ans	12
6.2.1	Contexte et cadre prédictif	12
6.2.2	Spécifications OLS évaluées	13
6.2.3	Métriques d'évaluation	13
6.2.4	Bilan quantitatif	13
6.2.5	Analyse visuelle	14
6.2.6	Conclusion et limites.	16
7	Emploi	16
7.1	Analyse exploratoire (EDA)	17
7.1.1	Visualisation des variables clés	17
7.1.2	Corrélations entre variables d'emploi et taux d'intérêt	17
7.1.3	Corrélations différées (<i>lags</i>)	19
7.1.4	Taux d'intérêt et chômage : un regard différé	20
7.2	Modélisation prédictive : prévoir le chômage à 12 mois	20
7.2.1	Cadre prédictif	20
7.2.2	Résultats OLS	21
7.2.3	Validation hors échantillon	21
7.2.4	Conclusion et limites	22
8	Consommation	22
8.1	Analyse Exploratoire des Données (EDA)	22
8.1.1	Préparation et Prétraitement des Données	22
8.1.2	Analyse Exploratoire des Données (Données Nominale)	23
8.1.3	Analyse des Décalages Temporels (Données Nominale)	24
8.1.4	Tests de Causalité de Granger (Données Nominale Différenciées)	25
8.1.5	Ajustement des Ventes en Termes Réels	25

8.1.6	Analyse sur les Données Réelles	26
8.2	Modèles prédictifs	27
8.2.1	Modélisation par Régression Linéaire Simple (Données Réelles)	27
8.3	Résultats et discussion	27
9	Conclusion	28

1 Introduction générale

Ces dernières années, l'économie mondiale a beaucoup changé. Depuis 2022, les prix ont fortement augmenté dans de nombreux pays, ce qu'on appelle l'inflation. Pour essayer de la ralentir, les banques centrales, comme la Réserve fédérale américaine, ont relevé leurs taux d'intérêt – après une longue période où ces taux étaient très bas.

La Réserve fédérale américaine, ou Fed, est la banque centrale des États-Unis. Elle est responsable de la régulation de la politique monétaire du pays, et l'un de ses principaux objectifs est de maintenir la stabilité des prix tout en soutenant la croissance économique. Pour ce faire, la Fed ajuste les taux d'intérêt afin de maîtriser l'inflation et stimuler ou freiner l'économie en fonction des besoins.

Mais ces taux ne sont pas qu'un chiffre technique : ils ont un impact direct sur notre vie de tous les jours. Ils influencent les décisions des ménages (acheter une maison ou non, emprunter ou attendre), des entreprises (investir, embaucher) et même des gouvernements. En bref, quand les taux montent ou baissent, c'est toute l'économie qui réagit.

2 Problématique

À la lumière de ces évolutions économiques, une question essentielle se pose : **Et si nos décisions étaient, bien plus que nous ne le pensons, guidées par les taux d'intérêt ?**

Ce rapport cherche à explorer cette idée, en étudiant de manière rigoureuse comment les taux d'intérêt peuvent peser sur différents aspects de la vie économique aux États-Unis. Nous allons analyser trois domaines clés, chacun soulevant une question concrète :

- **L'emploi** : Quand les taux d'intérêt augmentent, les entreprises empruntent moins. Cela se traduit-il par moins d'embauches, voire plus de chômage ?

En d'autres termes : est-ce que des taux élevés freinent vraiment l'emploi ?

- **La consommation** : Le coût du crédit influence nos décisions d'achat, tout comme les taux sur nos livrets influencent notre envie d'épargner.

Comment ces taux modifient-ils nos choix de consommation et d'épargne ?

- **L'immobilier** : Acheter un logement dépend souvent d'un prêt immobilier. Or, plus les taux sont élevés, plus ce rêve devient difficile à réaliser.

Est-ce que les taux peuvent décider, en quelque sorte, où l'on peut se permettre de vivre ?

3 Objectifs du rapport

L'objectif de ce rapport est double :

- Quantifier, à l'aide d'outils statistiques et économétriques, l'ampleur et la nature des liens entre les taux d'intérêt et les variables économiques clés relatives à l'emploi, la consommation et l'immobilier.
- Utiliser des modèles prédictifs pour simuler des scénarios alternatifs ("What if"), permettant ainsi d'évaluer les conséquences potentielles de différentes trajectoires futures des taux d'intérêt sur ces mêmes axes.

4 Méthodologie générale

Pour mener à bien cette étude, nous adopterons une méthodologie rigoureuse issue de la science des données (data science). Celle-ci comprendra la collecte et la préparation minutieuse de données, une analyse exploratoire

approfondie, l'interprétation des résultats à la lumière de la théorie économique et du contexte actuel, et enfin la construction et la validation de modèles prédictifs . Ce rapport technique détaillera chaque étape de cette démarche.

5 Données utilisées

Notre analyse repose sur plusieurs séries temporelles économiques mensuelles (et hebdomadaires pour certaines), issues de la base FRED (Federal Reserve Economic Data). Ces données couvrent différents volets de l'économie américaine, notamment l'emploi, la consommation, l'épargne et l'immobilier. Les séries ont été importées depuis des fichiers CSV à l'aide d'une fonction `load_fred_csv`, puis transformées et nettoyées pour permettre une analyse rigoureuse.

5.1 Séries utilisées

- **Taux de chômage (UNRATE)** : taux mensuel de chômage (en %) de la population active.
- **Créations d'emplois non-agricoles (PAYEMS)** : indicateur du volume d'emplois dans le secteur privé hors agriculture.
- **Taux d'épargne personnelle (PSAVERT)** : part mensuelle du revenu disponible épargnée par les ménages.
- **Taux directeur effectif (FEDFUNDS)** : taux interbancaire à court terme déterminé par la Réserve fédérale.
- **Taux hypothécaire fixe 30 ans (MORTGAGE30US)** : taux hebdomadaire agrégé par moyenne mensuelle.
- **Indice Case-Shiller des prix immobiliers (CSUSHPINSA)** : mesure mensuelle de l'évolution des prix du logement.
- **Permis de construire (PERMIT)** : nombre mensuel de permis de construire résidentiels.
- **Mises en chantier (HOUST)** : nombre de nouveaux chantiers immobiliers résidentiels démarrés.
- **Ventes de maisons neuves (HSN1FNSA)** : nombre de ventes mensuelles de logements neufs.
- **Indice des Prix à la Consommation (CPI)** : mesure mensuelle de l'évolution moyenne des prix des biens et services consommés par les ménages.
- **Emplois non agricoles (nfp)** : Nombre total d'emplois dans le secteur non agricole aux États-Unis, un indicateur clé de la création d'emplois (variable PAYEMS).
- **Taux directeur effectif (fed_rate)** : Taux d'intérêt effectif des fonds fédéraux, principal outil de politique monétaire de la Fed (variable FEDFUNDS).
- **Taux d'épargne personnelle (personal_savings)** : Taux d'épargne personnelle des ménages américains (variable PSAVERT).

5.2 Collecte de Données Complémentaires par Web Scraping

En complément des données issues de la base FRED, plusieurs scripts Python ont été développés pour collecter automatiquement des données économiques et financières pertinentes à partir de diverses sources web. Cette démarche de web scraping a permis d'enrichir le jeu de données pour l'analyse.

- **Indice des Prix à la Consommation (CPI) - Minneapolis FED :**
 - Un script (`scrape_minneapolisfed.py`) a été utilisé pour extraire les données historiques de l'Indice des Prix à la Consommation (CPI) à partir du site de la Réserve Fédérale de Minneapolis (<https://www.minneapolisfed.org/about-us/monetary-policy/inflation-calculator/consumer-price-index-1913->).
 - Le script parcourt le tableau HTML de la page, récupère les années, les valeurs annuelles moyennes du CPI, et les taux d'inflation annuels.

- [illegible]

Ces scripts permettent d'automatiser la collecte de données actualisées et spécifiques qui ne sont pas toujours disponibles via des API directes ou des fichiers CSV préexistants, enrichissant ainsi la base informationnelle du rapport.

5.3 Prétraitement des données

Plusieurs étapes de transformation ont été appliquées avant modélisation :

- **Alignement temporel et interpolation** : Le taux hypothécaire, disponible à fréquence hebdomadaire, a été converti en série mensuelle par moyenne mensuelle (**resample** sur la date).
- **Création de variables dérivées** : Pour chaque série principale, nous avons calculé l'*évolution mensuelle en pourcentage* à l'aide de la formule suivante :

$$\text{croissance}_t = \left(\frac{X_t - X_{t-1}}{X_{t-1}} \right) \times 100$$

Cela concerne notamment :

- `cs_growth` : croissance mensuelle de l'indice Case-Shiller
- `building_permits_growth`, `new_home_sales_growth`, `housing_starts_growth`
- `fed_rate_growth`, `mortgage_rate_growth`
- `unemployment_rate_growth`, `nfp_growth`, `personal_saving_growth`
- **Nettoyage** : Les premières lignes devenues NaN à la suite du calcul de pourcentages de variation (`pct_change`) ont été supprimées.
- **Standardisation (uniquement pour visualisation)** : Afin de faciliter la lecture des graphiques multi-variés (par exemple les séries en superposition ou certaines heatmaps), les variables ont été temporairement standardisées (centrées réduites) pour éviter que des différences d'échelle n'écrasent les autres séries sur les visualisations.

Remarque : Cette standardisation n'a pas été utilisée pour le calcul des corrélations (les coefficients de corrélation de Pearson sont déjà indépendants des unités) ni pour l'entraînement des modèles prédictifs, afin de préserver l'interprétation économique des valeurs.

Ce prétraitement assure à la fois la cohérence temporelle des séries, la comparabilité des évolutions relatives, et la fiabilité statistique pour les étapes de modélisation.

6 Immobilier

Cette section vise à analyser l'impact des taux d'intérêt sur le marché immobilier, en particulier sur les prix, l'activité de construction et la capacité des ménages à accéder à la propriété. L'objectif est de tester l'hypothèse suivante : *est-ce que les taux peuvent décider, en quelque sorte, où l'on peut se permettre de vivre ?*

Autrement dit, dans quelle mesure les décisions des banques centrales, via les taux directeurs, influencent-elles les conditions de financement immobilier.

Nous nous appuyons pour cela sur l'analyse menée dans le notebook Jupyter `real_estate.ipynb`, couvrant la période allant de janvier 1985 à décembre 2024. Nous examinons à la fois les prix (indice Case-Shiller), les taux (directeurs et hypothécaires), ainsi que plusieurs indicateurs d'activité (permis de construire, ventes de maisons neuves, mises en chantier).

6.1 Analyse exploratoire (EDA)

6.1.1 Visualisation des variables clés

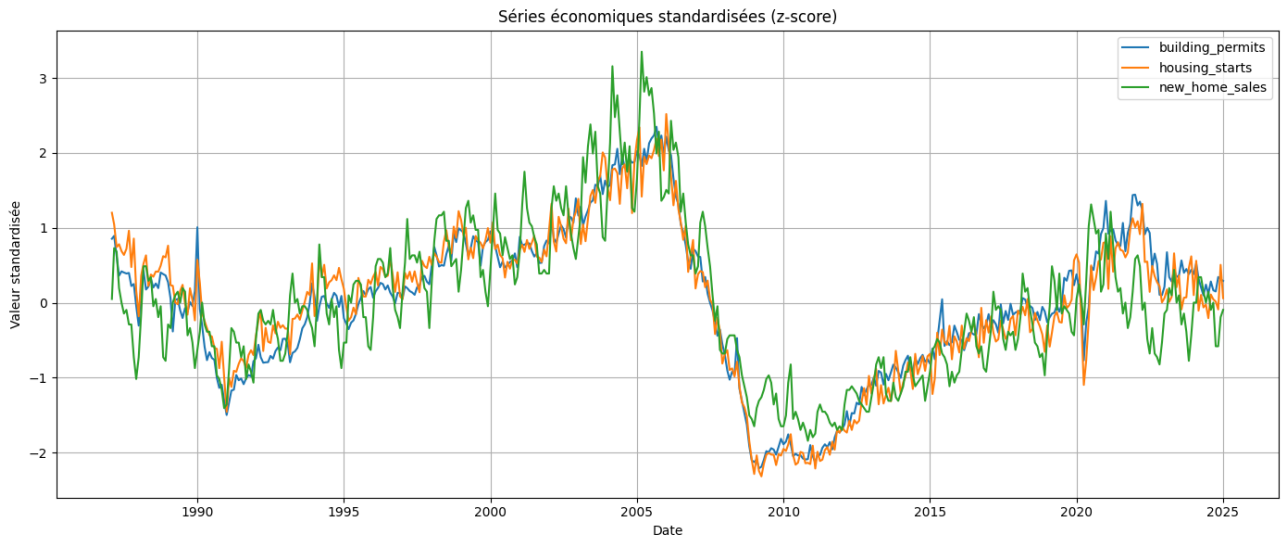


FIGURE 1 – Permis de construire, mises en chantier et ventes de maisons neuves

On remarque que ces trois séries suivent une dynamique très proche. Par souci de lisibilité et pour éviter la redondance dans les visualisations suivantes, nous choisirons de représenter uniquement la série la plus stable : `building_permits`.

