Mémoire présenté devant l'ENSAE Paris
pour l'obtention du diplôme de la filière Actuariat
et l'admission à l'Institut des Actuaires
le 06/03/2025

Par : **Tanguy Le Guern**

Titre : **Impact du contexte économique sur la loi de rachat et les taux d'affaires nouvelles d'un portefeuille emprunteur**

Confidentialité : ☒ NON ☐ OUI (Durée : ☐ 1 an ☐ 2 ans)

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus

Membres présents du jury de la filière

Nom : Nicolas BARADEL

Membres présents du jury de l'Institut des Actuaires

Entreprise : Accenture France

Signature :

Directeur du mémoire en entreprise :

Noms : Amélie SIRIEYS & Léa KARIM
Signature :

Autorisation de publication et de mise en ligne sur un site de diffusion de documents actuariels *(après expiration de l'éventuel délai de confidentialité)*

Signature du responsable entreprise

Secrétariat :

Signature du candidat

Bibliothèque :

Ecole Nationale de la Statistique et de l'Administration Economique (ENSAE)
5, avenue Henry Le Chatelier - 91120 PALAISEAU, FRANCE

Résumé

L'environnement économique actuel, caractérisé par une inflation élevée et des taux d'intérêt fluctuants, pose des défis considérables pour le secteur de l'assurance emprunteur. Ces conditions économiques volatiles influencent significativement le comportement des assurés, notamment en termes de souscription de nouveaux contrats et de rachats anticipés. Face à cette situation, les méthodes traditionnelles d'évaluation des risques et de provisionnement des assureurs se révèlent de moins en moins adaptées.

L'objectif de ce mémoire est de proposer une approche novatrice pour appréhender au plus juste l'impact du contexte économique sur les portefeuilles d'assurance emprunteur. L'idée centrale est d'exploiter des techniques avancées de modélisation des séries temporelles pour anticiper l'évolution des taux d'affaires nouvelles et des lois de rachat en fonction de divers scénarios économiques.

Dans un premier temps, le mémoire présente le cadre de l'étude en détaillant les spécificités de l'assurance emprunteur et en analysant le contexte économique actuel en France. Ensuite, une description approfondie des données économiques utilisées et de la démarche méthodologique adoptée est fournie, incluant les traitements effectués sur les données et les analyses statistiques préliminaires.

Le cœur de l'étude réside dans la construction de modèles prédictifs pour le taux d'affaires nouvelles et la loi de rachat. Ces modèles s'appuient sur des techniques de modélisation des séries temporelles, permettant de capturer les relations complexes entre les variables économiques et le comportement des assurés.

Enfin, l'impact du contexte économique sur un portefeuille d'assurance emprunteur est évalué à travers une étude de cas concrète. Un portefeuille fictif est projeté sur cinq ans selon sept scénarios économiques distincts, centrés sur différentes hypothèses d'inflation. Pour chaque scénario et chaque année de projection, le Best Estimate (BE) est calculé, offrant ainsi une vision complète de la sensibilité du portefeuille aux variations économiques.

Cette approche vise à fournir aux assureurs des outils plus robustes pour anticiper l'évolution de leurs engagements dans un contexte économique incertain, leur permettant ainsi d'ajuster leurs stratégies de provisionnement et de tarification de manière plus précise et réactive.

Mots clefs: Emprunteur, Économie, Inflation, Taux d'emprunt, Rachat, Affaires nouvelles

Abstract

The current economic environment, characterized by high inflation and fluctuating interest rates, presents significant challenges for the borrower insurance sector. These volatile economic conditions strongly influence policyholder behavior, particularly regarding the subscription of new contracts and early redemptions. In this context, traditional risk assessment and reserving methods used by insurers are proving increasingly inadequate.

The objective of this thesis is to propose an innovative approach to accurately assess the impact of the economic context on borrower insurance portfolios. The central idea is to leverage advanced time series modeling techniques to anticipate the evolution of new business rates and surrender laws under various economic scenarios.

First, the thesis introduces the study framework by detailing the specificities of borrower insurance and analyzing the current economic context in France. Subsequently, a comprehensive description of the economic data used and the adopted methodological approach is provided, including data processing and preliminary statistical analyses.

The core of the study lies in the construction of predictive models for new business rates and surrender laws. These models rely on time series modeling techniques, enabling the capture of complex relationships between economic variables and policyholder behavior.

Finally, the impact of the economic context on a borrower insurance portfolio is evaluated through a concrete case study. A fictional portfolio is projected over five years according to seven distinct economic scenarios, centered on different inflation assumptions. For each scenario and projection year, the Best Estimate (BE) is calculated, offering a comprehensive view of the portfolio's sensitivity to economic variations.

This approach aims to provide insurers with more robust tools to anticipate the evolution of their commitments in an uncertain economic environment, enabling them to adjust their reserving and pricing strategies more precisely and proactively.

Keywords: Borrower, Economy, Inflation, Borrowing rates, Surrender, New business

Note de synthèse

Mise en contexte et problématique

L'assurance emprunteur constitue un mécanisme fondamental dans le paysage financier français, offrant une protection essentielle aux emprunteurs et aux prêteurs dans le cadre des crédits immobiliers et à la consommation. Ce secteur connaît des évolutions significatives ces dernières années, influencées par des changements réglementaires et économiques.

Le marché de l'assurance emprunteur a connu une croissance soutenue, reflétée par l'augmentation des encours de crédits à l'habitat et des cotisations d'assurance. Les données de la Banque de France révèlent une progression constante des encours de crédits à l'habitat accordés aux particuliers résidents depuis 2000, avec une croissance moyenne annuelle de 7,3%. Cependant, en 2023, cette croissance s'est considérablement ralentie à 1%. Le graphique gauche de la *Figure 6* illustre cette évolution depuis 2018, montrant ainsi visuellement cette tendance à la hausse suivie du récent ralentissement.

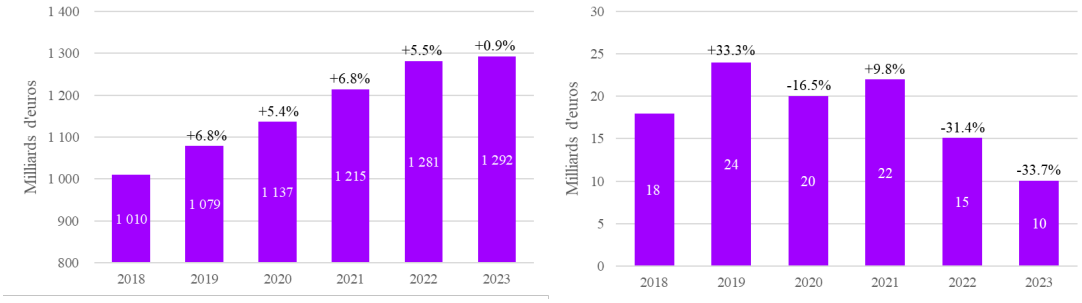


FIGURE 1 – Évolution des crédits à l'habitat en France : Encours total (gauche) et Flux annuel des nouveaux crédits (droite) (*Banque de France* [3])

La tendance des encours de crédits s'explique en partie par l'évolution des nouveaux crédits accordés aux particuliers. Les encours englobent les crédits existants et nouveaux, donc une baisse des nouveaux crédits ralentit la croissance des encours totaux, même si elle reste positive grâce au stock existant. Le graphique droit de la *Figure 6* illustre l'évolution annuelle des nouveaux crédits à l'habitat en France depuis 2018.

Le contexte réglementaire a évolué avec l'introduction de plusieurs lois visant à accroître la concurrence et à améliorer la protection des consommateurs, notamment la *Convention AERAS* (2006), la *Loi Lagarde* (2010), la *Loi Hamon* (2014), l'*amendement Bourquin* (2017), et la *Loi Lemoine* (2022). Ces changements législatifs ont modifié la dynamique du marché, favorisant la délégation d'assurance et la résiliation infra-annuelle des contrats.

Économiquement, la France fait face à des défis influençant le secteur de l'assurance emprunteur. L'instabilité politique, les fluctuations des taux d'intérêt, l'inflation, et l'évolution de la dette

publique sont autant de facteurs qui impactent le marché du crédit et, par extension, l'assurance emprunteur. L'interconnexion de ces variables crée un environnement complexe et dynamique pour les acteurs du secteur.

La problématique centrale de cette étude est d'analyser et de modéliser l'impact du contexte économique sur le marché de l'assurance emprunteur en France. L'objectif est de développer des modèles prédictifs pour les taux d'affaires nouvelles et les lois de rachat, en tenant compte des variables macroéconomiques pertinentes. L'approche utilise des modèles statistiques avancés, notamment les séries temporelles, pour capturer les tendances et les saisonnalités du marché. L'étude quantifie également l'impact de différents scénarios économiques sur un portefeuille d'assurance emprunteur fictif, en utilisant le calcul du Best Estimate (BE) comme métrique d'évaluation.

Cette approche vise à mieux comprendre les dynamiques actuelles du marché et à fournir des outils pour la gestion des risques et la prise de décision stratégique dans le secteur de l'assurance emprunteur.

Démarche mise en œuvre

La démarche mise en œuvre pour analyser l'impact du contexte économique sur le marché de l'assurance emprunteur en France s'articule autour de plusieurs étapes clés.

Initialement, l'étude débute par une collecte et une préparation des données économiques, puisées auprès de sources telles que la Banque de France et l'INSEE. Ces données sont agrégées et harmonisées pour garantir leur cohérence temporelle et qualitative. Une analyse statistique préliminaire est ensuite menée sur les taux d'affaires nouvelles et de rachat, examinant les tendances centrales, la dispersion ainsi que leurs corrélations avec les variables économiques exogènes.

La modélisation des séries temporelles constitue le cœur de l'étude. Après une introduction aux concepts clés, l'exploration se porte sur des modèles avancés tels que ARIMA, SARIMA, et SARIMAX. Un processus de modélisation est développé, intégrant des tests de stationnarité et l'incorporation de variables exogènes. Cette méthodologie est ensuite appliquée séparément aux taux d'affaires nouvelles et à la loi de rachat. Une analyse comparative des modèles permet de sélectionner les plus performants sur la base de critères statistiques.

Enfin, pour évaluer concrètement l'impact sur un portefeuille d'assurance emprunteur, un portefeuille fictif représentatif est construit. Ce portefeuille est projeté selon différents scénarios économiques, et le BE est calculé et analysé pour quantifier l'impact financier des variations économiques.

Cette approche systématique et multidimensionnelle permet d'explorer en profondeur les dynamiques du marché de l'assurance emprunteur et d'évaluer de manière tangible l'influence des facteurs économiques sur la performance des portefeuilles.

Constructions des modèles d'affaires nouvelles et de rachat

La *Figure 7* illustre les processus de modélisation par séries temporelles des taux d'affaires nouvelles et de la loi de rachat.

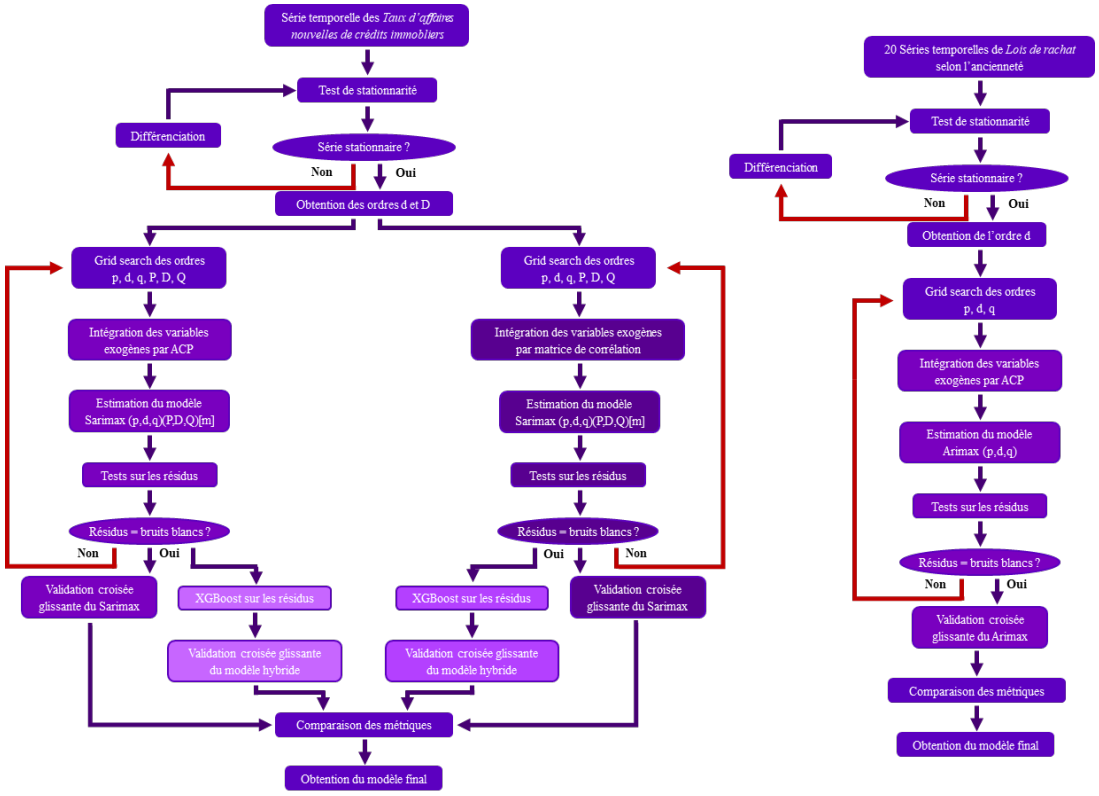


FIGURE 2 – Schémas des processus de modélisation du taux d'affaires nouvelles (gauche) et de la loi de rachat (droite)

Pour les taux d'affaires nouvelles, le modèle SARIMAX(4,0,1)(0,1,1)[12] avec ACP pour la sélection des variables exogènes s'avère optimal. Il capture efficacement les tendances et dynamiques complexes de la série temporelle, offrant une bonne précision prédictive. L'utilisation de l'ACP permet de réduire la dimensionnalité des données tout en préservant l'information pertinente, améliorant ainsi la stabilité et l'interprétabilité du modèle.

La loi de rachat utilise des modèles ARIMAX optimisés pour chaque année d'ancienneté (1 à 20 ans, les années 21 à 25 sont supposées à taux de rachat nuls). Cette approche individualisée saisit les nuances des comportements de rachat à différents stades du prêt, offrant une compréhension fine des dynamiques.

Les deux modélisations démontrent une capacité prédictive satisfaisante, validée par des métriques d'erreur (RMSE, MAE, MAPE) et des critères de sélection (AIC, BIC) favorables. Cependant, des limites sont identifiées, notamment la sensibilité aux changements structurels du marché et la difficulté à anticiper des événements économiques exceptionnels.

Ces outils peuvent aider les assureurs à mieux anticiper les dynamiques du marché de l'assurance emprunteur face aux variations économiques, offrant une vision intégrée des flux entrants et sortants du portefeuille.

Visualisation de l'impact du contexte économique sur un portefeuille d'assurance emprunteur

Pour évaluer l'impact des fluctuations économiques, notamment de l'inflation, sur la dynamique d'un portefeuille fictif d'assurance emprunteur, un modèle à six états est conçu (*Figure 8*). Chaque assuré peut évoluer entre les états suivants : valide, fin de prêt, rachat, incapacité, invalidité ou décès. Les transitions entre ces états sont régies par des lois spécifiques, permettant d'évaluer l'évolution du portefeuille sous différents scénarios économiques.
Les principales tables d'incidence utilisées sont :

- **Loi de rachat et d'affaires nouvelles** : Déterminées par séries temporelles et adaptée aux scénarios d'inflation ;
- **Loi d'entrée** : Un benchmark interne est utilisé pour l'incapacité, tandis que la table BCAC 2013 est employée pour l'invalidité ;
- **Loi de décès** : Les tables TH02/TF02 sont appliquées aux valides, la table BCAC 2002 à l'incapacité, et la table BCAC 2013 à l'invalidité ;
- **Loi de maintien** : Déduite pour les valides et repose sur les tables BCAC 2013 pour l'incapacité et l'invalidité.

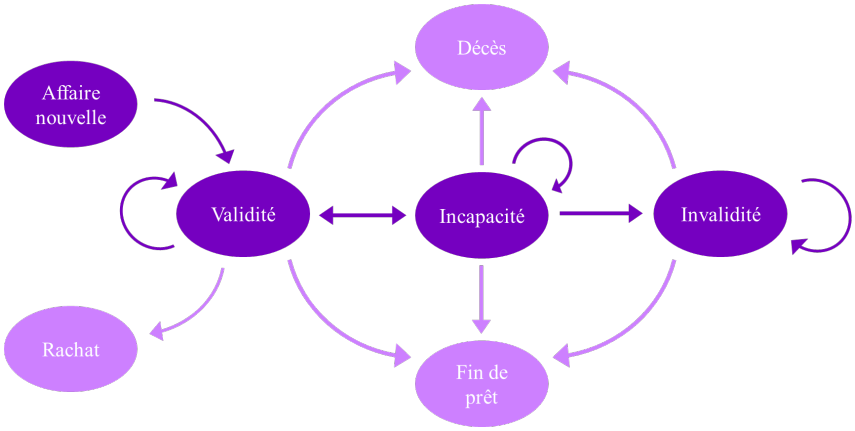


FIGURE 3 – Modèle à états

La projection du portefeuille s'étend sur la période du 31/01/2024 au 31/12/2028 et s'articule en trois étapes :

1. Vieillissement mensuel : les 15 000 model points sont actualisés en fonction des tables d'incidence (valide, incapable, invalide, décédé ou fin de prêt) ;
2. Application des lois économiques :
   - Les assurés valides sont soumis à la loi de rachat, ajustée selon le scénario économique considéré ;
   - De nouveaux assurés sont intégrés chaque mois, en fonction des taux d'affaires nouvelles projetés ;
3. Vieillissement des nouveaux assurés : les nouveaux entrants suivent ensuite le même processus de vieillissement et de transition que les assurés du portefeuille initial.

L'étude explore sept scénarios d'inflation pour la période 2023-2028, permettant d'évaluer l'impact de différentes trajectoires économiques sur le marché de l'assurance emprunteur et notamment sur les taux d'affaires nouvelles et de rachat. Ces scénarios sont divisés en trois catégories :

1. **Scénario standard** : Basé sur les prévisions de l'EIOPA et de la BCE, avec une inflation stabilisée à 2% à long terme ;
2. **Scénarios à la baisse** : Trois scénarios envisageant une inflation décroissante, avec des valeurs long terme de 1.3%, 0.65%, et 0% ;
3. **Scénarios à la hausse** : Trois scénarios prévoyant une inflation croissante, avec des valeurs long terme de 3%, 4%, et 5%.

Tous les scénarios partent d'une inflation de 4.9% au 31/12/2023 et convergent vers leurs valeurs cibles respectives au 31/12/2028. La *Figure 9* illustre graphiquement ces scénarios d'inflation :

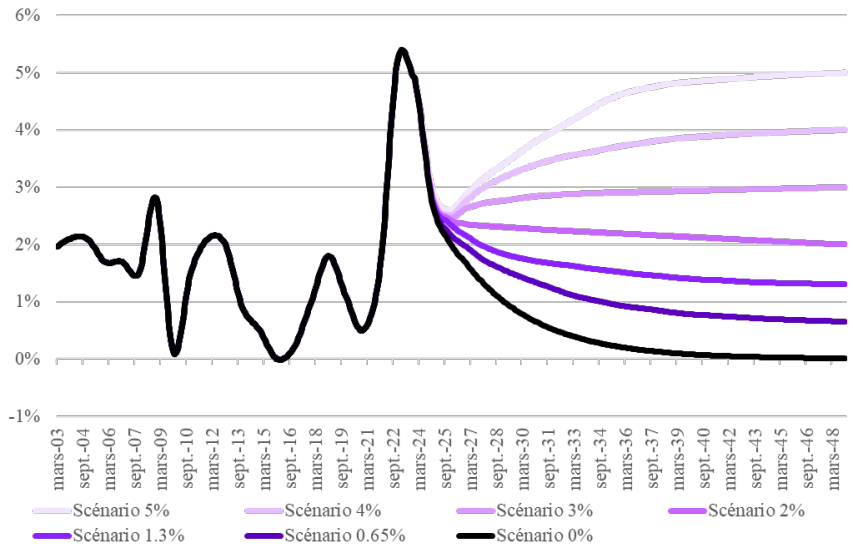


FIGURE 4 – Scénarios d'inflation

Pour projeter le portefeuille jusqu'à fin 2028 selon les sept scénarios d'inflation, des données économiques doivent être générées par modélisation. La méthodologie employée s'inspire de celle exposée *précédemment* pour les taux d'affaires nouvelles et de rachat et utilise un générateur de scénarios économiques (GSE) basé sur ces scénarios. Chaque variable économique est modélisée par séries temporelles en fonction de l'inflation, utilisant la méthode SARIMAX pour capturer les relations complexes.
Les projections révèlent des tendances distinctes selon les scénarios d'inflation :

Taux d'affaires nouvelles :

- Une relation inverse est observée entre l'inflation et les taux d'affaires nouvelles. Ainsi les scénarios à inflation élevée entraînent une diminution progressive des nouveaux entrants tandis que les scénarios à faible inflation favorisent une augmentation des taux d'affaires nouvelles ;
- Malgré les variations liées à l'inflation, une tendance générale à la hausse des taux d'affaires nouvelles est constatée, contrastant avec le déclin observé depuis 2022.

Taux de rachat : L'impact de l'inflation varie selon l'ancienneté du prêt :

- Pour les prêts récents : Une relation inverse entre inflation et taux de rachat est observée. Une baisse de l'inflation et des taux d'intérêt entraîne une augmentation des taux de rachat ;
- Pour les prêts anciens : La tendance s'inverse, avec une baisse de l'inflation et des taux d'intérêt correspondant à une diminution des taux de rachat.

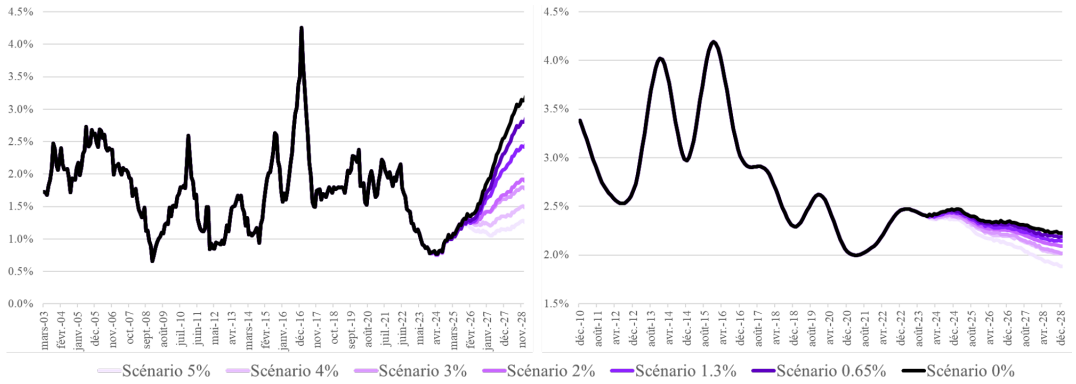


FIGURE 5 – Projection jusqu'en 2028 des taux d'affaires nouvelles (gauche) et des taux de rachat pour la 1ère année d'ancienneté du prêt (droite)

Ces interactions complexes, illustrées par la *Figure 10*, mettent en lumière que l'inflation, bien qu'importante, n'est qu'un facteur parmi d'autres influençant les flux entrants et sortants du portefeuille. Cette dynamique des flux entrants et sortants influence directement les BE du portefeuille, tant pour les primes que pour les sinistres.

La projection du portefeuille amène à une diminution des effectifs pour tous les scénarios d'inflation, mais à des rythmes différents :

- Scénario standard (2%) : L'effectif diminue de manière progressive et régulière, atteignant 12 184 assurés en 2028 ;
- Scénarios d'inflation faible (0%, 0,65%, 1,3%) : Une baisse rapide des effectifs est observée au début de la projection, en raison de taux de rachat élevés. Cependant, en fin de période (2028), ces scénarios aboutissent à des effectifs relativement mieux maintenus, grâce à un volume plus important d'affaires nouvelles ;
- Scénarios d'inflation élevée (3%, 4%, 5%) : La réduction des effectifs est globalement plus limitée en raison d'une baisse des taux de rachat. Le scénario extrême à 5% se démarque par une meilleure stabilisation des effectifs, atteignant 12 220 assurés en 2028, soit un niveau supérieur au scénario standard.

Ces tendances sont le résultat de l'interaction entre les taux d'affaires nouvelles, qui alimentent le portefeuille, et les taux de rachat, qui réduisent les effectifs.

Il est possible d'observer dans le *Tableau 3* que les BE de primes comme de sinistres du portefeuille projeté diminuent (en valeur absolue) progressivement dans tous les scénarios d'inflation, reflétant la baisse des effectifs due à la réduction des affaires nouvelles et aux rachats. Cependant, les scénarios d'inflation plus élevée ne conduisent pas systématiquement à des BE de primes moins élevés (en valeur absolue), ce qui souligne la complexité des interactions entre l'inflation et le comportement des assurés. En parallèle, les BE de sinistres, bien qu'affectés par les variations économiques, présentent des écarts limités entre les différents scénarios d'inflation.

en M€	BE de primes Déflatés					BE de sinistres Déflatés				
	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028
Inflation 0%	-132.27	-118.76	-105.54	-93.89	-83.02	94.48	86.92	77.77	70.26	63.65
Inflation 0.65%	-132.53	-118.92	-105.56	-94.15	-83.53	95.10	87.58	78.57	72.61	65.05
Inflation 1.3%	-132.37	-118.71	-105.63	-94.00	-82.88	95.07	87.20	78.59	71.82	64.02
Inflation 2%	-132.77	-119.31	-105.93	-93.88	-82.21	95.13	87.64	78.64	70.90	63.13
Inflation 3%	-132.25	-118.86	-105.49	-94.33	-83.44	94.28	88.14	78.59	72.06	64.08
Inflation 4%	-132.53	-119.29	-106.00	-94.25	-82.89	94.93	87.13	78.34	71.34	63.50
Inflation 5%	-132.56	-119.82	-106.23	-94.37	-82.79	94.42	87.29	78.25	71.24	63.03

TABLE 1 – BE de primes et de sinistres du portefeuille selon la date de projection et le scénario d'inflation

Les BE globaux (*Tableau 4*) combinent ces dynamiques, révélant des variations significatives selon les scénarios. Si une faible inflation tend à réduire les BE globaux (en valeur absolue), l'absence d'inflation ou une inflation élevée peut, dans certains cas, conduire à des BE globaux plus élevés (en valeur absolue) grâce à des facteurs compensatoires comme la stabilité du portefeuille ou l'effet de l'actualisation.

en M€	BE Déflatés				
	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028
Inflation 0%	-37.79	-31.84	-27.78	-23.63	-19.37
Inflation 0.65%	-37.43	-31.33	-26.99	-21.54	-18.48
Inflation 1.3%	-37.30	-31.50	-27.04	-22.18	-18.86
Inflation 2%	-37.64	-31.67	-27.29	-22.98	-19.07
Inflation 3%	-37.98	-30.72	-26.90	-22.28	-19.35
Inflation 4%	-37.60	-32.16	-27.66	-22.91	-19.39
Inflation 5%	-38.14	-32.52	-27.99	-23.13	-19.76

TABLE 2 – BE du portefeuille selon la date de projection et le scénario d'inflation

Cette étude met en lumière les effets directs de l'inflation sur les primes et sinistres, ainsi que les interactions subtiles influençant la performance globale du portefeuille d'assurance emprunteur.

Malgré les avancées méthodologiques et les résultats obtenus, certaines limites subsistent, notamment la disponibilité et la qualité des données concernant les comportements des emprunteurs en période de chocs économiques. L'intégration de modèles plus complexes, tels que des réseaux de neurones, pourrait enrichir l'analyse en capturant des interactions non linéaires.

Les perspectives offertes par ce mémoire sont nombreuses. L'approche développée pourrait être étendue à d'autres segments d'assurance ou à des portefeuilles réels pour affiner les stratégies de tarification et de gestion des risques. De plus, l'analyse pourrait être enrichie par l'intégration de scénarios macroéconomiques prospectifs, incluant des variables telles que le chômage ou les politiques de logement.

En conclusion, ce mémoire contribue à une meilleure compréhension des dynamiques du marché de l'assurance emprunteur dans un contexte économique en mutation, offrant aux assureurs une base solide pour anticiper les effets des variations macroéconomiques et ajuster leurs stratégies face aux incertitudes futures.

Executive summary

Context and issues

Loan insurance is a fundamental mechanism in the French financial landscape, offering essential protection to borrowers and lenders in the context of real estate and consumer loans. This sector has undergone significant change in recent years, influenced by regulatory and economic developments.

The creditor insurance market has enjoyed sustained growth, reflected in the increase in outstanding home loans and insurance premiums. Banque de France data show a steady rise in outstanding home loans to resident individuals since 2000, with average annual growth of 7.3%. By 2023, however, this growth had slowed considerably to 1.0%. The left-hand chart in *Figure 6* illustrates this evolution since 2018, visually showing this upward trend followed by the recent slowdown.

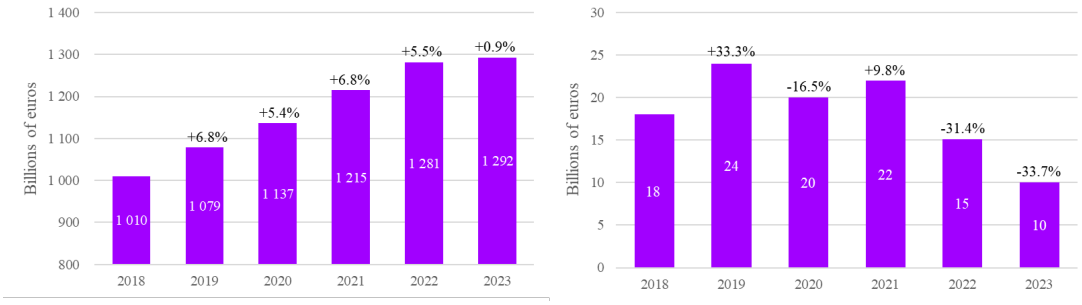


FIGURE 6 – Housing loans in France : Total outstanding (left) and annual flow of new loans (right) (*Banque de France* [3])

The trend in outstanding loans is partly explained by the evolution of new loans granted to individuals. Outstandings include both existing and new loans, so a fall in new loans slows the growth of total outstandings, even if it remains positive thanks to the existing stock. The right-hand chart in *Figure 6* illustrates the annual evolution of new housing loans in France since 2018.

The regulatory context has evolved with the introduction of several laws aimed at increasing competition and improving consumer protection, notably the *Convention AERAS* (2006), the *Loi Lagarde* (2010), the *Loi Hamon* (2014), the *amendment Bourquin* (2017), and the *Loi Lemoine* (2022). These legislative changes have altered market dynamics, favoring insurance delegation and infra-annual contract termination.

Economically, France is facing challenges that are having an impact on the loan insurance sector. Political instability, interest rate fluctuations, inflation and public debt trends are all factors impacting the credit market and, by extension, creditor insurance. The interconnection of these variables creates a complex and dynamic environment for industry players.

The central issue of this study is to analyze and model the impact of the economic context on the French loan insurance market. The aim is to develop predictive models for new business rates and surrender laws, taking into account relevant macroeconomic variables. The approach uses advanced statistical models, notably time series, to capture market trends and seasonality. The study also quantifies the impact of different economic scenarios on a fictitious loan insurance portfolio, using Best Estimate (BE) calculation as a valuation metric.

This approach aims to better understand current market dynamics and provide tools for risk management and strategic decision-making in the creditor insurance sector.

Approach

The approach used to analyze the impact of the economic context on the French loan insurance market is based on several key stages.

Initially, the study begins with the collection and preparation of economic data, drawn from sources such as the Banque de France and INSEE. These data are aggregated and harmonized to ensure temporal and qualitative consistency. A preliminary statistical analysis is then carried out on new business and buy-back rates, examining central trends, dispersion and correlations with exogenous economic variables.

Time series modeling is at the heart of the study. After an introduction to key concepts, the exploration focuses on advanced models such as ARIMA, SARIMA and SARIMAX. A modeling process is developed, incorporating stationarity tests and the incorporation of exogenous variables. This methodology is then applied separately to new business rates and the redemption law. A comparative analysis of the models enables us to select the best performers on the basis of statistical criteria.

Finally, to assess the concrete impact on a loan insurance portfolio, a fictitious representative portfolio is constructed. This portfolio is projected according to different economic scenarios, and the BE is calculated and analyzed to quantify the financial impact of economic variations.

This systematic, multi-dimensional approach enables us to explore the dynamics of the creditor insurance market in depth, and to assess in a tangible way the influence of economic factors on portfolio performance.

Building new business and buyout models

Figure 7 illustrates the time-series modeling processes for new business rates and the law of redemption.

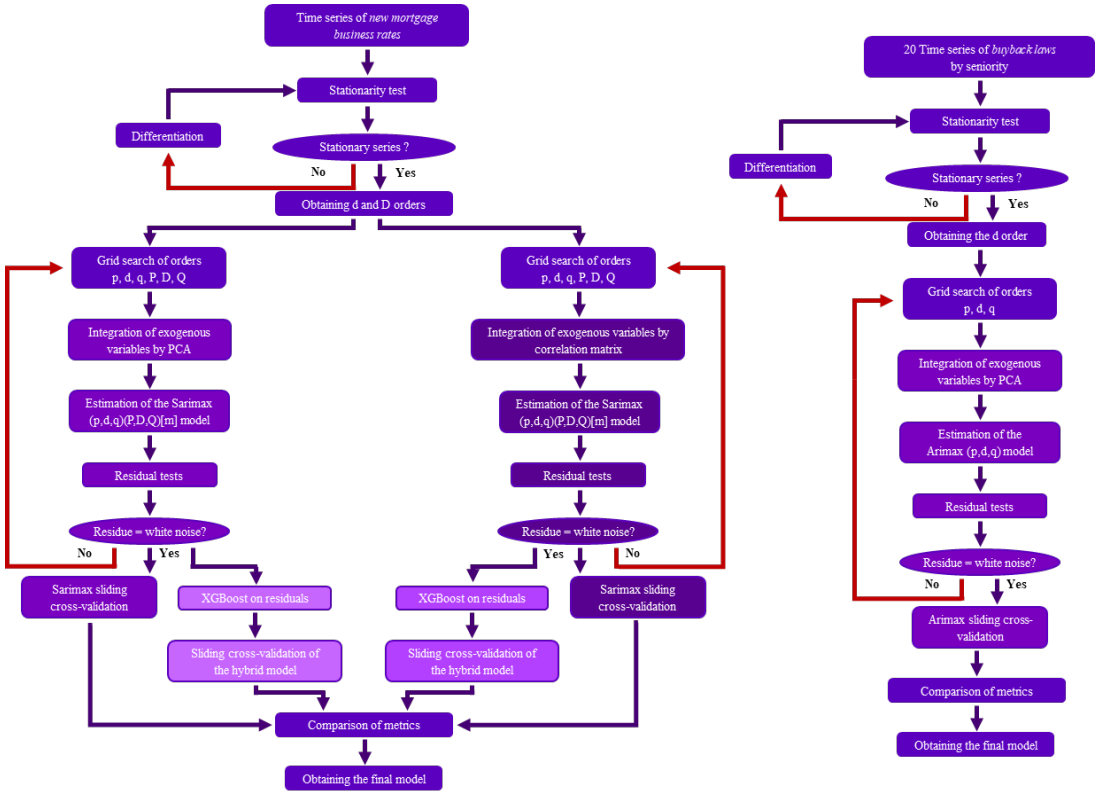


FIGURE 7 – Process diagrams for modeling the new business rate (left) and the law of redemption (right)

For new business rates, the SARIMAX(4,0,1)(0,1,1)[12] model with PCA for the selection of exogenous variables proves optimal. It effectively captures the complex trends and dynamics of the time series, offering good predictive accuracy. The use of PCA reduces the dimensionality of the data while preserving relevant information, improving the stability and interpretability of the model.

The redemption law uses ARIMAX models optimized for each year of age (1 to 20 years, years 21 to 25 are assumed to have zero redemption rates). This individualized approach captures the nuances of redemption behavior at different stages of the loan, providing a fine-tuned understanding of dynamics.

Both models demonstrate satisfactory predictive capacity, validated by favorable error metrics (RMSE, MAE, MAPE) and selection criteria (AIC, BIC). However, limitations have been identified, notably sensitivity to structural changes in the market and difficulty in anticipating exceptional economic events.

These tools can help insurers better anticipate the dynamics of the loan insurance market in the face of economic variations, offering an integrated view of portfolio inflows and outflows.

Visualizing the impact of the economic context on a loan insurance portfolio

To assess the impact of economic fluctuations, particularly inflation, on the dynamics of a fictitious loan insurance portfolio, a six-state model has been designed (*Figure 8*). Each policyholder can move between the following states : valid, end of loan, surrender, incapacity, disability or death. The transitions between these states are governed by specific laws, making it possible to evaluate the evolution of the portfolio under different economic scenarios.
The main impact tables used are :

- **Redemption and new business rates** : Determined by time series and adapted to inflation scenarios ;
- **Entry law** : An internal benchmark is used for disability, while the BCAC 2013 table is used for invalidity ;
- **Death law** : The TH02/TF02 tables are applied to the able-bodied, the BCAC 2002 table to incapacity, and the BCAC 2013 table to disability ;
- **Loi de maintien** : Deducted for the able-bodied and based on the BCAC 2013 tables for incapacity and disability.

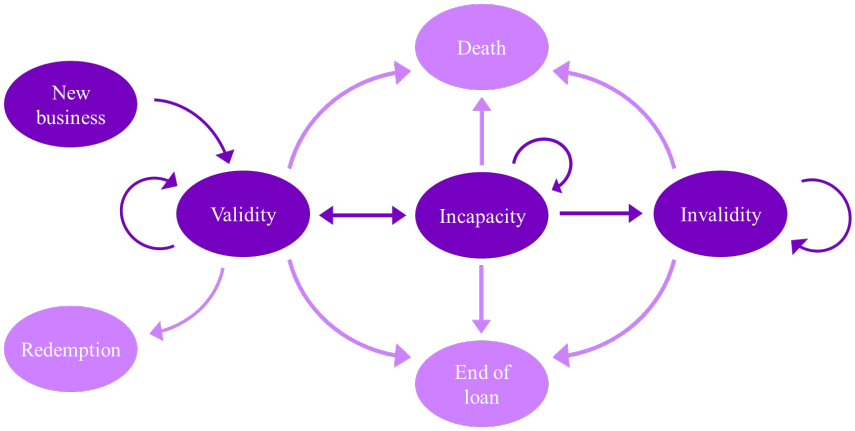


FIGURE 8 – Status model

The portfolio projection covers the period from 01/31/2024 to 12/31/2028 and is divided into three stages :

1. Monthly aging : the 15,000 model points are updated according to the incidence tables (able-bodied, unable-bodied, disabled, deceased or end of loan) ;
2. Application of economic laws :
   - Valid policyholders are subject to the redemption law, adjusted according to the economic scenario considered ;
   - New policyholders are integrated each month, based on projected new business rates ;
3. Aging of new policyholders : new entrants then follow the same aging and transition process as policyholders in the initial portfolio.

The study explores seven inflation scenarios for the period 2023-2028, assessing the impact of different economic trajectories on the creditor insurance market, and in particular on new business and surrender rates. These scenarios are divided into three categories :

1. **Standard scenario** : Based on EIOPA and ECB forecasts, with inflation stabilized at 2% over the long term ;
2. **Downward scenarios** : Three scenarios envisaging falling inflation, with long-term values of 1.3%, 0.65%, and 0% ;
3. **Upward scenarios** : Three scenarios forecasting rising inflation, with long-term values of 3%, 4%, and 5%.

All scenarios start from an inflation rate of 4.9% on 12/31/2023 and converge towards their respective target values on 12/31/2028. *Figure 9* graphically illustrates these inflation scenarios :

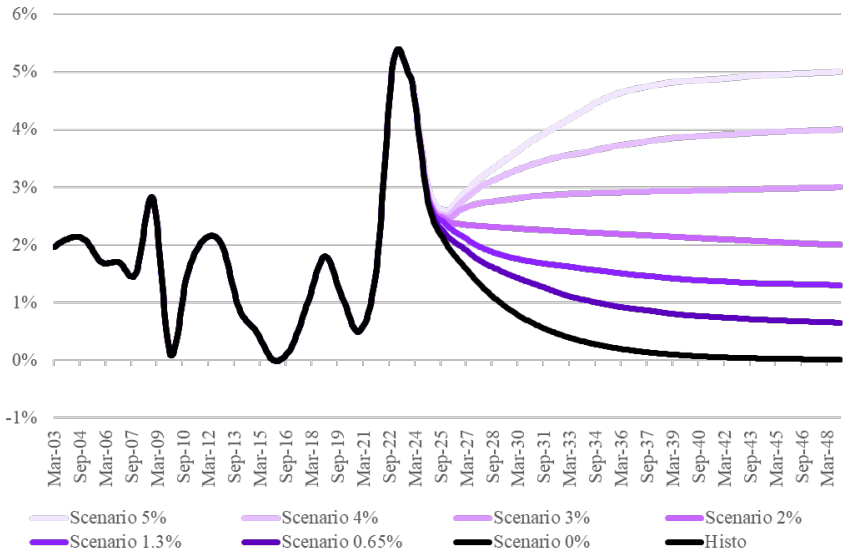


FIGURE 9 – Inflation scenarios

To project the portfolio to the end of 2028 according to the seven inflation scenarios, economic data must be generated by modeling. The methodology employed is based on that described above for new business and redemption rates, and uses an economic scenario generator (ESG) based on these scenarios. Each economic variable is time-series modeled against inflation, using the SARIMAX method to capture complex relationships.
The projections reveal distinct trends according to the inflation scenarios :

New business rate :

- An inverse relationship is observed between inflation and new business rates. Thus, high-inflation scenarios lead to a gradual decrease in new entrants, while low-inflation scenarios favor an increase in new business rates ;
- Despite inflation-related variations, a general upward trend in new business rates is observed, contrasting with the decline observed since 2022.

Redemption rate : The impact of inflation varies according to the age of the loan :

- For recent loans : An inverse relationship between inflation and repurchase rates is observed. A fall in inflation and interest rates leads to an increase in redemption rates ;
- For older loans : The trend is reversed, with a fall in inflation and interest rates corresponding to a fall in repurchase rates.

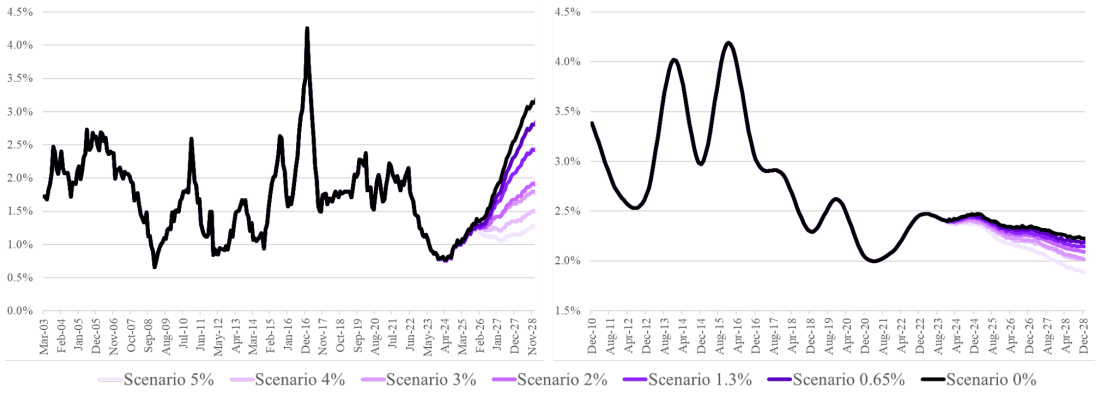


FIGURE 10 – Projection to 2028 of new business rates (left) and redemption rates for the 1st year of loan maturity (right)

These complex interactions, illustrated in *Figure 10*, highlight the fact that inflation, while important, is only one of several factors influencing the inflow and outflow of the portfolio. The dynamics of these inflows and outflows have a direct impact on the portfolio's BE, in terms of both premiums and claims.

The portfolio projection leads to a reduction in headcount for all inflation scenarios, but at different rates :

- Standard scenario (2%) : The workforce declines gradually and steadily, reaching 12,184 policyholders in 2028 ;
- Low inflation scenarios (0%, 0.65%, 1.3%) : A rapid decline in membership is observed at the start of the projection, due to high surrender rates. However, by the end of the period (2028), these scenarios result in a relatively more stable headcount, thanks to a higher volume of new business ;
- High inflation scenarios (3%, 4%, 5%) : Overall headcount reductions are more limited, thanks to lower redemption rates. The extreme 5% scenario is characterized by a better stabilization of the headcount, reaching 12,220 policyholders in 2028, a higher level than the standard scenario.

These trends are the result of the interaction between new business rates, which feed the portfolio, and surrender rates, which reduce the number of employees.

It can be seen from the *Table 3* that both the premium and claims EBs of the projected portfolio decrease progressively (in absolute terms) in all inflation scenarios, reflecting the decline in the number of policyholders due to the reduction in new business and surrenders. However, higher inflation scenarios do not systematically lead to lower premium EBs (in absolute terms), underscoring the complexity of the interactions between inflation and policyholder behavior. At the same time, claims EBs, although affected by economic variations, show limited differences between the different inflation scenarios.

en M€	Premium BE Deflated					BE of claims Deflated				
	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028
Inflation 0%	-132.27	-118.76	-105.54	-93.89	-83.02	94.48	86.92	77.77	70.26	63.65
Inflation 0.65%	-132.53	-118.92	-105.56	-94.15	-83.53	95.10	87.58	78.57	72.61	65.05
Inflation 1.3%	-132.37	-118.71	-105.63	-94.00	-82.88	95.07	87.20	78.59	71.82	64.02
Inflation 2%	-132.77	-119.31	-105.93	-93.88	-82.21	95.13	87.64	78.64	70.90	63.13
Inflation 3%	-132.25	-118.86	-105.49	-94.33	-83.44	94.28	88.14	78.59	72.06	64.08
Inflation 4%	-132.53	-119.29	-106.00	-94.25	-82.89	94.93	87.13	78.34	71.34	63.50
Inflation 5%	-132.56	-119.82	-106.23	-94.37	-82.79	94.42	87.29	78.25	71.24	63.03

TABLE 3 – Portfolio premium and claims BE by projection date and inflation scenario

Overall BE (*Table 4*) combine these dynamics, revealing significant variations across scenarios. While low inflation tends to reduce global BE (in absolute terms), high or no inflation can, in some cases, lead to less negative global BE, thanks to compensating factors such as portfolio stability or the effect of discounting.

en M€	BE Deflated				
	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028
Inflation 0%	-37.79	-31.84	-27.78	-23.63	-19.37
Inflation 0.65%	-37.43	-31.33	-26.99	-21.54	-18.48
Inflation 1.3%	-37.30	-31.50	-27.04	-22.18	-18.86
Inflation 2%	-37.64	-31.67	-27.29	-22.98	-19.07
Inflation 3%	-37.98	-30.72	-26.90	-22.28	-19.35
Inflation 4%	-37.60	-32.16	-27.66	-22.91	-19.39
Inflation 5%	-38.14	-32.52	-27.99	-23.13	-19.76

TABLE 4 – BE portfolio by projection date and inflation scenario

This study highlights the direct effects of inflation on premiums and claims, as well as the subtle interactions influencing the overall performance of the loan insurance portfolio.

Despite the methodological advances and results obtained, certain limitations remain, notably the availability and quality of data concerning borrower behavior during periods of economic shock. The integration of more complex models, such as neural networks, could enrich the analysis by capturing non-linear interactions.

The prospects offered by this thesis are numerous. The approach developed could be extended to other insurance segments or to real portfolios to refine pricing and risk management strategies. In addition, the analysis could be enriched by integrating forward-looking macroeconomic scenarios, including variables such as unemployment or housing policies.

In conclusion, this thesis contributes to a better understanding of the dynamics of the creditor insurance market in a changing economic context, providing insurers with a solid basis for anticipating the effects of macroeconomic variations and adjusting their strategies in the face of future uncertainties.

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué à la réalisation de ce mémoire.

Tout d'abord, je remercie chaleureusement mes tutrices de stage, Mesdames SIRIEYS Amélie et KARIM Léa, pour leur encadrement bienveillant, leur disponibilité, et leurs précieux conseils tout au long de la réalisation de ce mémoire.

Un grand merci à tous mes collègues pour leur accueil et leur aide au quotidien. En particulier, je tiens à mentionner Madame BONNIFAIT Coralie, Messieurs BESSON Guillaume et GASSIOT Pablo, avec qui les échanges ont été d'une grande richesse et ont grandement facilité la réalisation de ce mémoire.

Je remercie également Monsieur BERTHELE Emmanuel pour ses précieux conseils et ses orientations, qui ont été d'une grande aide dans la réalisation de ce travail.

Je souhaite exprimer ma gratitude à l'Université Gustave Eiffel qui m'a ouvert les portes du monde de l'actuariat à travers mon master, et m'a permis de développer mes premières compétences dans ce domaine passionnant.

Je saisis cette occasion pour adresser mes profonds remerciements aux responsables et aux professeurs de l'ENSAE pour les enseignements qui m'ont été dispensés lors de cette année de formation.

Enfin, je remercie du fond du cœur ma famille et mes amis pour leur soutien indéfectible, leur compréhension et leurs encouragements tout au long de mes études.

Je souhaite adresser une mention particulière pour Madame LA ROSA Manon, avec qui j'ai eu le plaisir de partager mon master et mon mastère spécialisé en actuariat. Je souhaite également remercier Maître LE GUERN Pauline et Monsieur MASSY Loïc pour l'expertise précieuse qu'ils m'ont apportée.

À tous, je vous exprime ma reconnaissance et ma gratitude pour avoir rendu cette expérience possible et enrichissante.

Glossaire

ACP	*Analyse en Composantes Principales*
ADF	*Augmented Dickey-Fuller*
AERAS	*Assurer et Emprunter avec un Risque Aggravé de Santé*
AR	*Auto-régressif*
ARIMA	*Autoregressive Integrated Moving Average*
ARMA	*Autoregressive Moving Average*
BCAC	*Bureau Commun des Assurances Collectives*
BCE	*Banque Centrale Européenne*
BE	*Best Estimate*
CCSF	*Comité Consultatif du Secteur Financier*
CP	*Composantes Principales*
CRD	*Capital Restant Dû*
CSP	*Catégorie Socio-Professionnelle*
CSS	*Conditional Sum of Squares*
DF	*Dickey-Fuller*
EIOPA	*European Insurance and Occupational Pensions Authority*
GSE	*Générateur de scénarios économiques*
ICC	*Indice du coût de la construction*
ICP-F	*Indice du coût de production dans la construction*
INSEE	*Institut National de la Statistique et des Études Économiques*
IPC	*Indice des Prix à la Consommation*
IPCH	*Indice des Prix à la Consommation Harmonisé*
IPT	*Invalidité Permanente Totale*
ITT	*Incapacité Temporaire de Travail*
MA	*Moyenne Mobile*
ML	*Maximum Likelihood*
OFDT	*Observatoire français des drogues et des tendances addictives*
PIB	*Produit Intérieur Brut*
PTIA	*Perte Totale et Irréversible d'Autonomie*
RSS	*Sum of Squared Residuals*
SARIMA	*Seasonal Autoregressive Integrated Moving Average*
SARIMAX	*Seasonal Autoregressive Integrated Moving Average with eXogenous variables*

Table des matières

Introduction **1**

1 Cadre de l'étude **2**
1.1 L'assurance emprunteur . . . 2
1.1.1 Principes généraux de l'assurance emprunteur . . . 2
1.1.2 Aspects spécifiques de l'assurance emprunteur . . . 6
1.1.3 Principales lois et règlements en vigueur . . . 6
1.1.4 Évolution et tendances du marché de l'assurance emprunteur en France . 9
1.1.5 Conclusion . . . 15
1.2 Contexte économique en France . . . 15
1.2.1 Facteurs Géopolitiques . . . 15
1.2.2 Croissance Économique . . . 16
1.2.3 Interconnexion du taux d'intérêt, de l'inflation et de la dette . . . 23
1.2.4 Marché de l'emploi . . . 26
1.2.5 Conclusion . . . 28

2 Présentation des données économiques et de la démarche **29**
2.1 Description de la base de données . . . 29
2.1.1 Sources des données . . . 29
2.1.2 Données du benchmark interne à d'Accenture . . . 30
2.1.3 Données de la Banque de France . . . 31
2.1.4 Données de l'INSEE . . . 31
2.2 Traitements effectués sur les données . . . 34
2.2.1 Retraitement et contrôle . . . 34
2.2.2 Agrégation des bases . . . 34
2.2.3 Filtres effectués sur les données . . . 35
2.2.4 Variables retenues . . . 36
2.3 Analyses statistiques . . . 38
2.3.1 Objet de l'étude . . . 38
2.3.2 Analyse de la tendance centrale et de la dispersion . . . 39
2.3.3 Décomposition chronologique . . . 41
2.3.4 Comparaison avec les variables économiques . . . 43
2.3.5 Synthèse . . . 50
2.4 Conclusion . . . 51

3 DÉMARCHE THÉORIQUE : Construction des lois d'Affaires Nouvelles et de Rachat **52**
3.1 Introduction aux séries temporelles . . . 53
3.1.1 Caractéristiques fondamentales . . . 53
3.1.2 Concepts clés . . . 54
3.2 Modèles de séries temporelles . . . 55

3.2.1 Modèles AR (Auto-regressive) et MA (Moving Average) . . . 55
3.2.2 Modèle ARMA . . . 57
3.2.3 Principe de stationnarité, différenciation et tests . . . 58
3.2.4 Modèle ARIMA . . . 65
3.2.5 Modèle SARIMA (Seasonal Auto-regressive Integrated Moving Average) . . . 66
3.2.6 Modèle SARIMAX (SARIMA avec variables exogènes) . . . 67
3.2.7 Synthèse . . . 68
3.3 Processus de modélisation . . . 69
3.3.1 Préparation des données . . . 69
3.3.2 Tests de stationnarité, différenciation et transformation des données . . . 70
3.3.3 Identification des ordres p, d, q (ARIMA) et P, D, Q (composante saisonnière) 71
3.3.4 Intégration des variables exogènes . . . 71
3.3.5 Base *train* et *test* . . . 75
3.3.6 Mise en place du modèle SARIMAX(p,d,q)(P,D,Q)[m][1] . . . 76
3.3.7 Analyse et Tests statistiques des résidus . . . 78
3.3.8 Projection, validation croisée et analyse vis à vis de la base test . . . 80
3.3.9 Critères de sélection et métriques d'erreur (AIC, BIC, MSE, MAE, RMSE, MAPE) . . . 81
3.3.10 Synthèse . . . 83

4 APPLICATION : Construction des taux d'Affaires Nouvelles 84
4.1 Remise en contexte . . . 84
4.2 Application de la méthodologie au taux d'affaires nouvelles . . . 85
4.2.1 Préparation des données . . . 86
4.2.2 Tests de stationnarité, Différenciation et transformation des données . . . 86
4.2.3 Identification des ordres p, d, q (ARIMA) et P, D, Q (composante saisonnière) 89
4.2.4 Intégration des variables exogènes . . . 90
4.2.5 Base train et test . . . 93
4.2.6 Mise en place du modèle SARIMAX(p,d,q)(P,D,Q)[m] . . . 93
4.2.7 Analyse et Tests statistiques des résidus . . . 93
4.2.8 Modélisation hybride . . . 93
4.2.9 Projection, validation croisée et analyse vis à vis de la base test . . . 94
4.2.10 Analyse comparative des modèles et résultats du modèle optimal . . . 95
4.3 Limites et perspectives d'amélioration . . . 101
4.4 Conclusion sur la modélisation du taux d'affaires nouvelles . . . 101

5 APPLICATION : Construction de la loi de Rachat 102
5.1 Remise en contexte . . . 102
5.2 Application de la méthodologie à la loi de rachat . . . 103
5.2.1 Préparation des données . . . 104
5.2.2 Tests de stationnarité et transformation des données . . . 105
5.2.3 Identification des ordres p, d, q (ARIMA) . . . 106
5.2.4 Intégration des variables exogènes . . . 107
5.2.5 Base train et test . . . 108
5.2.6 Mise en place du modèle ARIMAX(p,d,q) . . . 108
5.2.7 Analyse et Tests statistiques des résidus . . . 108
5.2.8 Modélisation hybride . . . 109
5.2.9 Projection, validation croisée et analyse vis à vis de la base test . . . 109
5.2.10 Analyse comparative des modèles et résultats du modèle optimal . . . 110
5.3 Limites et perspectives d'amélioration . . . 116
5.4 Conclusion sur la modélisation de la loi de rachat . . . 116

6 Impact du contexte économique sur un portefeuille d'assurance emprunteur 118
6.1 Présentation du portefeuille fictif . . . 118
6.1.1 Sources et méthodologie de construction du portefeuille . . . 119
6.1.2 Caractéristiques détaillées du portefeuille . . . 119
6.2 Méthodologie de projection . . . 120
6.2.1 Tables d'incidence . . . 121
6.2.2 Explication des 7 scénarios économiques basés sur l'inflation . . . 126
6.2.3 Projection des variables économiques, des taux d'affaires nouvelles et des lois de rachat . . . 127
6.2.4 Fonctionnement du modèle de projection du portefeuille . . . 130
6.2.5 Résultats de la projection du portefeuille . . . 130
6.3 Calcul du Best Estimate (BE) . . . 134
6.3.1 Définition du BE dans le contexte de l'étude . . . 134
6.3.2 Méthode de calcul pour chaque année et scénario . . . 134
6.3.3 Analyse comparative des résultats . . . 140
6.4 Conclusion du chapitre . . . 147

Conclusion 148

A CHAPITRE 1 : Cadre de l'étude 1

B CHAPITRE 2 : Présentation des données économiques et de la démarche 2
B.1 Méthode d'interpolation cubique d'Hermite . . . 2

C CHAPITRE 3 : DÉMARCHE THÉORIQUE : Construction du volume d'Affaires Nouvelles et de la loi de Rachat 5

D CHAPITRE 4 : APPLICATION : Construction de la loi d'Affaires Nouvelles 7
D.1 Méthode de Gradient Boosting . . . 7
D.1.1 Généralités . . . 7
D.1.2 Algorithme . . . 8
D.1.3 Gradient Boosted Regression Trees (GBRT) . . . 9
D.1.4 Régularisation . . . 10
D.1.5 Avantages et inconvénients . . . 10

E CHAPITRE 5 : APPLICATION : Construction de la loi de Rachat 11

F CHAPITRE 6 : Impact du contexte économique sur un portefeuille d'assurance emprunteur 16

Bibliographie 23

Table des figures

1 Évolution des crédits à l'habitat en France : Encours total (gauche) et Flux annuel des nouveaux crédits (droite) (*Banque de France* [3]) . . . iii
2 Schémas des processus de modélisation du taux d'affaires nouvelles (gauche) et de la loi de rachat (droite) . . . v
3 Modèle à états . . . vi
4 Scénarios d'inflation . . . vii
5 Projection jusqu'en 2028 des taux d'affaires nouvelles (gauche) et des taux de rachat pour la 1ère année d'ancienneté du prêt (droite) . . . viii
6 Housing loans in France : Total outstanding (left) and annual flow of new loans (right) (*Banque de France* [3]) . . . x
7 Process diagrams for modeling the new business rate (left) and the law of redemption (right) . . . xii
8 Status model . . . xiii
9 Inflation scenarios . . . xiv
10 Projection to 2028 of new business rates (left) and redemption rates for the 1st year of loan maturity (right) . . . xv

1.1 Répartition des cotisations des contrats d'assurance emprunteur selon l'objet du prêt en 2023 (*France Assureurs* [14]) . . . 10
1.2 Encours des crédits à l'habitat accordés aux particuliers résidents (*Banque de France* [3]) . . . 10
1.3 Flux annuel des crédits nouveaux à l'habitat aux particuliers résidents (*Banque de France* [3]) . . . 11
1.4 Cotisations des contrats d'assurance emprunteur (*France Assureurs* [14]) . . . 11
1.5 Répartition des cotisations des contrats d'assurance emprunteur en 2023 (*France Assureurs* [14]) . . . 12
1.6 Répartition des cotisations des contrats d'assurance emprunteur (*France Assureurs* [14]) . . . 12
1.7 Répartition des crédits (gauche) et des substitutions (droite) en portefeuille par CSP (*CCSF* [6]) . . . 13
1.8 Répartition des CSP par acteurs du marché assurantiel (*CCSF* [6]) . . . 14
1.9 Répartition des crédits (gauche) et des substitutions (droite) en portefeuille par capital assuré (*CCSF* [6]) . . . 14
1.10 Taux directeurs de la BCE (*BCE* [12]) . . . 16
1.11 Montant en milliards d'euros 2020 (gauche) et glissement annuel en % (droite) du PIB français (*INSEE* [29]) . . . 17
1.12 Taux de croissance du PIB et de la Dette Publique comparés aux taux d'intérêt annuel des crédits (*Banque de France* [3] & *INSEE* [21].[29]) . . . 18
1.13 IPCH[1] français sur Base 100 en 2015 (haut) et son glissement annuel en % (bas) (*INSEE* [17].[25]) . . . 19

1.14 Dépenses de consommation des ménages français en milliards d'euros (haut) avec et sans correction de l'inflation et son glissement annuel en % (bas) (*INSEE* [19]) 20
1.15 Conséquences de la tendance de la Dette publique (axe gauche) sur l'évolution du PIB français (axe droit) (*INSEE* [21].[29]) 22
1.16 Influence mutuelle des variables macro-économiques clés 23
1.17 Glissement annuel en % des flux et taux d'intérêt annuels des crédit nouveaux à l'habitat (*Banque de France* [3]) 24
1.18 Glissement annuel en % de l'IPCH français (axe gauche) et du taux d'intérêt annuel des crédit nouveaux à l'habitat (axe droit) (*Banque de France* [3] & *INSEE* [25]) 24
1.19 ICPH et indice du salaire horaire français en Base 100 2015 (gauche) et leur glissement annuel en % (droite) (*INSEE* [22].[25]) 27
1.20 Taux de chômage français au sens du BIT[1] (gauche) et glissements annuels en du taux de chômage et de l'IPCH (droite) (*INSEE* [25].[26]). 27

2.1 Taux d'affaires nouvelles (axe gauche) et son glissement annuel en % (axe droit) (*Banque de France* [3]) 40
2.2 Taux de rachat des crédits immobiliers par ancienneté du prêt (2010-2023) 40
2.3 Décomposition de séries chronologiques appliquée aux Taux d'affaires nouvelles (*Banque de France* [3]) 41
2.4 Moyenne (axe gauche) et Ecart-type (axe droit) par mois des Taux d'affaires nouvelles (*Banque de France* [3]) 42
2.5 Corrélations des variables économiques exogènes avec les taux de rachat des crédits immobiliers par plage d'années d'ancienneté du crédit[1] 45
2.6 Glissements annuels en % comparés des taux d'affaires nouvelles et des taux d'intérêt annuel des crédits nouveaux à l'habitat des particuliers (*Banque de France* [3]) 46
2.7 Comparaison Taux de rachat des crédits immobiliers par plage d'années d'ancienneté du crédit et glissement annuel en % des taux d'intérêt annuel des crédits nouveaux à l'habitat des particuliers 47
2.8 Glissements annuels en % comparés des taux d'affaires nouvelles (axe gauche) et de l'IPC (axe droit) (*Banque de France* [3] & *INSEE* [27]) 48
2.9 Comparaison Taux de rachat des crédits immobiliers par plage d'années d'ancienneté du crédit et glissement annuel en % de l'IPC 49

3.1 Évolution de la variance cumulée en fonction des composantes principales 74
3.2 Validation croisée de la série temporelle 81

4.1 Schéma du processus de modélisation du taux d'affaires nouvelles 85
4.2 Taux d'affaires nouvelles de crédits immobiliers (*Banque de France* [3]) 86
4.3 Différenciation des Taux d'affaires nouvelles de crédits immobiliers (*Banque de France* [3]) 88
4.4 Matrice de corrélation des variables exogènes retenues 90
4.5 Tendance, ACF et histogramme des résidus du modèle SARIMAX (4,0,1)(0,1,1)[12] sur les trois Bases Trains de la validation croisée 97
4.6 QQ-Plot des résidus du modèle SARIMAX (4,0,1)(0,1,1)[12] sur les trois Bases Trains de la validation croisée 98
4.7 Projections du modèle SARIMAX (4,0,1)(0,1,1)[12] sur la période 2019-2021 . . . 99
4.8 Projections du modèle SARIMAX (4,0,1)(0,1,1)[12] sur la période 2020-2022 . . . 99
4.9 Projections du modèle SARIMAX (4,0,1)(0,1,1)[12] sur la période 2021-2023 . . . 100

5.1 Schéma du processus de modélisation de la loi de rachat 103
5.2 Taux de rachat mensuels en % pour les 20 premières années d'ancienneté du prêt 105

5.3 Projections des modèles ARIMAX (4,0,2)(1,0,2)[12] sur la période 2021-22 114
5.4 Projections des modèles ARIMAX (4,0,2)(1,0,2)[12] sur la période 2022-23 115

6.1 Modèle à états 120
6.2 Scénarios d'inflation 127
6.3 Projection des taux d'affaires nouvelles jusqu'en 2028 128
6.4 Projection des taux de rachat jusqu'en 2028 129

A.1 Répartition de la dépense totale de consommation des ménages (corrigés de l'inflation) en milliards d'euros(*INSEE* [19]) 1
A.2 Taux d'épargne des ménages en % du revenu brut (gauche) et son glissement annuel en % (*INSEE* [28]) 1

B.1 Taux de rachat de crédits immobiliers selon l'ancienneté du crédit 4
B.2 Corrélations des variables économiques exogènes avec les taux de rachat des crédits immobiliers selon l'année d'ancienneté du crédit 4

C.1 Boîte à moustache de l'ensemble du jeu de données 5
C.2 Boîte à moustache de l'ensemble du jeu de données normalisé 6

E.1 Tendance, ACF et histogramme des résidus des modèles ARIMAX des 1 à 20 ans d'ancienneté de prêt sur les deux Bases Trains de la validation croisée 13
E.2 QQ-Plot des résidus des modèles ARIMAX des 1 à 10 ans d'ancienneté de prêt sur les deux Bases Trains de la validation croisée 14
E.3 QQ-Plot des résidus des modèles ARIMAX des 11 à 20 ans d'ancienneté de prêt sur les deux Bases Trains de la validation croisée 15

F.1 Schéma du processus de modélisation du GSE 16

Liste des tableaux

1 BE de primes et de sinistres du portefeuille selon la date de projection et le scénario d'inflation . . . ix
2 BE du portefeuille selon la date de projection et le scénario d'inflation . . . ix
3 Portfolio premium and claims BE by projection date and inflation scenario . . . xvi
4 BE portfolio by projection date and inflation scenario . . . xvi

1.1 Cotisations brutes de réassurance 2023 et variation 2023/2022 par organisme . . . 4
1.2 Cotisations brutes de réassurance 2023 et variation 2023/2022 par organisme . . . 5

2.1 Données collectées à partir du benchmark interne d'Accenture . . . 30
2.2 Données collectées auprès de la Banque de France [3].[4] . . . 31
2.3 Données collectées auprès de l'INSEE [18].[20] (Partie 1) . . . 32
2.4 Données collectées auprès de l'INSEE [21].[23].[24].[26].[27].[28].[29].[30] (Partie 2) 33
2.5 Données finales retenues pour la modélisation . . . 37
2.6 Statistiques descriptives des taux d'affaires nouvelles en France . . . 39
2.7 Corrélations des variables économiques exogènes[1] avec les taux d'affaires nouvelles 44

3.1 Valeurs critiques du test ADF pour les 3 différents marchés par niveaux de significativité[1]. (*INSEE* [16]) . . . 62
3.2 Valeurs critiques du test KPSS pour un marché avec et sans dérive par niveaux de significativité. (*INSEE* [16]) . . . 64
3.3 Valeurs critiques des tests ADF (à gauche) et KPSS (à droite) sur les CP différenciées 74
3.4 Valeurs critiques et p-valeurs du test ADF sur les CP différenciées . . . 74
3.5 Valeurs critiques et p-valeurs du test KPSS sur les CP différenciées . . . 75

4.1 Valeurs critiques du test ADF sur la variable Y . . . 87
4.2 Valeurs critiques du test KPSS sur la variable Y . . . 87
4.3 Valeurs critiques du test ADF sur la série différenciée Z . . . 88
4.4 Valeurs critiques du test KPSS sur la série différenciée Z . . . 89
4.5 Variables économiques exogènes retenues pour la modélisation . . . 91
4.6 Valeurs critiques des tests ADF (à gauche) et KPSS (à droite) sur les variables exogènes . . . 91
4.7 Valeurs critiques et p-valeurs des tests ADF (à gauche) et KPSS (à droite) sur les variables exogènes différenciées . . . 92
4.8 Résultats d'analyse des 4 modèles de l'étude . . . 95
4.9 Valeurs critiques des tests ADF (à gauche) et KPSS (à droite) sur les variables exogènes . . . 96
4.10 Valeurs critiques et p-valeurs des tests ADF (à gauche) et KPSS (à droite) sur les résidus . . . 97

5.1 Valeurs critiques des tests ADF (à gauche) et KPSS (à droite) sur les variables exogènes . . . 105

5.2 Résultats des tests ADF et KPSS pour chaque année d'ancienneté 106
5.3 Moyenne des résultats d'analyse de la validation croisée pour chaque année d'ancienneté 110
5.4 Valeurs critiques des tests ADF (à gauche) et KPSS (à droite) sur les variables exogènes 112
5.5 Valeurs critiques et p-valeurs des tests ADF (à gauche) et KPSS (à droite) sur les résidus par ancienneté 112

6.1 Loi d'entrée en incapacité 122
6.2 Table 2013 du BCAC : Loi de passage d'incapacité à invalidité (*BCAC* [5]) 122
6.3 Loi de décès TH00-02 (*Source [54]*) 123
6.4 Table 2002 du BCAC : Loi de décès en incapacité (*BCAC* [5]) 123
6.5 Table 2013 du BCAC : Loi de maintien en incapacité (*BCAC* [5]) 124
6.6 Table 2013 du BCAC : Loi de maintien en invalidité (*BCAC* [5]) 125
6.7 Exemple de mensualisation des maintiens en invalidité en nombre 125
6.8 Loi de maintien en invalidité 126
6.9 Client "moyen" du portefeuille 130
6.10 Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 2% 131
6.11 Focus sur le rachat et les affaires nouvelles des tableaux d'effectif selon les années et scénarios d'inflation 133
6.12 Mensualisation de la courbe des taux EIOPA annuelle 135
6.13 Table des taux "forwardés" 136
6.14 Table des déflateurs 136
6.15 BE de primes du portefeuille selon la date de projection et le scénario d'inflation 141
6.16 Comparaison des BE de primes du portefeuille selon la date de projection et le scénario d'inflation à 2% 141
6.17 Focus sur les effectifs d'incapables et d'invalides selon les années et scénarios d'inflation 143
6.18 BE de sinistres du portefeuille selon la date de projection et le scénario d'inflation 144
6.19 Comparaison des BE de sinistres du portefeuille selon la date de projection et le scénario d'inflation à 2% 145
6.20 BE du portefeuille selon la date de projection et le scénario d'inflation 145
6.21 Comparaison des BE du portefeuille selon la date de projection et le scénario d'inflation à 2% 146

B.1 Lexique des variables 3

E.1 Résultats d'analyse de la Base Train 2021-22 pour chaque année d'ancienneté . . 11
E.2 Résultats d'analyse de la Base Train 2022-23 pour chaque année d'ancienneté . . 12

F.1 Résultats d'analyse des validations croisées des Sarimax des variables économiques retenus pour le GSE 17
F.2 Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 0% 17
F.3 Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 0.65% 18
F.4 Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 1.3% 18
F.5 Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 3% 19

F.6 Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 4% 19
F.7 Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 5% 20

Introduction

Les fluctuations économiques actuelles, marquées par une inflation persistante et des variations importantes des taux d'intérêt, représentent un défi majeur pour les assureurs. Ces transformations économiques, en constante évolution, modifient non seulement le cadre dans lequel évolue le secteur de l'assurance emprunteur, mais influencent également le comportement des assurés. Les décisions de souscription, de maintien ou de rachat de contrats sont ainsi fortement affectées, rendant les méthodes traditionnelles de gestion et d'évaluation parfois obsolètes.

Dans ce contexte, il devient essentiel de développer des outils et des approches capables d'anticiper ces changements pour préserver la stabilité et la rentabilité des portefeuilles. Ce mémoire s'inscrit dans cette perspective en explorant les interactions complexes entre les conditions économiques et les paramètres clés d'un portefeuille d'assurance emprunteur.

Après une présentation des spécificités de ce type d'assurance et du cadre économique français, une attention particulière est portée aux méthodologies permettant d'interpréter et de modéliser ces interactions. La modélisation des séries temporelles constitue l'axe central de cette démarche, en offrant un cadre robuste pour analyser et prédire des phénomènes dynamiques comme les taux d'affaires nouvelles et les lois de rachat.

Enfin, cette réflexion se concrétise par une étude de cas appliquée à un portefeuille fictif d'assurance emprunteur. À travers divers scénarios économiques, centrés sur les variations de l'inflation, cette étude évalue l'impact de ces changements sur la projection des engagements futurs, mesurés par les Best Estimates (BE). Ces travaux visent à fournir des perspectives nouvelles et des outils concrets aux acteurs de l'assurance, pour les aider à affiner leurs stratégies face à un environnement de plus en plus incertain.

Chapitre 1

Cadre de l'étude

Ce chapitre introductif établit le cadre essentiel pour comprendre l'assurance emprunteur et son interaction avec le contexte économique actuel en France. Il commence par définir les notions clés de l'assurance emprunteur, en abordant ses principes généraux, ses spécificités, les lois et règlements en vigueur, ainsi que les évolutions et tendances du marché. Ensuite, le chapitre explore le contexte économique français, en examinant la croissance économique, les liens entre les taux d'intérêt, l'inflation et la dette, le marché de l'emploi, et les facteurs géopolitiques. Cette exploration vise à clarifier les enjeux économiques qui affectent l'assurance emprunteur.

1.1 L'assurance emprunteur

L'assurance emprunteur est un élément capital dans le processus d'obtention d'un prêt, particulièrement pour les crédits immobiliers. Elle offre une protection à la fois pour l'emprunteur et pour l'établissement prêteur en garantissant le remboursement du prêt en cas d'événements imprévus tels que le décès, l'invalidité ou l'incapacité de travail de l'emprunteur. Bien qu'elle ne soit pas légalement obligatoire, l'assurance emprunteur est souvent exigée par les banques comme condition d'octroi du prêt. Son rôle, ses mécanismes et son cadre réglementaire ont considérablement évolué ces dernières années, reflétant son importance croissante dans le paysage financier français.

1.1.1 Principes généraux de l'assurance emprunteur

L'assurance emprunteur, pilier essentiel de la sécurisation des prêts immobiliers, repose sur des principes généraux qui définissent ses garanties et ses modalités de distribution, assurant ainsi une protection optimale pour les emprunteurs et les prêteurs.

1.1.1.1 Définition de l'assurance emprunteur

L'assurance emprunteur couvre divers types de prêts : les prêts à la consommation, immobiliers et professionnels. Les prêts personnels ou prêts à la consommation sont destinés à financer des projets de petite envergure, comme l'achat d'une voiture, d'électroménager ou des travaux de rénovation. Les prêts immobiliers, quant à eux, sont spécifiquement conçus pour financer l'acquisition d'un bien immobilier, que ce soit une maison, un appartement ou un terrain. Ils sont généralement accordés sur des durées plus longues. Enfin, les prêts professionnels s'adressent aux entreprises et permettent de financer divers investissements nécessaires à leur développement, tels que l'achat de matériel, le lancement de campagnes publicitaires mais également pour financer l'achat d'un bien immobilier à usage professionnel, comme un local commercial, un bureau, ou un entrepôt.

Ce mémoire se focalise sur l'assurance emprunteur pour le prêt immobilier. Cette assurance

offre une protection multiple. Il s'agit d'une assurance limitée à la durée du prêt, qui protège l'emprunteur, ses héritiers et l'établissement prêteur en assurant le remboursement du prêt en cas de réalisation des risques définis dans le contrat. En cas de décès de l'emprunteur, l'assureur rembourse le capital restant dû et les intérêts courus à la banque.

1.1.1.2 Les garanties de l'assurance emprunteur

L'assurance emprunteur est généralement exigée par les établissements de crédit pour accorder un prêt immobilier, bien que ce ne soit pas une obligation légale en France. Les garanties Décès et Perte Totale et Irréversible d'Autonomie (PTIA) sont systématiquement requises, tandis que les garanties d'Incapacité Temporaire de Travail (ITT) et d'Invalidité Permanente Totale (IPT) sont souvent demandées pour les achats d'une résidence principale ou secondaire. Les garanties d'Invalidité Permanente Partielle (IPP) et de perte d'emploi ne sont quant à elles pas systématiquement exigées.

- **Garantie décès**

La garantie décès constitue un élément fondamental de l'assurance emprunteur, faisant partie de la couverture minimale pour tout prêt immobilier. Elle assure le remboursement du Capital Restant Dû (CRD) à la banque en cas de décès de l'emprunteur, protégeant ainsi les héritiers ou co-emprunteurs de la charge financière assurée du crédit. Le fonctionnement de cette garantie est encadré par plusieurs conditions. Elle prend en charge le CRD au moment du décès, selon le montant assuré et la quotité[1] choisie. Toutefois, certaines causes du décès, mentionnées dans les clauses d'exclusion, ne sont pas couvertes, comme la pratique de sports extrêmes ou le suicide durant la première année du contrat. De plus, la garantie s'applique jusqu'à un âge limite défini par le contrat, généralement entre 60 et 80 ans, et un montant maximum de capital assurable est souvent fixé. Enfin, certains contrats peuvent inclure une période initiale, nommée franchise, durant laquelle la garantie n'est pas effective, bien que cela soit rare pour la garantie décès. Cette structure permet à l'assureur de gérer efficacement les risques tout en offrant une protection essentielle aux emprunteurs et à leurs proches.

- **Garantie Perte Totale et Irréversible d'Autonomie (PTIA)**

La garantie Perte Totale et Irréversible d'Autonomie (PTIA) est un élément clé de l'assurance emprunteur, complémentaire à la garantie décès. Elle couvre les cas d'invalidité grave où l'assuré ne peut plus exercer d'activité professionnelle et nécessite une assistance permanente pour les actes quotidiens. En cas de reconnaissance de l'état de PTIA, l'assureur prend en charge le remboursement du CRD, de manière similaire à la garantie décès. Cette prise en charge est basée sur le montant du prêt au jour de la consolidation de l'état de PTIA, tel que défini dans le tableau d'amortissement, et dans la limite du capital assuré. L'évaluation de la PTIA est réalisée par l'assureur selon ses propres critères, qui peuvent différer de ceux de la sécurité sociale. Cette évaluation est généralement effectuée par des médecins experts mandatés par l'assureur. Par conséquent, une reconnaissance de PTIA par la sécurité sociale ne garantit pas automatiquement la prise en charge par l'assureur. La garantie PTIA est généralement soumise à des conditions d'âge et peut comporter certaines exclusions spécifiques, détaillées dans le contrat d'assurance.

- **Garantie Incapacité Temporaire de Travail (ITT)**

La garantie Incapacité Temporaire de Travail (ITT) couvre les échéances du prêt en cas d'arrêt de travail temporaire dû à une maladie ou un accident. L'assureur prend en charge les mensualités du prêt après une période de franchise qui commence à la date de survenance du sinistre. Cette

[1] La quotité en assurance emprunteur désigne le pourcentage du capital emprunté qui est couvert par l'assurance. Elle représente la part du prêt immobilier que l'assureur prendra en charge en cas de sinistre, tel qu'un décès, une invalidité ou une perte d'autonomie.

période de franchise peut varier selon les contrats mais est souvent fixée à 30, 60, 90, ou 180 jours. La prise en charge des échéances cesse dès la reprise du travail, le passage à la retraite ou le décès de l'assuré (cas où la garantie décès s'active). En cas de mi-temps thérapeutique, l'assureur peut continuer à verser une partie des remboursements pendant une période limitée, souvent 6 mois. Les modalités d'indemnisation et les conditions spécifiques peuvent varier selon les contrats.

- **Garantie Invalidité Permanente Totale (IPT)**

La garantie Invalidité Permanente Totale (IPT) intervient lorsque l'assuré est reconnu en invalidité permanente totale, c'est-à-dire une incapacité définitive à exercer une activité professionnelle à la suite d'une maladie ou d'un accident. Pour être considéré en IPT, l'assuré doit présenter une invalidité fonctionnelle et professionnelle avec un taux d'invalidité supérieur ou égal au barème médical mentionné dans le contrat d'assurance. L'assureur prend alors en charge le capital assuré, de manière similaire à la garantie ITT, basé sur le montant du prêt au jour de la consolidation de l'état d'invalidité. Cette garantie est souvent soumise à des conditions d'âge et peut comporter des exclusions spécifiques, qui sont définies dans le contrat d'assurance.