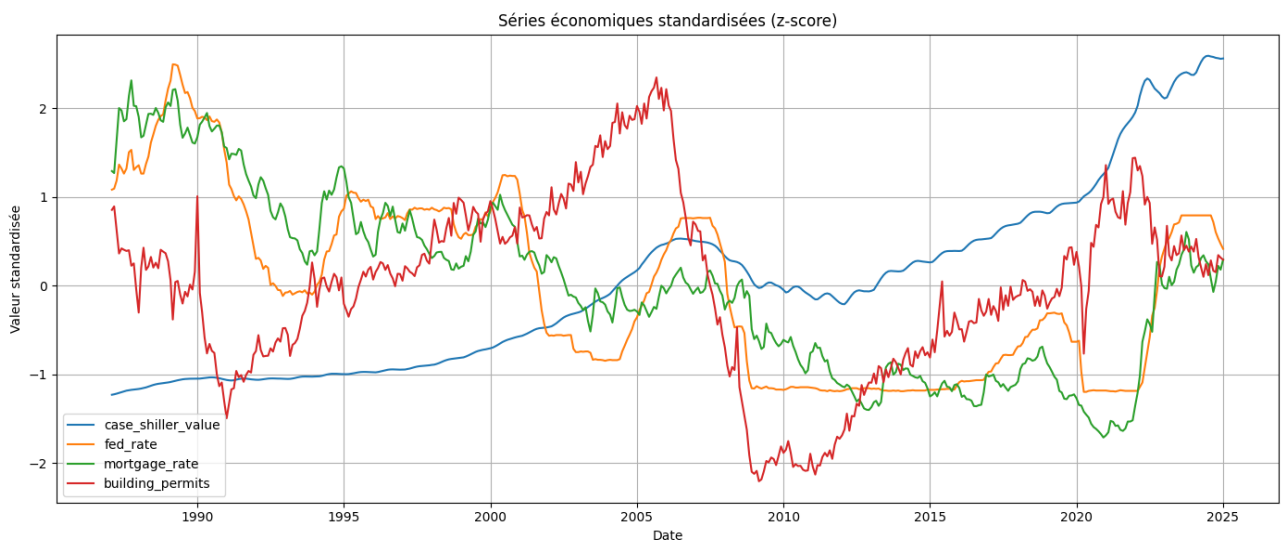


FIGURE 2 – Indice des prix immobiliers, taux directeur, taux hypothécaire et permis de construire

6.1.2 Corrélations entre variables immobilières et taux d'intérêt

Afin de mieux comprendre les relations entre les taux d'intérêt et la dynamique du marché immobilier, nous avons calculé plusieurs matrices de corrélation de Pearson, avec et sans décalage temporel (*lags*). Cette approche permet d'examiner à la fois les relations immédiates et les effets différés des politiques monétaires sur les prix et volumes immobiliers.

1. Corrélations internes au secteur immobilier (sans lag)

La Figure 3 illustre les corrélations contemporaines entre les principales variables immobilières.

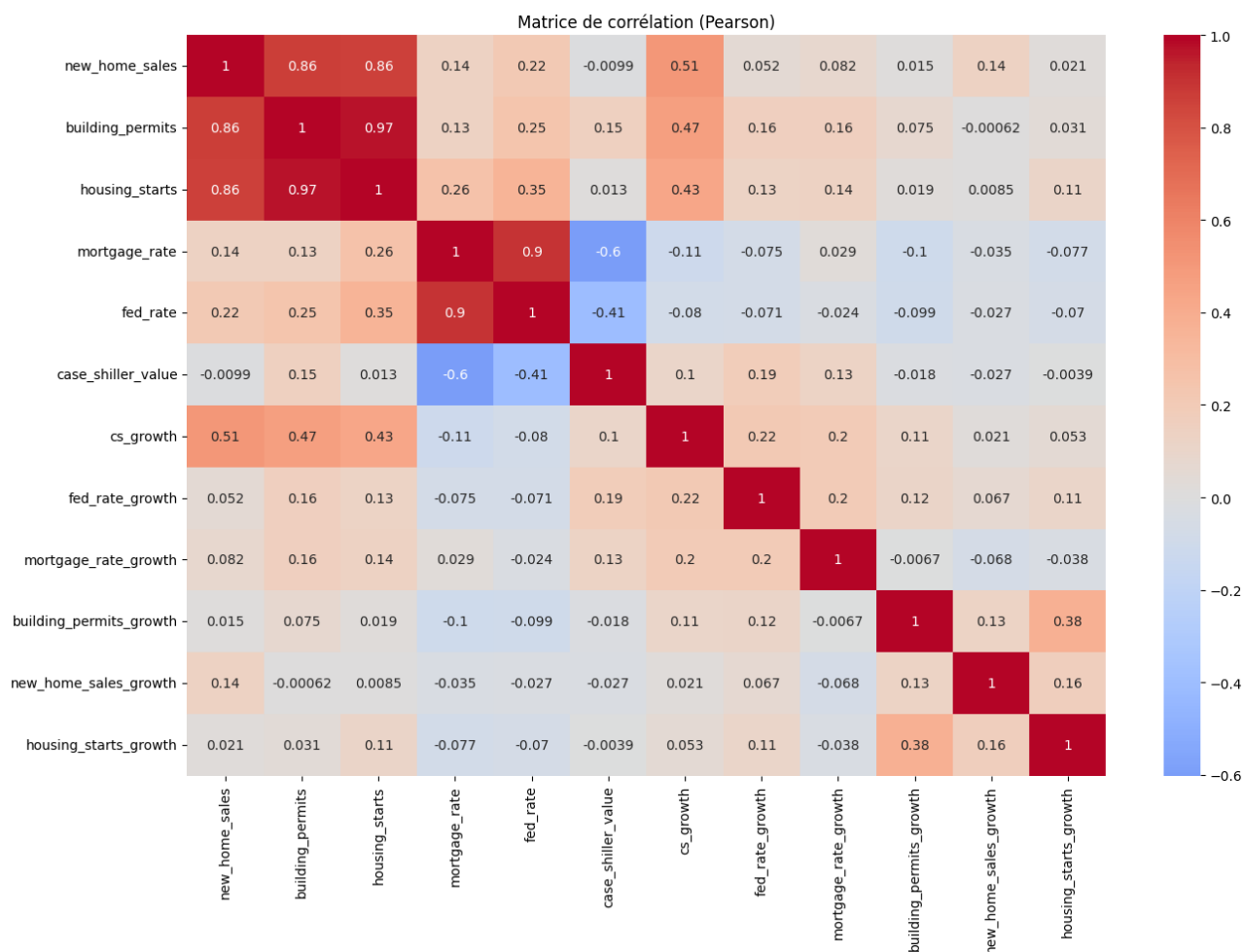


FIGURE 3 – Matrice de corrélation immobilière (sans décalage)

TABLE 1 – Corrélations internes au marché immobilier

Paires de variables	Corrélation (r)	Interprétation
new_home_sales vs building_permits	0.86	Très forte corrélation positive. Cela montre que la vente de maisons neuves suit de près l'activité de construction.
building_permits vs housing_starts	0.97	Les deux variables sont quasiment identiques. Inutile de les garder toutes les deux dans un modèle prédictif.
new_home_sales vs housing_starts	0.86	Relation également très forte. Ces trois indicateurs ont la même dynamique.

Conclusion : Ces variables racontent toutes la même histoire. Pour éviter la redondance et la multicolinéarité dans un modèle, il est recommandé de n'en conserver qu'une seule — par exemple **building_permits**, qui est moins bruyé.

2. Corrélations avec les taux d'intérêt (effets immédiats)

- **mortgage_rate** vs **case_shiller_value** : **-0.60**
→ Quand les taux hypothécaires augmentent, les prix immobiliers baissent. Cela renforce notre hypothèse initiale.
- **fed_rate** vs **case_shiller_value** : **-0.41**
→ Effet similaire, bien que plus faible. L'influence du taux directeur est moins directe que celle du taux hypothécaire.

3. Corrélations avec lags (effets différés)

Afin d'étudier les effets différés des taux d'intérêt sur l'activité et les prix immobiliers, nous avons appliqué plusieurs horizons de décalage temporel (*lags*) sur les variables de taux. Ces lags ont été regroupés selon quatre familles :

- **Short lags** : 3, 6 et 12 mois
- **Medium lags** : 16, 24 et 30 mois
- **Long lags** : 36, 42 et 48 mois
- **Very long lags** : 54, 60 et 66 mois

Pour des raisons de lisibilité et de concision, seules les **corrélations issues des lags longs (36–48 mois)** ont été affichées dans la Figure 4. Ce choix permet de mieux mettre en évidence les effets les plus différés, qui se sont révélés être les plus significatifs.

Par ailleurs, pour ne pas surcharger l'analyse avec l'ensemble des décalages croisés, nous avons choisi de ne considérer que les **corrélations entre variables initiales et versions décalées** (par exemple, **X_lag12** vs **Y**). En effet, une corrélation entre **X_lag3** et **Y_lag6** implique logiquement une corrélation entre **X** et **Y_lag3**. Ce filtrage permet de mieux identifier les relations de cause à effet temporelles sans se perdre dans des matrices trop volumineuses.