- **Garantie Perte d'Emploi**

La garantie perte d'emploi est une option de l'assurance emprunteur qui couvre les échéances du prêt en cas de licenciement involontaire, excluant les licenciements pour faute grave et les démissions. Pour en bénéficier, l'assuré doit être en contrat à durée indéterminée (CDI) et avoir une ancienneté minimale, définie au contrat dans l'entreprise. Cette garantie est soumise à des conditions spécifiques :

— Un délai de carence de 6 à 12 mois après la souscription ;

— Un délai de franchise de 3 à 6 mois après la perte d'emploi ;

— Une indemnisation limitée dans le temps, généralement entre 12 et 24 mois ;

— Un montant pris en charge plafonné à un pourcentage des mensualités du prêt, typiquement entre 30 % et 80 %.

1.1.1.3 La distribution des contrats de l'assurance emprunteur

- **Les acteurs**

Le marché de l'assurance emprunteur en France est dominé par les bancassureurs, cependant les assureurs alternatifs se font lentement une place sur le marché . Voici un aperçu, selon le rapport de l'Argus [1], des principaux acteurs et de leur classement sur le marché établi sur la base des cotisations encaissées en France, en brut de réassurance, y compris les acceptations, arrêtés au 31/12/2023 :

Bancassureurs (environ 85% du marché) :

Rang	Organisme	Cotisations brutes de réassurance 2023 (en millions d'euros)	Variation 2023/2022
1	Crédit Agricole Assurances	2 531.0	4.2 %
2	CNP Assurances	2 344.0	-13.1 %
3	Groupe des Assurances du Crédit Mutuel	1 802.0	5.3 %
4	BPCE Assurances	1 076.0	6.6 %
5	BNP Paribas Cardif	894.0	3.1 %

TABLE 1.1 – Cotisations brutes de réassurance 2023 et variation 2023/2022 par organisme

Assureurs alternatifs (environ 15% du marché) :

Rang	Organisme	Cotisations brutes de réassurance 2023 (en millions d'euros)	Variation 2023/2022
1	Axa France	449.0	-1.8 %
2	Generali Vie	367.0	15.4 %
3	Malakoff Humanis	155.8	0.3 %
4	Allianz France	150.8	6.8 %
5	MetLife	123.3	2.2 %
6	Swiss Life	104.2	1.8 %

TABLE 1.2 – Cotisations brutes de réassurance 2023 et variation 2023/2022 par organisme

Il est important de noter que malgré les efforts législatifs pour ouvrir le marché à la concurrence (à travers les lois *Lagarde*, *Hamon*, *Bourquin*, et plus récemment *Lemoine* qui sont abordées plus amplement par la suite dans ce mémoire), les bancassureurs conservent une position dominante. Cependant, les assureurs alternatifs gagnent progressivement des parts de marché, notamment grâce à des offres souvent plus compétitives et personnalisées. La part des assurances externes (hors bancassureurs) serait ainsi passée de 15.3% en 2021 à environ 16% en mai 2023, montrant une légère progression mais pas de bouleversement majeur du marché.

- **Les types de contrats**

Deux types de contrats sont disponibles sur le marché : les contrats groupes ou collectifs et les contrats individuels.

Contrats groupes

Les contrats d'assurance emprunteur de groupe sont des contrats collectifs à adhésion facultative souvent proposés par les établissements prêteurs. Lorsqu'ils sont proposés par les établissement bancaires, ces derniers négocient ces contrats avec un assureur, qui peut être externe ou une filiale du groupe bancaire, et les offrent à l'ensemble de leurs clients emprunteurs. En contrepartie de procédures simplifiées, l'assureur verse des commissions à l'établissement prêteur. L'adhésion à ces contrats est simplifiée : l'emprunteur n'a qu'un seul interlocuteur pour le prêt et l'assurance, avec des mensualités regroupées. Une évaluation médicale reste nécessaire, sous forme de questionnaire ou d'examen, selon le montant du prêt depuis la loi *Lemoine*. La tarification des contrats groupe est standardisée, mutualisant les risques entre tous les assurés. Le taux de prime, fixé principalement en fonction de l'âge à la souscription et du capital initial, tend à favoriser les assurés plus âgés au détriment des plus jeunes. Cette faible segmentation tarifaire permet une forte mutualisation des risques entre différents profils d'assurés. Ainsi, les contrats groupe offrent souvent des tarifs plus avantageux aux profils considérés comme plus risqués, grâce à la présence de profils moins risqués dans l'ensemble d'assurés.

Contrats individuels

Les contrats d'assurance emprunteur individuels, distincts des contrats groupes, lient directement les assurés aux assureurs, indépendamment de la banque prêteuse. Les primes des contrats individuels sont calculées en fonction des caractéristiques spécifiques de l'emprunteur, telles que l'âge, le sexe, la catégorie socio-professionnelle (CSP), le statut de fumeur, et la pratique de sports à risque. Ces contrats sont souvent commercialisés par des intermédiaires en assurances, comme les courtiers, qui ciblent les "bons risques". Les contrats individuels offrent des garanties mieux adaptées aux profils de risques favorables, comme les non-fumeurs de moins de 40 ans, offrant des tarifs souvent plus avantageux que les contrats collectifs. Ces contrats sont également une solution pour les personnes présentant des risques élevés qui ne parviennent pas à souscrire à une assurance collective. Bien que les contrats groupe tentent de segmenter leurs tarifs pour rester compétitifs, des différences significatives subsistent, rendant les contrats individuels plus attractifs pour ceux présentant des risques spécifiques liés à leur état de santé ou à leur profession.

1.1.2 Aspects spécifiques de l'assurance emprunteur

Dans le cadre de l'assurance emprunteur, les aspects spécifiques tels que le remboursement anticipé et la pluralité d'emprunteurs jouent un rôle crucial, influençant la manière dont les contrats sont structurés et gérés pour répondre aux besoins variés des emprunteurs.

1.1.2.1 Remboursement anticipé

Une particularité des prêts immobiliers est la possibilité pour l'emprunteur de rembourser son prêt par anticipation, que ce soit partiellement ou totalement. Ce remboursement anticipé survient souvent lors de la vente du bien immobilier, d'un rachat de crédit (qui implique la signature d'un nouveau contrat de prêt et la résiliation de l'assurance liée à l'ancien crédit), ou en cas d'amélioration des capacités financières de l'emprunteur. Toutefois, ce remboursement peut entraîner le paiement d'indemnités prévues dans le contrat de prêt pour compenser le manque à gagner de la banque sur les intérêts futurs non perçus. Un remboursement partiel modifie le tableau d'amortissement à la date de la renégociation, réduisant le montant des échéances et/ou raccourcissant la durée du prêt. En revanche, un remboursement total met fin au contrat de prêt et à l'assurance associée. L'assureur doit alors gérer le risque de remboursement anticipé, qui peut augmenter les taux de résiliation de son portefeuille d'assurés. Ce risque de remboursement (que nous nommerons risque de rachat dans la suite du mémoire) se divise en deux catégories : le rachat conjoncturel, qui dépendent de facteurs externes dynamiques tels que les taux d'intérêt du marché, les rendements espérés par les assurés, et la fiscalité liée à l'ancienneté des contrats, et le rachat structurel, lié aux caractéristiques stables des contrats et des assurés, indépendamment du marché.

1.1.2.2 Pluralité d'emprunteurs

Une autre particularité de l'assurance emprunteur intervient dans le cadre d'un co-emprunt. Chaque co-emprunteur a la possibilité de souscrire sa propre assurance, ce qui permet d'adapter les garanties et les quotités en fonction de la situation personnelle de chacun. La quotité représente le pourcentage du capital emprunté couvert par l'assurance pour chaque co-emprunteur ; deux co-emprunteurs peuvent donc choisir une quotité chacun qui leur est propre. Cette flexibilité permet de choisir des niveaux de garanties différents selon les besoins spécifiques, comme une couverture plus étendue pour un co-emprunteur ayant un emploi à risque. En cas de sinistre touchant l'un des co-emprunteurs, l'assurance interviendra selon la quotité et les garanties souscrites. Cependant, même si chaque co-emprunteur a sa propre assurance, ils restent solidaires pour le remboursement du prêt, ce qui signifie que si l'un d'eux n'est pas ou insuffisamment assuré, l'autre demeure responsable du remboursement total. Il est donc essentiel pour les co-emprunteurs d'évaluer soigneusement leurs besoins individuels et collectifs en matière d'assurance emprunteur afin d'assurer une protection adéquate tout en optimisant le coût global de l'assurance.

1.1.3 Principales lois et règlements en vigueur

Le marché de l'assurance emprunteur en France a évolué ces dernières années, passant d'une domination des bancassureurs et de leurs contrats collectifs à une ouverture accrue à la concurrence. Cela a été rendu possible grâce à plusieurs lois visant à réduire les coûts pour les consommateurs en élargissant leurs choix. Ces réformes ont également renforcé les droits des emprunteurs en leur facilitant les démarches de résiliation et de souscription. Pour comprendre cette évolution, il est essentiel d'examiner les réglementations récentes et leur impact sur les relations entre bancassureurs, assureurs alternatifs et emprunteurs, ainsi que sur la protection des consommateurs dans ce secteur.

1.1.3.1 La convention AERAS (2006)

La convention « s'Assurer et Emprunter avec un Risque Aggravé de Santé » (AERAS), signée le 6 juillet 2006, a été mise en place pour faciliter l'accès à l'assurance emprunteur pour les personnes présentant un risque aggravé de santé. Elle succède à la convention Belorgey de 2001, en corrigeant ses défauts et en élargissant son champ d'application. Cette démarche a débuté en 1991 avec une première convention pour les personnes séropositives et a évolué pour inclure une plus large gamme de pathologies.

Le processus de rachat des demandes d'assurance pour les personnes à risque se déroule en trois niveaux :

1. Le premier niveau concerne l'évaluation standard, où l'assureur examine le questionnaire médical et peut accepter ou refuser la demande ;
2. Si la demande est refusée, c'est-à-dire qu'il y a une présence de risque aggravé de santé, elle est automatiquement transmise au deuxième niveau, où les services médicaux des assureurs examinent le dossier. À ce stade, des surprimes ou des exclusions partielles peuvent être appliquées ;
3. Si le dossier est de nouveau refusé, il est transmis au troisième niveau, où un pool de réassureurs[1] étudie la demande.

Ce processus permet aux assureurs de s'engager à examiner tous les dossiers, à condition que le montant emprunté soit inférieur à 320 000€ et que l'emprunteur ait moins de 70 ans à la fin du prêt.

La convention AERAS a été renouvelée en 2011 pour intégrer des évolutions thérapeutiques, uniformiser les questionnaires de santé, et introduire des mesures pour limiter les surprimes, notamment pour les emprunteurs à revenus modestes. En pratique, 90% des emprunteurs trouvent une assurance dans des conditions standards, sans surprimes ni exclusions. Toutefois, seulement 20% des dossiers au troisième niveau reçoivent une proposition d'assurance.

La convention inclut également des dispositifs tels que la grille de référence AERAS, qui liste les pathologies pour lesquelles un accès à l'assurance emprunteur est accordé sans surprime après un certain délai sans rechute, et le droit à l'oubli, permettant aux anciens malades de ne pas déclarer certaines pathologies après un délai déterminé. Ces mesures visent à garantir un accès équitable à l'emprunt pour les personnes ayant des antécédents médicaux, tout en maintenant un équilibre entre la protection des emprunteurs et la gestion des risques par les assureurs.

1.1.3.2 La loi Lagarde (2010)

La loi Lagarde[2], entrée en vigueur le 1er septembre 2010, a été conçue pour renforcer la concurrence sur le marché de l'assurance emprunteur en France, en offrant aux emprunteurs la possibilité de choisir leur assurance de prêt, à condition que les garanties soient équivalentes à celles des contrats de groupe proposés par les banques. Précédemment, les emprunteurs avaient déjà la possibilité de souscrire auprès d'une assurance différente de celle proposée par la banque, cependant la banque était en droit d'imposer à l'emprunteur le contrat groupe qu'elle proposait. Suite à cette loi, les banques voulant toujours imposer leurs contrats de groupe doivent motiver

[1]Pool de réassureurs : regroupement de réassureurs partageant des risques spécifiques pour augmenter leur capacité à couvrir des pertes importantes. Ce regroupement permet une répartition des primes et des sinistres proportionnellement à la participation des acteurs, ce qui permet de gérer des risques complexes ou de grande ampleur.

[2]Loi Lagarde : Loi n°2010-737 du 1er juillet 2010

par écrit leur refus de délégation. De plus il est interdit aux banques de modifier les conditions du prêt en cas de choix d'une assurance externe, par exemple modifier les taux d'emprunt. Enfin, elle a introduit une fiche d'information standardisée pour faciliter la comparaison des offres.

Bien que la loi Lagarde ait pour objectif de favoriser la concurrence, son impact sur le marché de l'assurance emprunteur a été limité. La part des contrats en délégation d'assurance n'a que peu évolué, en partie en raison de la complexité des contrats, du manque d'information des emprunteurs, et de la prévalence des contrats de groupe proposés par les banques. Ces facteurs ont conduit à une stagnation, voire un recul, de l'utilisation des délégations d'assurance, incitant à l'adoption de nouvelles mesures législatives pour pallier les insuffisances de cette loi.

1.1.3.3 La loi Hamon (2014)

La loi Hamon[1], entrée en vigueur le 26 juillet 2014, vise à renforcer les droits des consommateurs en matière d'assurance emprunteur, en complément de la loi Lagarde. Son principal apport est de permettre aux emprunteurs de résilier à tout moment leur assurance de prêt immobilier dans les 12 mois suivant la signature du contrat, sans frais, à condition de souscrire une nouvelle assurance offrant des garanties à minima équivalentes.

Cette disposition cherche à remédier au faible recours à la délégation d'assurance instaurée par la loi Lagarde, en accordant aux emprunteurs davantage de temps pour comparer les offres et choisir l'assurance qui correspond le mieux à leurs besoins. L'objectif est de stimuler la concurrence sur le marché de l'assurance emprunteur, afin de réduire les coûts et d'améliorer le pouvoir d'achat des emprunteurs.

En outre, la loi Hamon impose de nouvelles obligations d'information, telles que l'affichage du taux annuel effectif de l'assurance (TAEA)[2], ainsi que la fourniture d'une fiche d'information standardisée dès la première simulation, dans le but de faciliter la comparaison des offres d'assurance. Elle établit également un délai de 10 jours ouvrés pour que la banque accepte ou refuse le nouveau contrat d'assurance, tout en interdisant sa capacité à modifier les conditions du prêt ou à résilier l'assurance en cas d'aggravation du risque sans l'accord de l'emprunteur.

1.1.3.4 L'amendement Bourquin (2017)

L'amendement Bourquin[3], signé le 1er mars 2017, renforce davantage les droits des emprunteurs en matière d'assurance de prêt en leur permettant de résilier leur contrat chaque année à la date anniversaire de la signature de l'offre de prêt. Cette résiliation doit être notifiée par lettre recommandée au moins deux mois avant la date d'échéance, et l'emprunteur peut ensuite souscrire un contrat alternatif à condition que les garanties soient de nouveau au moins équivalentes à celles du contrat initial.

Cette mesure a profondément bouleversé le marché de l'assurance emprunteur, incitant les acteurs, anciens et nouveaux, à s'adapter aux changements. Elle a également encouragé certains assureurs et mutuelles à lancer des appels d'offres afin d'entrer sur ce marché en pleine évolution. En somme, l'amendement Bourquin s'inscrit dans une démarche visant à accroître la concurrence et à offrir

[1]Loi Hamon : Loi n°2014-344 du 17 mars 2014

[2]Le Taux Annuel Effectif d'Assurance (TAEA) est un indicateur qui permet de calculer le coût total de l'assurance liée à un prêt immobilier sur une base annuelle.

[3]Amendement Bourquin : Loi n°2017-203 du 21 février 2017

davantage de flexibilité aux emprunteurs, tout en contribuant potentiellement à la réduction des coûts d'assurance.

1.1.3.5 La loi Lemoine (2022)

Dernièrement adoptée le 28 février 2022, la loi Lemoine[1] est entrée en vigueur le 1er juin 2022 pour les nouveaux contrats et le 1er septembre 2022 pour les contrats existants. Cette loi apporte des changements significatifs au marché de l'assurance emprunteur en France et vise à rendre l'accès à l'assurance emprunteur plus juste, plus simple et plus transparent.

Le premier pilier de cette loi est la libéralisation de la résiliation des contrats. Désormais, les emprunteurs peuvent mettre fin à leur assurance de prêt à tout moment, sans frais ni préavis. Cette flexibilité accrue s'applique à tous les contrats liés à des crédits immobiliers pour usage d'habitation ou mixte, à condition que le nouveau contrat offre des garanties équivalentes. L'assureur résilié dispose toujours d'un délai de 10 jours pour effectuer l'avenant nécessaire. Un autre aspect novateur est l'élimination du questionnaire médical pour certains types de prêts. Cette mesure concerne les prêts dont la part assurée ne dépasse pas 200 000€ et dont le terme survient avant le 60ème anniversaire de l'emprunteur. Cette disposition vise à faciliter l'accès au crédit pour les personnes ayant des antécédents médicaux. La loi renforce également la transparence en imposant aux assureurs de nouvelles obligations d'information. Ils doivent désormais communiquer clairement sur le droit de résiliation, les échéances du contrat et le coût total de l'assurance sur 8 ans. Le non-respect de ces obligations peut entraîner une amende de 15 000€ pour l'organisme d'assurance. Enfin, la loi Lemoine réforme le "droit à l'oubli" en réduisant de 10 à 5 ans le délai après lequel les anciens malades du cancer et de l'hépatite C n'ont plus à déclarer leur maladie, indépendamment de l'âge du diagnostic.

Cette mesure, comme les précédentes, visent à stimuler davantage la concurrence sur le marché de l'assurance emprunteur et à faciliter l'accès au crédit immobilier pour un plus large éventail de profils d'emprunteurs. Cependant, certains aspects de la loi, notamment concernant le calcul de l'encours cumulé pour les prêts multiples, nécessitent encore des clarifications.

1.1.4 Évolution et tendances du marché de l'assurance emprunteur en France

Dans un marché en constante évolution, l'assurance emprunteur en France se transforme sous l'effet des réformes législatives et des dynamiques concurrentielles, influençant les encours de crédits, les cotisations, et le paysage réglementaire.

1.1.4.1 Évolution des encours des crédits

Selon le rapport annuel 2023 de France Assureur [14] et comme le montre la *Figure 1.1*, le marché de l'assurance emprunteur est étroitement lié à celui des crédits immobiliers, les encours détenus par les ménages étant nettement plus importants que ceux des crédits à la consommation et des prêts professionnels.

[1]Loi Lemoine : Loi n°2022-270 du 28 février 2022

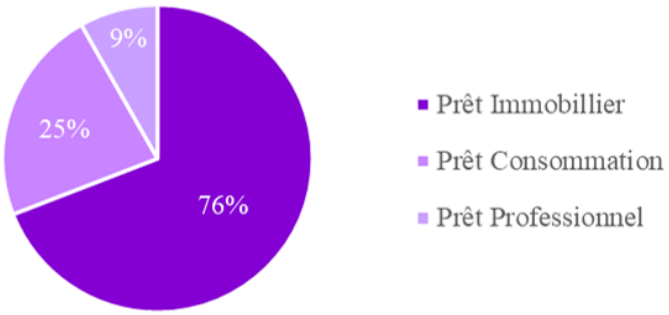


FIGURE 1.1 – Répartition des cotisations des contrats d'assurance emprunteur selon l'objet du prêt en 2023 (*France Assureurs* [14])

Les encours de crédits accordés aux particuliers, tels que rapportés par la Banque de France, ont connu une croissance constante depuis l'an 2000, avec une progression moyenne annuelle de 7.3%. Cependant, en 2023, cette croissance s'est considérablement ralentie, atteignant à peine 1%. Cette évolution des encours de crédits a naturellement entraîné une augmentation mécanique des montants des cotisations perçues par les contrats d'assurance emprunteur.

La *Figure 1.2* illustre cette évolution à travers l'historique des montants d'encours de crédits des ménages depuis 2018, montrant ainsi visuellement cette tendance à la hausse suivie du récent ralentissement.

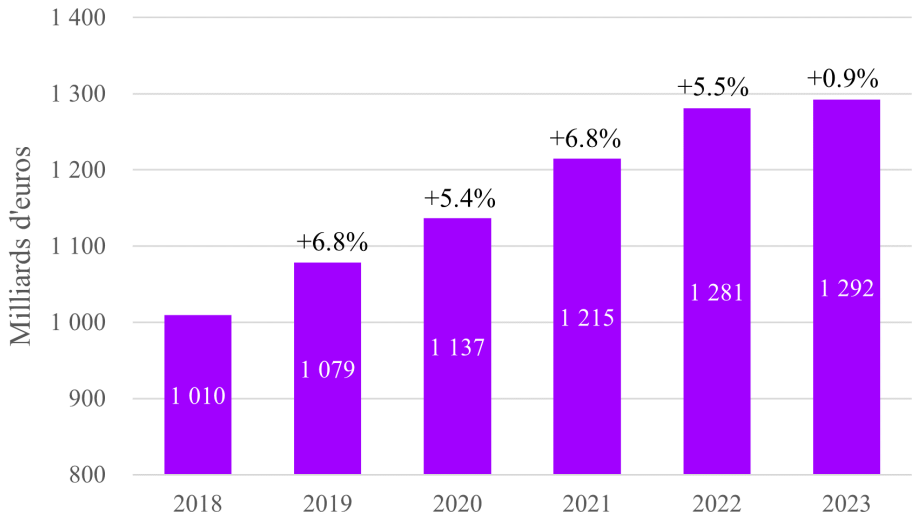


FIGURE 1.2 – Encours des crédits à l'habitat accordés aux particuliers résidents (*Banque de France* [3])

Cette tendance des encours de crédits peut en partie s'expliquer par l'évolution de la quantité de nouveaux crédits accordés aux particuliers. Il est important de noter que les encours incluent à la fois les crédits existants et les nouveaux crédits. Ainsi, une baisse de la quantité de nouveaux crédits se traduit par un ralentissement de la croissance des encours totaux, même si celle-ci reste positive en raison du stock existant. La *Figure 1.3* présente l'évolution annuelle des nouveaux crédits à l'habitat accordés aux particuliers sur le marché français depuis 2018.

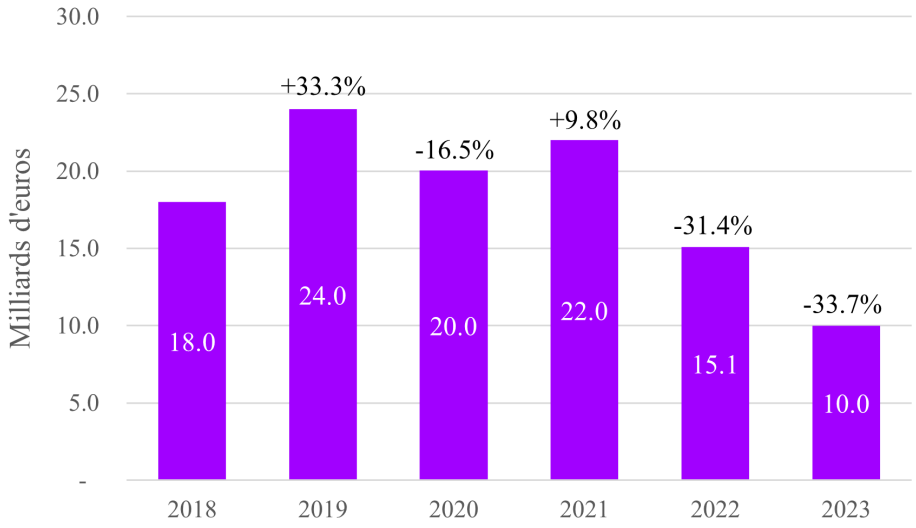


FIGURE 1.3 – Flux annuel des crédits nouveaux à l'habitat aux particuliers résidents (*Banque de France* [3])

1.1.4.2 Évolution des cotisations des crédits

Les cotisations des contrats d'assurance emprunteur ont totalisé 11.8 milliards d'euros sur le marché français en 2023. Ce marché a connu une croissance continue depuis le début des années 2000. Cependant, depuis la pandémie de Covid-19 en 2020 et en conséquence de la baisse de la quantité de nouveaux crédits accordés aux particuliers, un léger ralentissement de cette croissance est notable. La *Figure 1.4* illustre ainsi l'évolution des cotisations des contrats d'assurance emprunteur sur le marché français depuis 2018.

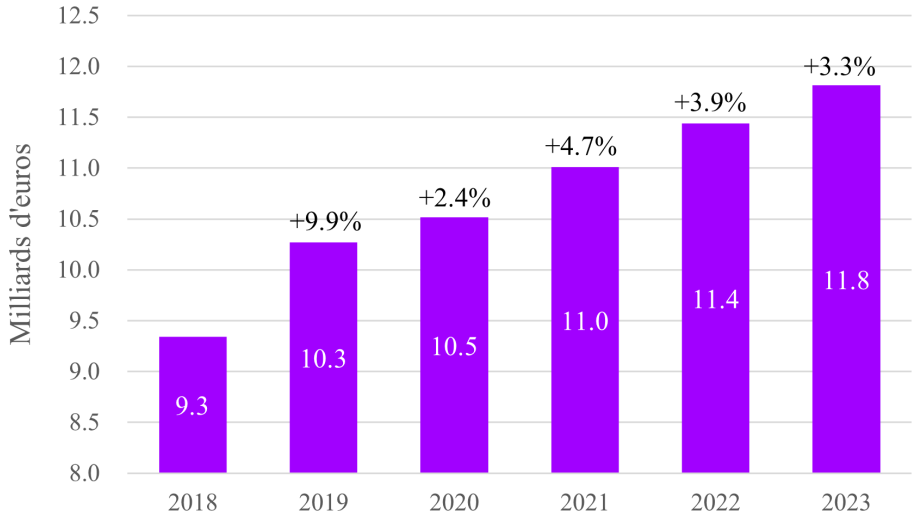


FIGURE 1.4 – Cotisations des contrats d'assurance emprunteur (*France Assureurs* [14])

Pour l'ensemble des types de prêts, les primes associées aux garanties décès représentaient la majeure partie du chiffre d'affaires en 2023, atteignant 68.6%. En comparaison, les garanties incapacité-invalidité constituaient 29.8% du total. La *Figure 1.5* illustre cette répartition des cotisations d'assurance emprunteur selon l'objet du prêt.

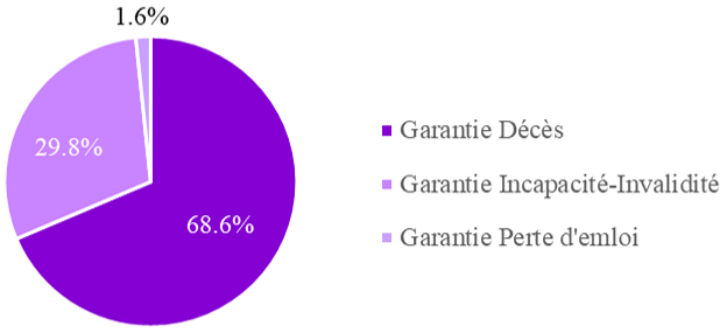


FIGURE 1.5 – Répartition des cotisations des contrats d'assurance emprunteur en 2023 (*France Assureurs* [14])

1.1.4.3 Impacts des réformes sur le paysage de l'assurance

Comme expliqué précédemment dans la partie sur *Les acteurs* de l'assurance emprunteur, le marché de l'assurance emprunteur en France est principalement dominé par les bancassureurs, qui détiennent 84.6% du marché en 2023.

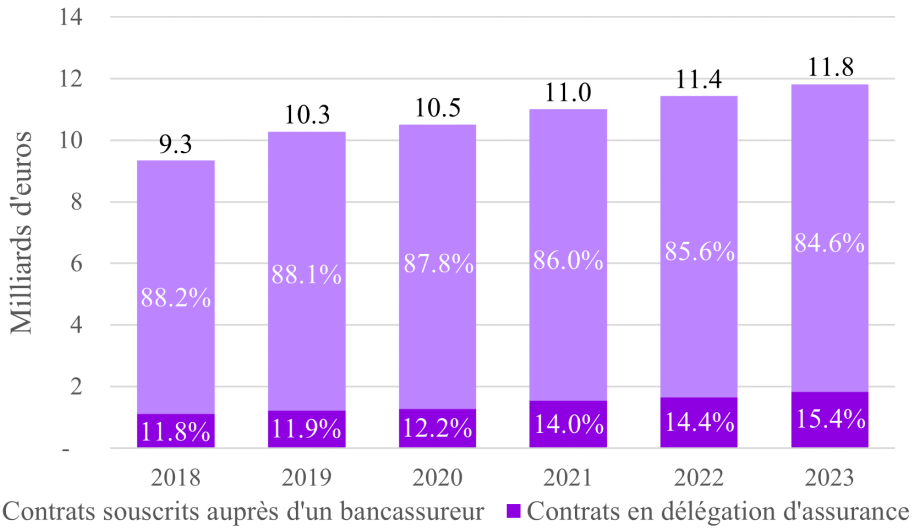


FIGURE 1.6 – Répartition des cotisations des contrats d'assurance emprunteur (*France Assureurs* [14])

L'*Histogramme 1.6* met en évidence une évolution notable du marché de l'assurance emprunteur en France. Il permet de constater une diminution progressive de la part de marché des bancassureurs, tandis que celle des contrats en délégation d'assurance augmente. Cette transformation du paysage de l'assurance emprunteur est le fruit d'une série de mesures législatives adoptées au cours de la dernière décennie, comprenant la *Loi Lagarde* (2010), la *Loi Hamon* (2014), l'*amendement Bourquin* (2017), et plus récemment la *Loi Lemoine* (2022). Malgré ces changements législatifs significatifs, la part de marché des bancassureurs reste prédominante. Cela suggère que, bien que ces lois aient ouvert la voie à plus de concurrence, la transformation d'un marché historiquement dominé par les établissements prêteurs est un processus lent et complexe.

Selon le bilan emprunteur du Comité Consultatif du Secteur Financier (CCSF) de 2023 [6], la loi Lemoine poursuit deux objectifs principaux :

1. simplifier l'accès à la substitution d'assurance pour renforcer la liberté de choix des emprunteurs ;
2. rendre cette substitution accessible au plus grand nombre.

Bien que les données de la *Figure 1.6* montrent que le premier objectif ait été atteint (7% d'augmentation entre 2022 et 2023 pour la part sur le marché des contrats en délégation d'assurance passant de 14.4% à 15.4%), le second peut être évalué en examinant les profils des clients souscripteurs.

La *Figure 1.7* montre que plus de la moitié des crédits immobiliers dans les portefeuilles bancaires sont souscrits par des clients de la CSP2, qui regroupe les employés, techniciens et fonctionnaires de catégorie B. En revanche, 27% des crédits sont détenus par des clients de CSP1, incluant les cadres et professions libérales, tandis que 18% proviennent de la CSP3, qui englobe d'autres professions. Cependant, cette répartition n'est pas reflétée dans les substitutions d'assurance, où les clients de CSP1 sont sur-représentés, représentant 58% des substitutions, tandis que les CSP2 et CSP3 sont sous-représentés. Cela indique que la substitution d'assurance emprunteur bénéficie principalement aux populations les plus aisées.

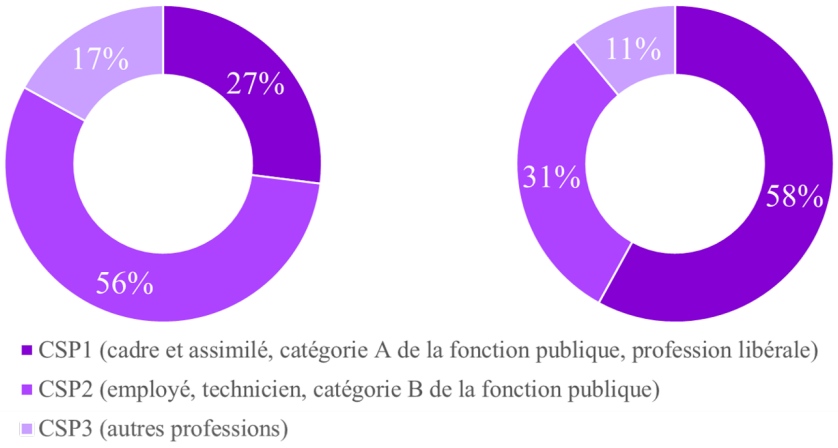


FIGURE 1.7 – Répartition des crédits (gauche) et des substitutions (droite) en portefeuille par CSP (*CCSF* [6])

La disparité observée dans les substitutions d'assurance emprunteur s'explique principalement par la stratégie tarifaire des contrats alternatifs. Ces derniers se différencient essentiellement par leurs prix, utilisant la CSP comme critère de segmentation, contrairement aux contrats bancaires traditionnels. Cette approche avantage particulièrement les clients de la CSP1 (cadres, professions libérales), qui bénéficient souvent de tarifs plus attractifs.

D'après les retours des acteurs du marché ayant participé à l'étude du *CCSF* [6], ces clients aisés sont généralement plus enclins à envisager un changement d'assurance, malgré le risque potentiel de tension avec leur banque, et disposent d'un pouvoir de négociation supérieur. En revanche, les employés et ouvriers (CSP2 et CSP3) sont moins susceptibles de changer d'assurance, non seulement en raison d'un gain tarifaire potentiellement moindre, mais aussi d'une plus grande réticence à modifier leur contrat existant.

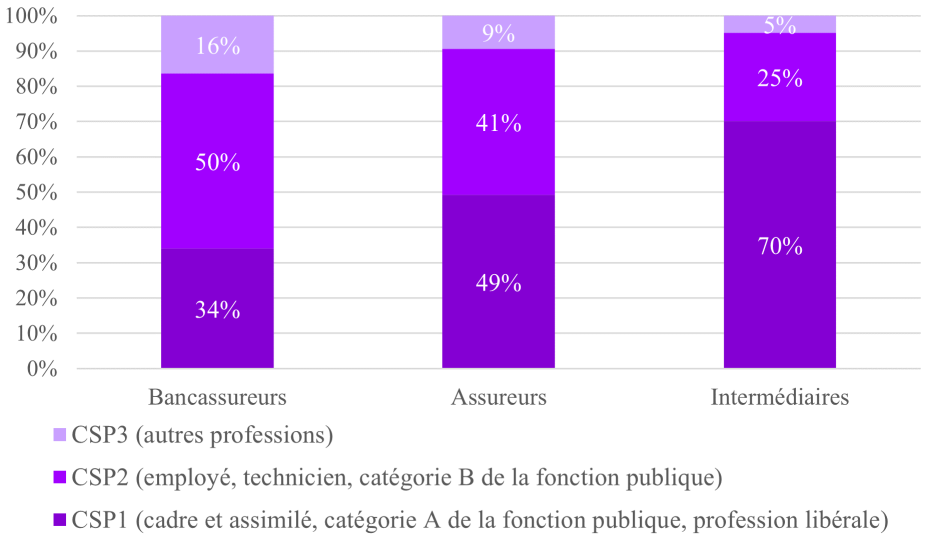


FIGURE 1.8 – Répartition des CSP par acteurs du marché assurantiel (*CCSF* [6])

La répartition des catégories socioprofessionnelles dans les portefeuilles d'assurance emprunteur est influencée par la clientèle et les types de crédits de chaque acteur. La répartition des CSP par acteurs du marché assurantiel de la *Figure 1.8* illustre bien que les contrats alternatifs externes se concentrent souvent sur les crédits à capitaux élevés, qui attirent des emprunteurs plus aisés. Ces prêts génèrent des primes d'assurance plus élevées, ce qui les rend plus rentables pour les distributeurs. De plus, ces emprunteurs sont généralement mieux informés sur les options d'assurance et bénéficient d'écarts tarifaires plus significatifs, renforçant ainsi leur intérêt pour les contrats alternatifs. Par conséquent, la substitution d'assurance révèle une nette sur-représentation des emprunteurs ayant des capitaux assurés élevés. Les crédits entre 200 000€ et 500 000€ représentent 37% des substitutions, alors qu'ils ne constituent que 25% des crédits en portefeuille. Cette disparité est encore plus marquée pour les crédits dépassant 500 000€, qui représentent 10% des substitutions mais seulement 2% des crédits en portefeuille (*Figure 1.9*).

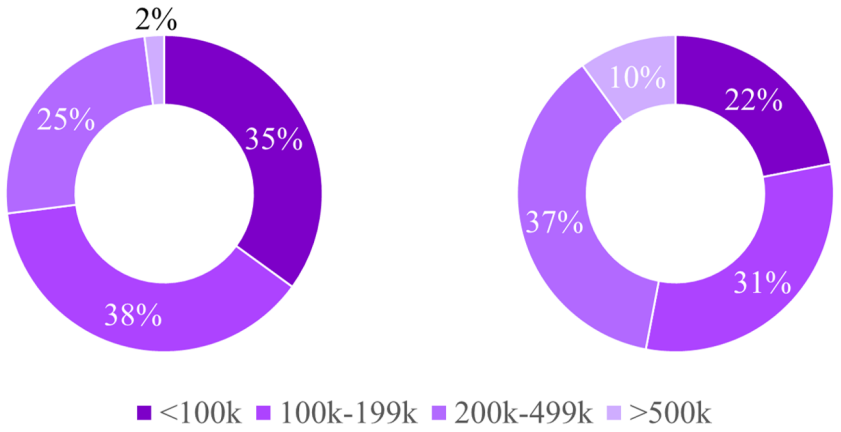


FIGURE 1.9 – Répartition des crédits (gauche) et des substitutions (droite) en portefeuille par capital assuré (*CCSF* [6])

Ces chiffres et conclusions tirés du bilan emprunteur du CCSF de 2023 montrent que, bien que la loi Lemoine ait simplifié l'accès à la substitution d'assurance, cette possibilité profite principalement aux CSP les plus élevées, tandis que les catégories moins aisées restent sous-représentées.

1.1.5 Conclusion

Le marché de l'assurance emprunteur en France connaît une évolution progressive, marquée par des changements législatifs visant à accroître la concurrence et la liberté de choix des consommateurs. Cependant, ces transformations s'inscrivent dans un contexte économique plus large qui influence inévitablement les dynamiques du secteur. Pour mieux comprendre les enjeux et les perspectives de ce marché, il est essentiel d'examiner le contexte économique global de la France et les chiffres clés qui l'illustrent.

1.2 Contexte économique en France

Avec un marché de l'assurance emprunteur en pleine transformation, il est impératif de replacer ces évolutions dans le contexte économique global de la France. La dynamique économique actuelle est influencée par divers facteurs, allant de la croissance du Produit Intérieur Brut (PIB) aux fluctuations de l'emploi et des revenus des ménages. En parallèle, les tensions géopolitiques et les politiques européennes jouent un rôle déterminant dans la structuration de l'économie française. Pour saisir pleinement les enjeux et les perspectives du marché, il est essentiel d'examiner ces éléments économiques clés et leur impact sur le secteur de l'emprunteur.

1.2.1 Facteurs Géopolitiques

Les facteurs géopolitiques jouent un rôle crucial dans l'évolution économique mondiale, exacerbant les incertitudes et influençant directement la stabilité économique. Après le rebond économique post-COVID-19, la guerre en Ukraine et d'autres crises ont assombri les perspectives économiques, posant des défis majeurs pour la croissance et la gestion de l'inflation, en particulier en France et dans les régions proches du conflit.

1.2.1.1 Instabilité politique

Après l'effondrement économique mondial en 2020 dû aux mesures de confinement pour contenir le COVID-19, la production française comme mondiale a connu une reprise économique sans précédent en 2021 (cf *Figure 1.11*), en ligne avec la tendance mondiale. Cette reprise a été soutenue par une augmentation de la demande intérieure (cf *Figure 1.14*) et des exportations, ainsi que par l'assouplissement progressif des restrictions sanitaires et les progrès de la vaccination. Cependant, l'économie française a également vu son endettement public augmenter considérablement (cf *Figure 1.15*) en raison des mesures de soutien budgétaire et monétaire mises en place pour stabiliser l'économie.

En 2022, la France, comme de nombreux autres pays européens, a été confrontée à des défis économiques supplémentaires en raison de la guerre en Ukraine, entraînant à la baisse la croissance mondiale et française à 2.5% pour 2022 et 1.1% pour 2023 (cf *Figure 1.11*). Ces événements ont entraîné une hausse des prix de l'énergie et des perturbations des chaînes d'approvisionnement, ralentissant ainsi la croissance économique et augmentant les pressions inflationnistes. La France doit naviguer entre la gestion de l'inflation et le soutien à la croissance économique dans un contexte de tensions géopolitiques et d'incertitudes économiques, tout en cherchant à maintenir la stabilité économique et à soutenir la reprise.

1.2.1.2 Politiques Monétaires et Fiscales

Les politiques monétaires de la Banque Centrale Européenne (BCE) ont eu un impact significatif sur l'économie française, notamment à travers le resserrement monétaire qui a freiné la croissance. Depuis l'invasion de l'Ukraine par la Russie et comme il est possible de le voir *Figure*

1.10, la BCE a relevé ses taux d'intérêt à plusieurs reprises, passant d'un niveau négatif à un taux record de 4% entre juillet 2022 et septembre 2023. Cette politique a été mise en place pour lutter contre l'inflation, qui avait atteint un pic de 10.6% en octobre 2022 dans la zone euro, avant de redescendre à 2.6% au deuxième trimestre 2024 (cf *Figure 1.13*).

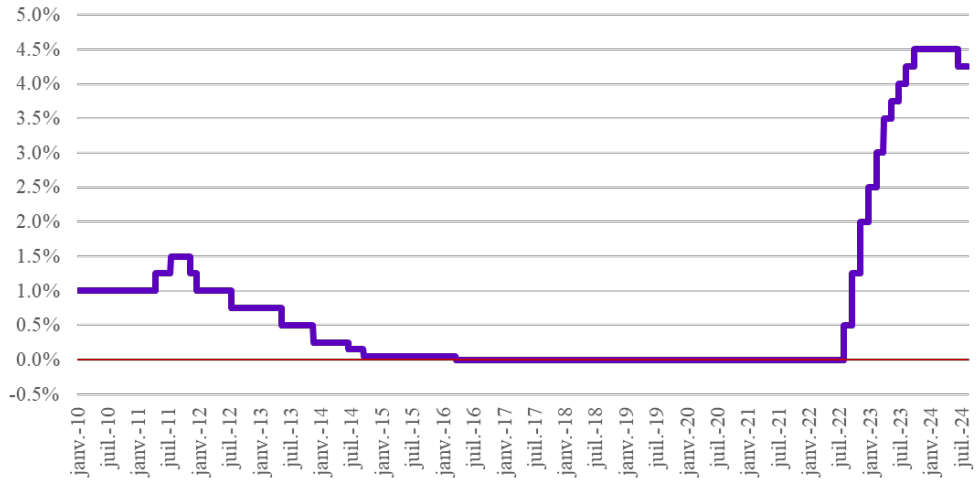


FIGURE 1.10 – Taux directeurs de la BCE (*BCE* [12])

Ce resserrement monétaire a eu pour effet de rendre le crédit plus coûteux, ce qui a freiné les investissements et réduit la demande de nouveaux prêts dans la zone euro et en France (cf *Figure 1.3*). En France, comme dans d'autres pays de la zone euro, cela a contribué à une stagnation de la croissance économique (cf *Figure 1.11*), car les entreprises et les ménages ont réduit leurs dépenses en raison de l'augmentation des coûts d'emprunt.

La BCE a également indiqué que bien que l'impact le plus sévère de sa politique restrictive ait probablement déjà été ressenti, elle prévoit que cette politique continuera à ralentir la croissance au cours des prochains trimestres. En conséquence, la France, comme d'autres économies européennes, doit naviguer entre la gestion de l'inflation et le soutien à la croissance économique dans un contexte de tensions géopolitiques et d'incertitudes économiques.

1.2.2 Croissance Économique

La croissance économique française, pilier fondamental de la prospérité nationale, se situe à l'intersection des facteurs géopolitiques et des politiques monétaires et fiscales évoqués précédemment. Son évolution, liée à des indicateurs tels que le PIB, les taux d'intérêt, l'inflation et la dette publique, impacte directement le secteur de l'emprunt et, par extension, l'assurance emprunteur. Ces indicateurs influencent les conditions de crédit et la demande de prêts, façonnant ainsi le marché de l'assurance emprunteur. Par exemple, les fluctuations des taux directeurs de la BCE et par extensions des taux d'intérêt bancaires affectent le coût des emprunts, stimulant ou freinant la demande de crédits immobiliers et, par conséquent, d'assurances emprunteur. Avant d'explorer cet impact spécifique, il est essentiel de comprendre comment le marché économique fonctionne et s'auto-influence. Les interactions complexes entre les secteurs de l'économie, les politiques gouvernementales et les tendances mondiales créent un écosystème dynamique où chaque élément peut avoir des répercussions sur l'ensemble.

1.2.2.1 Produit Intérieur Brut (PIB)

Le Produit Intérieur Brut (PIB) est un indicateur économique clé qui mesure la valeur totale des biens et services produits par un pays sur une période donnée, généralement un an. Il reflète la santé économique globale d'une nation et sa capacité à générer de la richesse. Le PIB est crucial pour comprendre le contexte économique car il permet d'évaluer la croissance économique, de comparer les performances entre pays, et d'orienter les politiques économiques. Une augmentation du PIB indique généralement une expansion économique, tandis qu'une diminution peut signaler une récession. Ainsi, le PIB influence directement les décisions d'investissement, les politiques gouvernementales, et le marché de l'emploi, ayant des répercussions sur divers secteurs, y compris celui de l'assurance emprunteur.

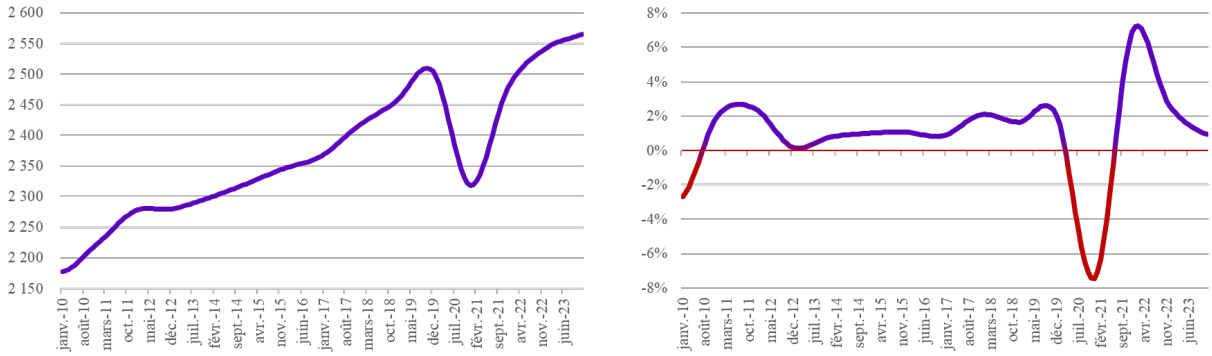


FIGURE 1.11 – Montant en milliards d'euros 2020 (gauche) et glissement annuel en % (droite) du PIB français (*INSEE* [29])

En 2023, la France a enregistré une croissance annuelle du PIB de 1.1%. Cette progression visible en *Figure 1.11*, bien que modeste par rapport aux années précédentes (2.5% en 2022 et 6.4% en 2021), s'inscrit dans un contexte de reprise économique après les impacts sévères de la pandémie de COVID-19. En 2020, le PIB avait chuté de 7.9% en raison des mesures de confinement et des perturbations économiques. Depuis cette chute, l'économie française a progressivement rebondi, atteignant environ 2600 milliards d'euros en 2023, en tenant compte de la valeur de l'euro en 2020 (*Figure 1.11*). Cette croissance de 1.1% montre que l'économie française est sur la voie de la reprise post-pandémie. Cependant, elle demeure relativement faible comparée à la croissance moyenne de 1.58% observée sur les 20 années précédant la pandémie (*INSEE* [29]). Cette performance pourrait indiquer une stagnation ou un ralentissement économique, susceptible d'affecter l'emploi et les revenus à long terme.

Cette situation de reprise économique s'inscrit dans une période complexe de forte inflation, de taux d'intérêt en hausse et d'une dette publique historiquement élevée.

1.2.2.2 Taux d'intérêt

Le taux d'intérêt est le prix qu'il faut payer pour emprunter de l'argent ou la rémunération reçue pour prêter de l'argent. Exprimé en pourcentage du montant emprunté ou prêté, il représente le coût du crédit pour l'emprunteur et le rendement pour le prêteur. Il existe différents types de taux d'intérêt, notamment les taux nominaux (fixés dans le contrat) et les taux réels (corrigés de l'inflation), ainsi que des taux fixes ou variables selon les termes du contrat. Les taux d'intérêt jouent un rôle fondamental dans l'économie, influençant directement le PIB par le biais de la consommation, de l'investissement et de la gestion de la dette publique. Lorsqu'ils augmentent, le coût de l'emprunt pour les ménages et les entreprises s'accroît, ce qui tend à réduire les dépenses de consommation et les investissements. En conséquence, une hausse des taux d'intérêt peut

freiner la croissance économique en diminuant la demande globale, car les prêts deviennent plus coûteux et moins accessibles.

Le différentiel entre le taux d'intérêt et le taux de croissance du PIB est essentiel pour la soutenabilité de la dette publique. Si le taux d'intérêt est supérieur au taux de croissance du PIB, cela peut entraîner une augmentation du ratio dette/PIB, posant des défis pour la stabilité budgétaire à long terme. Un ratio de dette élevé limite la capacité de l'État à investir dans des projets de croissance, ce qui peut freiner davantage l'expansion économique et créer un cercle vicieux.

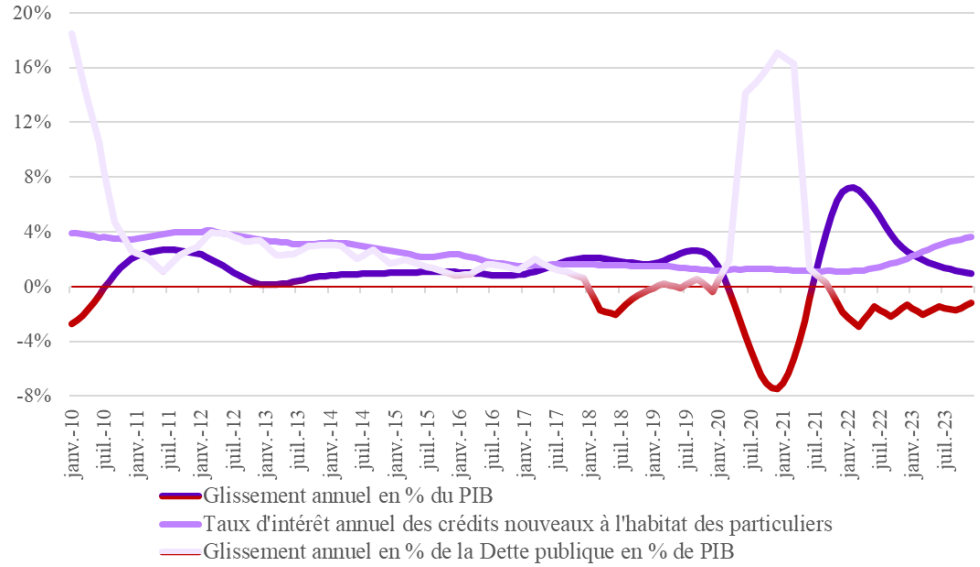


FIGURE 1.12 – Taux de croissance du PIB et de la Dette Publique comparés aux taux d'intérêt annuel des crédits (*Banque de France* [3] & *INSEE* [21].[29])

Les données de l'INSEE et de la Banque de France de la *Figure 1.12* illustrent bien que lorsque le taux d'intérêt dépasse l'évolution du PIB, le ratio dette/PIB tend à augmenter, comme cela a été particulièrement observé lors de la pandémie, où l'évolution du PIB était en déclin. Les taux d'intérêt élevés peuvent ainsi peser lourdement sur l'économie, limitant les projets d'investissement des entreprises et la consommation des ménages, ce qui peut entraîner une contraction économique. Ceux-ci jouent un rôle central dans l'économie, influençant directement le PIB à travers la consommation, l'investissement et la gestion de la dette publique. Le secteur immobilier illustre parfaitement cette dynamique.

Par exemple, *"une hausse des taux directeurs de la BCE de 0.5% peut entraîner une augmentation des taux de prêts immobiliers de 3% à 3.5%. Cette variation apparemment minime réduit significativement le pouvoir d'emprunt des ménages : un emprunt de 200 000€ sur 20 ans à 3% avec des mensualités de 1 109€ ne permettrait plus d'emprunter que 188 000€ à 3.5% pour la même mensualité"* (*Imkiz* [15]). Cette diminution du pouvoir d'achat immobilier a des répercussions en chaîne sur l'économie. Elle entraîne une baisse de la demande de crédits, ralentissant le secteur immobilier et de la construction, ce qui freine la croissance du PIB. Les ménages déjà endettés voient une part plus importante de leurs revenus absorbée par le remboursement de leurs emprunts, réduisant leur consommation dans d'autres secteurs. Les entreprises, face à des coûts d'emprunt plus élevés, peuvent être amenées à réduire leurs investissements.
Cet effet domino souligne l'importance des décisions de politique monétaire et leur impact sur l'ensemble de l'économie à travers le marché de l'emprunt. Il met en évidence la nécessité pour les autorités monétaires de calibrer soigneusement leurs décisions pour maintenir un équilibre

entre croissance économique et stabilité des prix.

1.2.2.3 Inflation

L'inflation est un phénomène économique caractérisé par une hausse généralisée des prix des biens et services sur une période donnée, entraînant une diminution du pouvoir d'achat. Elle se manifeste par une augmentation globale des prix, affectant aussi bien les biens d'usage courant, comme les produits alimentaires et l'énergie, que les biens durables, tels que l'électronique et les voitures, ainsi que les services, comme les loyers et les assurances. Cette hausse des prix peut être causée par divers facteurs :

— Une augmentation de la demande globale qui dépasse l'offre ;

— Des chocs d'offre, tels qu'une hausse des coûts de production ;

— Une expansion excessive de la masse monétaire, typiquement entre 30% et 80%.

L'inflation est généralement mesurée par l'indice des prix à la consommation (IPC), qui suit l'évolution des prix d'un panier de biens et services représentatif de la consommation des ménages. En France, l'INSEE est responsable du calcul de cet indice. L'IPC permet de mesurer les variations du coût de la vie pour les consommateurs et de la valeur de la monnaie dans ses aspects les plus concrets pour les ménages.

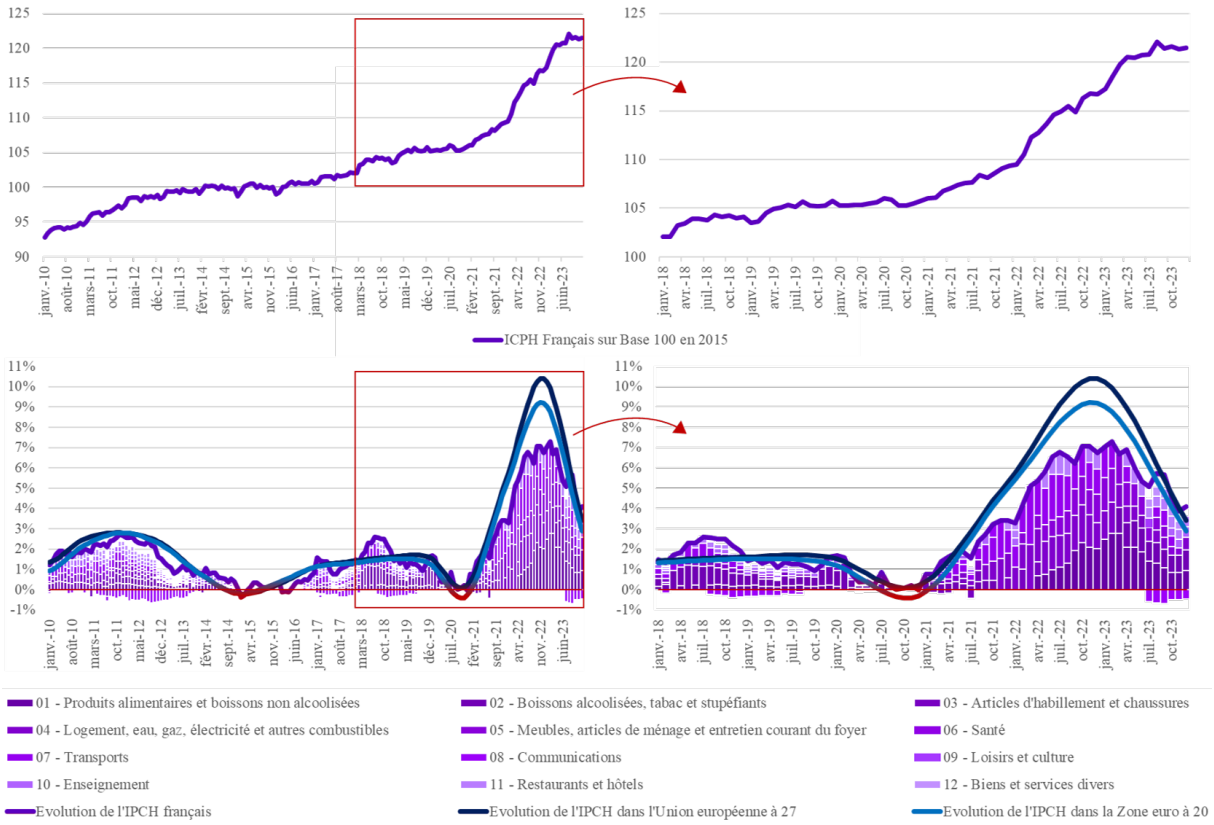


FIGURE 1.13 – IPCH[1] français sur Base 100 en 2015 (haut) et son glissement annuel en % (bas) (*INSEE* [17].[25])

[1] L'indice des prix à la consommation harmonisé (IPCH) est un indicateur économique clé utilisé dans l'Union européenne pour mesurer l'inflation et évaluer la stabilité des prix. Il permet de vérifier le respect des critères de convergence économique établis par le traité de Maastricht pour les pays candidats à l'adoption de l'euro.

Cette *Figure 1.13* ci-dessus de l'IPCH et de son glissement annuel montrent que, sur la période de 2010 à 2023, l'inflation a connu une croissance constante, avec quelques brèves périodes de décroissance, notamment aux premiers trimestres de 2015 et 2016, et une stagnation en 2020 due aux confinements qui ont limité les dépenses des ménages. Il est important de noter que ces périodes de croissance nulle ou négative de l'inflation ne sont pas considérées comme de la désinflation, car elles n'ont pas duré plus de trois trimestres consécutifs. Jusqu'à 2015, l'inflation était globalement généralisée à tous les postes de consommation spécifiés dans la *Figure 1.13*. Depuis, elle a tendance à toucher plus sensiblement certains secteurs spécifiques, tels que les matières premières (énergie, produits alimentaires), le logement, le tabac, l'alcool, les transports, et l'enseignement. Il est également notable que, depuis la pandémie de COVID-19, l'évolution de l'inflation en France est restée inférieure à celle de l'Union Européenne dans son ensemble.

L'inflation influence la croissance du PIB de plusieurs manières. Une inflation modérée peut stimuler la consommation, car les consommateurs anticipent des hausses de prix futures et préfèrent acheter maintenant plutôt que plus tard. Cependant, une inflation élevée peut éroder le pouvoir d'achat des ménages. En effet, l'une des conséquences directes de l'inflation est la baisse du pouvoir d'achat, ce qui signifie que les consommateurs peuvent acheter moins avec la même somme d'argent qu'auparavant, réduisant ainsi la consommation, qui est une composante majeure du PIB. Ce phénomène est particulièrement observable à partir du premier trimestre 2022 sur la *Figure 1.14*, 3ème graphique, illustrant le glissement annuel de l'IPCH français en fonction de celui du pourvoir d'achat.

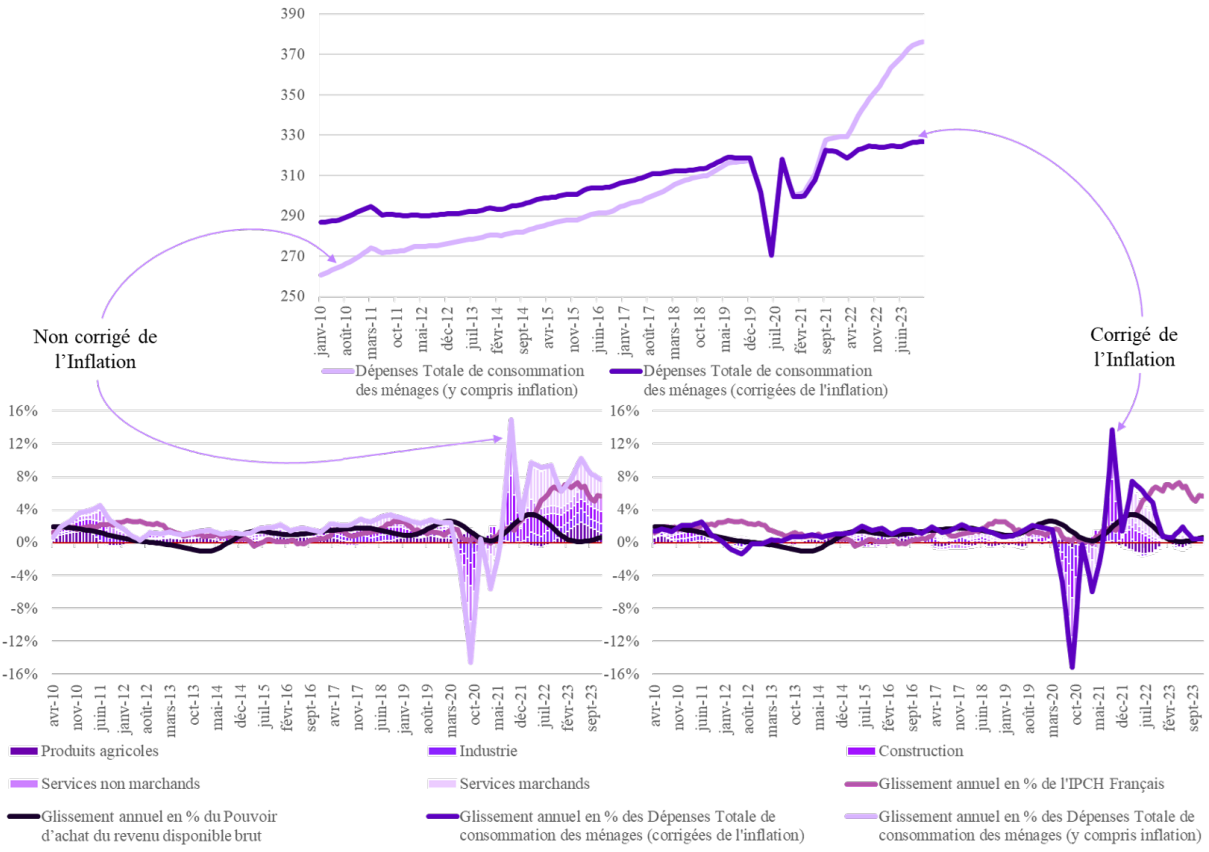


FIGURE 1.14 – Dépenses de consommation des ménages français en milliards d'euros (haut) avec et sans correction de l'inflation et son glissement annuel en % (bas) (*INSEE* [19])

Les *données de l'INSEE* illustrent bien cet impact de l'inflation sur la consommation des ménages. Les principales dépenses des ménages se concentrent sur l'industrie (énergie, eau, produits manu-

facturés) et les services marchands (commerce, transport, hébergement, restauration, information, communication, services financiers, immobiliers, aux entreprises et aux ménages) (*cf Annexe A.1*). La *Figure 1.14* montre que l'inflation peut biaiser l'évaluation de la consommation des ménages. Corrigée de l'inflation, la consommation des ménages a augmenté de 13.9% sur 14 ans, contre 44.4% sans correction. Une inflation autour de 2% permet généralement une augmentation de la consommation, comme observé avant la pandémie de COVID-19.

La situation de 2020 a été marquée par une pause économique due aux confinements, entraînant une stagnation de l'inflation et une augmentation de l'épargne des ménages (*cf Annexe A.2*), qui ont profité de cette période pour économiser, augmentant ainsi brièvement leur pouvoir d'achat. Cependant, les politiques monétaires accommodantes adoptées pendant la pandémie pour soutenir l'économie ont entraîné une hausse intense de l'inflation. Cette inflation, déjà présente, s'est aggravée avec le déclenchement du conflit russo-ukrainien en février 2022. La guerre a exacerbé l'inflation importée, augmentant les prix des biens importés et exerçant une pression à la hausse sur les prix intérieurs, en particulier dans les secteurs des produits agricoles et énergétiques. Les perturbations des chaînes d'approvisionnement mondiales, déjà fragilisées par la pandémie, se sont intensifiées, tandis que les prix de l'énergie ont fortement grimpé en raison des tensions géopolitiques. Cela a entraîné une baisse du pouvoir d'achat et, par conséquent, une diminution de la consommation.

En résumé, l'inflation est un indicateur clé de la santé économique, reflétant des changements dans le pouvoir d'achat et influençant les décisions économiques des ménages et des entreprises. Dans le secteur de l'emprunt, l'inflation joue un rôle complexe. Pour les emprunteurs ayant des prêts à taux fixe, une période d'inflation peut être avantageuse, car une augmentation des salaires peut réduire le poids relatif des mensualités : "*À titre d'exemple, pour une mensualité de 1000€ consentie avec un salaire de 3500€, si l'inflation se poursuit à un rythme annuel de 2.9% sur 3 ans et que le salaire progresse au même rythme, l'emprunteur remboursera 1000€ avec un salaire de 3813€, soit un endettement qui passerait de 30% à environ 26%.*" (*Pretto* [43]) Cependant, l'inflation peut également durcir les conditions d'accès au crédit. Les banques peuvent exiger des apports plus importants pour compenser les risques liés à l'incertitude économique. De plus, les taux d'intérêt nominaux peuvent augmenter, ce qui impacte les crédits à taux variable en alourdissant les mensualités.

Ainsi, dans un contexte inflationniste, emprunter peut présenter des avantages, notamment pour ceux qui anticipent une hausse de leurs revenus. Néanmoins, les emprunteurs doivent rester vigilants face aux évolutions du marché du crédit et aux politiques des banques qui s'adaptent à cet environnement économique.

1.2.2.4 Dette

La dette publique représente l'ensemble des emprunts contractés par les administrations publiques (État, collectivités locales, organismes de sécurité sociale) pour financer leurs dépenses lorsque leurs recettes sont insuffisantes. Elle constitue une obligation financière de l'État envers ses créanciers, qu'ils soient nationaux ou internationaux. La dette publique joue un rôle capital dans l'économie d'un pays et a un impact direct sur le PIB. Le ratio dette/PIB est un indicateur clé qui mesure l'endettement d'un pays par rapport à la taille de son économie. Un ratio élevé peut signaler des risques pour la soutenabilité de la dette, ce qui affecte la confiance des investisseurs[1] et les conditions de financement. En effet, un niveau élevé de dette publique peut freiner la

[1] "Investisseurs" désigne les acteurs qui achètent des titres de dette publique, tels que les institutions financières, les fonds d'investissement, les banques centrales, les particuliers et les organismes internationaux.

croissance économique. Selon les économistes Carmen Reinhart et Kenneth Rogoff [55], lorsque la dette dépasse un certain seuil, souvent située autour de 90% du PIB, la croissance économique tend à ralentir de manière significative, comme l'illustre bien les *données de l'INSEE* début 2012. Cela s'explique par l'effet de la dette sur les taux d'intérêt, les dépenses publiques et la confiance des investisseurs.

La dette publique est souvent utilisée pour financer des déficits budgétaires, notamment en période de crise économique. Si elle est bien gérée, elle peut soutenir la croissance économique en finançant des investissements publics productifs. Cependant, une mauvaise gestion de la dette peut entraîner des déséquilibres budgétaires et limiter la capacité de l'État à stimuler l'économie. Par conséquent, la gestion de la dette publique doit être soigneusement calibrée pour maintenir un équilibre entre le soutien de la croissance économique et la soutenabilité budgétaire à long terme.

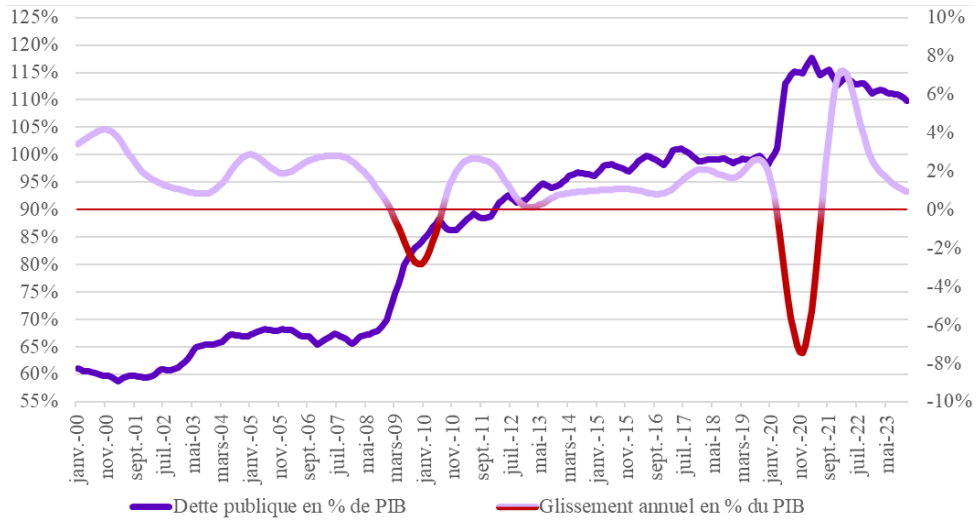


FIGURE 1.15 – Conséquences de la tendance de la Dette publique (axe gauche) sur l'évolution du PIB français (axe droit) (*INSEE* [21].[29])

Comme l'illustre la *Figure 1.15*, la dette publique française a connu une augmentation significative au cours des dernières décennies, passant de 60% du PIB au début des années 2000 à plus de 110% récemment, soit passant de 1750 milliards d'euros de dettes publiques en 2000 à 2800 milliards d'euros à début 2024 (en euro 2020) . Cette hausse est souvent attribuée aux crises économiques majeures, telles que la crise des subprimes de 2008 et la crise du COVID-19 en 2020, qui ont nécessité des dépenses publiques accrues pour stabiliser l'économie. Les politiques monétaires et budgétaires jouent un rôle capital dans la gestion de la dette, influençant les taux d'intérêt et les conditions de financement.

En résumé, bien que la dette publique puisse être un outil utile pour stimuler l'économie en période de crise, elle nécessite une gestion prudente pour éviter des déséquilibres financiers à long terme. Cette gestion de la dette publique a des répercussions directes sur les emprunteurs particuliers. En effet, lorsque l'État emprunte massivement, cela peut influencer les taux d'intérêt sur le marché, affectant ainsi le coût des emprunts pour les ménages et les entreprises. Par exemple, une augmentation de la dette publique peut entraîner une hausse des taux d'intérêt, rendant les prêts immobiliers ou à la consommation plus coûteux pour les particuliers. De plus, la

perception de la stabilité financière d'un pays peut influencer la confiance des prêteurs[1], impactant potentiellement les conditions d'octroi de crédit aux emprunteurs individuels.

1.2.3 Interconnexion du taux d'intérêt, de l'inflation et de la dette

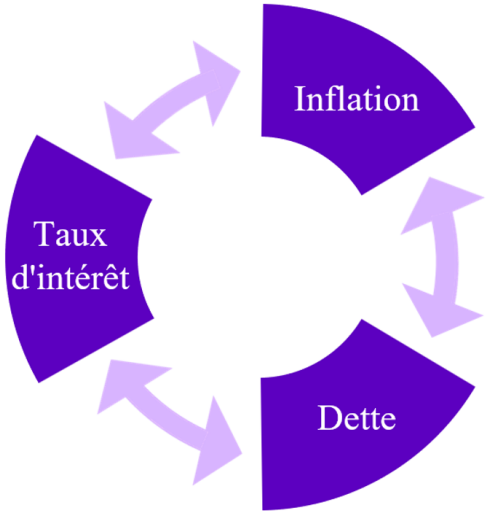


FIGURE 1.16 – Influence mutuelle des variables macro-économiques clés

La dette, l'inflation et les taux d'intérêt forment une triade macro-économique dont les interactions complexes façonnent l'environnement économique global et influencent directement les conditions d'emprunt pour les particuliers et les entreprises. Cette interconnexion dynamique se manifeste de plusieurs manières, impactant à la fois les politiques économiques nationales et les décisions financières individuelles des emprunteurs.

1.2.3.1 Taux d'intérêt et Inflation

Les taux d'intérêt sont un outil essentiel utilisé par les banques centrales, comme la Banque Centrale Européenne, pour maîtriser l'inflation. Comme le montre la *Figure 1.17* inspirée des données de la Banque de France, en augmentant les taux d'intérêt, les banques centrales rendent le crédit plus coûteux, ce qui réduit généralement la demande de prêts et par conséquent, l'investissement des entreprises et certains types de consommation des ménages, notamment pour les achats importants nécessitant un emprunt. En effet, sur la *Figure 1.17*, il est nettement visible qu'une décroissance du taux d'emprunt entraine une augmentation du nombre de crédits, et inversement, une augmentation des taux d'intérêt fait chuter le nombre de nouveaux crédit.

[1] "Prêteurs" désigne les acteurs qui fournissent des fonds aux emprunteurs, incluant des institutions financières telles que les banques et les organismes de crédit, ainsi que des investisseurs individuels ou des entreprises qui consentent des prêts.

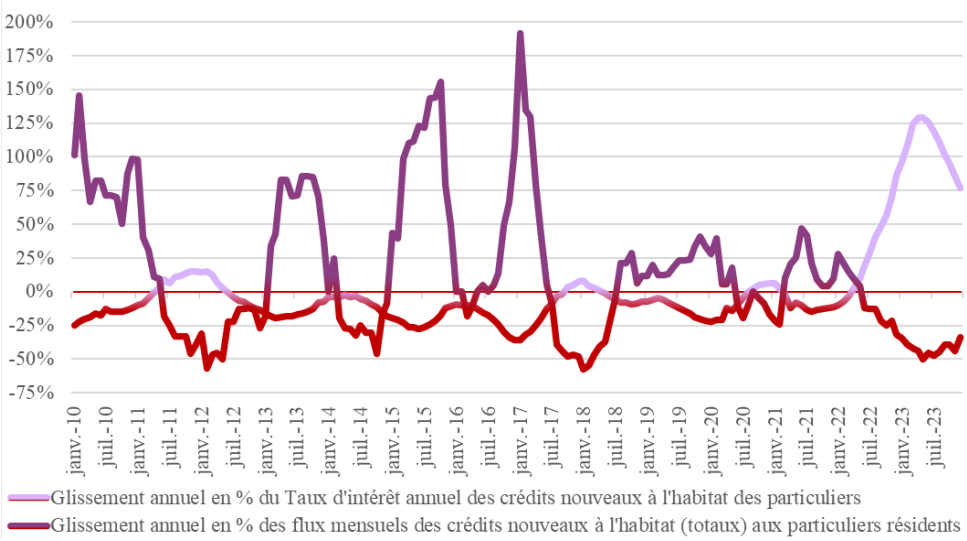


FIGURE 1.17 – Glissement annuel en % des flux et taux d'intérêt annuels des crédit nouveaux à l'habitat (*Banque de France* [3])

Cette diminution de la demande de crédit et des dépenses associées contribue à freiner l'inflation à moyen et long terme. En effet, les ménages et les entreprises sont moins enclins à emprunter pour des projets d'envergure lorsque les coûts d'emprunt sont élevés. Cependant, il est important de noter que l'impact sur les dépenses quotidiennes est généralement indirect et moins immédiat. Ces effets se manifestent plutôt à travers des changements dans les comportements d'épargne et d'investissement à plus long terme.

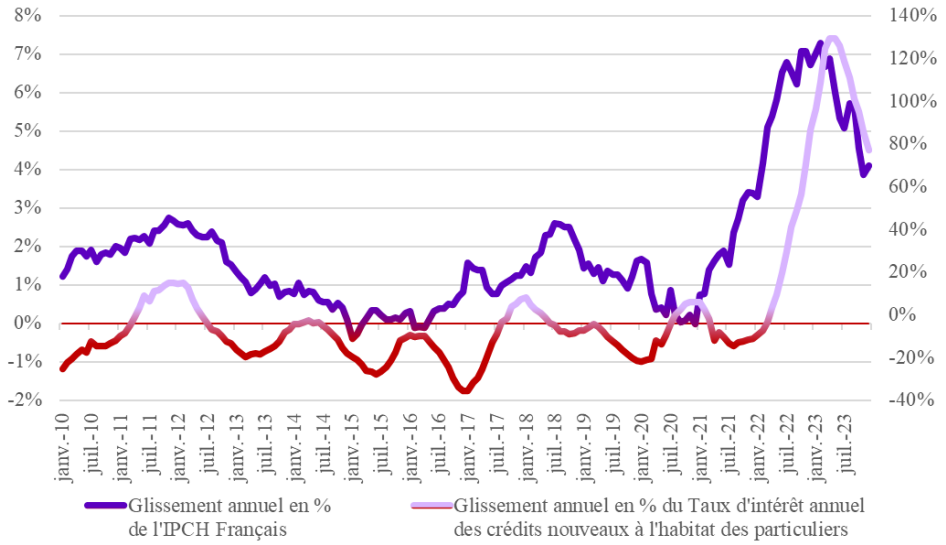


FIGURE 1.18 – Glissement annuel en % de l'IPCH français (axe gauche) et du taux d'intérêt annuel des crédit nouveaux à l'habitat (axe droit) (*Banque de France* [3] & *INSEE* [25])

Tout comme les taux d'intérêt sont utilisés pour maîtriser l'inflation, l'inflation elle-même influence les taux d'intérêt, créant une interaction dynamique. Une hausse de l'inflation comme montrée sur la *Figure 1.18* des données de l'INSEE, incite généralement les banques centrales à augmenter les taux d'intérêt pour contenir la pression inflationniste. Cette augmentation rend les emprunts plus coûteux, non seulement pour les ménages comme mentionné précédemment, mais aussi pour les entreprises. En conséquence, les investissements des entreprises peuvent diminuer, freinant

la croissance économique et les gains de productivité à long terme. Ainsi, l'interaction entre l'inflation et les taux d'intérêt affecte l'ensemble de l'économie, influençant les décisions financières tant des particuliers que des entreprises, et jouant un rôle fondamental dans le cycle économique global.

1.2.3.2 Inflation et Dette Publique

L'inflation a un impact ambigu sur la dette publique, notamment sur le ratio de la dette par rapport au PIB. Lorsque l'inflation augmente, le PIB nominal[1] augmente également. Cela peut réduire le ratio dette/PIB, rendant la dette plus facile à gérer, car la valeur réelle de la dette diminue par rapport à un PIB en croissance. En effet, l'inflation érode la valeur réelle de la dette, ce qui peut sembler avantageux pour les gouvernements, car les recettes fiscales augmentent avec l'inflation, réduisant proportionnellement la charge du service de la dette existante. Comme le note un article du Trésor français [8] : *"On entend souvent dire que l'inflation est favorable aux finances publiques, avec essentiellement deux arguments. Le premier est qu'une hausse des prix des biens à la consommation entraîne mécaniquement un relèvement des recettes de TVA ; une hausse des salaires élève mécaniquement les cotisations sociales versées ; et ainsi de suite pour chaque prélèvement sur une assiette en valeur. Le second argument est que la dette publique est fixée en euros, de sorte que si on la divise par un PIB dont la valeur augmente, le rapport dette/PIB diminue."*

Cependant, si l'inflation est trop élevée, elle peut rendre la dette publique moins soutenable en entraînant une hausse des taux d'intérêt nominaux. Cette augmentation des taux d'intérêt accroît les charges d'intérêt sur la dette publique, rendant son service plus coûteux à long terme. Les marchés financiers ajustent les taux d'intérêt pour compenser l'inflation anticipée, ce qui détériore la dynamique du ratio de la dette. De plus, une inflation non anticipée peut augmenter les rendements obligataires car les investisseurs exigent une prime de risque plus élevée, et réduire la marge de manœuvre budgétaire des gouvernements en augmentant les dépenses indexées sur l'inflation.

1.2.3.3 Taux d'intérêt et Dette Publique

Les taux d'intérêt et la dette publique s'influencent mutuellement, créant un cycle qui affecte l'ensemble de l'économie, y compris les emprunteurs individuels. Lorsque les taux d'intérêt sont bas, le service de la dette publique est moins coûteux, permettant à l'État d'investir dans des projets publics ou de stimuler l'économie. Cependant, une augmentation des taux accroît les charges d'intérêt, réduisant la capacité de l'État à financer des initiatives de croissance.