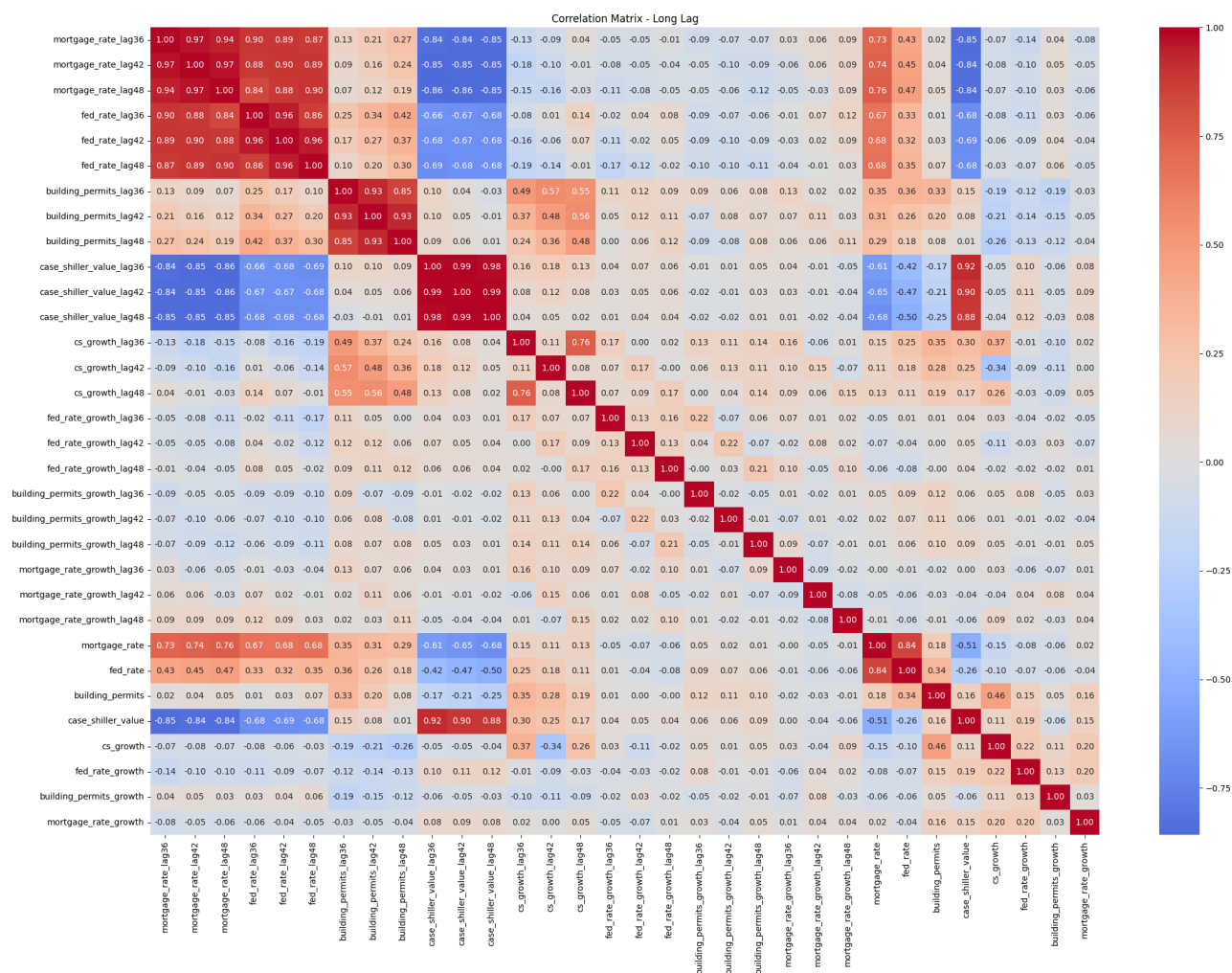


FIGURE 4 – Matrice de corrélation avec décalages (lags) jusqu'à 48 mois

TABLE 2 – Corrélations différées entre taux et prix immobiliers

Paire de variables	r	Interprétation
mortgage_rate_lag12 vs case_shiller_value	-0.70	L'effet des taux devient plus marqué après un an : le coût du crédit agit avec un décalage sur les prix.
mortgage_rate_lag24 vs case_shiller_value	-0.79	Corrélation négative encore plus forte à deux ans.
mortgage_rate_lag36 vs case_shiller_value	-0.85	Effet différé maximal observé.
fed_rate_lag12 vs case_shiller_value	-0.53	L'effet différé du taux directeur est significatif.
fed_rate_lag24 vs case_shiller_value	-0.65	Corrélation plus marquée à deux ans.
fed_rate_lag36 vs case_shiller_value	-0.69	Corrélation forte mais toujours inférieure à celle des taux hypothécaires.

Conclusion : Le marché immobilier ne réagit pas immédiatement aux hausses de taux. L'effet est visible avec un décalage de 12 à 36 mois. Cela confirme le mécanisme économique attendu : hausse des taux → crédit plus cher → baisse de la demande → ajustement des prix.

4. Réaction différée de la politique monétaire

— building_permits_lag12 vs fed_rate : +0.41

→ Lorsque l'activité immobilière était élevée un an plus tôt (hausse des permis), la Fed tend à relever ses taux aujourd'hui. Cela reflète une réponse monétaire (retardée) à une économie en surchauffe.

5. Inertie des séries temporelles

- `fed_rate` vs `fed_rate_lag3` : 0.98
- `mortgage_rate` vs `mortgage_rate_lag3` : 0.98
- `case_shiller_value` vs `case_shiller_value_lag12` : 0.99
- `cs_growth` vs `cs_growth_lag12` : 0.70

Cela montre que ces séries sont très autocorrélées, ce qui motive l'intégration de variables retardées ou de modèles autorégressifs.

6. Corrélations secondaires à surveiller

- `mortgage_rate_growth` vs `cs_growth` : 0.20
- `fed_rate_growth` vs `cs_growth` : 0.22
- `fed_rate_growth` vs `building_permits` : 0.16
- `mortgage_rate` vs `building_permits` : 0.16
- `building_permits` vs `cs_growth_lag6` : 0.54

Ces relations sont plus faibles, mais elles peuvent révéler des effets indirects ou non linéaires à explorer dans des modèles plus complexes.

7. Synthèse

- Les taux d'intérêt ont un effet fort, mais différé, sur les prix immobiliers.
- L'impact est surtout visible entre 12 et 36 mois après une hausse des taux.
- Les variables immobilières sont très corrélées entre elles, suggérant qu'une seule peut suffire dans un modèle.
- La forte inertie des séries justifie l'introduction de lags dans la modélisation.

En résumé : *Les résultats confirment que les taux d'intérêt — notamment les taux hypothécaires — influencent significativement le marché immobilier, avec un effet différé dans le temps. Une hausse des taux réduit la capacité d'emprunt des ménages, freine la demande, puis finit par peser sur les prix. Ce mécanisme valide notre hypothèse initiale : les taux d'intérêt peuvent, en effet, orienter nos choix résidentiels et impacter profondément la dynamique du secteur immobilier.*

6.1.3 Taux d'intérêt et prix immobiliers : un regard différé

Les corrélations ci-dessous mettent en évidence le rôle clef des taux d'intérêt *décalés et inversés* dans l'explication de la dynamique des prix immobiliers. Ces deux séries (indice Case-Shiller et taux hypothécaire inversé décalé de 24 et 36 mois) constituent donc la base de nos modèles prédictifs de la section suivante.

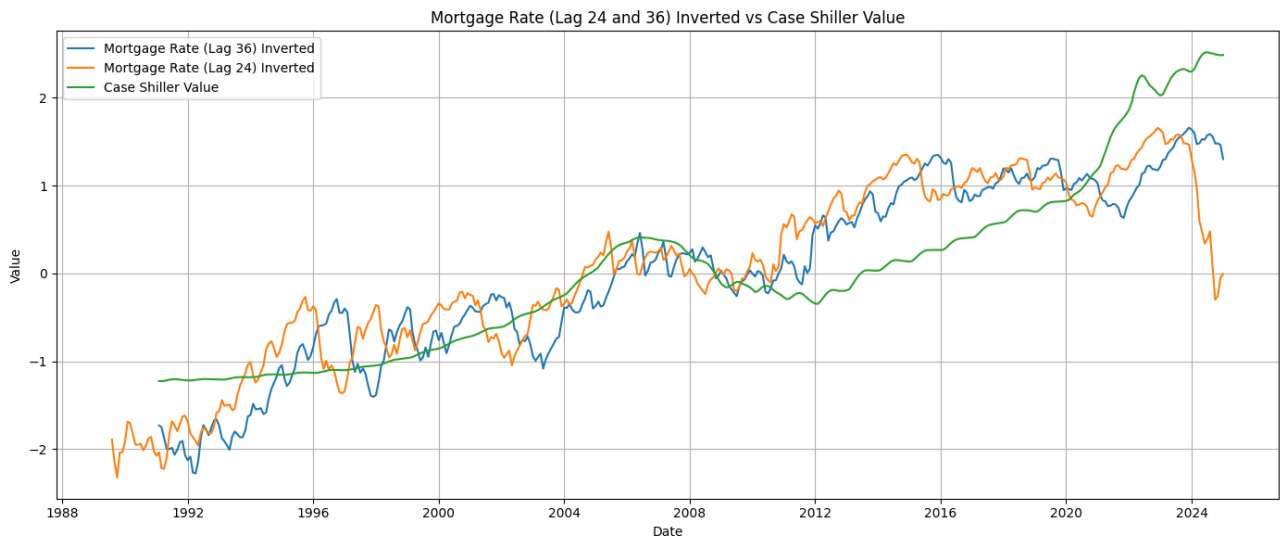


FIGURE 5 – Indice Case-Shiller et taux hypothécaire inversé (lags 24 et 36 mois)

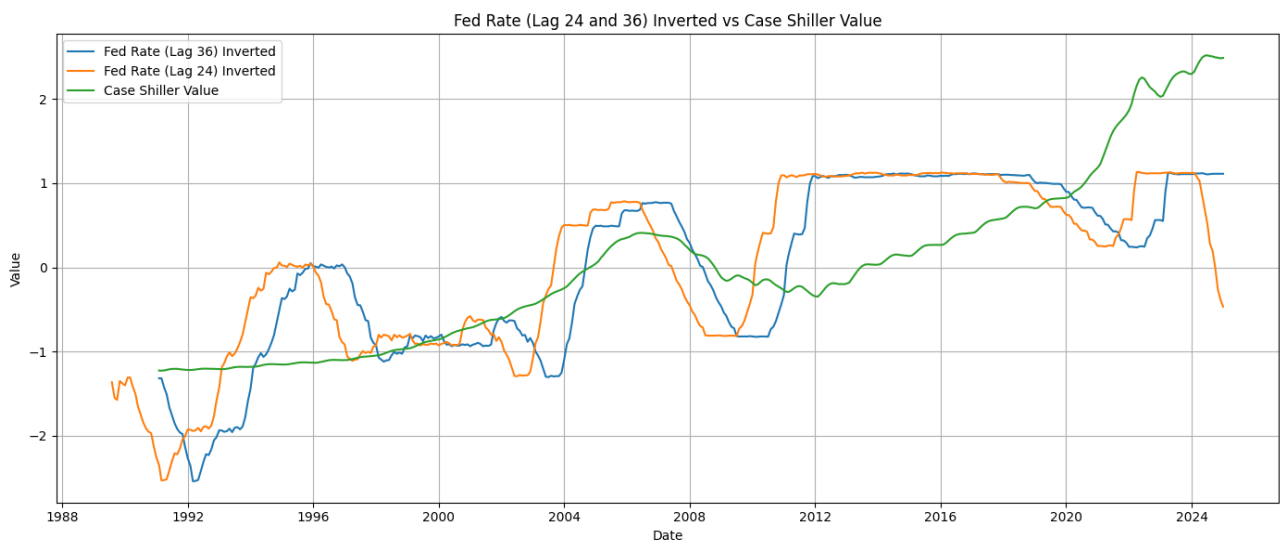


FIGURE 6 – Indice Case-Shiller et taux directeur inversé (lags 24 et 36 mois)

6.2 Modélisation prédictive : prévoir les prix immobiliers à 3 ans

6.2.1 Contexte et cadre prédictif

L'objectif est de tester la capacité de modèles simples à **prédire l'indice Case-Shiller (CS)** à un horizon **fixe de 36 mois**. Ce choix impose de n'utiliser *aucune* information plus récente que $t - 36$, excluant ainsi les lags de courte période (1–24 mois) pourtant très corrélés à CS.

Les modèles mobilisent donc uniquement :

- des variables macroéconomiques (*fed_rate*, *mortgage_rate*) retardées de 36 mois ;
- ou l'historique lointain de CS lui-même (autoregressions AR- x avec $x \geq 36$).

L'évaluation des modèles est réalisée selon une approche par validation *walk-forward* qui simule une utilisation en conditions réelles. Pour chaque intervalle, les 20% de la fin des données sont utilisés pour faire la prédiction. Les performances sont moyennées sur 7 fenêtres de test glissantes, comme indiqué dans le Tableau 4. La Figure 8

illustre les prédictions du modèle A par rapport aux valeurs réelles sur plusieurs de ces périodes, notamment : 1989-2000, 2000-2010, 2005-2015 et 2015-2025.

6.2.2 Spécifications OLS évaluées

TABLE 3 – Spécifications testées

Sigle	Forme	Régressors à $t - 36$
A	Full macro	$fed_rate_{t-36}, mortgage_rate_{t-36}, CS_{t-36}$
B	Fed & CS	$fed_rate_{t-36}, CS_{t-36}$
C	Fed only	fed_rate_{t-36}
D	Mortgage & CS	$mortgage_rate_{t-36}, CS_{t-36}$
E	Mortgage only	$mortgage_rate_{t-36}$
AR- x	Autoregressif	$CS_{t-36}, \dots, CS_{t-x}, x \in [36, 66]$

Remarque : Les modèles autoregressifs AR- x successifs utilisent un nombre croissant de lags de l'indice Case-Shiller uniquement, à partir de $x = 36$ jusqu'à $x = 66$, soit 31 modèles. Le premier AR utilise seulement CS_{t-36} , le deuxième ajoute CS_{t-37} , etc.