Parallèlement, une dette publique élevée peut entraîner une hausse des taux d'intérêt, les prêteurs exigeant une prime de risque plus importante. Cette situation augmente non seulement le coût du service de la dette pour l'État, mais impacte également directement les emprunteurs particuliers et les entreprises. Les prêts hypothécaires, automobiles et professionnels deviennent plus coûteux, freinant les investissements et les dépenses des ménages.

De plus, des taux d'intérêt élevés réduisent la demande globale en laissant moins de revenus disponibles aux ménages pour la consommation. Bien que cela puisse aider à contrôler l'inflation, cela peut aussi ralentir la croissance économique si la situation perdure. Pour les emprunteurs, cela se traduit par une diminution de leur pouvoir d'achat et une plus grande difficulté à accéder au crédit.

En somme, la gestion de la dette publique et des taux d'intérêt par les autorités a un impact direct

[1] "PIB nominal" représente la valeur totale des biens et services produits dans un pays en incluant l'effet de l'inflation.

sur la capacité d'emprunt et les décisions financières des citoyens et des entreprises. L'équilibre entre croissance économique et soutenabilité de la dette publique dépend fortement des niveaux de taux d'intérêt, influençant ainsi l'environnement économique global dans lequel évoluent tous les acteurs économiques.

1.2.3.4 Synthèse

En résumé, les taux d'intérêt, l'inflation et la dette publique forment un triangle économique complexe dont les interactions ont des répercussions directes sur le marché de l'emprunteur. Les banques centrales ajustent les taux d'intérêt pour contrôler l'inflation, ce qui influence non seulement le coût de la dette publique, mais aussi les conditions d'emprunt pour les particuliers et les entreprises. Lorsque l'inflation augmente, les taux d'intérêt tendent à suivre, rendant les emprunts plus coûteux pour tous. Parallèlement, la gestion de la dette publique peut influencer les taux du marché, affectant ainsi la capacité d'emprunt des ménages et des entreprises. L'emprunteur se trouve donc au cœur de ces dynamiques macro-économiques : les variations de taux d'intérêt impactent directement le coût de ses crédits, tandis que l'inflation érode le pouvoir d'achat de ses revenus et de son épargne. La soutenabilité de la dette publique, quant à elle, influence l'environnement économique global dans lequel l'emprunteur évolue, affectant ses perspectives d'emploi et d'investissement. Ainsi, comprendre ces interactions est décisif pour les emprunteurs dans leur prise de décisions financières à court et long terme.

1.2.4 Marché de l'emploi

L'inflation a un impact significatif sur le chômage, les salaires, et par extension, sur le marché de l'emprunt, créant une dynamique complexe dans l'économie. Lorsque l'inflation augmente, elle érode le pouvoir d'achat des salaires réels. Cette augmentation est particulièrement problématique pour les bas salaires qui consacrent une grande partie de leur revenu à des biens essentiels (*cf Annexe A.1*).

Face à cette situation, deux scénarios se présentent, chacun ayant des implications pour les emprunteurs :

1. Augmentation des salaires : En réponse à l'inflation, on peut observer une augmentation des salaires pour maintenir le pouvoir d'achat (phénomène visible sur la *Figure 1.19* pour la période post-covid). Cependant, cela peut alimenter une spirale "prix-salaires", exacerbant l'inflation. Dans ce contexte, les banques centrales tendent à relever les taux d'intérêt, rendant les emprunts plus coûteux. Pour les emprunteurs, cela se traduit par des mensualités plus élevées, surtout pour les prêts à taux variables.

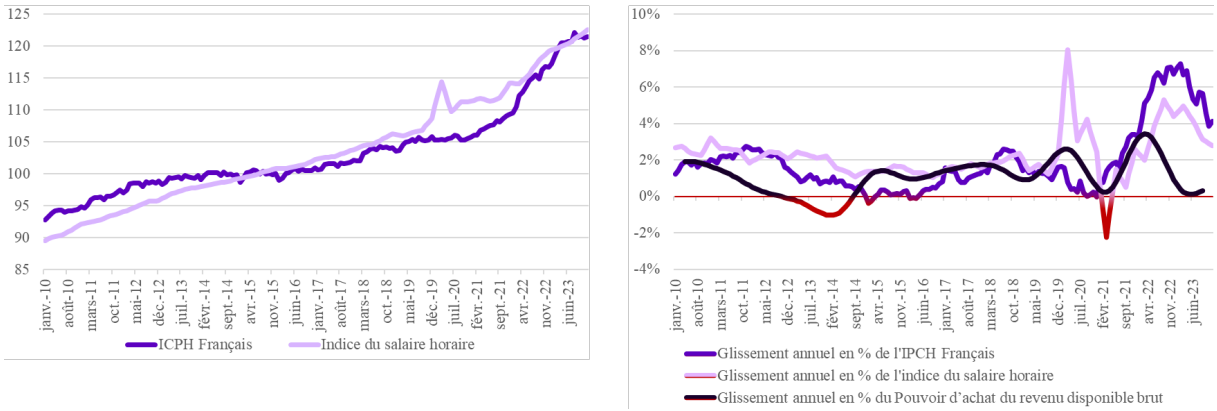


FIGURE 1.19 – ICPH et indice du salaire horaire français en Base 100 2015 (gauche) et leur glissement annuel en % (droite) (*INSEE* [22].[25])

2. Stagnation des salaires : Les entreprises peuvent être réticentes à augmenter les salaires au même rythme que l'inflation, par crainte de perdre en compétitivité. Cela peut entraîner une stagnation ou une baisse des salaires réels, réduisant le pouvoir d'achat des ménages (cf *Figure 1.14*). Cette situation peut aggraver le chômage en diminuant la demande pour les biens et services. Pour les emprunteurs, si les salaires ne suivent pas l'inflation, le remboursement des dettes devient plus difficile, pouvant conduire à une augmentation des défauts de paiement.

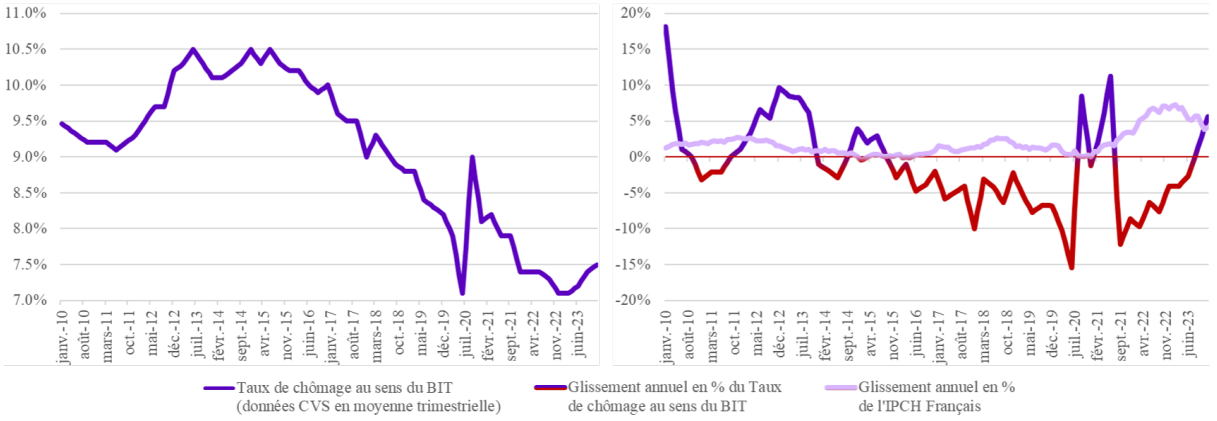


FIGURE 1.20 – Taux de chômage français au sens du BIT[1] (gauche) et glissements annuels en du taux de chômage et de l'IPCH (droite) (*INSEE* [25].[26]).

Dans les deux cas, le marché de l'emprunt est affecté. Un chômage élevé peut amener les banques à devenir plus prudentes, exigeant des garanties plus importantes ou des taux d'intérêt plus élevés. Cela rend l'accès au crédit plus difficile pour les ménages et les entreprises, freinant les investissements et la consommation, ce qui peut exacerber le ralentissement économique.

[1] Une personne au chômage au sens Bureau International du Travail (BIT) est une personne âgée de 15 ans ou plus qui répond simultanément à trois conditions :

— être sans emploi durant une semaine donnée,
— être disponible pour prendre un emploi dans les deux semaines,
— avoir cherché activement un emploi au cours des quatre dernières semaines ou en avoir trouvé un qui commence dans moins de trois mois.

Il est possible de voir des périodes où la progression des salaires dépasse l'inflation, améliorant le pouvoir d'achat et stimulant la consommation. Cependant, pour éviter les effets négatifs d'une spirale inflationniste, les ajustements salariaux et les politiques monétaires doivent être gérés avec soin, équilibrant croissance économique et stabilité des prix.

Ainsi, la relation complexe entre inflation, salaires et chômage influence directement les conditions du marché de l'emprunt, affectant la capacité des individus et des entreprises à accéder au crédit et à gérer leurs dettes. Cette dynamique souligne l'importance d'une gestion économique équilibrée pour maintenir un marché de l'emprunt stable et accessible.

1.2.5 Conclusion

L'assurance emprunteur, adossée aux crédits, est fortement influencée par le contexte économique, notamment les taux d'intérêt et l'inflation. Ces facteurs impactent non seulement le coût des emprunts, mais aussi la capacité des emprunteurs à souscrire et maintenir leurs assurances. Dans un contexte de resserrement monétaire, où les taux d'intérêt augmentent pour maîtriser l'inflation, le coût du crédit s'élève, ce qui peut réduire la demande de nouveaux prêts et, par conséquent, d'assurances emprunteur. Cela peut également influencer les décisions de rachat de prêts, car les emprunteurs cherchent à renégocier ou refinancer leurs dettes pour bénéficier de meilleures conditions.

Les politiques monétaires et fiscales jouent un rôle non négligeable dans ce contexte car elles déterminent les conditions de financement et la stabilité économique. Les mesures de soutien budgétaire, bien que nécessaires pour stabiliser l'économie, ont conduit à une augmentation de l'endettement public, ce qui peut influencer les perspectives économiques à long terme et, par extension, le marché de l'assurance emprunteur.

Pour la suite de ce mémoire, prendre en compte ces variables économiques sera essentiel pour développer des modèles précis loi de rachat de crédits et de taux d'affaires nouvelles. Ces modèles seront utilisés pour projeter l'évolution d'un portefeuille emprunteur sur une période de 5 ans, en tenant compte de divers scénarios de chocs d'inflation. Cette approche permettra d'anticiper les tendances du marché et d'optimiser les stratégies de tarification et de gestion des risques en fonction des différents contextes économiques simulés.

En conclusion, l'étude de l'impact du contexte économique sur l'assurance emprunteur ouvre la voie à une meilleure compréhension des dynamiques du marché et à l'élaboration de stratégies robustes face aux incertitudes économiques.

Chapitre 2

Présentation des données économiques et de la démarche scientifique

Suite à l'exploration du contexte économique et des dynamiques du marché de l'assurance emprunteur abordées dans le chapitre précédent, il est maintenant important de se concentrer sur les données concrètes qui servent de fondement à cette analyse. Ce chapitre est dédié à la présentation et à l'analyse des données essentielles qui sont utilisées pour modéliser les lois de rachat de crédits et le taux d'affaires nouvelles de crédits dans le chapitre suivant.

L'objectif de cette section est de fournir une vue d'ensemble complète et détaillée des informations quantitatives pertinentes pour l'étude. Des données historiques et contemporaines relatives aux variables économiques jugées essentielles sont examinées, telles que l'inflation, les taux d'emprunt et d'autres indicateurs macro-économiques significatifs. En outre, l'analyse porte sur les données spécifiques au marché de l'assurance emprunteur, notamment les tendances des nouvelles souscriptions et les taux de rachat de crédits. Cette analyse des données permet :

- D'identifier les tendances et les motifs significatifs au sein du marché de l'assurance emprunteur ;
- De mettre en lumière les corrélations potentielles entre les variables économiques et les comportements des emprunteurs ;
- D'établir une base pour la conception des modèles de projection qui sont développés dans le chapitre suivant.

Les sections suivantes présentent d'abord les sources et la nature des données utilisées, puis procèdent à une analyse descriptive de ces données, mettant en avant les tendances et les relations clés qui orientent la future modélisation.

2.1 Description de la base de données

Cette section présente les sources de données sélectionnées pour l'étude et justifie leur pertinence dans le cadre de l'analyse du marché de l'assurance emprunteur.

2.1.1 Sources des données

Les données utilisées dans cette étude proviennent principalement d'une base interne à l'entreprise et de deux institutions françaises reconnues :

- Un benchmark interne à l'entreprise Accenture, représentatif du marché français de l'assurance emprunteur ;
- La Banque de France ;

— L'Institut National de la Statistique et des Études Économiques (INSEE).

Ces sources ont été choisies pour leur fiabilité, leur exhaustivité et leur pertinence dans le contexte économique français. Il est important de souligner que toutes les données collectées sont à l'échelle nationale française. Cette uniformité dans la maille géographique des données permet d'assurer une cohérence dans l'analyse et de dresser un portrait fidèle du marché de l'assurance emprunteur à l'échelle du pays. L'utilisation de données nationales offre également l'avantage de capturer les tendances et les dynamiques macro-économiques qui influencent le secteur de l'assurance emprunteur dans son ensemble, sans être biaisées par des particularités régionales.

Le benchmark interne d'Accenture apporte une valeur ajoutée significative à cette étude. En effet, il fournit des données spécifiques au marché de l'assurance emprunteur, complétant ainsi les statistiques macroéconomiques fournies par la Banque de France et l'INSEE. Ce benchmark, étant représentatif du marché français, permet d'ancrer l'analyse dans la réalité opérationnelle du secteur et d'apporter des connaissances précieuses sur les tendances et les pratiques du marché.

De plus, dans le cadre de cette étude, une attention particulière est portée, dans la mesure du possible, à la sélection de données Corrigées des Variations Saisonnières (CVS[1]) plutôt que des données brutes. Les séries CVS permettent de neutraliser les fluctuations périodiques liées aux saisons ou aux jours ouvrables, facilitant ainsi l'identification des tendances de fond et des évolutions structurelles. Ce choix vise à alléger les futures modélisations en réduisant la complexité des ajustements nécessaires et en améliorant la lisibilité des tendances à long terme. L'utilisation de données CVS est particulièrement pertinente pour l'analyse du marché de l'assurance emprunteur, où les effets saisonniers peuvent masquer les véritables dynamiques du marché.

2.1.2 Données du benchmark interne à d'Accenture

Les données issues du benchmark interne d'Accenture fournissent des informations sur les taux de rachat des crédits immobiliers. Ces données sont particulièrement pertinentes pour l'analyse du marché de l'assurance emprunteur :

Données	Unité	Format
Taux de rachat des crédits immobiliers	%	Brut

Table 2.1 – Données collectées à partir du benchmark interne d'Accenture

— *Taux de rachat des crédits immobiliers* : Cette variable représente le pourcentage de crédits immobiliers rachetés et est donc utilisée ici comme base pour la modélisation de la loi de rachat en assurance emprunteur. Elle est essentielle pour comprendre la dynamique du marché de l'assurance emprunteur, car les rachats de crédits ont un impact direct sur les contrats d'assurance associés.

Ces données de taux de rachat présentent deux dimensions importantes :

1. **Année d'entrée en crédit** : Cette dimension permet d'analyser comment les cohortes de crédits, selon leur année d'origine, se comportent en termes de rachat.
2. **Année d'ancienneté du crédit** : Cette dimension offre une perspective sur l'évolution des taux de rachat au fil de la durée de vie du crédit.

[1]La Correction des Variations Saisonnières et des effets de jours ouvrables (CVS-CJO) est un traitement statistique de la série brute (i.e. initiale) mensuelle ou trimestrielle qui vise à en éliminer les composantes cycliques (saisonnalité, nombre de jours ouvrables, année bissextile) pour permettre de mieux discerner son évolution fondamentale (tendance et chocs conjoncturels, notamment). (*Ministère de l'Intérieur et de Outre-Mer* [37])

L'analyse croisée de ces deux dimensions offre une perspective détaillée et dynamique des comportements de rachat. Elle permet d'observer si certaines périodes de souscription ou certains moments de la durée du prêt sont plus propices aux rachats.

Ces données de benchmark apportent une valeur significative à l'étude. Elles fournissent une représentation fidèle du marché français de l'assurance emprunteur et révèlent des tendances dans les comportements de rachat. De plus, elles servent à calibrer et valider les modèles de prévision, tout en contextualisant les autres données macro-économiques et financières.

L'intégration de ces informations dans l'analyse globale permet une compréhension plus fine des dynamiques du marché de l'assurance emprunteur. Elle met en lumière les comportements de rachat, un élément principal dans la gestion des portefeuilles d'assurance emprunteur.

2.1.3 Données de la Banque de France

Les données obtenues de la Banque de France incluent plusieurs indicateurs clés pour l'analyse du marché de l'assurance emprunteur :

Données	Unité	Format
Encours des crédits à l'habitat accordés aux particuliers résidents	Mds€	Brut
Flux mensuels des crédits nouveaux à l'habitat aux particuliers résidents	Mds€	CVS
Taux d'intérêt annuel des crédits nouveaux à l'habitat des particuliers	%	Brut
Encours des prêts immobiliers à taux zéro	Mds€	Brut

TABLE 2.2 – Données collectées auprès de la Banque de France [3].[4]

— *Encours des crédits à l'habitat accordés aux particuliers résidents* et *Flux mensuels des crédits nouveaux à l'habitat aux particuliers résidents* : Ces deux variables sont essentielles pour la modélisation du taux d'affaires nouvelles de crédits immobiliers en assurance emprunteur. Le rapport entre les flux de crédits nouveaux et l'encours total permet de calculer ce taux, reflétant ainsi la dynamique du marché du crédit immobilier à l'échelle française ;

— *Taux d'intérêt annuel des crédits nouveaux à l'habitat des particuliers* : Cet indicateur abordé en *Partie 1.2.2.2*, fournit des informations essentielles sur les conditions de crédit, qui influencent directement la demande de prêts et, par conséquent, le marché de l'assurance emprunteur ;

— *Encours des prêts immobiliers à taux zéro* : Cette donnée offre un aperçu des politiques de soutien au crédit et peut aider à comprendre certaines dynamiques spécifiques du marché de l'emprunt.

Ces quatre variables sont particulièrement pertinentes pour l'étude du marché de l'emprunteur. Les encours de crédits et les flux mensuels des crédits nouveaux servent de base principale pour la modélisation du taux d'affaires nouvelles, tandis que les taux d'intérêt et l'encours des prêts à taux zéro fournissent des informations complémentaires essentielles pour affiner cette modélisation et comprendre les tendances du marché.

2.1.4 Données de l'INSEE

Les données recueillies auprès de l'INSEE comprennent plusieurs indicateurs clés qui sont essentiels pour comprendre le contexte économique dans lequel évolue le marché de l'assurance emprunteur :

Données					Unité	Format
Volume de construction	Nombre de logements	Autorisés	Tous logements		Nombre	CVS
			Collectif et Résidence			
			Individuel			
		Commencés	Tous logements			
			Collectif et Résidence			
			Individuel			
	Surfaces de logements	Autorisés	Tous logements		M^2	
			Collectif et Résidence			
			Individuel			
		Commencés	Tous logements			
			Collectif et Résidence			
			Individuel			
Volume de commercialisation	Nombre de logements	Mis en vente aux particuliers	Toutes constructions	Tous logements	Nombre	Brut
				Collectif et Résidence		
				Individuel		
			Construction sur existant	Tous logements		
				Collectif et Résidence		
				Individuel		
			Nouvelles constructions	Tous logements		
				Collectif et Résidence		
				Individuel		
		Réservés à la vente par des particuliers	Toutes constructions	Tous logements		
				Collectif et Résidence		
				Individuel		
			Construction sur existant	Tous logements		
				Collectif et Résidence		
				Individuel		
			Nouvelles constructions	Tous logements		
				Collectif et Résidence		
				Individuel		
		Annulés après réservation par des particuliers	Toutes constructions	Tous logements		
				Collectif et Résidence		
				Individuel		
			Construction sur existant	Tous logements		
				Collectif et Résidence		
				Individuel		
			Nouvelles constructions	Tous logements		
				Collectif et Résidence		
				Individuel		
	Encours de logements disponibles à la vente aux particuliers		Toutes constructions	Tous logements		
				Collectif et Résidence		
				Individuel		
			Construction sur existant	Tous logements		
				Collectif et Résidence		
				Individuel		
			Nouvelles constructions	Tous logements		
				Collectif et Résidence		
				Individuel		
	Délai moyen d'écoulement de l'encours de logements disponibles à la vente aux particuliers		Toutes constructions	Tous logements	Nombre de trimestres	
				Collectif et Résidence		
				Individuel		
			Construction sur existant	Tous logements		
				Collectif et Résidence		
				Individuel		
			Nouvelles constructions	Tous logements		
				Collectif et Résidence		
				Individuel		

TABLE 2.3 – Données collectées auprès de l'INSEE [18].[20] (Partie 1)

Données			Unité	Format
Indice des prix des logements	Neufs	Tous logements	Base 100 - 2015[1]	CVS
	Anciens	Tous logements		
		Maisons		
		Appartements		
Indice du coût de production dans la construction (ICP-F)			Base 100 - 2021	Brut
Indice du coût de la construction (ICC)			Base 100 - 2021	Brut
Taux de chômage au sens du BIT			%	CVS
Dette publique en % de PIB			%	Brut
Évolution annuelle de l'Indice des prix à la consommation (IPC)			%	Brut
Produit Intérieur Brut en Mds€ 2020			Mds€ 2020	Brut
Évolution annuelle relative du Pouvoir d'achat du revenu disponible brut			%	Brut
Taux d'épargne des ménages en % du revenu brut			%	Brut

TABLE 2.4 – Données collectées auprès de l'INSEE [21].[23].[24].[26].[27].[28].[29].[30] (Partie 2)

— *Volume de construction et Volume de commercialisation* : Ces indicateurs reflètent directement l'activité du secteur immobilier, un moteur principal de la demande en assurance emprunteur ;

— *Indice du coût de production dans la construction (ICP-F) et Indice du coût de la construction (ICC)* : Ces indices permettent de suivre l'évolution des coûts dans le secteur de la construction, influençant les prix de l'immobilier et donc les montants empruntés ;

— *Indice des prix des logements* : Cet indicateur est essentiel pour comprendre l'évolution du marché immobilier et son impact sur les montants des prêts ;

— *Taux de chômage au sens du BIT* : Comme discuté dans en *Partie 1.2.4*, le taux de chômage est un indicateur important de la santé économique et peut influencer la capacité des ménages à emprunter ;

— *Dette publique en % de PIB* : Cet indicateur, abordé précédemment en *Partie 1.2.2.4*, peut influencer les politiques économiques et monétaires, affectant indirectement les conditions de crédit ;

— *Évolution annuelle de l'indice des prix à la consommation* : Cette mesure de l'inflation, déjà évoquée en *Partie 1.2.2.3*, est fondamentale pour comprendre l'évolution du pouvoir d'achat et son impact sur la capacité d'emprunt des ménages ;

— *Produit Intérieur Brut* : Comme mentionné en *Partie 1.2.2.1*, le PIB est un indicateur clé de la santé économique globale, influençant indirectement le marché de l'emprunt ;

— *Évolution annuelle relative du Pouvoir d'achat du revenu disponible brut* : Cet indicateur est déterminant pour évaluer la capacité des ménages à emprunter et à rembourser leurs prêts ;

— *Taux d'épargne des ménages en % du revenu brut* : Le taux d'épargne peut influencer la propension des ménages à emprunter ou à autofinancer leurs projets.

Ces indicateurs, dont certains ont été introduits dans le Chapitre 1, ont maintenant pour objectif d'être utilisés de manière plus spécifique pour analyser le marché de l'assurance emprunteur. Ils permettent de contextualiser les conditions économiques qui influencent directement la demande de crédit et, par extension, le marché de l'assurance emprunteur. L'utilisation de ces données vise à établir des corrélations entre les tendances macro-économiques et les dynamiques spécifiques du marché de l'assurance emprunteur, fournissant ainsi une base pour la modélisation et l'analyse qui suivent.

[1] L'indice base 100 en 2015 est une méthode de normalisation des données où la valeur de référence (100) est attribuée à l'année 2015. Toutes les autres valeurs de la série sont exprimées en pourcentage par rapport à cette référence, permettant ainsi de mesurer facilement l'évolution relative d'une grandeur au fil du temps.

2.2 Traitements effectués sur les données

Après avoir collecté les données pertinentes auprès des différentes sources, l'étape suivante consiste à les préparer pour l'analyse et la modélisation. Cette phase de traitement est essentielle pour garantir la qualité et la fiabilité des résultats futurs. Le processus de traitement des données vise plusieurs objectifs :

— Assurer la cohérence temporelle et la comparabilité des différentes séries de données ;
— Agréger les données lorsque nécessaire pour obtenir une granularité uniforme ;
— Vérifier et traiter les valeurs manquantes ou aberrantes ;
— Harmoniser les formats et les unités de mesure ;
— Créer des variables dérivées si nécessaire pour l'analyse.

Dans les sections suivantes, les différentes étapes de ce processus de traitement sont détaillées, ainsi que les choix méthodologiques effectués pour optimiser la qualité des données.

2.2.1 Retraitement et contrôle

Dans le cadre de cette étude, les données collectées auprès de l'INSEE et de la Banque de France se sont avérées particulièrement fiables et complètes. Sur l'ensemble des séries temporelles utilisées, seul un nombre négligeable de valeurs manquantes a été identifié (uniquement 4 occurrences). Pour traiter ces rares cas, une méthode d'imputation simple par la moyenne a été appliquée. Bien que cette approche puisse être considérée comme basique, son impact sur la qualité globale des données est minime étant donné le faible nombre de valeurs concernées.

Il convient de souligner que l'utilisation de données en open data provenant d'institutions nationales telles que l'INSEE et la Banque de France présentent un avantage significatif en termes de fiabilité et de précision. Ces sources offrent des séries de données généralement bien maintenues et régulièrement mises à jour, ce qui minimise les problèmes de valeurs manquantes ou aberrantes souvent rencontrés dans d'autres types de bases de données.

Dans un souci de cohérence et de facilité de comparaison, les indices ICP-F (Indice du coût de production dans la construction) et ICC (Indice du coût de la construction), initialement en base 100 en 2021, ont été convertis en base 100 en 2015. Cette transformation n'a aucun impact sur la nature ou la qualité des données elles-mêmes, mais permet une harmonisation avec les autres séries temporelles utilisées dans l'étude.

2.2.2 Agrégation des bases

Les données collectées présentent des fréquences variées :

— Mensuelles : pour les *Flux mensuels des crédits nouveaux à l'habitat*, les *Taux d'intérêt annuel des crédits nouveaux à l'habitat* et les données afférentes aux *Volumes de construction* ;
— Trimestrielles : pour les *Encours des prêts immobiliers à taux zéro*, les *ICP-F* et *ICC*, les *Indice des prix des logements* , les *Taux de chômage* et les données associées aux *Volumes de commercialisation* ;
— Annuelles : pour les *Taux de rachat par ancienneté*, pour le *PIB*, la *Dette publique*, les *IPC*, les *Évolution annuelle relative du Pouvoir d'achat*, les *Taux d'épargne des ménages*, .

Pour former une base de données unique et cohérente, il était nécessaire d'harmoniser ces différentes fréquences. Le choix a été fait de convertir toutes les données à une fréquence mensuelle, et ce pour plusieurs raisons :

1. Préserver le maximum d'information disponible, en particulier pour les séries déjà mensuelles ;
2. Obtenir un volume de données suffisant pour les futures modélisations, ce qui contribue à l'apprentissage et à la robustesse des modèles utilisés ;
3. Maintenir une granularité fine permettant de capturer les variations à court terme du marché.

Pour convertir les données trimestrielles et annuelles en données mensuelles, la méthode d'interpolation cubique d'Hermite est utilisée. Cette méthode présente plusieurs avantages :

1. Elle produit une courbe lisse et continue entre les points de données connus, évitant les discontinuités abruptes ;
2. Elle préserve la monotonie locale des données, ce qui signifie qu'elle respecte les tendances à la hausse ou à la baisse entre les points connus ;
3. Elle n'introduit pas d'oscillations artificielles entre les points de données, un problème parfois rencontré avec d'autres méthodes d'interpolation polynomiale.

Cette méthode (détaillée en *Annexe B.1*) permet de générer des valeurs mensuelles plausibles et cohérentes à partir des données trimestrielles et annuelles, tout en préservant les caractéristiques essentielles des séries temporelles originales.

2.2.3 Filtres effectués sur les données

Pour faire suite aux retraitements initiaux et agrégations décrits dans les deux sous-sections précédentes, la démarche se poursuit avec une étape de sélection et de filtrage des données. Cette étape vise à constituer une base finale cohérente et adaptée aux objectifs de modélisation du taux d'affaires nouvelles et de taux de rachat.

2.2.3.1 Alignement temporel des données

Étant donné que le rapport des données *Encours des crédits à l'habitat accordés aux particuliers résidents* et *Flux mensuels des crédits nouveaux à l'habitat* sert de variable dépendante y à prédire pour les futures modélisations de *Taux d'affaires nouvelles*, ces données sont utilisées comme référence pour définir la période d'étude. L'historique de ces données mensuelles remontant à mars 2003, toutes les autres variables sont donc alignées sur cette date de début. Cette approche permet d'assurer une cohérence temporelle entre toutes les variables et d'éviter les problèmes liés aux données manquantes en début de série.

Cependant, il est important de noter que les données sur les *Taux de rachat par ancienneté* r_j ne sont disponibles qu'à partir de décembre 2010. Cette particularité n'affecte pas l'alignement global des données, mais elle implique que lors de la modélisation future intégrant ces taux de rachat, les autres variables doivent être tronquées pour correspondre à cette période plus courte. Cette troncature permet de maintenir la cohérence temporelle de l'ensemble des données utilisées dans le modèle.

2.2.3.2 Sélection des variables basée sur l'historique

La sélection des variables pour cette étude est effectuée en tenant compte de deux horizons temporels distincts, correspondant aux deux principaux aspects de cette modélisation :

1. Pour la modélisation du *Taux d'affaires nouvelles*, uniquement les variables disposant d'un historique remontant au moins à mars 2003 sont conservées. Cette approche permet de préserver l'intégrité des séries temporelles sur une période étendue et d'éviter la création artificielle de données ;

2. En ce qui concerne la modélisation des *Taux de rachat*, un ensemble de données plus restreint est utilisé, commençant en décembre 2010. Cette date correspond au début de la disponibilité des données sur les *Taux de rachat par ancienneté* r_j.

Cette décision impacte notamment les données relatives au *Volume de commercialisation*. Ainsi, seules les données portant sur les *"Nouvelles constructions"* sont retenues, car elles présentent un historique suffisant. En revanche, les données sur les *"Constructions sur existant"* et sur le *"Nombre de logements Annulés après réservation par des particuliers"*, qui s'arrêtent en mars 2017, sont exclues de l'analyse pour éviter d'introduire des discontinuités dans les séries temporelles.

Cette stratégie de sélection à deux niveaux permet de maximiser l'utilisation des données historiques disponibles pour chaque aspect du modèle, tout en maintenant la cohérence temporelle au sein de chaque ensemble de variables. Ainsi, une base est assurée pour les analyses tout en tenant compte des contraintes spécifiques à chaque type de modélisation.

2.2.3.3 Élimination des redondances

Un filtrage supplémentaire est effectué pour éviter la redondance d'informations et prévenir la création de corrélations artificielles. Les variables qui représentent des sommes d'autres variables déjà présentes dans la base sont systématiquement écartées. Par exemple, dans le cas des données sur les logements, seules les variables séparées pour les logements *"Individuels"* et *"Collectif et Résidence"* sont conservées, tandis que la variable agrégée les regroupant est supprimée. Cette approche de sélection et de filtrage des données vise à obtenir une base finale robuste, cohérente et exempte de redondances, fournissant ainsi une fondation solide pour les analyses et modélisations à venir.

2.2.4 Variables retenues

À l'issue des processus de traitement, d'agrégation et de filtrage décrits précédemment, un ensemble final de variables est sélectionné pour la suite de l'étude. Ces variables sont choisies en fonction de leur pertinence pour l'analyse du marché de l'assurance emprunteur, de la qualité et de la complétude de leurs données historiques, ainsi que de leur capacité à fournir des informations uniques et non redondantes.

Au total, 31 variables sont retenues pour la suite de l'analyse. Ces variables couvrent un large éventail d'indicateurs économiques et spécifiques au marché de l'immobilier et du crédit, offrant ainsi une vue d'ensemble complète des facteurs susceptibles d'influencer le taux d'affaires nouvelles en assurance emprunteur.

	Données					Unité	Format
y	**Taux d'affaires nouvelles de crédits immobiliers**[1]					%	Brut
r_j	**Taux de rachat des crédits immobiliers**					%	Brut
x_1	**Taux d'intérêt annuel des crédits nouveaux à l'habitat des particuliers**					%	Brut
x_2	**Encours des prêts immobiliers à taux zéro**					Mds€	Brut
x_3	**Indice des prix des logements**		Neufs	Tous logements		Base 100 - 2015	CVS
x_4			Anciens	Maisons			
x_5				Appartements			
x_6	**Indice du coût de production dans la construction (ICP-F)**					Base 100 - 2015	Brut
x_7	**Indice du coût de la construction (ICC)**					Base 100 - 2015	Brut
x_8	**Taux de chômage au sens du BIT**					%	CVS
x_9	**Dette publique en % de PIB**					%	Brut
x_{10}	**Évolution annuelle de l'IPC**					%	Brut
x_{11}	**Produit Intérieur Brut en Mds€ 2020**					Mds€ 2020	Brut
x_{12}	**Évolution annuelle relative du Pouvoir d'achat du revenu disponible brut**					%	Brut
x_{13}	**Taux d'épargne des ménages en % du revenu brut**					%	Brut
x_{14}	**Volume de construction**	Nombre de logements	Autorisés	Collectif et Résidence		Nombre	CVS
x_{15}				Individuel			
x_{16}			Commencés	Collectif et Résidence			
x_{17}				Individuel			
x_{18}		Surfaces de logements	Autorisés	Collectif et Résidence		M^2	
x_{19}				Individuel			
x_{20}			Commencés	Collectif et Résidence			
x_{21}				Individuel			
x_{22}	**Volume de commercialisation**	Nombre de logements	Mis en vente aux particuliers	Nouvelles constructions	Collectif et Résidence	Nombre	Brut
x_{23}					Individuel		
x_{24}			Réservés à la vente par des particuliers	Nouvelles constructions	Collectif et Résidence		
x_{25}					Individuel		
x_{26}		Encours de logements disponibles à la vente aux particuliers		Nouvelles constructions	Collectif et Résidence		
x_{27}					Individuel		
x_{28}		Délai moyen d'écoulement de l'encours de logements disponibles à la vente aux particuliers		Nouvelles constructions	Collectif et Résidence	Nombre de trimestres	
x_{29}					Individuel		

TABLE 2.5 – Données finales retenues pour la modélisation

Le *Tableau 2.5* présente la liste exhaustive des 31 variables soigneusement sélectionnées pour cette étude. Cette sélection assure que l'analyse se concentre sur les indicateurs les plus pertinents et fiables, tout en évitant la redondance d'information. Ces variables constituent la base de données finale sur laquelle s'appuient toutes les analyses et modélisations ultérieures, offrant ainsi une fondation solide et efficace pour les étapes suivantes de l'étude.

Après avoir établi cette base de données finale, il est essentiel d'approfondir la compréhension de ces variables et de leurs caractéristiques. La section suivante se consacre donc à une analyse descriptive détaillée de ces 31 variables retenues. Cette analyse vise à mettre en lumière les tendances, les distributions et les relations potentielles entre ces indicateurs, fournissant ainsi un aperçu du contexte économique et du marché de l'assurance emprunteur qu'ils représentent. Cette étape est fondamentale pour orienter les choix de modélisation ultérieurs et pour interpréter de manière pertinente les résultats qui en découleront.

[1] Cette variable créée est le rapport des données *Flux mensuels des crédits nouveaux à l'habitat* et *Encours des crédits à l'habitat accordés aux particuliers résidents*. Elle correspond donc à un taux de crédits nouveaux à l'habitat

2.3 Analyses statistiques

Dans l'optique de développer une méthode de modélisation robuste pour le taux d'affaires nouvelles et la loi de rachat en assurance emprunteur, il est primordial de commencer par une étude descriptive des variables explicatives potentielles. Cette analyse préliminaire vise à identifier et comprendre les facteurs qui pourraient influencer significativement l'évolution du taux d'affaires nouvelles . Cette étape est nécessaire car elle permet de :

— Mettre en lumière les tendances et les patterns au sein de chaque variable ;

— Identifier les relations potentielles entre les différentes variables explicatives ;

— Détecter d'éventuelles anomalies ou particularités dans les données qui pourraient impacter la modélisation ;

— Guider le choix des variables les plus pertinentes pour la construction du modèle final.

L'analyse descriptive qui suit se concentre principalement sur les variables clés de l'étude : les *taux d'affaires nouvelles de crédits immobiliers* et les *taux de rachats de crédits immobiliers*. Les caractéristiques statistiques, l'évolution temporelle et les possibles interactions avec d'autres variables économiques pertinentes sont examinées en détail. Cette exploration des données sert de fondement solide pour les étapes de modélisation ultérieures, assurant ainsi une approche éclairée et rigoureuse dans la prédiction du taux d'affaires nouvelles et des taux de rachats par ancienneté en assurance emprunteur. Bien que l'accent soit mis sur ces variables centrales, les autres variables retenues sont également mentionnées lorsque cela s'avère pertinent pour contextualiser ou approfondir la compréhension des flux et taux de rachat de crédits.

2.3.1 Objet de l'étude

Cette étude se concentre sur deux séries de données principales :

1. Les *taux d'affaires nouvelles en France*, qui couvrent une période de près de 21 ans, de mars 2003 à décembre 2023. Cette série temporelle étendue offre une perspective approfondie sur l'évolution du marché du crédit immobilier en France ;
2. Les *taux de rachat par année d'ancienneté de prêt*, qui s'étendent de décembre 2010 à décembre 2023 et couvrent des anciennetés de prêt allant de 1 à 25 ans. Cette série fournit des informations cruciales sur la dynamique des rachats de crédit, un aspect essentiel du marché de l'assurance emprunteur.

Ces deux séries de données sont examinées en parallèle, chacune apportant des éclairages uniques sur différents aspects du marché du crédit immobilier et de l'assurance emprunteur.

L'intérêt de la période étudiée réside dans sa richesse en termes de contextes économiques variés. Elle englobe notamment :

— La crise financière de 2008 et ses répercussions ;

— La période de taux bas qui a suivi ;

— La crise sanitaire de 2020 et ses impacts économiques ;

— La récente remontée des taux d'intérêt.

Cette diversité de scénarios économiques permet d'observer comment le marché du crédit immobilier et les comportements de rachat réagissent face à différentes conditions macro-économiques. L'analyse conjointe de ces deux séries de données offre ainsi une base pour la modélisation du taux d'affaires nouvelles en assurance emprunteur et pour la compréhension des dynamiques de rachat dans ce secteur.

2.3.2 Analyse de la tendance centrale et de la dispersion

L'analyse de la tendance centrale et de la dispersion des taux d'affaires nouvelles et des taux de rachat permet d'obtenir une vue d'ensemble de ces indicateurs clés du marché de l'assurance emprunteur. Cette étude statistique révèle les caractéristiques principales de leur distribution, offrant ainsi des informations sur leur comportement et leur variabilité au fil du temps.

2.3.2.1 Analyse des taux d'affaires nouvelles

L'examen des statistiques descriptives des taux d'affaires nouvelles en France révèle des informations sur la distribution et la variabilité de ces données.

Statistique	Valeur
Moyenne	1,775%
Médiane	1,778%
Écart-type	0,570%
Minimum	0,661%
Maximum	4,259%
Amplitude	3,598%
Variance	0,003%

TABLE 2.6 – Statistiques descriptives des taux d'affaires nouvelles en France

Les statistique descriptives du *Tableau 2.6* indiquent que la moyenne des taux d'affaires nouvelles en France s'établit à 1,775% du volume total, tandis que la médiane est de 1,778%. La proximité entre ces deux mesures de tendance centrale suggère une distribution relativement symétrique, indiquant une répartition équilibrée des taux autour de la valeur centrale.

En termes de dispersion, l'écart-type de 0,570% témoigne d'une variabilité modérée des taux mensuels. Cette variabilité est également mise en évidence par l'amplitude de 3,598 points de pourcentage entre le taux minimum de 0,661% et le maximum de 4,259%.

La variance faible de 0,003% confirme une dispersion relativement limitée des données, soulignant une certaine stabilité du taux d'affaires nouvelles sur la période étudiée.

Ces résultats, appuyés de la *Figure 2.1*, mettent en lumière la nature relativement stable du taux d'affaires nouvelles en assurance emprunteur en France. La proximité entre la moyenne et la médiane, ainsi que la variance faible, suggèrent une certaine constance dans l'évolution de ce taux. Cependant, l'écart entre les valeurs minimales et maximales indique qu'il existe des périodes de fluctuations plus importantes. Cette stabilité relative, ponctuée de variations occasionnelles, souligne l'importance d'une analyse plus approfondie des tendances temporelles et des facteurs influençant ces fluctuations pour mieux comprendre et potentiellement prédire l'évolution du taux d'affaires nouvelles en assurance emprunteur.

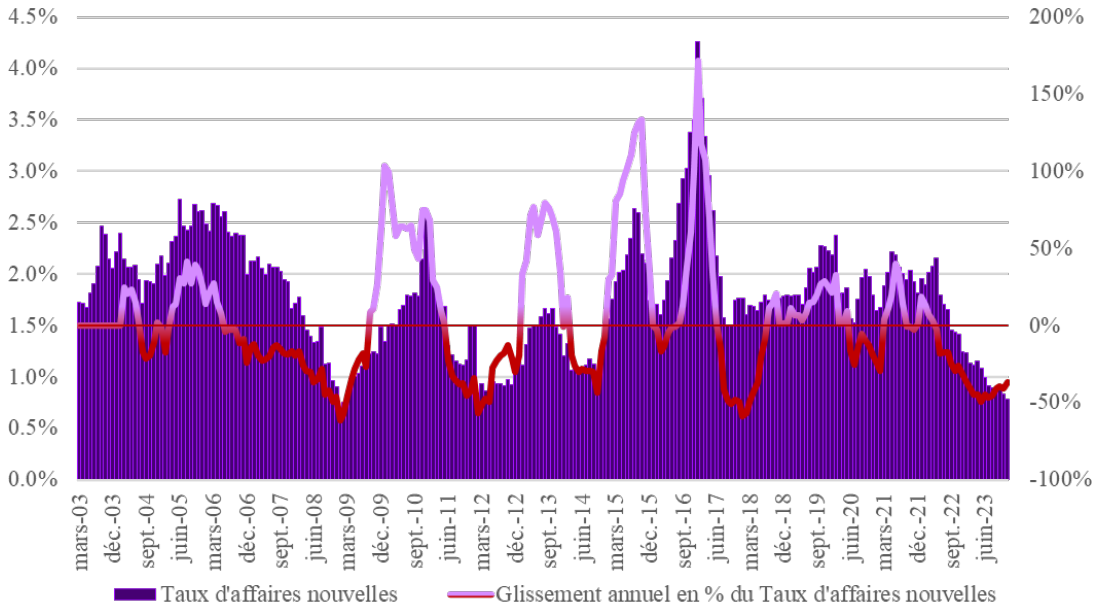


FIGURE 2.1 – Taux d'affaires nouvelles (axe gauche) et son glissement annuel en % (axe droit) (*Banque de France* [3])

2.3.2.2 Analyse des taux de rachat

D'autre part l'analyse des taux de rachat des crédits immobiliers entre 2010 et 2023 révèle des tendances intéressantes en fonction de l'ancienneté des prêts.

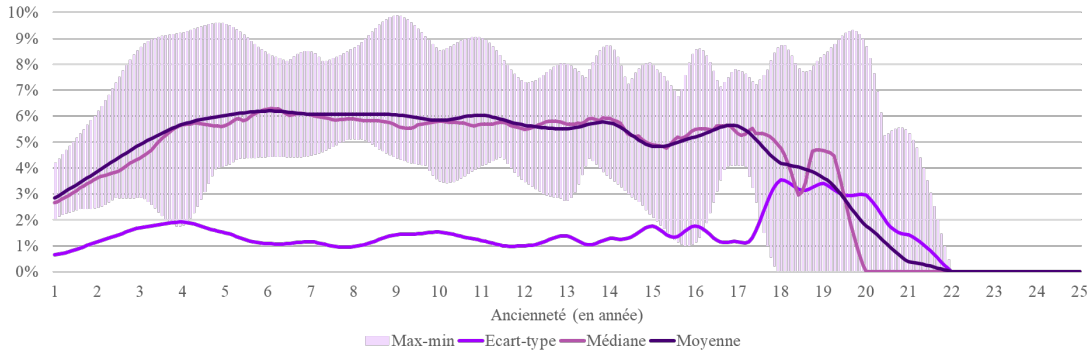


FIGURE 2.2 – Taux de rachat des crédits immobiliers par ancienneté du prêt (2010-2023)

Les taux de rachat moyens varient considérablement selon l'ancienneté du prêt, allant de 2,83% pour les prêts d'un an à un pic de 6,20% pour les prêts de 6 ans d'ancienneté, avant de diminuer progressivement pour les prêts plus anciens. La médiane suit une tendance similaire, avec un maximum de 6,28% également à 6 ans d'ancienneté.

La dispersion des taux de rachat, mesurée par l'écart-type, est particulièrement élevée pour les prêts entre 3 et 5 ans d'ancienneté (atteignant 1,91% pour les prêts de 4 ans), indiquant une plus grande variabilité des comportements de rachat dans ces premières années. Pour les prêts plus anciens, notamment au-delà de 20 ans, les taux de rachat deviennent négligeables, avec des moyennes et médianes proches ou égales à zéro.

L'amplitude des taux de rachat est considérable, avec des minimums proches de 0% pour plusieurs

anciennetés et des maximums atteignant 9,87% (pour les prêts de 9 ans d'ancienneté). Cette large amplitude souligne la volatilité du marché des rachats de crédits et suggère l'influence potentielle de facteurs externes sur les décisions de rachat.

Ces résultats, illustrés dans la *Figure 2.2*, mettent en évidence la dynamique complexe des rachats de crédits immobiliers en France. La concentration des taux de rachat les plus élevés entre 3 et 10 ans d'ancienneté suggère une période critique pour les assureurs, où le risque de perte de contrats est le plus élevé. La variabilité observée d'une année à l'autre (comme le montre le graphique en *Annexe B.1*) souligne l'importance d'une analyse des facteurs influençant ces fluctuations pour mieux anticiper et gérer le risque de rachat dans les portefeuilles d'assurance emprunteur.

2.3.3 Décomposition chronologique

Après avoir examiné les tendances centrales et la dispersion des taux d'affaires nouvelles et des taux de rachat, une analyse plus approfondie des taux d'affaires nouvelles est nécessaire. Cette focalisation est justifiée par la complexité et la richesse d'information contenue dans cette série temporelle unique, contrairement aux multiples séries de taux de rachat (25 séries temporelles pour les 25 années d'ancienneté de taux de rachat) qui nécessiteraient une analyse beaucoup plus étendue et complexe.

L'analyse des taux d'affaires nouvelles bénéficie grandement d'une décomposition chronologique. Cette méthode permet de séparer la série temporelle en ses composantes fondamentales, offrant ainsi une meilleure compréhension des dynamiques du marché du crédit immobilier. La *Figure 2.3*, obtenue grâce à la formule R `decompose()` illustre cette décomposition en quatre éléments distincts.

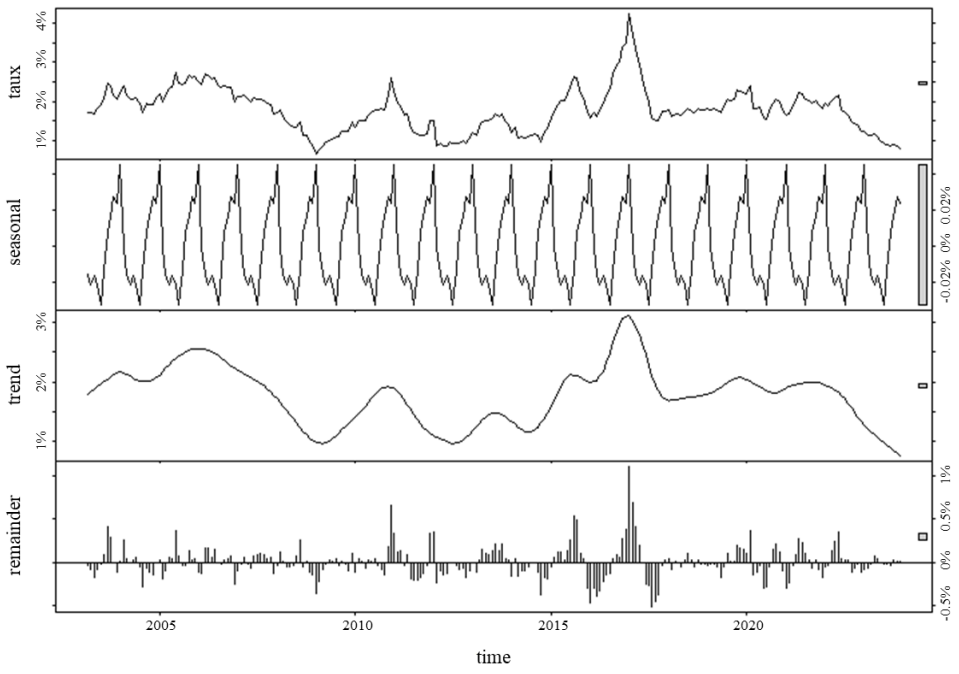


FIGURE 2.3 – Décomposition de séries chronologiques appliquée aux Taux d'affaires nouvelles (*Banque de France* [3])

Dans les sections suivantes, chaque composante de la décomposition chronologique est examinée

en détail, permettant une analyse séparée des différents aspects des taux d'affaires nouvelles.

2.3.3.1 Flux (Données brutes) :

Le *premier graphique* représente les données brutes, qui montrent l'évolution des taux d'affaires nouvelles sur la période donnée de mars 2003 à décembre 2023. On peut observer une forte baisse en 2008, suivi de deux pics en 2012 et 2015 revenant au statut-quo des années 2000. Une forte hausse est visible en 2017, puis d'une stabilisation avec une légère baisse récente.

2.3.3.2 Seasonal (Composante saisonnière) :

Le *deuxième graphique* met en évidence la composante saisonnière des taux d'affaires nouvelles, révélant des variations cycliques régulières au sein de chaque année. Pour mieux quantifier ces variations saisonnières, une analyse mensuelle des taux moyens est réalisée, dont les résultats sont présentés dans l'*Histogramme 2.4*.

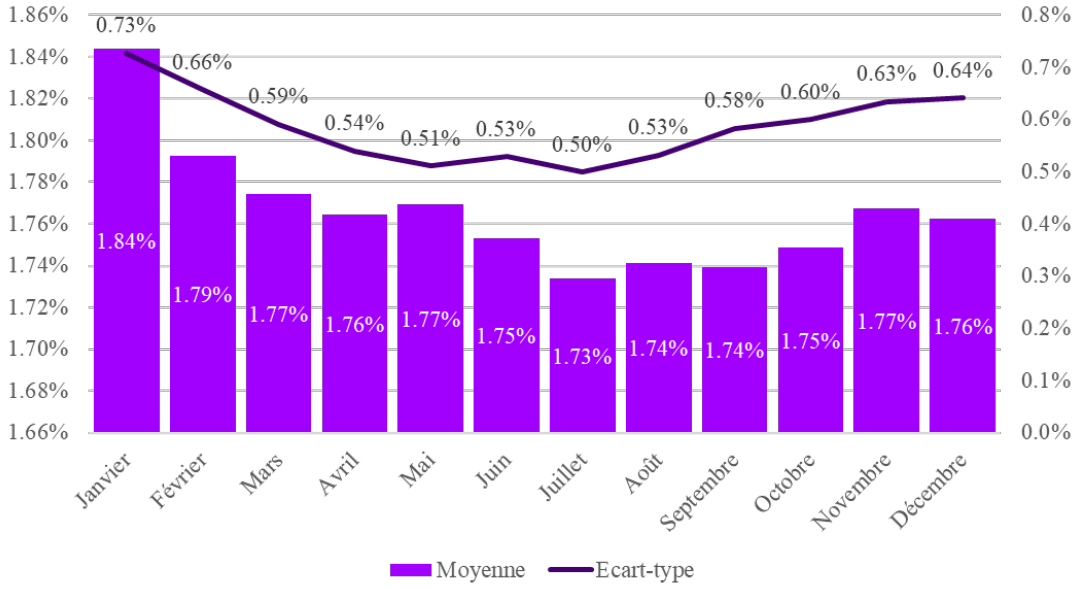


FIGURE 2.4 – Moyenne (axe gauche) et Ecart-type (axe droit) par mois des Taux d'affaires nouvelles (*Banque de France* [3])

L'analyse des taux moyens et de leur variabilité mensuelle montre des tendances notables :

- Pics saisonniers : Le mois de janvier présente la moyenne la plus élevée (1,84%), suivi de près par les mois de novembre à mars (1,76% et 1,79%), suggérant une hausse de l'activité en début et en fin d'année ;
- Creux saisonnier : Les mois de juillet et août affichent les moyennes les plus basses (1,73% et 1,74%), ce qui peut refléter un ralentissement des flux durant cette période estivale ;
- Variabilité : L'écart-type le plus élevé est observé en janvier (0,73%), ce qui pourrait indiquer une volatilité accrue en début d'année, possiblement liée aux incertitudes économiques ou aux ajustements de début d'année ;
- Stabilité relative : Entre juin et octobre, les écarts-types restent plus faibles, suggérant une certaine stabilité des taux durant cette période.

Ces variations mensuelles peuvent s'expliquer par des facteurs tels que :

- Les cycles de consommation et de dépenses saisonnières ;

— Les ajustements budgétaires ou politiques financières en début d'année ;

— La période estivale, souvent marquée par une réduction d'activité.

Comprendre ces fluctuations est essentiel pour optimiser les prévisions et les stratégies, en adaptant la gestion des risques et les politiques tarifaires en fonction des tendances saisonnières anticipées.

2.3.3.3 Trend (Tendance à long terme) :

La tendance à long terme, illustrée dans le *troisième graphique*, offre un aperçu des mouvements structurels du marché. On distingue une décroissance soutenue jusqu'en 2008, suivie d'une période de forte variation jusqu'à 2018 et d'une période stable en déclin jusqu'à maintenant. Cette évolution pourrait refléter des changements dans les politiques de crédit, l'environnement économique global ou les dynamiques du marché immobilier.

2.3.3.4 Remainder (Composante résiduelle) :

Le *dernier graphique* présente les résidus, c'est-à-dire les variations non expliquées par la tendance ou la saisonnalité. Les pics et les creux observés dans cette composante peuvent correspondre à des événements exceptionnels ou à des chocs économiques spécifiques, nécessitant une analyse plus approfondie pour être pleinement compris. Cette identification des événements marquants pouvant potentiellement expliquer ces variations sont abordées dans la *Sous-partie 2.3.4*, permettant ainsi une compréhension plus complète des facteurs influençant les taux d'affaires nouvelles au-delà des tendances générales et des cycles saisonniers.

2.3.4 Comparaison avec les variables économiques

Après avoir examiné en détail la décomposition chronologique des taux d'affaires nouvelles, il est essentiel d'élargir la perspective en explorant les relations entre ces taux, les taux de rachat, et les 29 variables économiques exogènes disponibles. Cette analyse comparative permet non seulement d'identifier les facteurs les plus fortement corrélés avec les variations des taux d'affaires nouvelles et de rachat, mais aussi d'apporter un éclairage supplémentaire sur les fluctuations majeures observées dans la décomposition chronologique.

L'examen des variables les plus corrélées peut potentiellement expliquer certains des pics et des creux notables identifiés dans la composante résiduelle des taux d'affaires nouvelles, tout en offrant des informations sur les dynamiques qui influencent les taux de rachat.

2.3.4.1 Corrélation avec les variables économiques exogènes

L'analyse des corrélations entre les taux d'affaires nouvelles et les variables économiques exogènes, obtenues avec la formule R `cor()`, révèle des relations significatives, comme le montre le *Tableau 2.7* ci-dessous :

	Données					Corrélations avec les Taux d'affaires nouvelles
y	**Flux mensuels des crédits nouveaux à l'habitat aux particuliers résidents**					100 %
x_1	**Taux d'intérêt annuel des crédits nouveaux à l'habitat des particuliers**					**-86,0 %**
x_2	**Encours des prêts immobiliers à taux zéro**					58,7 %
x_3	**Indice des prix des logements**		**Neufs**	**Tous logements**		**62,3 %**
x_4			Anciens	Maisons		45,5 %
x_5				Appartements		56,9 %
x_6	**Indice du coût de production dans la construction (ICP-F)**					53,0 %
x_7	**Indice du coût de la construction (ICC)**					52,9 %
x_8	**Taux de chômage au sens du BIT**					-8,3 %
x_9	**Dette publique en % de PIB**					**66,7 %**
x_{10}	**Évolution annuelle de l'IPC**					-25,7 %
x_{11}	**Produit Intérieur Brut en Mds€ 2020**					**63,5 %**
x_{12}	**Évolution annuelle relative du Pouvoir d'achat du revenu disponible brut**					18,6 %
x_{13}	**Taux d'épargne des ménages en % du revenu brut**					27,1 %
x_{14}	**Volume de construction**	Nombre de logements	Autorisés	Collectif et Résidence		35,9 %
x_{15}				Individuel		-31,2 %
x_{16}			Commencés	Collectif et Résidence		51,0 %
x_{17}				Individuel		-39,2 %
x_{18}		Surfaces de logements	Autorisés	Collectif et Résidence		24,8 %
x_{19}				Individuel		-33,4 %
x_{20}			Commencés	Collectif et Résidence		38,3 %
x_{21}				Individuel		-43,9 %
x_{22}	**Volume de commercialisation**	Nombre de logements	Mis en vente aux particuliers	Nouvelles constructions	Collectif et Résidence	32,9 %
x_{23}					Individuel	-43,1 %
x_{24}			Réservés à la vente par des particuliers	Nouvelles constructions	Collectif et Résidence	59,0 %
x_{25}					Individuel	-36,8 %
x_{26}		Encours de logements disponibles à la vente aux particuliers		Nouvelles constructions	Collectif et Résidence	52,4 %
x_{27}					Individuel	-35,3 %
x_{28}		Délai moyen d'écoulement de l'encours de logements disponibles à la vente aux particuliers		Nouvelles constructions	Collectif et Résidence	5,2 %
x_{29}					Individuel	-1,0 %

TABLE 2.7 – Corrélations des variables économiques exogènes[1] avec les taux d'affaires nouvelles

Les corrélations les plus importantes sont observées avec :

- Le taux d'intérêt annuel des crédits nouveaux à l'habitat des particuliers (-86,0%) ;
- La dette publique en pourcentage du PIB (66,7%) ;
- Le Produit Intérieur Brut en milliards d'euros 2020 (63,5%) ;
- L'indice des prix des logements neufs (62,3%).

Il est intéressant de noter que ces trois premières variables avaient déjà été identifiées dans le *Chapitre 1.2.2* comme étant des indicateurs pro-actifs dans l'évolution du marché économique. Leur forte corrélation avec les taux d'affaires nouvelles confirme leur rôle central dans la dynamique du marché du crédit à l'habitat, renforçant ainsi l'analyse préliminaire effectuée précédemment.

Ces fortes corrélations suggèrent un rôle potentiellement décisif de ces variables dans l'explication des variations des taux d'affaires nouvelles. De plus, des corrélations positives sont observées pour ce qui touche aux volumes de constructions et commercialisations de logements collectifs, et négatives pour les logements individuels. Cette observation semble cohérente avec le fait que, en volume, le collectif est bien plus important que l'individuel et entraîne donc une demande de

[1] Le lexique des indices x_i utilisés dans cette analyse sont disponibles en *Annexe B.1.*

prêts plus élevée pour une même construction ou un même espace utilisé.

		x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15
Ancienneté	1 à 5 ans	-14%	-67%	-55%	-68%	-68%	-66%	-63%	58%	-35%	-70%	-36%	13%	-53%	22%	6%
	6 à 10 ans	-7%	-67%	-51%	-63%	-67%	-58%	-56%	58%	-36%	-59%	-28%	14%	-59%	22%	-1%
	11 à 15 ans	-33%	-10%	-6%	-19%	-17%	-25%	-24%	-9%	-11%	-29%	3%	57%	-14%	47%	39%
	16 à 20 ans	-52%	94%	91%	81%	89%	84%	84%	-88%	80%	55%	80%	28%	59%	-2%	-21%

		x16	x17	x18	x19	x20	x21	x22	x23	x24	x25	x26	x27	x28	x29
Ancienneté	1 à 5 ans	28%	6%	18%	5%	24%	3%	46%	44%	41%	66%	-5%	24%	-41%	-47%
	6 à 10 ans	21%	-2%	17%	-3%	17%	-5%	42%	40%	34%	58%	1%	30%	-28%	-32%
	11 à 15 ans	52%	41%	46%	38%	52%	37%	61%	28%	68%	60%	-17%	-43%	-69%	-66%
	16 à 20 ans	10%	-30%	-4%	-21%	4%	-29%	-20%	-67%	12%	-58%	51%	-68%	22%	22%

FIGURE 2.5 – Corrélations des variables économiques exogènes avec les taux de rachat des crédits immobiliers par plage d'années d'ancienneté du crédit[1]

En ce qui concerne les taux de rachat, l'analyse des corrélations avec les variables économiques exogènes, observable dans le *Tableau 2.5* ci-dessus et obtenue avec la formule R `cor()`, révèle des tendances intéressantes selon l'ancienneté des prêts :

- Pour les prêts de 1 à 10 ans d'ancienneté :
  - Une forte corrélation négative est observée avec l'encours des prêts à taux zéro (-67%), les indices des prix des logements (neufs et anciens, autour de -60% à -68%), et l'évolution annuelle de l'IPC (-70% à -59%) ;
  - Une corrélation positive notable apparaît avec le taux de chômage (58%).
- Pour les prêts de 11 à 15 ans :
  - Les corrélations sont généralement plus faibles, mais on note une corrélation positive significative avec l'évolution du pouvoir d'achat (57%) et les volumes de construction et de commercialisation des logements neufs (entre 40% et 68%).
- Pour les prêts de 16 à 20 ans :
  - On observe un renversement des tendances avec des corrélations fortement positives avec les indices de prix des logements (80% à 91%), la dette publique (80%), et le PIB (80%) ;
  - Une forte corrélation négative apparaît avec le taux de chômage (-88%).

Ces résultats suggèrent que les facteurs influençant les décisions de rachat varient considérablement selon l'ancienneté du prêt. Les prêts plus récents semblent plus sensibles aux conditions économiques immédiates, tandis que les prêts plus anciens sont davantage influencés par les tendances économiques à long terme et l'évolution du marché immobilier.

Il est particulièrement intéressant de noter le changement de signe des corrélations pour de nombreuses variables entre les prêts récents et anciens, ce qui pourrait indiquer des stratégies de rachat différentes selon la maturité du prêt et le contexte économique global.

Parmi les variables économiques examinées, l'inflation (mesurée par l'évolution annuelle de l'IPC) et les taux d'intérêt méritent une attention particulière pour deux raisons distinctes. D'une part, l'inflation est au cœur des scénarios de chocs économiques analysés dans le *Chapitre 6*, ce qui justifie son examen approfondi. D'autre part, les taux d'intérêt se révèlent être une variable clé

[1] Les corrélations des taux de rachat pour toutes les années d'ancienneté de crédit avec les variables exogènes utilisées dans cette analyse sont disponibles en *Annexe B.2.*

fortement corrélée non seulement avec les taux d'affaires nouvelles et les taux de rachat, mais aussi avec l'inflation elle-même, comme démontré dans le *Chapitre 1*. Cette analyse préliminaire de ces deux indicateurs économiques permet de mieux appréhender les dynamiques sous-jacentes du marché du crédit immobilier et leurs potentielles répercussions sur l'assurance emprunteur, préparant ainsi le terrain pour les analyses plus poussées des chapitres suivants.

2.3.4.2 Focus sur la relation entre les taux d'affaires nouvelles, les taux de rachat et les taux d'intérêt des crédits nouveaux

Une analyse plus approfondie de la relation entre les taux d'intérêt et les taux d'affaires nouvelles révèle une dynamique particulièrement intéressante, illustrée dans la *Figure 2.6*. Cette figure met en évidence une dualité marquée entre l'évolution des taux d'intérêt et celle des taux d'affaires nouvelles.

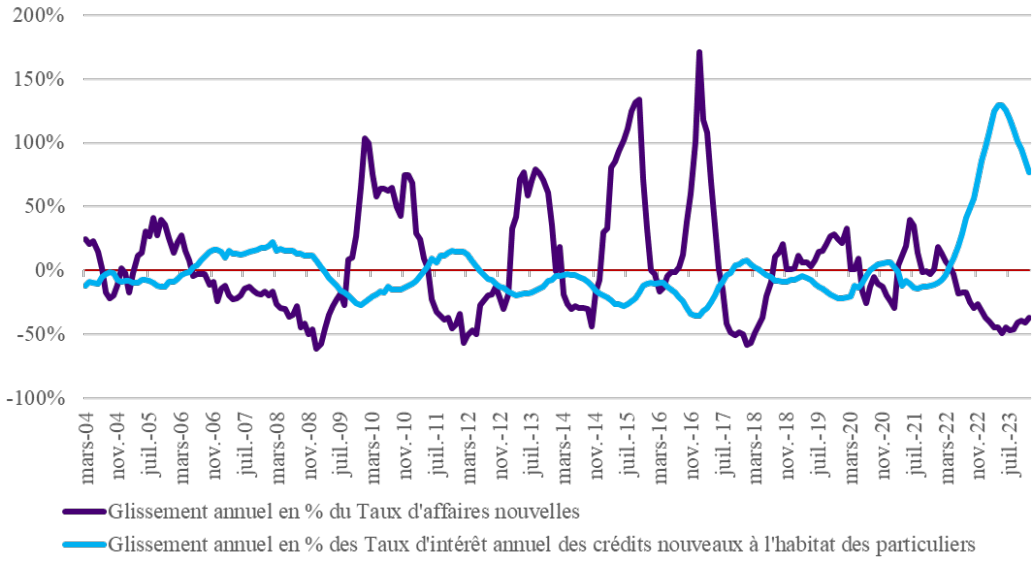


FIGURE 2.6 – Glissements annuels en % comparés des taux d'affaires nouvelles et des taux d'intérêt annuel des crédits nouveaux à l'habitat des particuliers (*Banque de France* [3])

Il est clairement observable que lorsque les taux d'intérêt sont en baisse, les taux d'affaires nouvelles connaissent une forte augmentation. Cette tendance s'explique par l'attractivité accrue des emprunts immobiliers lorsque le coût du crédit diminue. À l'inverse, dès que les taux d'intérêt amorcent une remontée, ou même lorsque leur baisse ralentit, un net ralentissement du marché du crédit immobilier est constaté.

Cette relation inverse est particulièrement visible dans le contexte actuel. Depuis 2022, face à une hausse significative des taux d'intérêt, une baisse marquée des taux d'affaires nouvelles est observée. Cette situation illustre de manière frappante l'impact considérable des taux d'intérêt sur le dynamisme du marché immobilier et, par extension, sur le taux d'affaires nouvelles en assurance emprunteur.

Cette observation renforce l'importance du taux d'intérêt comme variable explicative clé dans la modélisation des taux d'affaires nouvelles. Cependant, pour obtenir une compréhension plus complète de la dynamique du marché, il est essentiel d'examiner non seulement l'impact des taux d'intérêt sur les nouveaux crédits, mais aussi leur influence sur les comportements de rachat des crédits existants. Cette analyse permet de mieux appréhender les stratégies de gestion des risques

et de développement commercial dans le secteur de l'assurance emprunteur.

Dans cette optique, la *Figure 2.7* offre un éclairage complémentaire en illustrant la relation complexe entre les taux de rachat des crédits immobiliers, selon différentes plages d'ancienneté, et l'évolution des taux d'intérêt des crédits nouveaux. Cette analyse révèle des dynamiques intéressantes qui corroborent et enrichissent les corrélations observées précédemment.

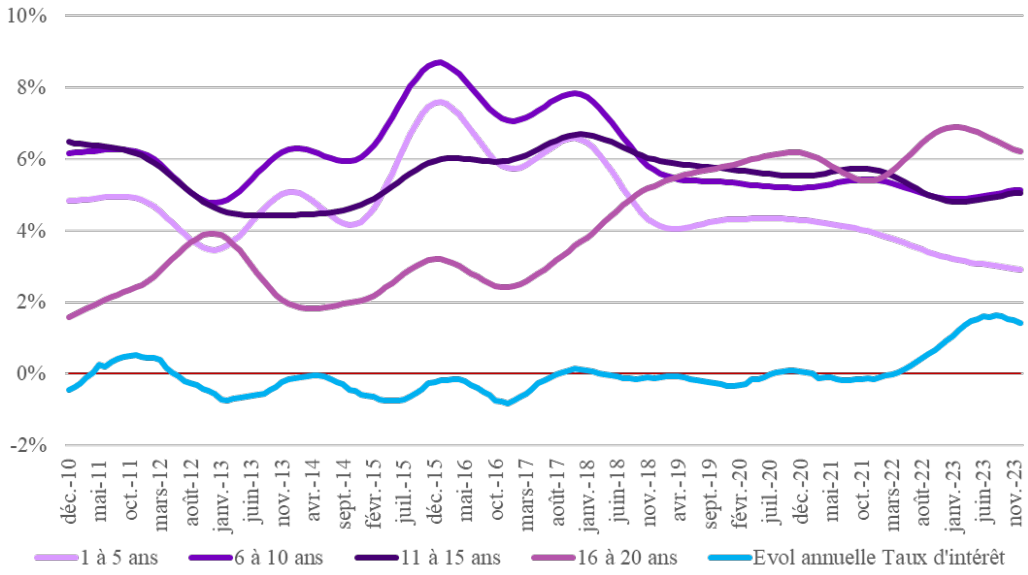


FIGURE 2.7 – Comparaison Taux de rachat des crédits immobiliers par plage d'années d'ancienneté du crédit et glissement annuel en % des taux d'intérêt annuel des crédits nouveaux à l'habitat des particuliers

Pour les prêts récents (1 à 5 ans et 6 à 10 ans), on observe une relative stabilité des taux de rachat, oscillant respectivement entre 5-7% et 4-6%. Ces catégories montrent une sensibilité modérée aux variations des taux d'intérêt, avec une corrélation négative faible à modérée (-14% à -7%) comme indiqué précédemment *Tableau 2.5*. Cette stabilité relative pourrait s'expliquer par d'autres facteurs influençant les décisions de rachat pour les prêts récents, tels que les coûts de transaction ou les pénalités de remboursement anticipé.

Les prêts d'ancienneté moyenne (11 à 15 ans) présentent une dynamique plus volatile, avec un pic notable autour de 2015 suivi d'une baisse progressive. Cette catégorie montre une corrélation négative plus prononcée avec les taux d'intérêt (-33%), suggérant une sensibilité accrue aux opportunités de refinancement lorsque les taux baissent.

Les prêts les plus anciens (16 à 20 ans) affichent la corrélation négative la plus forte avec les taux d'intérêt (-52%). Cela se reflète dans leur comportement sur le graphique, où l'on observe une tendance inverse marquée par rapport à l'évolution des taux d'intérêt, particulièrement visible après 2015. Un point d'inflexion notable se situe autour de 2014-2015, où un pic de rachats est constaté pour toutes les catégories d'ancienneté, coïncidant avec une période de baisse significative des taux d'intérêt. Cette observation souligne l'impact considérable des conditions de marché favorables sur les décisions de rachat, toutes anciennetés confondues.

À partir de 2020, alors que les taux d'intérêt amorcent une remontée, une tendance générale à la baisse des taux de rachat est visible, et cela est particulièrement prononcée pour les prêts les plus anciens. Cette tendance illustre clairement l'effet dissuasif de la hausse des taux sur les décisions

de rachat.

Ces observations renforcent l'importance de considérer non seulement le niveau absolu des taux d'intérêt, mais aussi leur évolution relative dans le temps, comme facteur clé influençant les comportements de rachat. Cette analyse souligne également la nécessité d'une approche différenciée dans la modélisation et la gestion des risques de rachat selon l'ancienneté des prêts, étant donné leurs sensibilités variables aux fluctuations des taux d'intérêt.

2.3.4.3 Focus sur la relation entre les taux d'affaires nouvelles, les taux de rachat et l'inflation

L'inflation, caractérisée dans ce mémoire par l'évolution de l'Indice des Prix à la Consommation (IPC), joue un rôle similaire aux taux d'intérêt dans son influence sur les taux d'affaires nouvelles, bien que son impact soit moins prononcé. Il est important de noter que l'IPC est utilisé ici comme proxy pour mesurer l'inflation, car il représente l'évolution du niveau général des prix à la consommation et est largement reconnu comme un indicateur fiable de l'inflation.

Cette relation est d'abord mise en évidence par le coefficient de corrélation entre l'inflation et les taux d'affaires nouvelles, qui s'établit à -25,7%. Bien que moins forte que la corrélation observée avec les taux d'intérêt (-86,0%), cette valeur négative indique néanmoins une tendance inverse entre l'inflation et les taux d'affaires nouvelles.

La *Figure 2.8* illustre cette relation de manière plus détaillée. On peut observer que, malgré le fait que l'inflation soit toujours positive sur la période étudiée, ses variations ont un impact notable sur les taux d'affaires nouvelles. De manière générale, une augmentation du taux d'inflation semble correspondre à une hausse des taux d'affaires nouvelles, tandis qu'une baisse de l'inflation est associée à une diminution de ces taux.

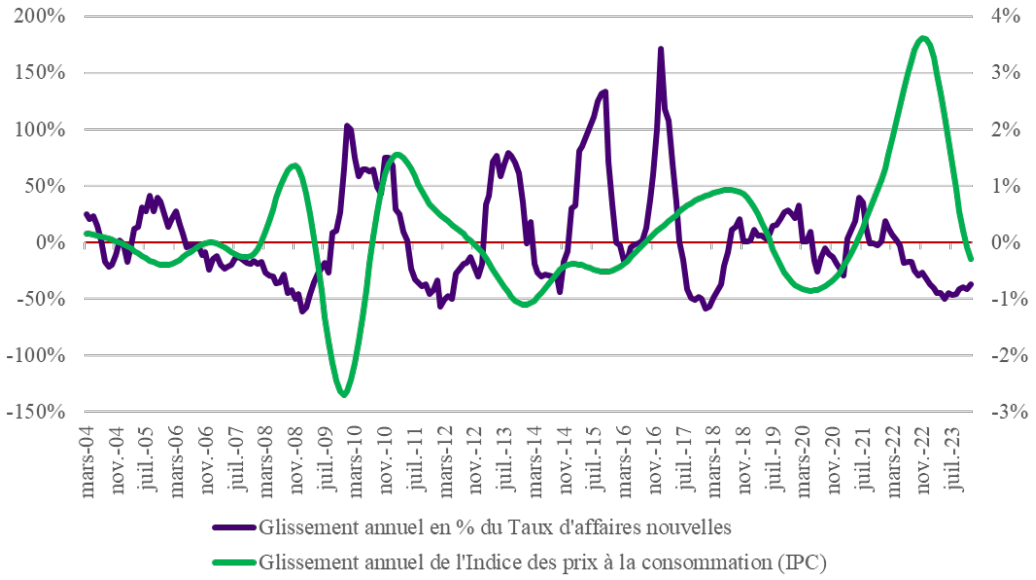


FIGURE 2.8 – Glissements annuels en % comparés des taux d'affaires nouvelles (axe gauche) et de l'IPC (axe droit) (*Banque de France* [3] & *INSEE* [27])

Cette dynamique peut s'expliquer par plusieurs facteurs :

— L'anticipation de la hausse des prix immobiliers en période d'inflation élevée peut inciter les ménages à accélérer leurs projets d'achat immobilier ;

— L'inflation peut éroder la valeur réelle des dettes, rendant l'emprunt plus attractif pour les acheteurs potentiels ;

— En période de forte inflation, l'immobilier est souvent perçu comme une valeur refuge, ce qui peut stimuler la demande de crédits.

Cependant, il est important de noter que cette relation n'est pas aussi directe et prononcée que celle observée avec les taux d'intérêt. D'autres facteurs économiques et réglementaires peuvent influencer cette dynamique, soulignant la complexité du marché du crédit immobilier.

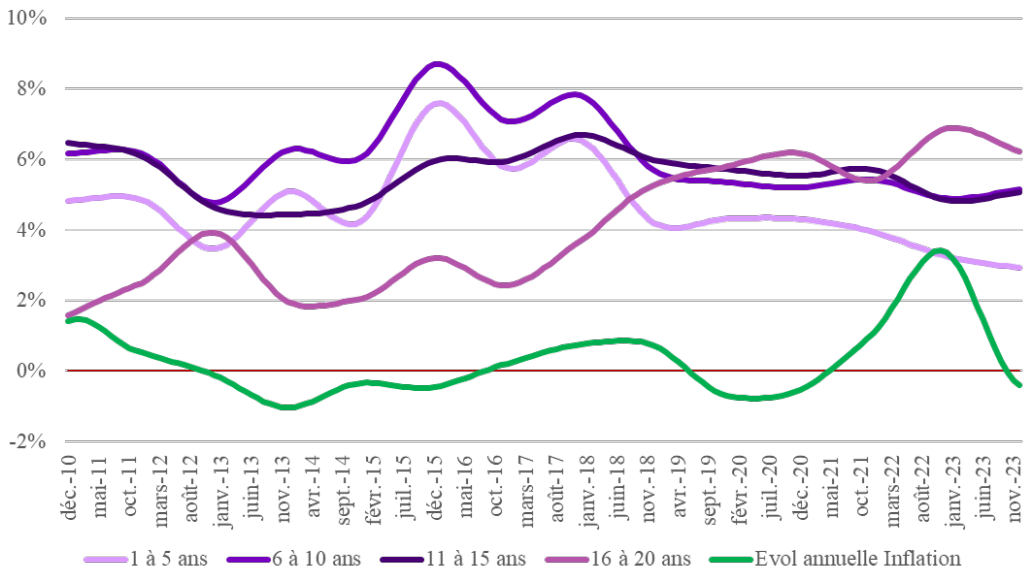


FIGURE 2.9 – Comparaison Taux de rachat des crédits immobiliers par plage d'années d'ancienneté du crédit et glissement annuel en % de l'IPC

D'autre part, la *Figure 2.9* révèle des relations distinctes entre l'inflation et les taux de rachat selon l'ancienneté des prêts :

- Prêts récents (1 à 5 ans) :
  Ces prêts montrent une forte corrélation négative (-70%) avec l'inflation. Sur le graphique, on observe que leurs taux de rachat restent relativement stables autour de 4-6%, malgré les fluctuations de l'inflation. Cette stabilité suggère que les emprunteurs récents sont moins sensibles aux variations de l'inflation, probablement en raison des conditions avantageuses de leurs prêts actuels.
- Prêts de 6 à 10 ans :
  Avec une corrélation négative de -59% avec l'inflation, ces prêts montrent plus de volatilité dans leurs taux de rachat. Les pics observés vers 2015 et 2017 coïncident avec des périodes de faible inflation, suggérant que les emprunteurs sont plus enclins à racheter leurs prêts lorsque l'inflation est basse.
- Prêts de 11 à 15 ans :
  La corrélation négative plus modérée (-29%) avec l'inflation se reflète dans des fluctuations plus importantes des taux de rachat. Cette catégorie semble plus réactive aux changements économiques, avec des pics de rachat correspondant à des périodes de stabilité de l'inflation.
- Prêts anciens (16 à 20 ans) :
  Contrairement aux autres catégories, ces prêts affichent une corrélation positive (55%) avec l'inflation. Leurs taux de rachat, bien que généralement plus bas, semblent augmenter légèrement lors des périodes de hausse de l'inflation, notamment après 2020.

Cette analyse révèle que l'impact de l'inflation sur les décisions de rachat varie considérablement selon l'ancienneté du prêt. Les prêts récents semblent moins affectés par l'inflation, tandis que les prêts plus anciens montrent une sensibilité accrue, avec des comportements de rachat qui s'inversent pour les prêts les plus anciens.

Cette analyse de la relation entre l'inflation et les taux d'affaires nouvelles apporte un éclairage supplémentaire sur les dynamiques du marché et renforce l'importance de prendre en compte les indicateurs macro-économiques dans la modélisation et la prévision des taux d'affaires nouvelles en assurance emprunteur et des taux de rachat de crédit.