6.2.3 Métriques d'évaluation

Pour chaque fenêtre *walk-forward*, trois indicateurs sont calculés :

— **MAE** (Mean Absolute Error) :

$$MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|$$

— R^2 (coefficient de détermination) :

$$R^2 = 1 - \frac{\sum_t (y_t - \hat{y}_t)^2}{\sum_t (y_t - \bar{y})^2}$$

— **Trend Accuracy** (%) :

$$TA = \frac{1}{T-1} \sum_{t=2}^T \mathbf{1}[\text{sign}(y_t - y_{t-1}) = \text{sign}(\hat{y}_t - \hat{y}_{t-1})] \times 100$$

6.2.4 Bilan quantitatif

TABLE 4 – Performances moyennes sur les 7 fenêtres de test

Modèle	MAE ↓	R^2 ↑	Trend Acc. (%) ↑
A (full macro)	22.0	− 9.8	52.9
B (fed+CS)	23.3	− 12.6	54.7
C (fed)	40.0	− 51.1	40.4
D (mortg.+CS)	24.0	− 16.6	43.5
E (mortg.)	33.4	− 39.8	44.0
Best AR- x (AR-39)	31.1	− 31.2	62.1

Remarques.

- **MAE** : le modèle A réduit l'erreur de 40 % par rapport au meilleur AR- x .
- **Trend Accuracy** : le signal directionnel du taux directeur (modèle B) est utile, mais la simple mémoire longue (AR-39) reste la plus fiable pour les retournements (62 %).
- R^2 **négatifs** : à 36 mois, aucun modèle ne parvient à expliquer la variance des prix.

6.2.5 Analyse visuelle

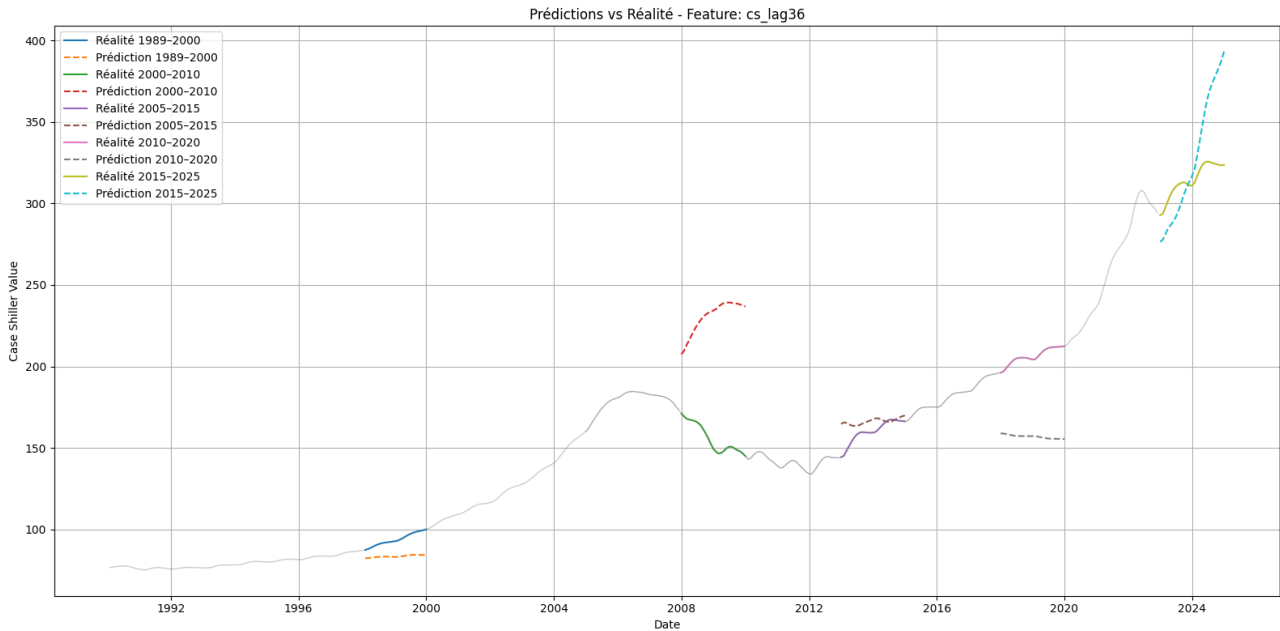


FIGURE 7 – Baseline AR-36 : prédiction à partir de CS_{t-36} uniquement

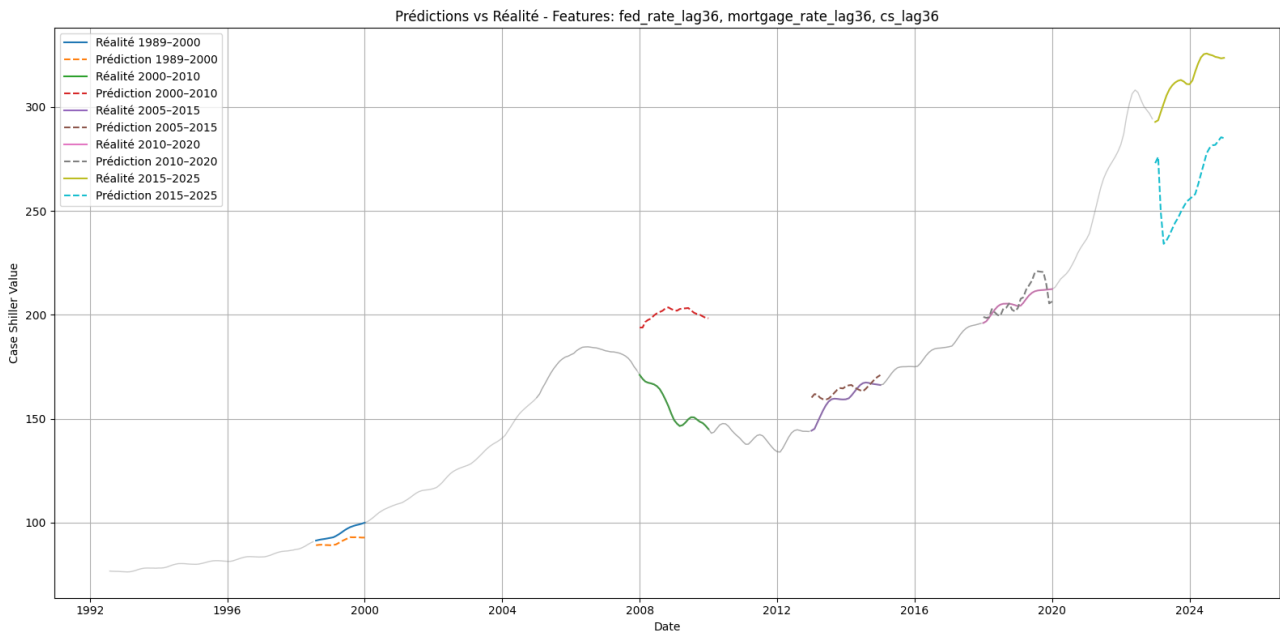


FIGURE 8 – Modèle A : fed_rate_{t-36} , $mortgage_rate_{t-36}$ et CS_{t-36}

On observe que le modèle A suit mieux le niveau global (2000-2024) que l'AR-36, lequel prolonge simplement la tendance passée. Les chocs de 2008 et 2020 restent toutefois mal anticipés dans les deux cas.

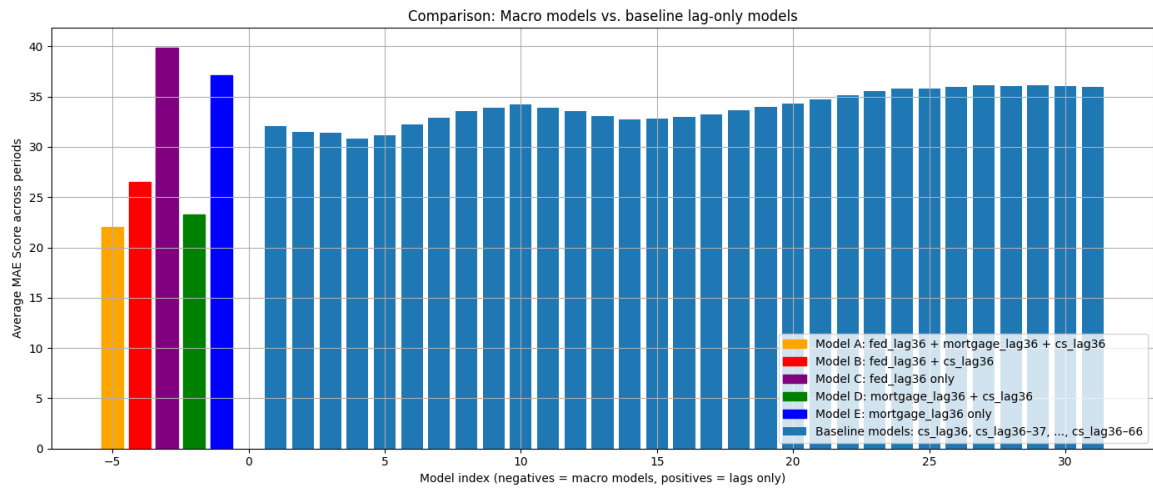


FIGURE 9 – Barplot des MAE par modèle

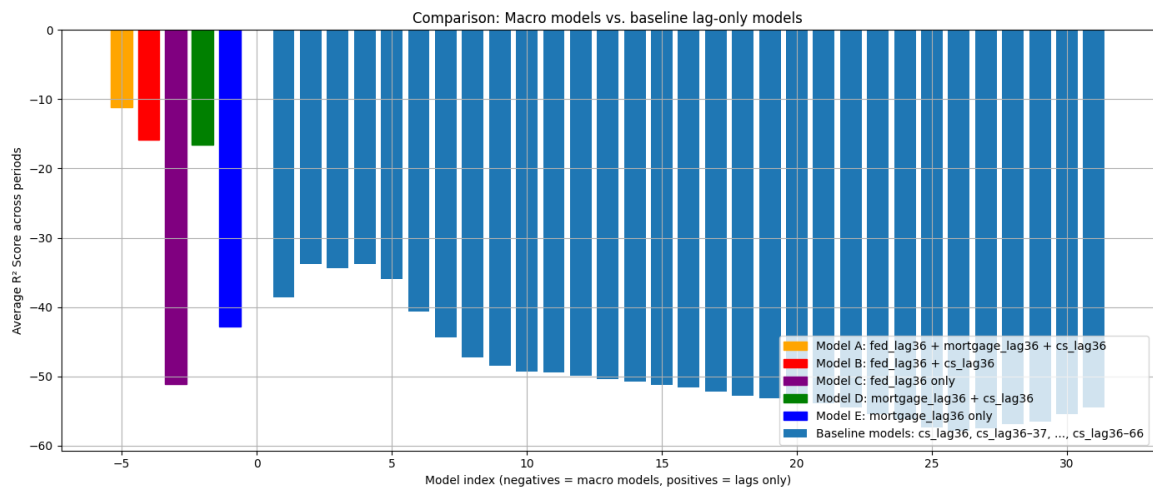


FIGURE 10 – Barplot des R^2 (négatifs = sous-performance vs moyenne)

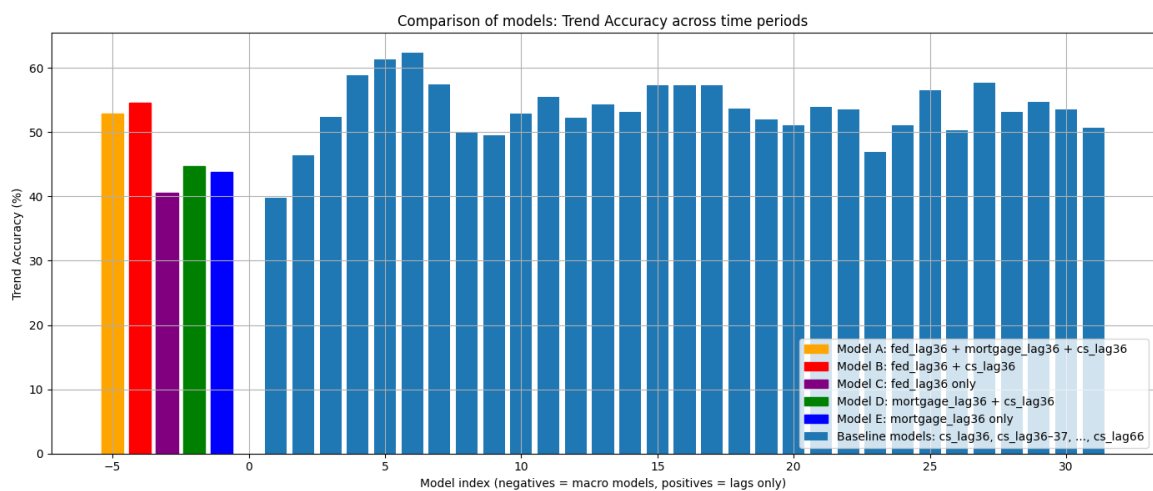


FIGURE 11 – Barplot de la Trend Accuracy

6.2.6 Conclusion et limites.

- **Horizon de prédiction exigeant.** Prédire à 36 mois rend la tâche extrêmement difficile. Un horizon plus court permettrait d'utiliser des informations plus récentes (lags plus faibles), mieux corrélées avec l'indice.
- **Simplicité des modèles.** Les modèles testés sont volontairement linéaires et peu profonds. Des approches non linéaires (forêts aléatoires, réseaux neuronaux) ou dynamiques (VAR/VECM) pourraient améliorer la qualité des prédictions.
- **Instabilité.** On remarque que pendant les moments de crises les modèles ont du mal, et c'est cela qui dégrade leurs performance, par exemple période post-covid ou crise de 2008. Pour résoudre ce problème il faudrait pouvoir prédire la crise 3 années en avant ce qui est quasi-impossible à faire avec des modèles assez simple et des données limitées.