2.3.5 Synthèse

L'analyse statistique des taux d'affaires nouvelles et des taux de rachat en France a révélé plusieurs aspects clés du marché du crédit immobilier et de l'assurance emprunteur :

1. Une forte variabilité des taux d'affaires nouvelles, avec une moyenne de 1.775% et une amplitude de 3,598 points de pourcentage entre le taux minimum de 0,661% et le maximum de 4,259%., indiquant des fluctuations importantes du marché ;
2. Une décomposition chronologique mettant en évidence :
   — Une tendance à la baisse jusqu'en 2008, suivie d'une période de forte variation jusqu'en 2018, puis d'une stabilisation et d'un léger déclin ;
   — Une forte saisonnalité, avec un pic en début d'année et un creux en juin ;
   — Des composantes résiduelles suggérant l'influence d'événements exceptionnels.
3. Des corrélations significatives avec des variables économiques, notamment :
   — Une forte corrélation négative avec les taux d'intérêt (-86,0%) ;
   — Une corrélation modérée avec l'inflation, mesurée par l'IPC (-25,7%).
4. Une relation inverse marquée entre les taux d'intérêt et les taux d'affaires nouvelles, ainsi qu'un impact notable de l'inflation sur les décisions d'emprunt ;
5. Des variations significatives des taux de rachat selon l'ancienneté des prêts :
   — Des taux de rachat plus élevés pour les prêts de 6 à 10 ans d'ancienneté ;
   — Une stabilité relative pour les prêts récents (1 à 5 ans) ;
   — Une tendance à la baisse pour les prêts les plus anciens (16 à 20 ans).
6. Des corrélations distinctes entre les taux de rachat et les variables économiques selon l'ancienneté des prêts :
   — Des corrélations négatives avec les taux d'intérêt :
     - Une faible corrélation négative pour les prêts récents (-14% pour 1 à 5 ans) ;
     - Une corrélation négative modérée pour les prêts de moyenne durée (-33% pour 11 à 15 ans) ;
     - Une forte corrélation négative pour les prêts les plus anciens (-52% pour 16 à 20 ans).
   — Des corrélations variables avec l'inflation :
     - Une forte corrélation négative avec l'inflation pour les prêts récents (-70% pour 1 à 5 ans) ;
     - Une corrélation positive avec l'inflation pour les prêts les plus anciens (55% pour 16 à 20 ans).

Ces observations soulignent l'importance de prendre en compte les facteurs macro-économiques et les variations saisonnières dans la modélisation et la prévision des taux d'affaires nouvelles en assurance emprunteur.

2.4 Conclusion

Ce chapitre permet d'établir une base pour l'analyse et la modélisation des taux d'affaires nouvelles en France, ainsi que des taux de rachat associés. À travers une sélection et un traitement des données provenant de sources telles que la Banque de France et l'INSEE, un ensemble de variables pertinentes est constitué, couvrant une période riche en événements économiques.

L'analyse statistique révèle la complexité et la dynamique du marché du crédit immobilier et de l'assurance emprunteur. La décomposition chronologique des taux d'affaires nouvelles met en lumière des patterns saisonniers marqués et une tendance à long terme reflétant les évolutions structurelles du marché. L'étude des corrélations avec diverses variables économiques souligne l'importance des taux d'intérêt et, dans une moindre mesure, de l'inflation dans les fluctuations des taux d'affaires nouvelles.

Parallèlement, l'analyse des taux de rachat met en évidence des comportements distincts selon l'ancienneté des prêts. Les corrélations observées entre les taux de rachat, les taux d'intérêt et l'inflation varient considérablement en fonction de la maturité des prêts, soulignant la nécessité d'une approche différenciée dans la modélisation des comportements de rachat.

Ces observations fournissent des indications précieuses sur les facteurs influençant à la fois les taux d'affaires nouvelles et les taux de rachat en assurance emprunteur. Elles soulignent notamment l'importance de prendre en compte non seulement les tendances à long terme, mais aussi les variations saisonnières, les chocs économiques ponctuels et l'ancienneté des prêts dans toute tentative de modélisation.

Les résultats de cette analyse préparatoire ouvrent la voie au chapitre suivant, qui se concentre sur la modélisation proprement dite du taux d'affaires nouvelles et des lois de rachat de crédit. Les éléments clés identifiés, notamment l'importance des taux d'intérêt et de l'inflation, la saisonnalité du marché, ainsi que les différences de comportement selon l'ancienneté des prêts, servent de base pour développer des modèles précis et robustes.

Cette transition vers la modélisation permet d'exploiter pleinement la richesse des données analysées, avec l'objectif de fournir des outils prédictifs pour le secteur de l'assurance emprunteur, tenant compte de la complexité et de la volatilité inhérentes au marché du crédit immobilier, ainsi que des dynamiques spécifiques des comportements de rachat.

Chapitre 3

DÉMARCHE THÉORIQUE : Construction des lois d'Affaires Nouvelles et de Rachat

Cette série de 3 chapitres s'appuie sur les analyses et les observations détaillées et présentées dans le *Chapitre 2*. Les variables économiques identifiées comme significatives et les tendances observées dans les taux d'affaires nouvelles et de rachat servent de base solide pour les modélisations qui suivent.

Le présent chapitre se concentre sur la démarche théorique qui sert à la modélisation de deux aspects centraux de l'assurance emprunteur : le taux d'affaires nouvelles et la loi de rachat de crédit. L'objectif est de développer des modèles prédictifs robustes permettant d'anticiper les tendances futures et d'analyser les sensibilités aux variables économiques clés.

Ce chapitre, ainsi que les deux suivants, s'articulent autour de trois axes principaux :

1. **Chapitre 3 : Théorie des séries temporelles et démarche méthodologique :** Ce chapitre présente les fondements théoriques et l'approche méthodologique adoptée pour la modélisation des séries temporelles dans le contexte de l'assurance emprunteur ;
2. **Chapitre 4 : Modélisation du taux d'affaires nouvelles :** Ce chapitre utilise les techniques de séries temporelles pour analyser et prédire les *Taux d'affaires nouvelles de crédits immobiliers*. L'approche vise à capturer les tendances, les saisonnalités et les autres *patterns* identifiés dans le chapitre précédent ;
3. **Chapitre 5 : Modélisation des taux de rachat :** Ce chapitre explore l'application des techniques de séries temporelles pour estimer les *Taux de rachat anticipé des crédits immobiliers*. L'approche ici diffère car il n'est pas question de saisonnalité dans ces données mais uniquement de tendance. De plus, les lois sont décomposées en fonction des années d'ancienneté du prêt, il en ressort donc 20 lois différentes pour caractériser cette modélisation.

L'objectif final de ces modélisations, qui sont développées dans les chapitres suivants, est de permettre des projections fiables des données, qui serviront de base à une étude de sensibilité. Cette étude se concentrera particulièrement sur l'impact de l'*Inflation* caractérisée ici par l'*Évolution de l'IPC*. Ces études d'impact seront analysées sur un calcul de BE (Best Estimate) sur un portefeuille fictif, permettant ainsi d'évaluer les implications de ces variables économiques.

Ce chapitre présente l'approche méthodologique adoptée pour la modélisation des séries temporelles dans le contexte de l'assurance emprunteur. Elle couvre les fondements théoriques, les techniques

d'analyse et les modèles spécifiques utilisés pour comprendre et prédire les dynamiques des taux d'affaires nouvelles et des taux de rachat. L'objectif est d'établir un cadre rigoureux permettant une modélisation précise et fiable, adaptée aux particularités du secteur de l'assurance emprunteur.

3.1 Introduction aux séries temporelles

Cette partie se concentre sur les aspects fondamentaux des séries temporelles, en abordant leurs caractéristiques essentielles et les concepts clés indispensables à leur compréhension et leur manipulation. L'approche structurée permet de :

1. Définir les *propriétés fondamentales* des séries temporelles, incluant l'ordre temporel, la dépendance temporelle et l'intervalle de temps ;
2. Examiner en détail les *concepts clés* tels que la stationnarité, l'auto-corrélation, la tendance, la saisonnalité, les cycles et le bruit.

Cette analyse des composantes et des caractéristiques des séries temporelles est primordiale pour le développement de modèles prédictifs précis et fiables, particulièrement dans le contexte de l'assurance emprunteur, de l'analyse des crédits immobiliers et des comportements de rachat.

Cette introduction aux séries temporelles sert de base théorique solide pour les analyses plus avancées et les modélisations ultérieures. Les concepts théoriques et les modèles sont principalement illustrés à travers l'exemple du taux d'affaires nouvelles, permettant ainsi une compréhension approfondie avant leur application. Il y figure également une présentation détaillée des différents modèles de séries temporelles, allant des modèles AR (Auto-regressive) et MA (Moving Average) de base jusqu'aux modèles plus complexes comme SARIMAX (Seasonal Auto-regressive Integrated Moving Average with eXogenous variables). Enfin, une méthodologie de modélisation complète est exposée, couvrant la préparation des données, les tests de stationnarité, et le processus de sélection du meilleur modèle.

Cette approche méthodique et rigoureuse permet d'établir un cadre solide pour la prévision du taux d'affaires nouvelles et l'analyse des lois de rachat en assurance emprunteur, tenant compte des spécificités de chaque type de données.

3.1.1 Caractéristiques fondamentales

Une série temporelle, également appelée série chronologique, est une suite de données mesurées ou enregistrées à des intervalles de temps réguliers. Elle permet d'étudier l'évolution d'une variable au fil du temps et de prévoir ses tendances futures. L'analyse des séries temporelles vise à identifier des modèles dans les données, fournissant des informations significatives sur des événements passés et futurs, tels que des variations saisonnières, des anomalies ou des événements uniques.

Les séries temporelles se caractérisent par trois propriétés fondamentales :

1. **Ordre temporel** : Les observations sont organisées chronologiquement, ce qui est essentiel pour analyser l'évolution d'une variable dans le temps. Dans le contexte des taux d'affaires nouvelles, cet ordre permet d'identifier des tendances, des cycles et des variations saisonnières qui pourraient autrement passer inaperçus. Cette caractéristique est primordiale pour la prévision, car elle aide à établir des relations entre les valeurs passées et futures, fondamentales pour la modélisation prédictive ;
2. **Dépendance temporelle** : Les valeurs actuelles peuvent être influencées par les valeurs passées, un phénomène souvent mesuré par l'auto-corrélation. Cette dépendance est particulièrement importante pour la modélisation, car elle permet d'utiliser les données historiques

pour faire des prévisions. Par exemple, si une augmentation des taux d'affaires nouvelles est observée au cours des mois précédents, il est probable que cette tendance se poursuive dans un avenir proche. Comprendre cette dépendance aide à construire des modèles plus précis et à anticiper les fluctuations du marché ;

3. **Intervalle de temps régulier** : Les observations sont généralement prises à intervalles réguliers, facilitant l'analyse. Dans le cas des taux d'affaires nouvelles, chaque observation représente le montant total de crédits accordés au cours d'un mois. Cette régularité est essentielle pour appliquer diverses techniques d'analyse et de modélisation, garantissant que les données soient comparables et cohérentes dans le temps.

L'analyse des séries temporelles utilise des techniques statistiques et de modélisation spécifiques pour comprendre les facteurs influençant leur évolution. Ces méthodes permettent d'exploiter pleinement les caractéristiques uniques des séries temporelles, offrant des outils puissants pour l'analyse et la prévision dans divers domaines.

3.1.2 Concepts clés

Une série temporelle est composée d'éléments structurels et non-structurels qui peuvent être décomposés en plusieurs composantes fondamentales :

— **Tendance ($\mathbf{T_t}$)** : Elle représente le mouvement à long terme de la série. La tendance peut être croissante, décroissante ou stable, reflétant des changements structurels tels que l'évolution économique ou démographique. Identifier la tendance est indispensable pour comprendre la direction générale d'une série et ajuster les modèles prédictifs en conséquence ;

— **Composante cyclique ($\mathbf{C_t}$)** : Elle correspond à des fluctuations irrégulières à moyen terme, souvent liées aux cycles économiques. Contrairement à la saisonnalité, les cycles ne suivent pas nécessairement un calendrier fixe et peuvent varier en durée et en amplitude. Comprendre ces cycles peut aider à anticiper les phases de croissance ou de récession ;

— **Composante saisonnière ($\mathbf{S_t}$)** : Elle représente des fluctuations périodiques régulières, généralement au sein d'une année. Ces variations peuvent être dues à des facteurs comme le climat, les vacances ou les cycles économiques annuels. La détection de la saisonnalité permet d'améliorer la précision des prévisions en ajustant les modèles pour tenir compte de ces variations régulières ;

— **Composante résiduelle ($\varepsilon_\mathbf{t}$)** : Aussi appelée "bruit", elle contient les variations aléatoires non expliquées par les autres composantes. Bien que le bruit ne puisse pas être modélisé directement, il est important de le reconnaître pour évaluer la précision et l'incertitude des prévisions.

Ces composantes peuvent être combinées selon différents schémas :

$$\begin{aligned} \text{Additif :} \quad & Y_t = T_t + C_t + S_t + \varepsilon_t \\ \text{Multiplicatif :} \quad & Y_t = T_t C_t S_t + \varepsilon_t \\ \text{Multiplicatif complet :} \quad & Y_t = T_t C_t S_t \varepsilon_t \end{aligned}$$

En plus de ces composantes, d'autres concepts clés sont essentiels pour l'analyse des séries temporelles :

— **La Stationnarité** : C'est une propriété indiquant que les caractéristiques statistiques de la série (moyenne, variance, auto-corrélation) sont indépendantes du temps et restent constantes au fil de celui-ci. La stationnarité est impérative pour l'application de nombreux modèles de séries temporelles ;

— **L'Auto-corrélation** : Elle mesure la corrélation d'une série avec elle-même à différents décalages temporels. L'auto-corrélation est essentielle pour identifier les *patterns* répétitifs et déterminer l'ordre des modèles auto-régressifs (AR).

Comprendre ces composantes et concepts permet de décomposer efficacement une série temporelle et d'appliquer des modèles appropriés pour l'analyse et la prévision des données.

3.2 Modèles de séries temporelles

Les composantes fondamentales des séries temporelles étant dorénavant introduites, il est essentiel d'explorer les différents modèles utilisés pour leur analyse et leur prévision. Ces modèles permettent de capturer les diverses caractéristiques des séries temporelles, telles que la tendance, la saisonnalité, et les dépendances temporelles, afin de produire des prévisions précises et fiables.

Cette section présente une progression des modèles de séries temporelles, depuis les modèles les plus élémentaires jusqu'aux modèles les plus sophistiqués :

- ***AR*** (*Auto-regressive*) : Un modèle auto-régressif de base ;
- **MA** (*Moving Average*) : Un modèle de moyenne mobile de base ;
- **ARMA** (*Auto-regressive Moving Average*) : Un modèle de base combinant les processus auto-régressifs et de moyenne mobile ;
- **ARIMA** (*Auto-regressive Integrated Moving Average*) : Une extension d'ARMA intégrant une composante de différenciation pour traiter les séries non-stationnaires ;
- **SARIMA** (*Seasonal ARIMA*) : Un modèle ARIMA augmenté pour prendre en compte les *patterns* saisonniers ;
- **SARIMAX** (*Seasonal ARIMA with eXogenous variables*) : Une version plus avancée de SARIMA permettant l'inclusion de variables exogènes dans le modèle.

Chacun de ces modèles offre des capacités croissantes pour capturer la complexité des séries temporelles. Ce chapitre examine donc leurs principes de base, leurs avantages et leurs limites, ainsi que leur pertinence quant à l'analyse du marché de l'assurance emprunteur.

3.2.1 Modèles AR (Auto-regressive) et MA (Moving Average)

- **AR (Auto-regressive)**

Le modèle Auto-Régressif (AR) est une approche qui modélise la valeur actuelle d'une série temporelle comme une combinaison linéaire de ses valeurs passées, plus un terme d'erreur. L'ordre du modèle AR(p), noté p, indique combien de valeurs passées sont utilisées pour prédire la valeur actuelle.

Formulation mathématique : Un modèle AR(p) s'écrit sous la forme :

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \varepsilon_t$$

Où :

- Y_t est la valeur de la série au temps t ;
- μ est une constante réelle ;
- $\phi_1, \phi_2, ..., \phi_p$ sont les paramètres du modèle ;
- $Y_{t-1}, Y_{t-2}, ..., Y_{t-p}$ sont les valeurs passées de la série (appelées "lags") ;
- ε_t est le terme d'erreur au temps t.

Caractéristiques principales :

- **Dépendance temporelle :** Le modèle AR capture explicitement la dépendance d'une observation sur ses valeurs passées ;
- **Stationnarité :** Pour être valide, le processus AR doit être stationnaire, ce qui impose certaines contraintes sur les paramètres ϕ ;
- **Fonction d'auto-corrélation :** L'ACF d'un processus AR décroît exponentiellement ou de manière sinusoïdale.

Applications : Dans le contexte des *Taux d'affaires nouvelles*, un modèle AR pourrait être utile pour capturer les tendances à court terme et les dépendances temporelles. Par exemple, le taux d'affaires nouvelles d'un mois donné pourrait être influencé par les taux des mois précédents, reflétant une certaine inertie du marché immobilier.

Limites : Le modèle AR ne peut pas capter directement les composantes saisonnières ou les tendances à long terme, ce qui peut limiter son efficacité pour des séries temporelles complexes comme celles des taux d'affaires nouvelles de crédits immobiliers. C'est pourquoi il est souvent combiné avec d'autres modèles (comme dans ARMA ou ARIMA) pour une modélisation plus complète.

- **MA (Moyenne Mobile)**

Le modèle de Moyenne Mobile (MA) est une approche qui modélise la valeur actuelle d'une série temporelle comme une combinaison linéaire des erreurs et/ou chocs aléatoires passés. L'ordre du modèle MA(q), noté q, indique combien de termes d'erreur passés sont utilisés dans le modèle.

Formulation mathématique : Un modèle MA(q) s'écrit sous la forme :

$$Y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + ... + \theta_q\varepsilon_{t-q}$$

Où :

- Y_t est la valeur de la série au temps t ;
- μ est la moyenne du processus ;
- $\varepsilon_t, \varepsilon_{t-1}, ..., \varepsilon_{t-q}$ sont les termes d'erreur ;
- $\theta_1, \theta_2, ..., \theta_q$ sont les paramètres du modèle.

Caractéristiques principales :

- **Mémoire limitée :** Contrairement au modèle AR, le modèle MA a une "mémoire limitée"[1], déterminée par l'ordre q ;
- **Stationnarité :** Les processus MA sont toujours stationnaires, ce qui est un avantage par rapport aux modèles AR ;
- **Fonction d'auto-corrélation :** L'ACF d'un processus MA est nulle après le lag q, ce qui aide à identifier l'ordre du modèle.

[1]La "mémoire limitée" du modèle MA fait référence à sa capacité à ne prendre en compte que les erreurs passées jusqu'à un certain point dans le temps, défini par l'ordre q du modèle. Contrairement au modèle AR, où l'influence des valeurs passées peut théoriquement s'étendre indéfiniment, le modèle MA a une influence qui se limite strictement aux q dernières périodes, ce qui signifie que les chocs ou erreurs au-delà de ce point n'ont plus d'impact direct sur la valeur actuelle de la série.

Applications : Dans le contexte des *Taux d'affaires nouvelles*, un modèle MA pourrait être utile pour capturer les effets à court terme des chocs ou des événements imprévus sur le marché du crédit. Par exemple, il pourrait modéliser l'impact persistant mais décroissant d'un changement soudain de politique monétaire sur les volumes de crédits.

Limites : Comme le modèle AR, le modèle MA seul peut être limité pour capturer toutes les dynamiques d'une série temporelle complexe. Il est souvent combiné avec d'autres modèles (comme dans ARMA ou ARIMA) pour une modélisation plus complète des séries temporelles financières.

Cette explication du modèle MA complète celle du modèle AR, fournissant ainsi une base solide pour comprendre les composantes fondamentales des modèles plus avancés comme ARMA, ARIMA, et leurs extensions saisonnières qui sont abordées par la suite.

3.2.2 Modèle ARMA

Le modèle ARMA (*Auto-regressive Moving Average*) est une combinaison des modèles AR et MA. Il permet de capturer à la fois la dépendance des valeurs actuelles sur les valeurs passées (composante AR) et l'effet des chocs aléatoires passés (composante MA).

Formulation mathématique : Un modèle ARMA(p,q) s'écrit sous la forme :

$$Y_t = \mu + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

Où :

- Y_t est la valeur de la série au temps t ;
- μ est une constante réelle ;
- $\phi_1, ..., \phi_p$ sont les paramètres de la partie AR ;
- $\theta_1, ..., \theta_q$ sont les paramètres de la partie MA ;
- ε_t est le terme d'erreur au temps t.

Caractéristiques principales :

- **Flexibilité :** ARMA combine les avantages des modèles AR et MA, offrant une plus grande flexibilité pour modéliser des séries temporelles complexes ;
- **Parcimonie :** Bien que le modèle ARMA introduise deux types de paramètres (AR et MA), il peut dans certains cas offrir une représentation plus parcimonieuse des données qu'un modèle AR ou MA pur de degré élevé. Par exemple, un processus ARMA(1,1) pourrait parfois mieux représenter une série temporelle qu'un AR(p) ou MA(q) avec p ou q élevés. Cependant, cette parcimonie dépend fortement de la structure des données et doit être évaluée au cas par cas ;
- **Stationnarité :** Le processus ARMA doit être stationnaire pour être valide.

Applications : Dans le contexte des *Taux d'affaires nouvelles*, un modèle ARMA pourrait être utile pour :

- Capturer les tendances à court terme et les dépendances temporelles (partie AR) ;
- Modéliser l'impact des chocs ou événements imprévus sur le marché du crédit (partie MA) ;

- Fournir des prévisions à court terme plus précises que les modèles AR ou MA seuls.

Limites :

- Le modèle ARMA suppose que la série est stationnaire, ce qui n'est pas toujours le cas pour les données financières ;
- Il ne peut pas directement modéliser les tendances à long terme ou les composantes saisonnières ;
- L'identification des ordres p et q appropriés peut être complexe et nécessite souvent une analyse des fonctions d'auto-corrélation (ACF) et d'auto-corrélation partielle (PACF).

Le modèle ARMA constitue une base importante pour des modèles plus avancés comme ARIMA et SARIMA, qui sont abordés dans les sections respectives *3.2.4* et *3.2.5* suivantes.

3.2.3 Principe de stationnarité, différenciation et tests

Avant d'aborder le modèle ARIMA (*Auto-regressive Integrated Moving Average*), il est essentiel de comprendre le concept de différenciation dans le cadre de la stationnarisation des données. La composante "Intégrée" (I) du modèle ARIMA fait référence à ce processus de différenciation de la série temporelle. Le cours de Q. Lajaunie [44] approfondit ces notions.

3.2.3.1 Définition de la stationnarisation

La stationnarité est une propriété fondamentale des séries chronologiques, essentielle pour de nombreuses méthodes de prévision statistique. Une série chronologique est considérée comme stationnaire lorsque ses propriétés statistiques, telles que la moyenne, la variance et l'auto-corrélation, demeurent constantes au fil du temps.

On distingue deux types de stationnarité :

1. **Stationnarité stricte (ou forte) :**
   La distribution de probabilité conjointe de n'importe quel ensemble de points de la série reste inchangée lorsqu'on la déplace dans le temps.
2. **Stationnarité faible (ou d'ordre deux) :**
   - La moyenne du processus est constante : $E(Y_t) = \mu \ \forall t$;
   - La variance du processus est constante : $Var(Y_t) = \mu \ \forall t$ et $\perp t$;
   - La covariance entre deux points de la série ne dépend que de l'écart temporel entre ces points : $Cov(Y_t, Y_{t+k}) = \gamma(k) \ \forall t$ et k.

La stationnarité faible est souvent suffisante pour la plupart des analyses pratiques. Elle permet d'estimer la moyenne du processus par la moyenne empirique et de centrer la série en soustrayant cette moyenne.

3.2.3.2 Processus TS (*Trend Stationary*) et TD (*Difference Stationary*)

Un processus de série chronologique est dit non-stationnaire lorsque ses moments statistiques, notamment la moyenne et la variance, varient dans le temps. Il existe deux types de processus non-stationnaires :

1. **Le processus TS**
   Ce processus est un type particulier de processus non-stationnaire qui peut être rendu stationnaire en retirant une tendance déterministe.
   Caractéristiques principales du processus TS :

- Forme générale : $Y_t = TD_t + u_t$
  Où TD_t est une tendance déterministe et u_t est un processus stationnaire ;
- La série Y_t n'est pas stationnaire car $E[TD_t]$ dépend du temps ;
- Y_t présente une tendance à converger vers TD_t, avec des écarts limités autour de cette tendance ;
- La stationnarité peut être obtenue en soustrayant la tendance déterministe : $u_t = Y_t - TD_t$;
- Les tendances peuvent être linéaires ou polynomiales.

L'identification et la modélisation correctes de la tendance sont impératives pour rendre la série stationnaire et permettre une analyse et des prévisions précises.

2. **Le processus DS**
   Ce processus est un type de série chronologique non-stationnaire qui ne peut pas être rendue stationnaire par simple suppression de la tendance. Ce concept, introduit par Granger et Newbold (1974), décrit un comportement à tendance stochastique.
   Caractéristiques principales du processus DS :
   - Modèle de base : La marche aléatoire, $Y_t = Y_{t-1} + u_t$, où u_t est un bruit blanc gaussien (noté $BB(0; \sigma^2)$)[1] ;
   - Effet permanent des chocs : Chaque innovation u_t a un impact permanent et cumulatif sur les valeurs futures de Y_t. Contrairement à un processus stationnaire où l'effet d'un choc s'estompe avec le temps, dans un processus DS, l'effet de chaque choc persiste indéfiniment et s'accumule avec les chocs subséquents ;
   - Variance croissante : La variance du processus augmente avec le temps, ce qui le rend non-stationnaire ;
   - Mémoire infinie : Les effets des chocs passés persistent indéfiniment ;
   - La stationnarité peut être obtenue par différenciation : $\Delta Y_t = Y_t - Y_{t-1} = u_t$;
   - Ordre d'intégration : Un processus DS qui devient stationnaire après d différenciations est dit intégré d'ordre d, noté $I(d)$.

   Le processus DS est important en économétrie car de nombreuses variables économiques présentent ce type de comportement, nécessitant des techniques spécifiques pour leur analyse et leur modélisation.

3.2.3.3 Définition et objectif de la différenciation

La différenciation est une technique employée pour transformer une série temporelle non-stationnaire en une série stationnaire. Ce processus vise à éliminer les tendances et les variations saisonnières, que ces tendances soient stochastiques (comme celles présentes dans une marche aléatoire) ou déterministes (qui suivent un modèle prédéterminé), rendant ainsi la série plus adaptée à la modélisation statistique.

On distingue deux types de différenciation :

— **Différenciation simple (d=1)** : Chaque valeur de la série est soustraite de la valeur précédente ;

— **Différenciation d'ordre supérieur (d>1)** : Le processus de différenciation est appliqué de manière itérative.

[1] Un bruit blanc gaussien $BB(0; \sigma^2)$ est une séquence de variables aléatoires indépendantes et identiquement distribuées selon une loi normale de moyenne 0 et de variance σ^2.

Mathématiquement, pour une série Y_t, la première différence s'exprime comme :

$$\Delta Y_t = Y_t - Y_{t-1}$$

Ainsi, la différenciation est une technique polyvalente pour rendre stationnaire une série temporelle, quelle que soit la nature de sa non-stationnarité. Cependant, le choix entre la différenciation et d'autres méthodes de stationnarisation devrait idéalement être guidé par les résultats de tests statistiques rigoureux. Dans cette optique, trois tests sont couramment utilisés : le *test de Dickey-Fuller (DF)* et le *test de Augmented Dickey-Fuller (ADF)*, qui sont des tests de racine unitaire, ainsi que le *test de Kwiatkowski-Phillips-Schmidt-Shin (KPSS)*, qui, contrairement aux deux premiers, est un test de stationnarité. Ces tests fournissent une base solide pour déterminer la méthode de stationnarisation appropriée qu'une simple inspection visuelle de la série.

3.2.3.4 Tests de racine unitaire

Il est souvent difficile de déterminer visuellement si une série temporelle est stationnaire en tendance (TS) ou stationnaire en différence (DS). Pour obtenir une conclusion fiable sur la nature de la stationnarité d'une série, il est nécessaire de recourir à des tests statistiques formels, notamment les tests de racine unitaire[1].

Le **test de Dickey-Fuller (DF)** est une méthode statistique utilisée pour détecter la présence d'une racine unitaire dans une série temporelle, indiquant ainsi sa non-stationnarité. Le test se base sur l'estimation d'un modèle auto-régressif AR(1) et examine si le coefficient auto-régressif est statistiquement différent de 1.

Les principales caractéristiques sont :

1. Modèle de base :

$$\begin{aligned} Y_t &= \phi Y_{t-1} + \varepsilon_t \\ Y_t - Y_{t-1} &= \phi Y_{t-1} - Y_{t-1} + \varepsilon_t \\ \Delta Y_t &= (\phi - 1) Y_{t-1} + \varepsilon_t \\ \Delta Y_t &= \pi Y_{t-1} + \varepsilon_t \end{aligned}$$

   où :

   - ϕ est le coefficient auto-régressif dont la significativité est testée ;
   - ε_t est le terme d'erreur.

2. Hypothèses :
   - $H_0 : \pi = 0$, la série a une racine unitaire (non-stationnaire) ;
   - $H_1 : \pi < 0$, la série est stationnaire.

3. Statistique de test :

$$\hat{t}_{DF} = \frac{\hat{\pi}}{SE(\hat{\pi})}$$

   où $SE(\pi)$[2] représente l'erreur standard de l'estimation du coefficient π.

[1] Une série temporelle a une racine unitaire si elle peut être représentée par un processus auto-régressif dont l'une des racines du polynôme caractéristique est égale à 1. Autrement dit, c'est un modèle où la valeur actuelle dépend des valeurs passées. Par exemple, un processus AR(1) s'écrit : $Y_t = \phi_1 Y_{t-1} + \varepsilon_t$.

[2] L'erreur standard (SE) est une mesure de la dispersion des estimations d'un paramètre par rapport à la valeur réelle. Elle est calculée comme suit : $SE(\hat{\pi}) = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$ où n est la taille de l'échantillon.

4. Interprétation :
   Si H_0 ne peut pas rejeté, c'est-à-dire si la statistique de test est supérieure à la valeur critique, la série est considérée comme intégrée d'ordre 1 (I(1)).

Le **test de Augmented Dickey-Fuller (ADF)** est une extension sophistiquée du test de Dickey-Fuller original. Il est conçu pour traiter des séries temporelles plus complexes, notamment celles qui présentent une tendance évidente ou une auto-corrélation d'ordre supérieur.

Les principales caractéristiques du test ADF sont :

1. Prise en compte des tendances :
   Si la série chronologique montre une tendance croissante ou décroissante au fil du temps, le test ADF peut inclure une tendance déterministe dans l'hypothèse alternative.
2. Auto-corrélation d'ordre supérieur :
   Le test ADF gère les processus auto-régressifs d'ordre p (AR(p)) avec $p > 1$, en incluant des termes de différence retardés dans le modèle de régression.
3. Modèle de base :
   - Pour un modèle AR(2) :

$$\begin{aligned} Y_t &= \beta_1 + \beta_2 t + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t \\ Y_t &= \beta_1 + \beta_2 t + (\phi_1 + \phi_2) Y_{t-1} - \phi_2 (Y_{t-1} - Y_{t-2}) + \varepsilon_t \\ Y_t - Y_{t-1} &= \beta_1 + \beta_2 t + (\phi_1 + \phi_2 - 1) Y_{t-1} - \phi_2 (Y_{t-1} - Y_{t-2}) + \varepsilon_t \\ \Delta Y_t &= \beta_1 + \beta_2 t + \pi Y_{t-1} + \gamma_1 \Delta Y_{t-1} + \varepsilon_t \end{aligned}$$

     où :
     - β_1 est une constante ;
     - $\beta_2 t$ représente la tendance temporelle ;
     - $\pi = \phi_1 + \phi_2 - 1$ est le coefficient auto-régressif dont la significativité est testée ;
     - $\gamma_1 \Delta Y_{t-1} = -\phi_2 \Delta Y_{t-1}$ est le terme de différence retardée ;
     - ε_t est le terme d'erreur.
   - Par conséquent pour un modèle AR(p) :

$$\Delta Y_t = \beta_1 + \beta_2 t + \pi Y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta Y_{t-i} + \varepsilon_t$$

     où :
     - $\pi = \sum_{i=1}^{p} \phi_i - 1$;
     - $\gamma_1 = -\sum_{k=1}^{i} \phi_k$
4. Hypothèses :
   - $H_0 : \pi = 0$, la série a une racine unitaire (non-stationnaire) ;
   - $H_1 : \pi < 0$, la série est stationnaire.
5. Flexibilité du modèle :
   Le test ADF peut s'adapter à différentes spécifications, avec ou sans constante, avec ou sans tendance, selon les caractéristiques observées de la série temporelle.

6. Statistique de test :
   3 statistiques de test sont faites à partir de 3 potentielles équations :

$$\Delta Y_t = \pi Y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta Y_{t-i} + \varepsilon_t \quad (1)$$

$$\Delta Y_t = \beta_1 + \pi Y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta Y_{t-i} + \varepsilon_t \quad (2)$$

$$\Delta Y_t = \beta_1 + \beta_2 t + \pi Y_{t-1} + \sum_{i=1}^{p} \gamma_i \Delta Y_{t-i} + \varepsilon_t \quad (3)$$

   Ces 3 équations portent respectivement sur :

   - Un marché aléatoire pure (1), sans constante ni tendance ;
   - Un marché aléatoire avec dérive (2), incluant un terme d'interception β_1 ;
   - Un marché aléatoire avec dérive et tendance (3), incluant une tendance β_1 et une tendance linéaire en temps $\beta_2 t$.

   Bien que la procédure soit essentiellement la même quelle que soit l'équation de test utilisée, les valeurs critiques des t-statistiques dépendent de l'inclusion ou non de l'interception ou de la tendance en temps, comme indiqué précédemment et dépendent également de la taille de l'échantillon T :

T	1%	5%	10%
marché (1)			
100	−2.60	−1.95	−1.61
250	−2.58	−1.95	−1.62
500	−2.58	−1.95	−1.62
∞	−2.58	−1.95	−1.62
marché (2)			
100	−3.51	−2.89	−2.58
250	−3.46	−2.88	−2.57
500	−3.44	−2.87	−2.57
∞	−3.43	−2.86	−2.57
marché (3)			
100	−4.04	−3.45	−3.15
250	−3.99	−3.43	−3.13
500	−3.98	−3.42	−3.13
∞	−3.96	−3.41	−3.12

TABLE 3.1 – Valeurs critiques du test ADF pour les 3 différents marchés par niveaux de significativité[1]. (*INSEE* [16])

[1] Le Niveau de significativité est le seuil de probabilité au-delà duquel il est considéré que les résultats observés sont suffisamment improbables sous l'hypothèse nulle pour justifier le rejet de cette hypothèse.

3 test statistiques F[2] supplémentaires, désignés ici (ψ_1, ψ_2, ψ_3), sont utilisées pour tester des hypothèses conjointes relatives aux valeurs respectives des coefficients et peuvent être utilisées pour déterminer si l'un des trois modèles est approprié pour le processus générant les données :

- ψ_1 teste $H_0 : \pi = \beta_1 = 0$ pour l'équation (2) pour déterminer si le processus inclut une dérive ;
- ψ_2 teste $H_0 : \pi = \beta_1 = \beta_2 = 0$ pour l'équation (3) pour déterminer si le processus inclut une dérive ainsi qu'une tendance temporelle ;
- ψ_3 teste $H_0 : \pi = \beta_2 = 0$ pour l'équation (3) pour déterminer si le processus inclut une tendance temporelle déterministe.

Ces 3 statistiques (ψ_1, ψ_2, ψ_3) sont calculées comme des tests F :

$$\psi_i = \frac{\left(RSS_{\text{restreint}} - RSS_{\text{non restreint}}\right)/r}{RSS_{\text{non restreint}}/(T-k)}$$

où :

- RSS (*Sum of Squared Residuals*) : correspond à la somme des carrés des écarts entre les valeurs observées et les valeurs prédites par un modèle :
  - $RSS_{\text{restreint}}$: La somme des carrés des résidus pour le modèle avec les hypothèses imposées (par exemple, en supposant qu'il n'y a pas de constante ou de tendance) ;
  - $RSS_{\text{non restreint}}$: La somme des carrés des résidus pour le modèle sans les restrictions, qui inclut tous les termes potentiels (constante, dérive, etc...).
- r : Le nombre de restrictions imposées dans le modèle restreint ;
- T : Le nombre total d'observations dans les données ;
- k : Le nombre de paramètres estimés dans le modèle non restreint.

Le test compare les résidus des deux modèles (restreint et non restreint) pour voir si l'ajout des termes (dérive, tendance) améliore significativement la précision du modèle.

7. Interprétation :
   Les statistiques ψ_i sont comparées aux valeurs critiques des tests de Dickey-Fuller. Si la valeur calculée est supérieure à la valeur critique, alors l'hypothèse nulle H_0 est rejetée. Cela signifie que les données sont mieux modélisées par le modèle non restreint (celui qui inclut par exemple une dérive ou une tendance).

3.2.3.5 Test de stationnarité

Contrairement aux tests de racine unitaire présentés précédemment, qui posent comme hypothèse nulle la non-stationnarité de la série, les tests de stationnarité adoptent une approche complémentaire en supposant la stationnarité comme hypothèse nulle. Cette inversion de l'hypothèse permet de renforcer les conclusions sur la nature de la série temporelle, en particulier lorsque les résultats des tests de racine unitaire sont ambigus, c'est-à-dire lorsqu'ils produisent des résultats contradictoires selon le choix du modèle, la spécification des variables, ou la présence de tendances et de saisonnalités non prises en compte, ce qui complique l'interprétation de la stationnarité de la série temporelle.

[2] Le test F est une méthode statistique utilisée pour comparer les variances de deux populations ou pour tester l'hypothèse nulle que tous les coefficients de régression sont égaux à zéro dans le contexte de la régression linéaire. Il permet de déterminer si les variances de deux échantillons sont significativement différentes ou si un modèle de régression est globalement significatif.

Le **test de Kwiatkowski-Phillips-Schmidt-Shin (KPSS)** est une méthode statistique alternative pour évaluer la stationnarité d'une série temporelle. Contrairement aux tests de Dickey-Fuller, le test KPSS considère la stationnarité comme hypothèse nulle.

Les principales caractéristiques sont :

1. Modèle de base :

$$Y_t = \mu + X_t + \nu_t$$
$$X_t = X_{t-1} + u_t, \quad u_t \sim N(0, \sigma^2)\ \forall t$$

   où :

   - μ est une constante ;
   - ν_t est une composante stationnaire ;
   - X_t est une marche aléatoire.

2. Hypothèses :
   - $H_0 : \sigma^2 = 0$, la série est stationnaire ;
   - $H_1 : \sigma^2 > 0$, la série est non-stationnaire.

3. Statistique de test :

$$KPSS = \frac{\sum_{t=1}^{T} S_t^2}{T^2 \hat{\sigma}_\infty^2}$$

   où :

   - $S_t = \sum_{i=1}^{t} \hat{\nu}_i$ est la somme partielle des résidus ;
   - $\hat{\sigma}_\infty^2$ est un estimateur de la variance à long terme[1] du processus.

   À l'instar du test ADF, les valeurs critiques des t-statistiques KPSS dépendent du type de marché aléatoire de la régression :

1%	5%	10%
marché sans dérive		
0.739	0.463	0.347
marché avec dérive		
0.216	0.146	0.119

TABLE 3.2 – Valeurs critiques du test KPSS pour un marché avec et sans dérive par niveaux de significativité. (*INSEE* [16])

4. Interprétation :
   Si la statistique KPSS est supérieure à la valeur critique du *Tableau 3.2*, H_0 est rejetée, indiquant que la série est non-stationnaire.

[1] "Long terme" implique que l'estimation de la variance prend en considération les comportements et les fluctuations des données sur une période suffisamment longue pour refléter leur dynamique réelle.

Le test KPSS offre une approche complémentaire aux tests de racine unitaire traditionnels, permettant de renforcer les conclusions sur la nature de la série temporelle, en particulier lorsque les résultats des autres tests sont ambigus.