7 Emploi

Dans cette section, nous analysons l'influence des taux d'intérêt sur le marché du travail américain, c'est-à-dire sur le **taux de chômage** (`unemployment_rate`) et la **création nette d'emplois** (`nfp`). L'hypothèse directrice est la suivante : *En relevant (ou abaissant) ses taux directeurs, la fed peut-elle réellement accélérer ou freiner la dynamique de l'emploi ?*

L'étude s'appuie sur le notebook `employment.ipynb`, couvrant la période de **janvier 1959 à mars 2025**.

Les analyses portent sur les **quatre variables** suivantes :

- `unemployment_rate` : taux de chômage mensuel aux États-Unis, mesurant la proportion de la population active sans emploi ;
- `fed_rate` : taux d'intérêt effectif des fonds fédéraux, principal outil de politique monétaire de la FED ;
- `nfp` : nombre total d'emplois dans le secteur non agricole (*Non-Farm Payrolls*), indicateur clé de la création d'emplois ;
- `personal_saving_rate` : taux d'épargne personnelle des ménages américains, exprimé en pourcentage du revenu disponible.

7.1 Analyse exploratoire (EDA)

7.1.1 Visualisation des variables clés

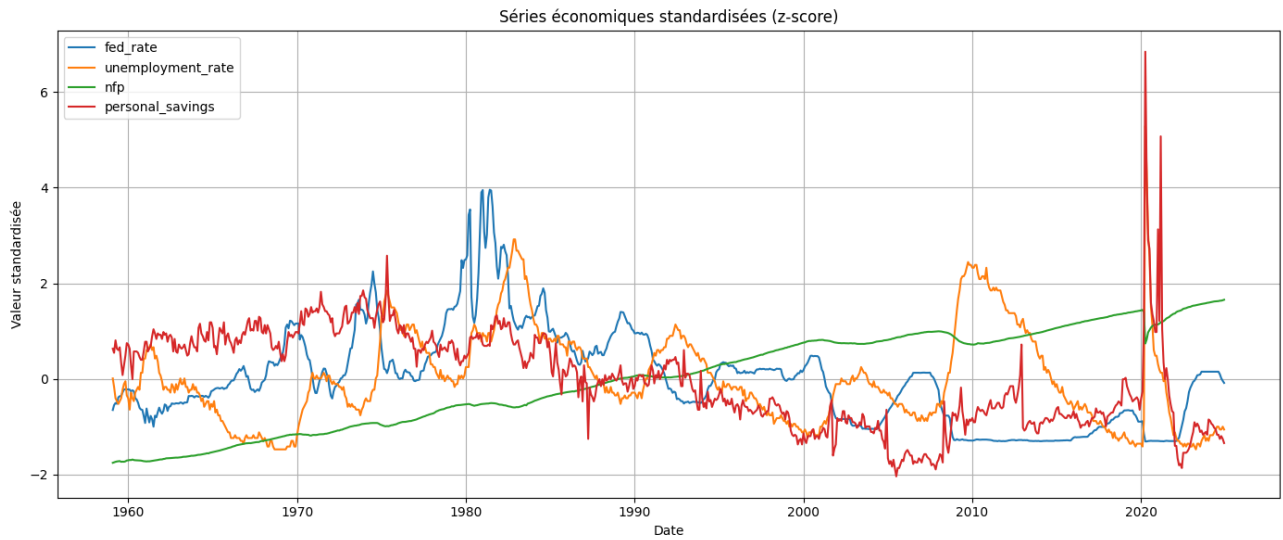


FIGURE 12 – Taux directeur, taux de chômage, emplois non-agricoles et taux d'épargne (séries standardisées)

Premiers enseignements.

- **Inertie** marquée du chômage (**unemployment_rate**) et surtout de l'emploi (**nfp**), aux profils très lissés.
- **Pics d'épargne** en 2020–2021 corrélés à la destruction d'emplois (confinements Covid).

7.1.2 Corrélations entre variables d'emploi et taux d'intérêt

1. Corrélations internes au marché du travail (sans lag).

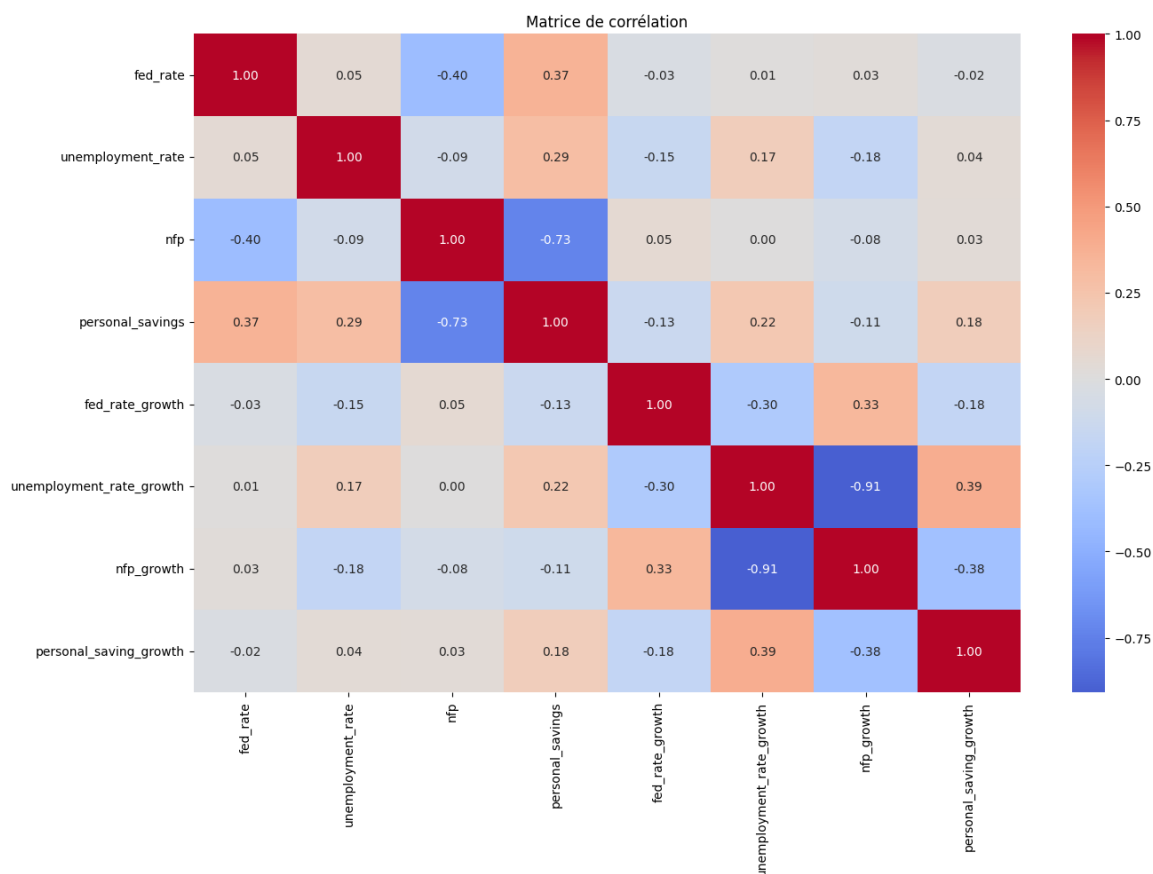


FIGURE 13 – Matrice de corrélation (sans décalage) – variables d’emploi

TABLE 5 – Corrélations internes au marché du travail

Paires de variables	r	Interprétation
unemployment_rate vs nfp	-0.91	Relation (quasi) mécanique : plus les créations d’emplois augmentent, plus le chômage baisse.
unemployment_rate vs personal_saving_rate	+0.29	En période de chômage élevé, les ménages épargnent davantage (corrélacion modérée).
nfp vs personal_saving_rate	-0.73	Davantage d’emplois est associé à moins d’épargne.

Conclusion : ces corrélations révèlent la *dualité* typique emploi-épargne : amélioration du marché du travail \Rightarrow baisse de l’épargne de précaution.

Corrélations avec les taux d’intérêt (effets immédiats).

- **fed_rate** vs **nfp** : -0.40 \rightarrow Resserrement monétaire modérément associé à une contraction de l’emploi.
- **fed_rate** vs **unemployment_rate** : +0.05 \rightarrow Effet instantané très faible \Rightarrow le chômage réagit lentement (voir *lags*).
- **fed_rate** vs **personal_saving_rate** : +0.37 \rightarrow Hausse des taux \Rightarrow hausse de l’épargne (rendements plus élevés / prudence).

7.1.3 Corrélations différées (lags)

Afin de capturer les effets retardés de la politique monétaire, quatre familles de décalages ont été testées :

- **Short lags** : 3, 6 et 12 mois
- **Medium lags** : 16, 24 et 30 mois
- **Long lags** : 36, 42 et 48 mois
- **Very long lags** : 54, 60 et 66 mois

Par souci de lisibilité, seuls les décalages « medium » (16–30 mois) sont présentés ci-après ; nous ne commentons que les *colonnes des variables non décalées* (règle de filtrage identique à la section Immobilier).

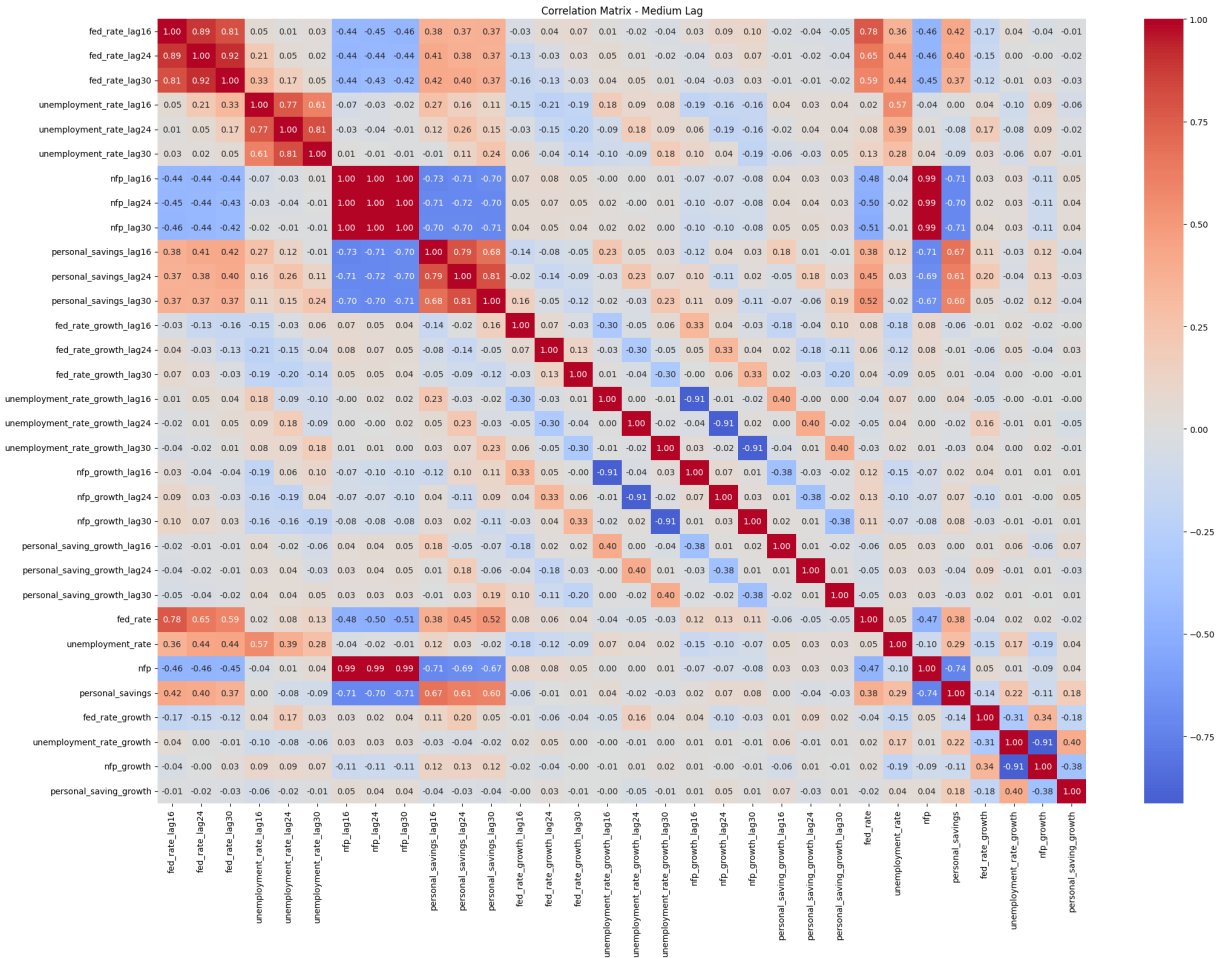


FIGURE 14 – Matrice de corrélation avec lags (16–30 mois)

TABLE 6 – Corrélations différées clés (*medium lags*)

Paire de variables	<i>r</i>	Interprétation
fed_rate_lag24 vs unemployment_rate	+0.44	Hausse des Fed Funds \Rightarrow montée du chômage ~2 ans plus tard.
fed_rate_lag24 vs rfp	+0.46	Signal faible et contre-intuitif : biais sectoriel (emplois qualifiés plus résilients).
personal_saving_rate_lag16 vs unemployment_rate	+0.32	L'épargne précautionniste précède souvent la dégradation du marché de l'emploi.

Conclusion : l'impact d'une hausse de taux sur le chômage n'apparaît qu'avec un *délai d'environ deux ans*.

Inertie des séries temporelles.

- `unemployment_rate` vs `unemployment_rate_lag3` : **0.90**; lag6 : **0.82**; lag12 : **0.67**.
- `nfp` vs `nfp_lag12` : ≈ 1 (quasi-parfaite jusqu'à 12 mois).
- `personal_saving_rate` vs `ps_lag12` : **0.73**.

Cette inertie justifie l'utilisation de composantes autorégressives (AR) dans les modèles prédictifs.

Synthèse.

- Les taux d'intérêt influencent le marché du travail, mais avec un **retard significatif** (~24 mois).
- En période de crise, chômage et épargne montent ensemble.
- Les séries d'emploi sont **hautement autocorrélées**, rendant les modèles AR très performants.

7.1.4 Taux d'intérêt et chômage : un regard différé

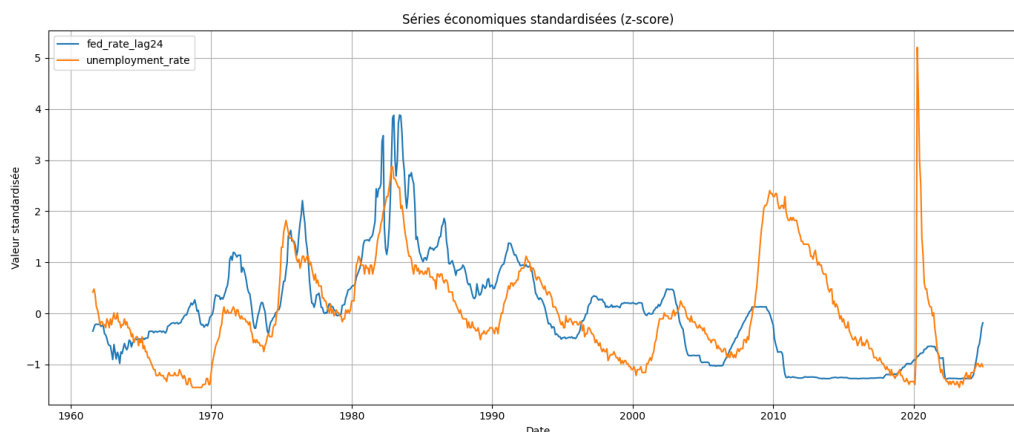


FIGURE 15 – Taux de chômage et taux directeur laggé de 24 mois

Le déphasage visuel confirme la corrélation de +0.44 : le point bas du chômage de 2019 suit la période de taux bas de 2015–2017, alors que la remontée attendue en 2024–2025 fait suite au resserrement historique de 2022.

7.2 Modélisation prédictive : prévoir le chômage à 12 mois

7.2.1 Cadre prédictif

Objectif : **prédire le taux de chômage à un horizon fixe de 12 mois**. Deux spécifications simples sont comparées :

TABLE 7 – Spécifications testées (horizon 12 mois)

Sigle	Forme	Régressors
AR	Autorégressif pur	$\{unemployment_t - 1, \dots, unemployment_t - 12\}$
Fed+Save+AR	AR + facteurs macro	AR (ci-dessus) + fed_rate_{t-12} , $personal_saving_t$

7.2.2 Résultats OLS

TABLE 8 – Régression OLS – chômage (modèle *Fed + Saving + AR*)

	Coef.	Std. Err.	<i>t</i>	<i>p</i> -val
Constante	4.474	0.145	30.92	0.000
personal_saving	0.069	0.017	4.02	0.000
fed_rate_lag24	0.176	0.016	10.76	0.000
<i>N</i>	767			
<i>R</i> ² (ajusté)	0.204			

Lecture rapide.

- $R^2 = 0.21$: modeste, mais **fed_rate_lag24** et **personal_saving** sont significatifs ($p < 1\%$).
- Coefficient positif sur **fed_rate_lag24** : +0.18 pp de chômage par point de taux, après 2 ans.

7.2.3 Validation hors échantillon

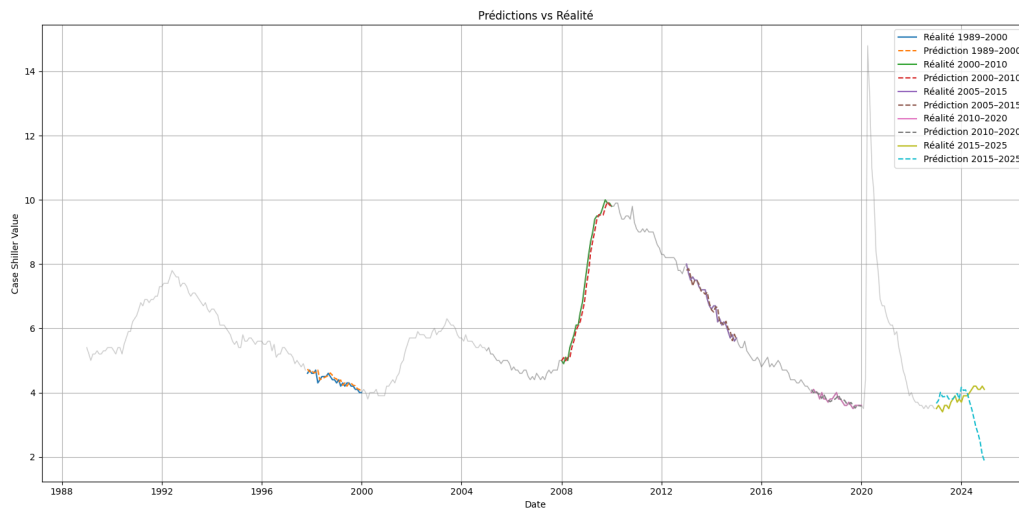


FIGURE 16 – Prédiction vs réalité (Fed+Save+AR)

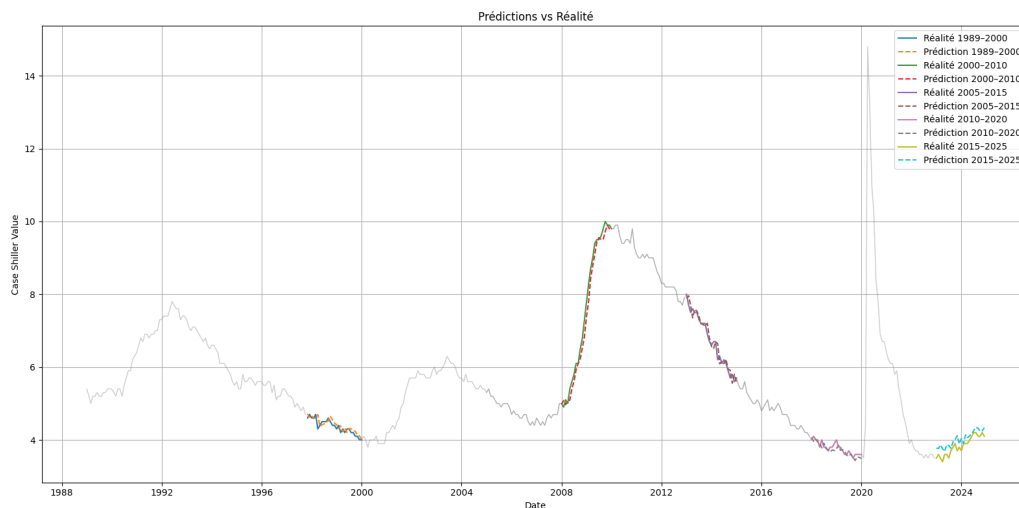


FIGURE 17 – Prédiction vs réalité (AR)

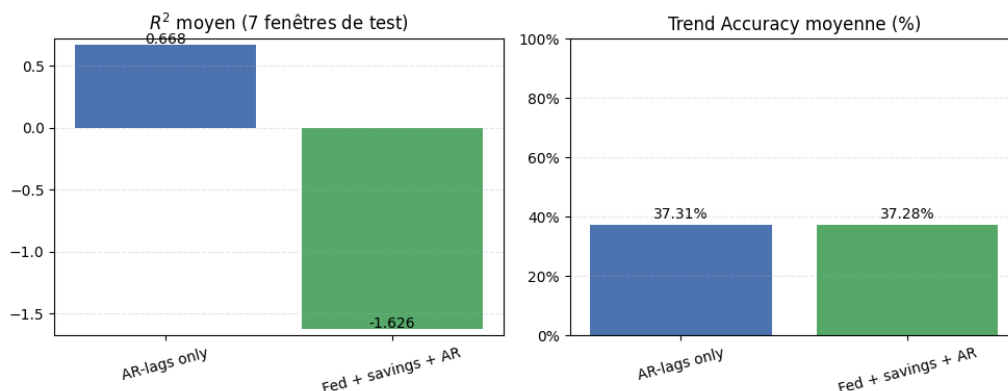


FIGURE 18 – Comparatif R^2 et Trend Accuracy – modèles AR vs Fed+Save+AR

Constats.

- **AR pur** : R^2 moyen = 0.67 ; Trend Accuracy = 37.3 %.
- **Fed+Save+AR** : R^2 moyen = -1.63 (dégradation) ; Trend Accuracy = 37.3 % (quasi identique).
- Les facteurs *fed* et *épargne* n'améliorent *pas* le pouvoir prédictif : l'autocorrélation du chômage suffit à expliquer la majeure partie de la dynamique à 12 mois.

7.2.4 Conclusion et limites

- **Rôle dominant de l'inertie.** À horizon 1 an, la meilleure prévision du chômage est fournie par... son propre passé.
- **Effet différé des taux.** Les Fed Funds n'exercent qu'un *impact mesurable à 24 mois* ; ils n'apportent donc aucune information supplémentaire pour un horizon 12 mois.
- **Perspectives.** Pour capter l'effet de politique monétaire, il faudrait soit allonger l'horizon de prévision, soit recourir à *modèles dynamiques multivariés* (VAR, VECM) capables de gérer simultanément plusieurs *lags*.

En résumé : *Le modèle autorégressif sur le taux de chômage surpasse les versions enrichies par les variables monétaires et d'épargne, preuve que l'information prédictive se trouve avant tout dans la forte inertie de la série elle-même.*

8 Consommation

Cette section a pour objectif d'explorer la relation entre les variations des taux d'intérêt et les décisions de consommation des ménages. Nous analysons différentes catégories de dépenses de consommation et les confrontons à l'évolution des taux d'intérêt directeurs sur la période allant de janvier 2001 à novembre 2024, en se basant sur les analyses effectuées dans le notebook Jupyter `consumption_analysis.ipynb`.

8.1 Analyse Exploratoire des Données (EDA)

8.1.1 Préparation et Prétraitement des Données

Les données utilisées comprennent le taux effectif des fonds fédéraux (Federal Funds Effective Rate) et les ventes au détail pour sept catégories : vêtements, stations-service, électronique, bijouterie, restaurants, supermarchés, et loisirs/jouets/jeux. Les étapes de prétraitement ont inclus :

- Conversion des dates au format `datetime`.

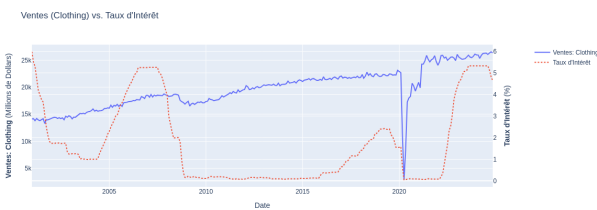
- Renommage des colonnes pour plus de clarté (par exemple, **FEDFUNDS** en **interest_rate**, et les colonnes de ventes en **sales_<catégorie>**).
- Définition de la colonne **DATE** comme index pour toutes les séries temporelles.
- Fusion des datasets sur la base des dates en un DataFrame unique.
- Sélection de la période d'analyse : du 1er janvier 2001 au 1er novembre 2024. Ce choix a été fait pour assurer la disponibilité des données pour la plupart des catégories, notamment **sales_supermarkets** qui débute en janvier 2001.
- Gestion des valeurs manquantes (**NaN**) par la méthode **ffill** (forward fill), qui propage la dernière observation valide. Le DataFrame résultant pour l'analyse des données nominales est **df_analysis**.

Une note importante concerne les ventes de bijoux (**sales_jewelry**), dont les données s'arrêtent en février 2021. Avec **ffill()**, les valeurs pour la période subséquente sont donc une imputation de la valeur de février 2021.

8.1.2 Analyse Exploratoire des Données (Données Nominales)

Visualisations Des visualisations graphiques ont été générées pour le taux d'intérêt et chaque catégorie de ventes nominales. Des graphiques superposant les ventes de chaque catégorie avec les taux d'intérêt ont également été produits pour une première inspection visuelle des relations.

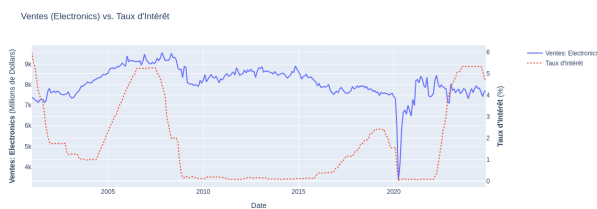
Ventes Nominales vs. Taux d'Intérêt



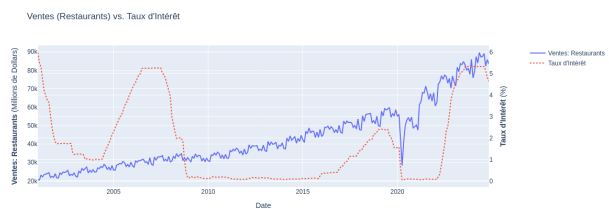
(a) Ventes nominales de Vêtements et Taux d'Intérêt.



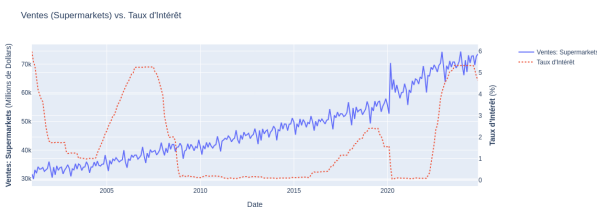
(b) Ventes nominales des Stations-Service et Taux d'Intérêt.



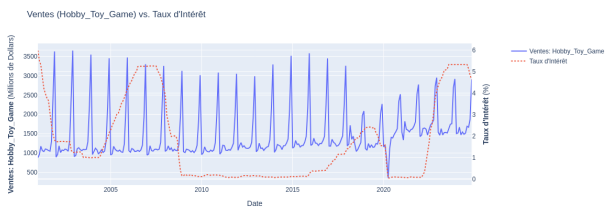
(c) Ventes nominales d'Électronique et Taux d'Intérêt.



(d) Ventes nominales des Restaurants et Taux d'Intérêt.



(e) Ventes nominales des Supermarchés et Taux d'Intérêt.



(f) Ventes nominales de Loisirs/Jouets/Jeux et Taux d'Intérêt.

FIGURE 19 – Comparaison des ventes nominales par catégorie avec le taux d'intérêt.

Matrice de Corrélation (Lag 0) La matrice de corrélation de Pearson a été calculée sur **df_analysis** pour les données nominales :

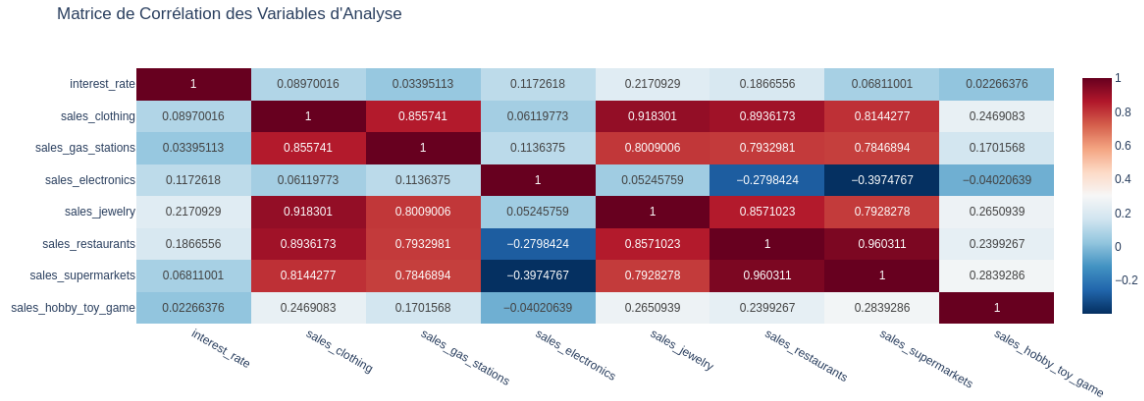


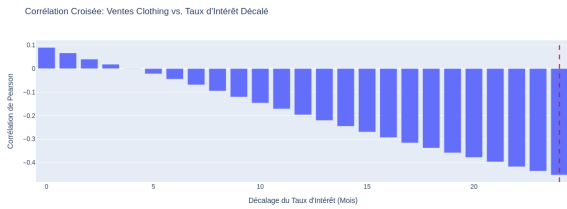
FIGURE 20 – Matrice de corrélation des variables nominales (Lag 0).

- Les corrélations linéaires directes (lag 0) entre les taux d'intérêt et les ventes se sont avérées généralement faibles.
- **sales_jewelry** : Corrélation de +0.217 (la plus élevée, positive).
- **sales_restaurants** : Corrélation de +0.187 (positive modérée).
- Autres catégories : corrélations positives faibles ou quasi nulles (ex : **sales_electronics** +0.117, **sales_clothing** +0.090).
- De fortes corrélations positives ont été observées entre plusieurs catégories de ventes (ex : **sales_restaurants** et **sales_supermarkets** : +0.960), indiquant une co-évolution due à des facteurs macroéconomiques communs.

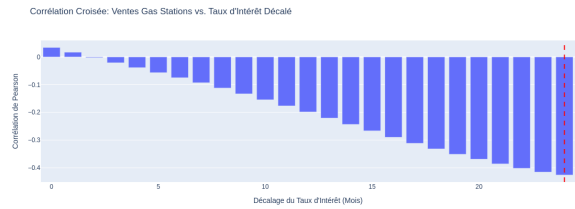
L'absence de corrélations négatives fortes suggère que l'impact direct et immédiat d'une hausse des taux sur une baisse des ventes n'est pas dominant au niveau mensuel agrégé pour les données nominales.

8.1.3 Analyse des Décalages Temporels (Données Nominales)

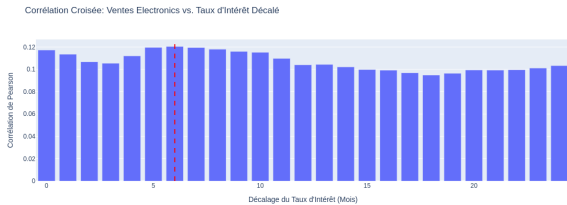
Cette analyse a cherché à identifier les décalages (jusqu'à 24 mois) pour lesquels la corrélation entre les taux d'intérêt passés et les ventes nominales actuelles est maximisée. Les graphiques illustrant ces analyses de décalage pour chaque catégorie sont présentés ci-dessous.



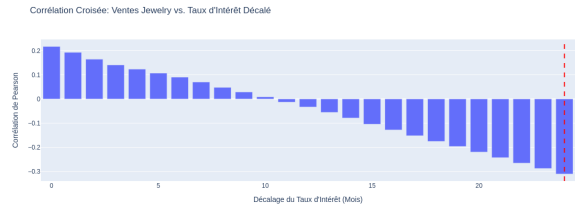
(a) Analyse de décalage : Ventes nominales de Vêtements vs. Taux d'Intérêt.



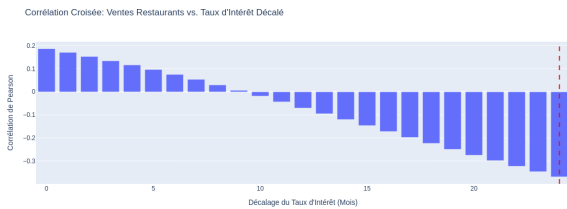
(b) Analyse de décalage : Ventes nominales des Stations-Service vs. Taux d'Intérêt.



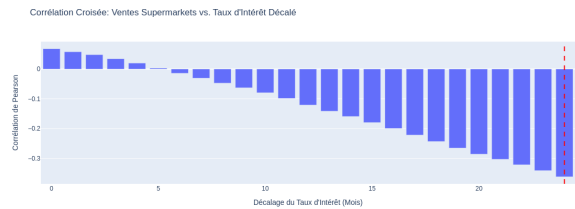
(c) Analyse de décalage : Ventes nominales d'Électronique vs. Taux d'Intérêt.



(d) Analyse de décalage : Ventes nominales de Bijouterie vs. Taux d'Intérêt.



(e) Analyse de décalage : Ventes nominales des Restaurants vs. Taux d'Intérêt.



(f) Analyse de décalage : Ventes nominales des Supermarchés vs. Taux d'Intérêt.

FIGURE 21 – Analyse de décalage des ventes nominales par catégorie avec le taux d'intérêt.

8.1.4 Tests de Causalité de Granger (Données Nominales Différenciées)

Les tests de causalité de Granger ont été effectués pour évaluer si les valeurs passées des taux d'intérêt ont un pouvoir prédictif sur les ventes. Les séries ont été préalablement différenciées (ordre 1) pour aider à assurer la stationnarité.

- **Hypothèse Nulle (H0)** : Les taux d'intérêt ne Granger-causent pas les ventes.
- Pour **toutes les catégories de ventes analysées**, l'hypothèse nulle a été rejetée (p-values des tests F < 0.05 pour divers lags). Cela indique que les variations passées des taux d'intérêt (différenciées) ont un pouvoir prédictif statistiquement significatif sur les variations actuelles des ventes (différenciées).
- Exemples de p-values minimales obtenues pour le test F : Vêtements (0.0128), Stations-Service (0.0106), Électronique (0.0000), Bijouterie (0.0004), Restaurants (0.0001), Supermarchés (0.0077), Loisirs/Jouets/Jeux (0.0007).

Ces tests renforcent l'idée que les taux d'intérêt sont une variable pertinente pour prédire les tendances de consommation, bien que cela indique une causalité prédictive et non structurelle.

8.1.5 Ajustement des Ventes en Termes Réels

Pour obtenir une mesure de la consommation en volume, les séries de ventes nominales ont été ajustées à l'aide de l'Indice des Prix à la Consommation (IPC, fichier `cpi-fed.csv`). L'indice CPI (`CPIAUCSL`) a été

normalisé et utilisé pour déflater les ventes. De nouvelles colonnes `sales_<catégorie>_real` ont été calculées dans un DataFrame `df_real`.

8.1.6 Analyse sur les Données Réelles

Matrice de Corrélation (Données Réelles vs. Taux d'Intérêt, Lag 0) L'ajustement à l'inflation a modifié les corrélations directes (lag 0) entre les taux d'intérêt et les ventes réelles :

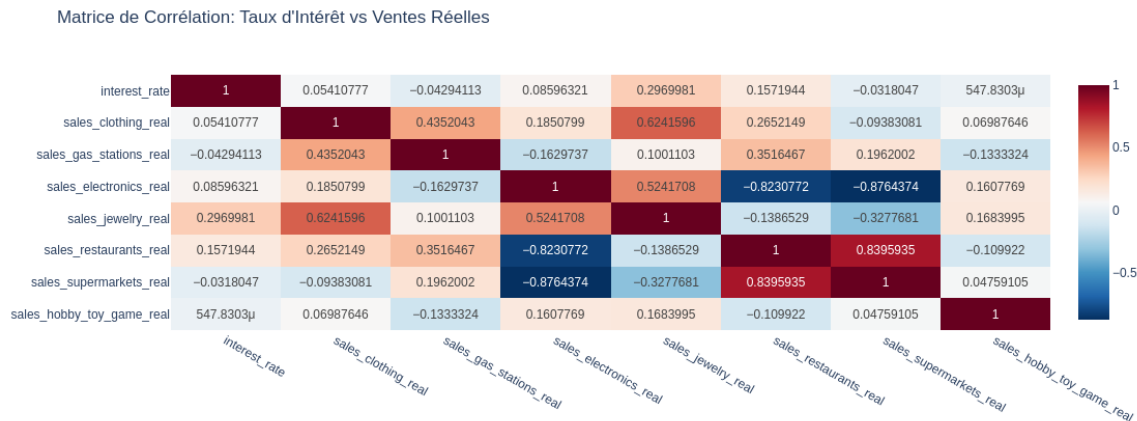


FIGURE 22 – Matrice de corrélation des variables réelles (Lag 0).

- Vêtements : +0.0541 (baisse par rapport au nominal +0.0897)
- Stations-Service : -0.0429 (inversion de signe par rapport au nominal +0.0339)
- Électronique : +0.0860 (légère baisse par rapport au nominal +0.1170)
- Bijouterie : +0.2970 (forte augmentation par rapport au nominal +0.2170)
- Restaurants : +0.1572 (baisse par rapport au nominal +0.1860)
- Supermarchés : -0.0318 (inversion de signe par rapport au nominal +0.0680)
- Loisirs/Jouets/Jeux : +0.0005 (baisse, quasiment nul)

L'inflation masquait ou altérait certaines relations. La corrélation pour la bijouterie est devenue notablement plus forte.

Analyse des Décalages (Données Réelles) Les décalages optimaux et les corrélations associées ont été recalculés avec les ventes réelles. Les résultats sont résumés dans le Tableau 9.

Points-clés :

1. La plupart des corrélations nominales s'affaiblissent après ajustement pour l'inflation, voire s'inversent (stations-service, supermarchés).
2. Les ventes de bijoux voient leur corrélation avec les taux augmenter (de 0.217 à 0.297), suggérant un effet inflation corrigé plus net pour le luxe.
3. Les relations perçues auparavant pouvaient être en partie dues à l'évolution générale des prix ; les séries réelles isolent la demande réelle des effets de niveau de prix.
4. Certains secteurs (électronique, restaurants) conservent une corrélation modérée mais réduite, indiquant que l'inflation masquait partiellement la vraie sensibilité.

Catégorie	Meilleur Lag (Réel)	Corrélation avec le Taux (Réel)	Corrélation Nominale	Changement de signe/magnitude
Vêtements	24	0.0541	0.0897	Baisse, reste positive
Stations-Service	23	-0.0429	0.0339	Inversion de signe
Électronique	24	0.0860	0.1170	Légère baisse
Bijouterie	0	0.2970	0.2170	Augmentation
Restaurants	24	0.1572	0.1860	Baisse
Supermarchés	24	-0.0318	0.0680	Inversion de signe
Jouets/Loisirs	22	0.0005	0.0226	Baisse, quasiment nul

TABLE 9 – Corrélations par catégorie

8.2 Modèles prédictifs

8.2.1 Modélisation par Régression Linéaire Simple (Données Réelles)

Des modèles de régression linéaire simple de la forme $ventes_reelles(t) = \alpha + \beta \cdot taux_interet(t - lag_{opt}) + \epsilon$ ont été estimés pour chaque catégorie de ventes réelles, en utilisant le décalage optimal (lag_{opt}) identifié pour les données réelles. Les résultats sont résumés dans le Tableau 10.

TABLE 10 – Résumé des Modèles de Régression Linéaire (Ventes Réelles vs. Taux d'Intérêt Décalé)

Catégorie	Lag Utilisé (Mois)	R-squared	β (Coefficient)	P-value (β)
Vêtements	24	0.069	-25.65	<0.0001
Stations-Service	23	0.091	-94.46	<0.0001
Électronique	24	0.118	+31.49	<0.0001
Bijouterie	0	0.088	+3.50	<0.0001
Restaurants	24	0.154	-215.86	<0.0001
Supermarchés	24	0.106	-81.29	<0.0001
Loisirs/Jouets/Jeux	22	0.008	+3.20	0.1591

La plupart des coefficients β sont statistiquement significatifs ($p\text{-value} < 0.05$), confirmant une influence des taux d'intérêt sur la consommation réelle. Les secteurs des Restaurants et Stations-Service montrent les sensibilités négatives les plus élevées en termes absolus. L'Électronique et la Bijouterie réagissent positivement aux hausses de taux (avec des décalages différents). Les R-squared restent modestes (entre 0.008 et 0.154), indiquant que d'autres facteurs expliquent une grande partie de la variation des ventes réelles. Le modèle simple ne capte pas bien le secteur Loisirs/Jouets/Jeux (coefficient non significatif).

8.3 Résultats et discussion

L'analyse approfondie menée dans le notebook `consumption_analysis.ipynb` confirme que les taux d'intérêt exercent une influence statistiquement significative sur la consommation des ménages dans la plupart des secteurs étudiés. Cette influence est cependant complexe, souvent retardée, et varie considérablement selon les catégories de biens et services.

Principaux Enseignements :

- **Importance de l'Ajustement à l'Inflation** : L'analyse des données en termes réels a fourni des perspectives plus claires, modifiant parfois la nature et l'ampleur des relations observées avec les données nominales.
- **Décalages Temporels** : L'impact des taux d'intérêt n'est généralement pas immédiat. Pour de nombreux secteurs (Vêtements, Restaurants, Supermarchés), un décalage de 24 mois maximise la corrélation négative avec les ventes réelles. L'électronique montre une corrélation positive avec un décalage de 24 mois, tandis que la bijouterie montre un effet positif instantané en termes réels.
- **Direction et Sensibilité de l'Impact** : Un impact négatif (hausse des taux → baisse des ventes réelles) est observé pour les Vêtements, Stations-Service, Restaurants et Supermarchés. Les Restaurants et Stations-Service montrent les sensibilités absolues les plus élevées. Un impact positif (hausse des taux → hausse des ventes réelles) est observé pour l'Électronique et la Bijouterie.
- **Pouvoir Explicatif Limité** : Bien que statistiquement significatifs, les taux d'intérêt seuls n'expliquent qu'une part modeste de la variance totale des ventes réelles (R-squared entre 0.008 et 0.154 dans les régressions simples). Cela souligne que de nombreux autres facteurs macroéconomiques jouent un rôle prépondérant.
- **Hétérogénéité Sectorielle** : La sensibilité et la nature de la réponse aux taux d'intérêt varient grandement d'un secteur à l'autre.
- **Causalité Prédictive** : Les tests de Granger ont confirmé que les taux d'intérêt passés contiennent des informations utiles pour prédire les ventes futures (après différenciation des séries).

En résumé, les taux d'intérêt sont un facteur pertinent dans l'analyse de la consommation, mais leur impact doit être évalué dans un contexte plus large. Des modèles plus sophistiqués, intégrant d'autres variables macroéconomiques (PIB, chômage, confiance des consommateurs, etc.) et les interdépendances entre secteurs (comme dans un modèle VAR/VECM), sont nécessaires pour une compréhension plus complète et un meilleur pouvoir explicatif.

9 Conclusion

Ce rapport a montré comment les taux d'intérêt façonnent — souvent avec du retard et par des chemins détournés — de nombreux aspects de l'économie américaine. Pour le dire simplement : **nos décisions, qu'on le veuille ou non, dépendent beaucoup des taux fixés par la Fed.** Les analyses s'appuient sur les séries FRED complétées par quelques données « fait-maison » (web scraping) sur l'inflation, les marchés et le crédit.

- **Immobilier.** Les taux hypothécaires pèsent lourdement sur les prix, mais avec un décalage de ~12–36 mois. Plus le crédit est cher, plus la demande recule, puis les prix fléchissent. Nos modèles à 36 mois l'anticipent partiellement : les taux passés donnent la bonne *direction*, mais pas toujours la bonne ampleur.
- **Consommation.** L'effet des taux est réel mais inégal : négatif (et logique) sur la restauration ou l'essence ; plus ambigu — voire positif — sur l'électronique ou la bijouterie. Ajuster les séries à l'inflation et regarder les lags (jusqu'à 24 mois) est indispensable pour y voir clair.
- **Emploi.** Le marché du travail réagit lentement. En moyenne, il faut ~2 ans pour qu'une hausse des Fed Funds se traduise par plus de chômage et moins de créations d'emplois. Les banques centrales doivent donc arbitrer entre ralentir l'inflation et préserver l'emploi, sachant que les effets ne sont pas instantanés.

En somme, les taux d'intérêt restent un levier central. Ils agissent certes avec délai, mais finissent toujours par se répercuter sur l'immobilier, la consommation et l'emploi.

Limites et perspectives

Ce travail a des limites. Les mécanismes économiques sont complexes ; nos modèles (OLS, régressions simples, AR, etc.) n'en captent qu'une part. Cela se voit dans des R^2 parfois faibles et dans les erreurs qui surviennent lors des phases de forte instabilité. Prédire les prix immobiliers trois ans à l'avance, par exemple, est un défi très ambitieux.

Malgré tout, un point ressort : les taux d'intérêt dirigés par la FED restent un levier majeur. Leur effet est réel, mais se manifeste surtout avec du retard et par des canaux indirects. On l'observe dans l'immobilier, la consommation et l'emploi.

Pour aller plus loin :

- tester des modèles plus poussés (VAR/VECM, méthodes non linéaires) ;
- inclure davantage les données collectées sur le Web dans les prévisions ;
- étudier d'autres canaux de transmission de la politique monétaire ;
- comparer ces résultats à ceux d'autres économies.