Ces 3 tests sont essentiels pour déterminer si une série temporelle nécessite une différenciation pour devenir stationnaire. En effet, le test KPSS est plus susceptible de commettre une erreur de type II, c'est-à-dire un faux négatif, lorsqu'une racine unitaire est présente. À l'inverse, le test ADF est plus enclin à commettre une erreur de type I, ou faux positif, en indiquant la stationnarité alors qu'une racine unitaire est effectivement présente. Cette distinction est nécessaire dans le cadre de la modélisation ARIMA et de ses variantes.

3.2.3.6 Rôle de la différenciation dans le modèle ARIMA

Dans le modèle ARIMA(p,d,q), le paramètre d représente l'ordre de différenciation. Il indique le nombre de fois où la série doit être différenciée pour atteindre la stationnarité. L'interprétation de ce paramètre est la suivante :

— $d = 0$: La série est déjà stationnaire ;

— $d = 1$: La série devient stationnaire après une différenciation ;

— $d = 2$: Deux différenciations sont nécessaires pour obtenir la stationnarité.

3.2.3.7 Importance de la composante intégrée (I)

La composante intégrée joue un rôle majeur dans l'analyse des séries temporelles. Elle permet d'étendre l'application des modèles ARMA à des séries non-stationnaires, élargissant ainsi considérablement le champ d'application de ces techniques d'analyse. Cette approche offre une plus grande flexibilité dans la modélisation des phénomènes économiques et financiers, qui présentent souvent des tendances et des variations saisonnières complexes.

3.2.4 Modèle ARIMA

Le modèle ARIMA est utilisé pour modéliser des séries temporelles qui présentent des tendances ou d'autres formes de non-stationnarité. Il combine les éléments des modèles auto-régressifs (AR) et de moyenne mobile (MA) avec une étape de différenciation pour éliminer la non-stationnarité.

Formulation mathématique : Un modèle ARIMA(p,d,q) s'écrit sous la forme :

$$\nabla^d Y_t = \mu + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

Où :

— Y_t est la valeur de la série au temps t ;

— μ est une constante réelle ;

— $\phi_1, ..., \phi_p$ sont les paramètres de la partie AR ;

— $\theta_1, ..., \theta_q$ sont les paramètres de la partie MA ;

— ε_t est le terme d'erreur au temps t ;

— d est le nombre de différenciations nécessaires pour rendre la série stationnaire.

Caractéristiques principales :

- **Différenciation** : La composante intégrée (I) du modèle ARIMA permet d'appliquer des transformations différentielles pour rendre stationnaire une série qui ne l'est pas initialement ;
- **Flexibilité** : Le modèle peut capturer à la fois les dépendances temporelles et les chocs aléatoires, tout en s'adaptant aux tendances à long terme ;
- **Stationnarité** : L'application de différenciation vise à éliminer les tendances non-stationnaires, rendant ainsi le processus plus apte à être modélisé par des techniques statistiques.

Applications : Dans le contexte des *Taux d'affaires nouvelles*, un modèle ARIMA pourrait être utilisé pour :

- Modéliser et prévoir les tendances à long terme après avoir éliminé les effets saisonniers ou cycliques ;
- Ajuster les prévisions en fonction des changements structurels du marché du crédit immobilier.

Limites :

- La sélection des ordres p, d, et q peut être complexe et nécessite souvent une analyse des fonctions d'auto-corrélation (ACF) et d'auto-corrélation partielle (PACF) ;
- Le modèle suppose que les relations linéaires capturent suffisamment bien la dynamique de la série temporelle, ce qui peut ne pas être le cas dans toutes les situations.

Le modèle ARIMA constitue une base solide pour l'analyse des séries temporelles complexes et sert souvent de point de départ pour développer des modèles plus avancés comme SARIMA et SARIMAX, qui intègrent des composants saisonniers ou exogènes.

3.2.5 Modèle SARIMA (Seasonal Auto-regressive Integrated Moving Average)

Le modèle SARIMA est conçu pour traiter les séries temporelles qui présentent à la fois des *patterns* non-saisonniers et saisonniers. Il combine les éléments du modèle ARIMA avec des composantes saisonnières.

Formulation mathématique :

Un modèle SARIMA(p,d,q)(P,D,Q)[m] s'écrit sous la forme :

$$\Phi_P(B^m)\phi_p(B)(1-B)^d(1-B^m)^D Y_t = \Theta_Q(B^m)\theta_q(B)\varepsilon_t$$

Où :

- (p, d, q) sont les ordres non-saisonniers (comme dans ARIMA) ;
- (P, D, Q) sont les ordres saisonniers ;
- m est la période saisonnière (par exemple, 12 pour des données mensuelles avec saisonnalité annuelle) ;
- B est l'opérateur de retard ($B^m Y_t = Y_{t-m}$) ;
- $\Phi_P, \phi_p, \Theta_Q, \theta_q$ sont des polynômes d'ordre P, p, Q, q respectivement.

Caractéristiques principales :

- **Composante saisonnière** : Capture les *patterns* qui se répètent à intervalles réguliers (par exemple, annuellement) ;
- **Différenciation saisonnière** : Permet de traiter la non-stationnarité saisonnière ;
- **Flexibilité** : Peut modéliser des structures complexes combinant tendances, cycles et saisonnalités.

Applications : Dans le contexte des *Taux d'affaires nouvelles*, un modèle SARIMA pourrait être particulièrement utile pour :

- Capturer les variations saisonnières dans l'activité de crédit immobilier (par exemple, les pics d'activité en certaines saisons) ;
- Modéliser simultanément les tendances à long terme et les *patterns* saisonniers ;
- Fournir des prévisions qui tiennent compte à la fois des tendances générales et des fluctuations saisonnières.

Avantages :

- Capacité à traiter des séries temporelles complexes avec des composantes saisonnières et non saisonnières ;
- Amélioration potentielle de la précision des prévisions par rapport aux modèles ARIMA non-saisonniers.

Limites :

- La complexité du modèle peut rendre l'estimation des paramètres plus difficile ;
- Nécessite une identification précise des ordres saisonniers et non-saisonniers ;
- Peut être sensible aux changements dans les *patterns* saisonniers au fil du temps.

Le modèle SARIMA offre une approche puissante pour analyser et prévoir les séries temporelles présentant des composantes saisonnières, ce qui le rend particulièrement pertinent pour l'étude des taux d'affaires nouvelles de crédits immobiliers qui peuvent être influencés par des facteurs saisonniers.

3.2.6 Modèle SARIMAX (SARIMA avec variables exogènes)

Le modèle SARIMAX est une extension du modèle SARIMA qui intègre des variables exogènes (ou explicatives) dans l'équation. Cette approche permet de prendre en compte non seulement les *patterns* internes de la série temporelle, mais aussi l'influence de facteurs externes.

Formulation mathématique :

Un modèle SARIMAX(p,d,q)(P,D,Q)[m] avec k variables exogènes peut s'écrire sous la forme :

$$\Phi_P(B^m)\phi_p(B)(1-B)^d(1-B^m)^D(Y_t - \beta X_t) = \Theta_Q(B^m)\theta_q(B)\varepsilon_t$$

Où :

- Tous les éléments de SARIMA restent identiques ;
- X_t est un vecteur de k variables exogènes au temps t ;

— β est un vecteur de coefficients pour les variables exogènes.

Caractéristiques principales :

— **Intégration de variables externes** : Permet d'inclure des facteurs économiques, politiques ou autres qui peuvent influencer la série temporelle ;

— **Amélioration potentielle des prévisions** : En incorporant des informations supplémentaires, le modèle peut offrir des prévisions plus précises ;

— **Analyse des relations** : Permet d'étudier l'impact des variables exogènes sur la série temporelle d'intérêt.

Applications : Dans le contexte des crédits immobiliers pour l'analyse des *Taux d'affaires nouvelles*, le modèle SARIMAX pourrait intégrer des variables telles que :

— Les taux d'intérêt annuels des crédits nouveaux à l'habitat ;

— Les indices des prix de l'immobilier ;

— L'évolution de l'indice des prix à la consommation ;

— Les politiques gouvernementales affectant le marché immobilier.

Avantages :

— Capacité à capturer des influences externes sur le marché du crédit immobilier ;

— Potentiel d'amélioration significative des prévisions par rapport aux modèles SARIMA ;

— Possibilité d'analyser l'impact relatif de différents facteurs sur les flux de crédits.

Limites :

— Complexité du modèle, nécessitant plus de données et de puissance de calcul ;

— Risque de sur-ajustement si trop de variables exogènes sont incluses ;

— Nécessité de disposer de prévisions fiables pour les variables exogènes lors de la réalisation de prévisions futures.

Le modèle SARIMAX étend les capacités du SARIMA en intégrant des variables exogènes, offrant ainsi une approche encore plus complète pour analyser et prévoir les séries temporelles complexes. Cette caractéristique le rend particulièrement adapté à l'étude du marché de l'assurance emprunteur, permettant de prendre en compte non seulement les composantes saisonnières, mais aussi l'influence de facteurs économiques externes tels que les taux d'intérêt ou l'inflation.

3.2.7 Synthèse

L'examen de cette gamme de modèles de séries temporelles, allant des plus simples (AR et MA) aux plus complexes (SARIMA et SARIMAX), permet de mettre en lumière les différentes approches disponibles pour l'analyse des séries chronologiques.

Chaque modèle présente des caractéristiques spécifiques :

1. Les modèles AR et MA offrent une base solide pour comprendre les dépendances temporelles simples ;
2. Le modèle ARMA combine ces approches pour une modélisation plus flexible ;
3. L'ARIMA introduit la notion de différenciation pour traiter les séries non stationnaires ;

4. Le SARIMA ajoute la capacité de capturer les composantes saisonnières ;
5. Enfin, le SARIMAX étend ces fonctionnalités en intégrant des variables exogènes, permettant de prendre en compte l'influence de facteurs externes sur la série temporelle.

Cette progression dans la complexité des modèles reflète la nécessité de s'adapter à des séries temporelles de plus en plus complexes, présentant des caractéristiques variées telles que des tendances, des saisonnalités, et des influences externes.

La compréhension de ces différents modèles est primordiale pour aborder l'analyse du taux d'affaires nouvelles en assurance emprunteur et des lois de rachat. Ces séries temporelles, comme observées dans les analyses précédentes, présentent des caractéristiques complexes incluant des tendances à long terme, des composantes saisonnières, et sont influencées par divers facteurs économiques externes.

La prochaine étape consiste à développer une méthodologie rigoureuse pour sélectionner et mettre en œuvre le modèle le plus approprié pour chacune de ces séries. Cette méthodologie comprend la préparation des données, l'identification des ordres appropriés pour les différentes composantes du modèle, la sélection des variables exogènes pertinentes, ainsi que les techniques de diagnostic et de validation du modèle.

L'objectif est de développer une approche structurée qui permet de choisir et d'ajuster le modèle le plus adapté, en tenant compte des spécificités de chaque série temporelle et du contexte économique dans lequel elles s'inscrivent. Cette démarche vise à établir un cadre robuste pour l'analyse et la prévision du taux d'affaires nouvelles et des lois de rachat, tout en prenant en considération les complexités du marché du crédit immobilier français.

3.3 Processus de modélisation

Après avoir présenté les différents modèles de séries temporelles, il est essentiel d'établir une méthodologie rigoureuse pour leur application pratique. Cette section détaille les étapes du processus de modélisation, depuis la préparation initiale des données jusqu'à l'évaluation finale du modèle.

La méthodologie proposée couvre un large éventail d'aspects, incluant les tests de stationnarité, l'identification des paramètres du modèle, l'intégration de variables exogènes, et les techniques de validation. Cette approche structurée vise à assurer une modélisation robuste et fiable, adaptée aux spécificités des séries temporelles étudiées dans le contexte de l'assurance emprunteur.

Chaque étape de cette méthodologie est expliquée en détail, fournissant ainsi un guide complet pour la construction et l'évaluation de modèles de séries temporelles complexes. Cette approche systématique permet d'aborder efficacement les défis spécifiques liés à la modélisation du taux d'affaires nouvelles et des lois de rachat en assurance emprunteur.

3.3.1 Préparation des données

En vue de la modélisation (S)ARIMAX, une préparation minutieuse des données s'avère obligatoire. Il est donc choisi de réaliser cette étape préliminaire l'environnement R.

Le processus de préparation des données se déroule comme suit :

1. **Séparation des variables :** Les données sont scindées en deux ensembles distincts. D'une part, la variable dépendante Y, et d'autre part, les variables explicatives X_i, englobant divers indicateurs économiques.

2. **Conversion en format série temporelle :** La variable Y est transformée en objet de série temporelle à l'aide de la fonction `ts()` de R. Cette transformation est essentielle pour capturer la structure temporelle inhérente aux données, elle permet de préserver l'ordre chronologique et la périodicité des observations, éléments déterminants pour l'analyse (S)ARIMAX. Elle facilite également l'application des fonctions d'auto-corrélation et la décomposition saisonnière, étapes essentielles dans l'identification des paramètres du modèle.

La préparation des données constitue ainsi le socle sur lequel repose la modélisation (S)ARIMAX, garantissant une base solide pour l'analyse de l'impact du contexte économique sur l'assurance emprunteur.

3.3.2 Tests de stationnarité, différenciation et transformation des données

Après la transformation de la variable Y en série temporelle, l'étape suivante est de vérifier sa stationnarité. Comme vu précédemment, pour obtenir une évaluation robuste, deux tests complémentaires sont appliqués : le *test ADF* et le *test KPSS*. Ces tests sont choisis pour leur complémentarité, car pour rappel ils présentent des hypothèses nulles inverses :

— Le test ADF pose comme hypothèse nulle la présence d'une racine unitaire ;

— Le test KPSS, à l'inverse, pose comme hypothèse nulle la stationnarité de la série.

Cette approche permet une analyse plus complète et robuste de la stationnarité de la série temporelle, en réduisant le risque d'erreur lié à l'utilisation d'un seul test.

En cas de rejet de l'hypothèse nulle du test ADF et à l'inverse de non-rejet de l'hypothèse nulle dans le test KPSS, la conclusion serait que la série temporelle est stationnaire. En d'autres termes, la variable Y ne présenterait pas de tendance stochastique ou de racine unitaire, indiquant que la série est déjà stationnaire dans sa forme actuelle. Dans ce cas, aucune différenciation ne serait nécessaire pour poursuivre l'analyse.

À l'inverse, en cas de non-rejet de l'hypothèse nulle du test ADF et de rejet de l'hypothèse nulle dans le test KPSS, la conclusion serait que la série temporelle est non-stationnaire. La variable Y présenterait alors des caractéristiques indiquant la présence d'une tendance stochastique ou d'une racine unitaire, rendant la série non-stationnaire dans sa forme actuelle. Ce cas de figure amènerait à devoir différencier la série pour la rendre stationnaire avant de poursuivre l'analyse.

Différenciation des données pour alléger le programme

En cas de non-stationnarité de la série originale, une différenciation de la variable Y est effectuée à l'aide de la fonction `diff(Y, lag = 1, differences = 1)`. Cette fonction calcule les différences entre les valeurs consécutives de la série temporelle, ce qui peut éliminer les tendances et rendre la série stationnaire. La syntaxe de la fonction est telle que `Y` est la série temporelle, `lag` spécifie l'intervalle entre les observations à soustraire (ici un intervalle mensuel), et `differences` indique le nombre de fois que la différenciation est appliquée (ici une fois). On obtient alors $\forall t \in [[2, T]]$:

$$Z_t = \Delta Y_t = Y_t - Y_{t-1}$$

Après cette transformation, de nouveaux tests *ADF* et *KPSS* sont réalisés sur la série différenciée pour vérifier si elle est devenue stationnaire.
Il est important de noter que cette procédure de différenciation et de test est répétée autant de fois que nécessaire jusqu'à ce que la série devienne stationnaire (dans la majorité des cas, la différenciation ne s'effectue pas plus de deux fois). Le nombre de différenciations requises

déterminera l'ordre d'intégration de la série temporelle originale.

Suite à la confirmation de la stationnarité de la série différenciée, il convient de poursuivre l'analyse en utilisant cette série transformée comme base pour la modélisation (S)ARIMAX. Il est important de noter qu'une fois le modèle ajusté et les prévisions effectuées, il est nécessaire de "dé-différencier" les résultats pour les interpréter dans l'échelle originale de la série temporelle.

3.3.3 Identification des ordres p, d, q (ARIMA) et P, D, Q (composante saisonnière)

Pour identifier les ordres optimaux du modèle ARIMA(p,d,q)(P,D,Q)[m], une approche systématique basée sur la recherche par grille (*grid search*) est employée. La recherche par grille est une technique d'optimisation qui consiste à évaluer un modèle pour différentes combinaisons de paramètres sur un espace prédéfini, afin de déterminer la combinaison optimale.

Dans ce contexte, la recherche par grille est appliquée aux paramètres suivants :

- p, d, q pour la composante non saisonnière du modèle ARIMA ;
- P, D, Q pour la composante saisonnière ;
- m pour la période saisonnière.

Pour chaque combinaison de ces paramètres, un modèle est ajusté aux données, et ses performances sont évaluées. Le choix final des ordres optimaux n'est pas effectué immédiatement après cette recherche par grille. Au contraire, les résultats de cette recherche servent de base pour une comparaison plus approfondie des modèles candidats.

La sélection définitive des paramètres optimaux est réalisée ultérieurement, en s'appuyant sur une analyse comparative des diverses métriques d'erreur et critères de sélection. Ces indicateurs de performance sont explicités en détail dans la section suivante, intitulée *Critères de sélection et métriques d'erreur*. Cette approche permet une évaluation plus complète et robuste des modèles, assurant ainsi la sélection du modèle le plus approprié pour les données en question.

3.3.4 Intégration des variables exogènes

Avant d'intégrer les variables exogènes dans le modèle (S)ARIMAX, une étape préliminaire de normalisation des données est effectuée à l'aide de la fonction `"scale"`. Cette normalisation est particulièrement pertinente dans le cas présent, car les variables exogènes présentent une grande diversité d'unités et d'échelles : certaines sont exprimées en pourcentage, d'autres sous forme d'indices, en mètres carrés (m^2), en quantités, ou encore en montants en milliards d'euros (*cf Annexe C.1*).

Le choix de la normalisation par standardisation, qui transforme les données afin qu'elles aient une moyenne de 0 et un écart-type de 1, est justifié par plusieurs raisons. D'une part, cette méthode permet de conserver les relations entre les variables tout en éliminant l'influence des différentes échelles. D'autre part, contrairement à la normalisation Min-Max qui contraint les valeurs dans un intervalle fixe [0,1], la standardisation est moins sensible aux valeurs aberrantes. Cela est crucial dans le contexte des modèles (S)ARIMAX où des fluctuations extrêmes peuvent affecter la performance du modèle. La formule de normalisation (ou standardisation) utilisée s'écrit ainsi :

$$X_i' = \frac{X_i - \bar{X_n}}{\widehat{\sigma}} = \frac{X_i - \frac{1}{n}\sum_{j=1}^{n} X_j}{\sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X_i})^2}}$$

où :

- X_i' est la valeur normalisée ;
- X_i est la valeur originale de la donnée ;
- n est le nombre de données de la variable ;
- $\bar{X_n}$ est la moyenne empirique de la variable ;
- $\widehat{\sigma}$ est l'écart-type empirique de la variable.

Cette normalisation vise à stabiliser les échelles des variables (*cf Annexe C.2*), faciliter l'interprétation des coefficients, améliorer la stabilité numérique du modèle, et préparer les données pour une éventuelle modélisation hybride incluant un Gradient Boosting. En ramenant toutes les variables à une échelle commune, il est possible d'éviter que certaines variables dominent le modèle uniquement en raison de leur grande amplitude numérique, assurant ainsi une comparaison plus proportionnelle de leur influence respective sur la variable dépendante.

Pour intégrer efficacement les variables exogènes tout en évitant les problèmes de redondance et de multicolinéarité, deux méthodes alternatives sont employées. Ces approches visent à garantir que la matrice des variables exogènes soit de rang plein, condition nécessaire pour son inversibilité dans le calcul du modèle (S)ARIMAX. En effet, pour qu'une matrice soit inversible, son noyau doit être réduit à 0, , ce qui signifie que le seul vecteur qui satisfait l'équation $Ax = 0$ est le vecteur nul. Cela implique que toutes ses colonnes doivent être linéairement indépendantes.

Les deux méthodes utilisées sont :

- La *Matrice de corrélation* : Cette approche permet d'identifier et d'éliminer les variables fortement corrélées, réduisant ainsi la redondance dans le jeu de données ;
- L'*Analyse en Composantes Principales (ACP)* : Cette technique de réduction de dimension transforme les variables originales en un ensemble de composantes principales (CP) non-corrélées, tout en préservant la majeure partie de l'information contenue dans les données.

Ces méthodes visent à optimiser l'utilisation des variables explicatives tout en évitant les problèmes de multicolinéarité et en réduisant la dimension des données. L'objectif est d'améliorer la précision et la robustesse du modèle (S)ARIMAX, tout en facilitant son interprétation et sa capacité prédictive.

Matrice de corrélation

La matrice de corrélation est un outil puissant pour identifier et éliminer les redondances dans les données, tout en évitant les combinaisons linéaires qui pourraient compromettre l'inversion de la matrice des variables exogènes lors du calcul du modèle (S)ARIMAX.

Pour cette étude, une matrice de corrélation est calculée pour les 29 variables exogènes et la variables Y. L'objectif est de réduire le nombre de variables explicatives en ne conservant que celles présentant une corrélation inférieure à 90% (en valeur absolue) entre elles. Lorsque deux variables sont fortement corrélées ($\geq 90\%$), seule celle présentant la plus forte corrélation avec la

variable dépendante est retenue.

Cette approche permet de réduire considérablement le nombre de variables explicatives tout en préservant l'information la plus pertinente pour le modèle. Cette réduction du nombre de variables permet non seulement d'améliorer la stabilité numérique du modèle (S)ARIMAX, mais aussi de simplifier son interprétation en se concentrant sur les variables les plus pertinentes et les moins redondantes.

Comme pour la variable Y, les variables exogènes sélectionnées sont soumises à des tests de stationnarité (ADF et KPSS) et différenciées si nécessaire. La méthodologie appliquée pour évaluer et assurer la stationnarité de ces variables est identique à celle utilisée pour la variable Y, décrite *précédemment*.

Cette approche permet de traiter les variables exogènes de manière cohérente avec la variable dépendante, assurant ainsi une base solide pour la modélisation (S)ARIMAX.

Analyse en Composantes Principales (ACP)

L'Analyse en Composantes Principales (ACP), telle qu'enseignée dans le cours d'analyse de données de l'Université Paris-Dauphine [41], est présentée ici comme une méthode alternative à la matrice de corrélation pour l'intégration des variables exogènes dans le modèle (S)ARIMAX. Cette technique statistique vise à réduire la dimension des données tout en préservant au maximum la variance des données exogènes et ne dépend pas de la variable Y.

L'ACP transforme les variables originales en un ensemble de nouvelles variables non-corrélées, appelées composantes principales. Ces composantes sont ordonnées de manière à ce que les premières retiennent la plus grande part de la variance totale présente dans l'ensemble des variables d'origine.

Le processus de l'ACP comprend plusieurs étapes clés :

— La standardisation des données (déjà effectuée en amont dans cette étude) ;
— Le calcul de la matrice de covariance ;
— Le calcul des valeurs propres et des vecteurs propres ;
— La formation et sélection des composantes principales.

Dans cette étude, l'ACP est réalisée à l'aide de la fonction `prcomp` de R [49]. L'application de l'ACP à l'ensemble de données permet d'obtenir un résumé montrant la proportion de variance expliquée par chaque composante principale. Comme la *Figure 3.1* ci-dessous permet de le constater, un palier de 90% de variance expliquée est rapidement atteint. Ajouter des composantes principales supplémentaires pour améliorer la variance expliquée n'est pas nécessairement efficace, car bien que cela ajoute de l'information au modèle, le gain d'information devient marginal par rapport à la complexité du modèle.

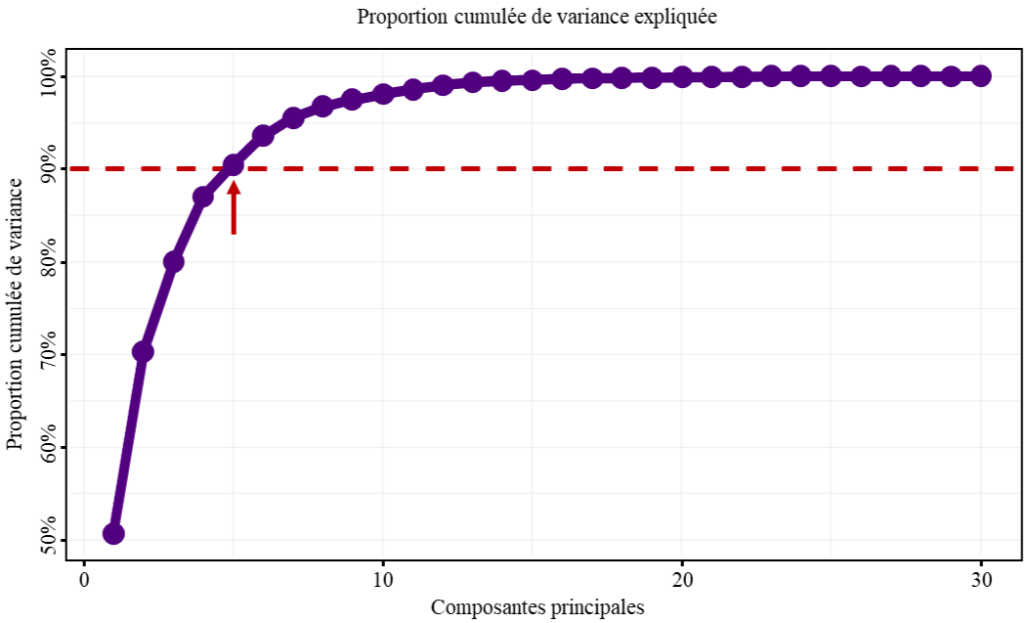


FIGURE 3.1 – Évolution de la variance cumulée en fonction des composantes principales

Suite à l'analyse du graphe de la variance cumulée, il est décidé pour le reste de l'étude de retenir les cinq premières composantes principales, qui expliquent cumulativement 90% de la variance totale des données. Ce choix permet de réduire considérablement la dimension des données tout en conservant une grande partie de l'information pertinente.

Comme pour la matrice de corrélation, il est également nécessaire de différencier les nouvelles variables exogènes, qui se présentent maintenant sous la forme de 5 composantes principales. Pour prouver leur stationnarité, des tests ADF et KPSS sont effectués à l'aide des fonctions `ur.df()` et `ur.kpss()`. Il en ressort 4 éléments d'analyse chacun :

- Le type de régression du test pour le test ADF, il est obtenu `"Test regression none"`, signifiant que la régression du test n'inclut ni constante (intercept), ni terme de tendance, soit un marché aléatoire pure ;
- Le type de test et le nombre de retards pour le test KPSS : `"Test is of type: mu with 5 lags"`, indiquant que le test est effectué avec une constante μ et 5 retards ;
- Les valeurs critiques pour les deux tests sont donc :

T	1%	5%	10%
marché (1)			
250	**-2.58**	**-1.95**	**-1.62**

10%	5%	2.5%	1%
0.347	**0.463**	**0.574**	**0.739**

TABLE 3.3 – Valeurs critiques des tests ADF (à gauche) et KPSS (à droite) sur les CP différenciées

- Les valeurs de la statistique de test et les p-valeurs pour les 5 composantes principales sont :

CP	1	2	3	4	5
Value of test-statistic	**-8.2197**	**-13.5273**	**-10.4625**	**-14.3075**	**-14.8696**
p-value	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**

TABLE 3.4 – Valeurs critiques et p-valeurs du test ADF sur les CP différenciées

CP	1	2	3	4	5
Value of test-statistic	**0.1305**	**0.1601**	**0.3292**	**0.1896**	**0.0387**
p-value	**0.1**	**0.1**	**0.1**	**0.1**	**0.1**

TABLE 3.5 – Valeurs critiques et p-valeurs du test KPSS sur les CP différenciées

L'analyse des résultats du test ADF révèle que les statistiques de test sont inférieures aux valeurs critiques à tous les seuils de significativité usuels (1%, 5% et 10%). De plus, les p-valeurs associées sont inférieures au seuil conventionnel de 0.05. Ces observations indiquent que l'hypothèse nulle de présence d'une racine unitaire est rejetée, suggérant que les CP différenciées sont stationnaires.

De même, l'analyse des résultats du test KPSS montre que les statistiques de test sont inférieures aux valeurs critiques à tous les seuils de significativité usuels. De plus, les p-valeurs associées sont supérieures au seuil conventionnel de 0.05. Ces observations indiquent que l'hypothèse nulle de stationnarité ne peut être rejetée.

En conclusion, les résultats des deux tests sur les séries différenciées convergent vers la même conclusion : **les variables exogènes sous formes de 5 CP deviennent stationnaires après différenciation**.

Cette approche présente plusieurs avantages significatifs pour la modélisation (S)ARIMAX. Bien que l'ACP et ses 5 composantes, induise une perte d'informations, elle permet une réduction substantielle de la dimension des données, offrant la possibilité de travailler avec un nombre réduit de variables tout en préservant l'essentiel de l'information contenue dans l'ensemble original. De plus, l'ACP élimine efficacement la multicolinéarité, car les composantes principales sont, par construction, non-corrélées entre elles. Cet aspect est particulièrement bénéfique pour la stabilité et l'interprétabilité du modèle (S)ARIMAX. Enfin, en se concentrant sur les composantes les plus informatives, cette méthode peut potentiellement améliorer la performance du modèle en réduisant le bruit présent dans les données, ce qui conduit à des prévisions plus précises et robustes.

L'utilisation de ces cinq composantes principales comme variables exogènes dans le modèle (S)ARIMAX offre ainsi une alternative robuste et efficace à la sélection basée sur la matrice de corrélation, tout en assurant une représentation condensée mais fidèle de l'information contenue dans l'ensemble original des variables explicatives.

3.3.5 Base *train* et *test*

La séparation du jeu de données en bases d'entraînement (*train*) et de test est une étape obligatoire dans le processus de modélisation. Cette division permet non seulement d'évaluer la performance du modèle sur des données non utilisées lors de son apprentissage, mais aussi de vérifier sa capacité de généralisation.

Dans le contexte des séries temporelles, la séparation ne se fait pas de manière aléatoire comme c'est souvent le cas pour d'autres types de données. Au contraire, elle respecte l'ordre chronologique des observations, en divisant simplement la série en "passé" (base d'entraînement) et "futur" (base de test). Il est important de noter que les valeurs "futures" utilisées pour tester le modèle proviennent de données les plus récentes déjà observées. Cela garantit que le modèle est évalué sur des données qui ont été observées et mesurées, permettant ainsi d'obtenir une mesure précise de sa performance.

Pour cette étude, le choix est fait de réserver les trois dernières années de données comme base

de test. Cette décision n'est pas arbitraire, mais répond à plusieurs considérations importantes :

- Période suffisamment longue : Une base de test de trois ans offre une fenêtre temporelle suffisamment étendue pour évaluer la robustesse du modèle face à différentes conditions de marché ;
- Diversité des scénarios : Cette durée permet au modèle d'être confronté à divers scénarios d'évolution du marché économique. Il ne se limite pas à une simple tendance à la hausse ou à la baisse, ce qui pourrait conduire à la validation d'un modèle trop spécialisé ou peu polyvalent ;
- Évaluation de la stabilité : Une période de test plus longue permet de mieux évaluer la stabilité des prévisions du modèle dans le temps, un aspect non négligeable pour les séries temporelles ;
- Prise en compte des cycles : Trois ans offrent la possibilité d'observer comment le modèle gère les éventuels cycles saisonniers ou annuels dans les données.

Cette approche de séparation des données vise donc à garantir une évaluation rigoureuse et réaliste du modèle (S)ARIMAX, en s'assurant qu'il soit capable de fournir des prévisions fiables dans diverses conditions de marché, et pas uniquement dans un contexte spécifique ou limité.

3.3.6 Mise en place du modèle SARIMAX(p,d,q)(P,D,Q)[m][1]

Après avoir préparé les données (différenciation de la variable dépendante Y et des variables exogènes X_i), l'étape suivante consiste à ajuster le modèle SARIMAX. Cette modélisation prend en compte à la fois les composantes auto-régressives, les moyennes mobiles, la saisonnalité, et l'influence des variables exogènes.

Ajustement du modèle SARIMAX

Le modèle SARIMAX (abordé *Partie 3.2.6*) est ajusté en utilisant la fonction `Arima` de R [45] avec la syntaxe suivante :

```
Arima(Z,
    order = c(p, d, q),
    seasonal = list(order = c(P, D, Q), period = m),
    xreg = Diff(X),
    method = "CSS-ML")
```

où :

- `Z` est la variable dépendante différenciée
- `(p, d, q)` sont les ordres non-saisonniers du modèle ARIMA
- `(P, D, Q)` sont les ordres saisonniers
- `m` est la période saisonnière
- `X` représente les variables exogènes différenciées

[1] Il convient de noter que la procédure décrite dans cette partie s'applique également à un modèle ARIMAX(p,d,q), à la différence près que pour ce dernier, la composante saisonnière n'est pas prise en compte. Ainsi, pour un modèle ARIMAX, les paramètres P, D, et Q sont omis, et la période saisonnière m ne sont pas considérée.

— La méthode `"CSS-ML"` combine l'estimation par somme conditionnelle des carrés et le maximum de vraisemblance

Méthode d'estimation CSS-ML

L'option `method = "CSS-ML"` détermine la méthode d'estimation utilisée pour ajuster le modèle SARIMAX. Il existe deux principales méthodes d'estimation possibles dans le cas des modèles ARIMA : CSS et ML (*RDocumentation* [46]).

La **méthode Conditional Sum of Squares (CSS)** est une approche d'estimation des paramètres d'un modèle SARIMAX qui vise à minimiser la somme des carrés des résidus. Cette technique est particulièrement utile pour sa rapidité de calcul et sa simplicité conceptuelle. L'objectif de cette méthode est de déterminer les valeurs optimales des paramètres ϕ (auto-régressifs) et θ (moyenne mobile) qui minimisent la somme des carrés des résidus conditionnels. Mathématiquement, cela se traduit par la résolution du problème d'optimisation suivant :

$$\min_{(\phi,\theta)} \sum_{t=1+p+q}^{T} (Z_t - \hat{Z}_t)^2$$

où $\hat{Z}_t$ est la valeur prédite par le modèle SARIMAX avec les paramètres actuels.

La méthode CSS procède comme suit :

1. Initialisation : Les conditions initiales (les premières observations $Z_1, Z_2, ..., Z_p$) sont considérées comme connues et fixes ;
2. Calcul itératif : Pour chaque combinaison de paramètres (ϕ, θ), les résidus sont calculés en comparant les valeurs observées aux valeurs prédites ;
3. Optimisation : La somme des carrés de ces résidus est minimisée pour trouver les valeurs optimales des paramètres.

La **méthode Maximum Likelihood (ML)** quant à elle est une approche statistique puissante pour l'estimation des paramètres d'un modèle SARIMAX. Elle repose sur le principe de maximisation de la probabilité que les données observées soient générées par le modèle estimé. Cette méthode vise à ajuster les paramètres ϕ et θ de manière à maximiser la fonction de vraisemblance. Cette fonction représente la probabilité conjointe d'observer la série temporelle étant donné un ensemble spécifique de paramètres.

Pour une série temporelle $Z_1, Z_2, ..., Z_T$, la fonction de vraisemblance est définie comme :

$$L(\phi, \theta) = P(Z_1, Z_2, \ldots, Z_T \mid \phi, \theta)$$

où :

— $L(\phi, \theta)$ est la fonction de vraisemblance ;

— $P(Z_1, Z_2, ..., Z_T \mid \phi, \theta)$ représente la probabilité conjointe des observations, conditionnellement aux paramètres ϕ et θ.

Cette formulation suppose que les résidus ε_t suivent une distribution normale $N(0, \sigma^2)$. (Supposition qui est testée dans la *sous-section suivante*)

L'objectif principal de la méthode ML est de maximiser la fonction de vraisemblance $L(\phi, \theta)$. En

pratique, pour des raisons de simplicité computationnelle[1], c'est le logarithme de la vraisemblance qui est maximisé, sans affecter la position du maximum :

$$\max_{(\phi,\theta)} L(\phi, \theta)$$

La procédure d'estimation ML suit généralement ces étapes :

1. Formulation de la fonction de vraisemblance basée sur le modèle SARIMAX spécifié ;
2. Transformation logarithmique de la fonction de vraisemblance ;
3. Utilisation d'algorithmes d'optimisation numérique pour trouver les valeurs des paramètres qui maximisent la log-vraisemblance ;
4. Calcul des erreurs standard et des intervalles de confiance pour les estimations des paramètres.

Les méthodes **CSS** et **ML** présentent chacune des avantages et des inconvénients distincts dans l'estimation des paramètres des modèles SARIMAX. La méthode CSS se distingue par sa rapidité et sa simplicité computationnelle, ne nécessitant pas d'hypothèses sur la distribution des erreurs. Cependant, elle peut manquer de précision, notamment en présence d'auto-corrélation des résidus ou lorsque les conditions initiales sont mal estimées. En revanche, la méthode ML offre généralement des estimations plus précises et est asymptotiquement efficace, mais au prix d'une complexité computationnelle et d'une sensibilité aux hypothèses sur la distribution des erreurs.

Face à ces caractéristiques complémentaires, la méthode **CSS-ML** est développée pour tirer parti des forces de chaque approche tout en atténuant leurs faiblesses respectives. Cette méthode hybride procède en deux étapes principales. Dans un premier temps, elle utilise la méthode CSS pour obtenir rapidement une estimation initiale des paramètres. Cette étape fournit un point de départ robuste pour la suite de l'estimation. Ensuite, ces estimations préliminaires sont affinées grâce à la méthode ML, permettant d'atteindre une plus grande précision dans l'estimation finale des paramètres.

L'approche **CSS-ML** s'avère particulièrement avantageuse dans le contexte de modèles complexes ou lorsque les données présentent des caractéristiques qui rendent difficile la convergence avec une seule méthode. En combinant la rapidité initiale de la CSS avec la précision finale de la ML, cette méthode offre un équilibre optimal entre efficacité computationnelle et qualité des estimations. Elle permet ainsi de surmonter les limites de chaque méthode, offrant une solution robuste et flexible pour l'estimation des paramètres dans les modèles SARIMAX.

3.3.7 Analyse et Tests statistiques des résidus

Après l'ajustement du modèle (S)ARIMAX, une analyse des résidus est menée pour évaluer la qualité et la validité du modèle. Cette analyse combine des tests statistiques rigoureux et des méthodes visuelles, offrant une évaluation complète des propriétés des résidus.

[1] Computationnel : Relatif à la complexité et à l'intensité des calculs informatiques nécessaires pour exécuter un algorithme ou une méthode statistique. Dans le contexte des modèles SARIMAX, cela fait référence à la puissance de calcul et au temps requis pour estimer les paramètres du modèle.

3.3.7.1 Test de Ljung-Box

La fonction R `checkresiduals()` [47] est utilisée comme point de départ de l'analyse. Cette fonction effectue automatiquement le test de Ljung-Box, un outil statistique qui évalue l'adéquation d'un modèle (S)ARIMAX en examinant l'auto-corrélation des résidus. Ce test permet de vérifier si le modèle a capturé efficacement toutes les dynamiques de la série temporelle.
Les caractéristiques principales sont :

1. L'Auto-corrélation L'auto-corrélation mesure la corrélation d'une série temporelle avec ses propres valeurs retardées. Pour une série de résidus, l'auto-corrélation au décalage k est donnée par :

$$r_k = \frac{\sum_{t=k+1}^{T} (e_t - \bar{e})(e_{t-k} - \bar{e})}{\sum_{t=1}^{T} (e_t - \bar{e})^2}$$

   où :

   - e_t sont les résidus à l'instant t ;
   - $\bar{e}$ est la moyenne des résidus ;
   - T est la taille de l'échantillon.

2. Hypothèses :
   - H_0 : Les résidus n'ont pas d'auto-corrélation (ils sont indépendants et suivent un bruit blanc) ;
   - H_1 : Les résidus présentent une auto-corrélation significative, indiquant que le modèle n'a pas capturé toutes les dynamiques de la série temporelle.

3. Statistique de test : La statistique de Ljung-Box est une version améliorée de la statistique de Box-Pierce[1], particulièrement adaptée aux petits échantillons :

$$Q_{LB} = T(T+2) \sum_{k=1}^{m} \frac{r_k^2}{T-k}$$

   où :

   - Q_{LB} est la statistique de Ljung-Box ;
   - T est la taille de l'échantillon ;
   - r_k est l'auto-corrélation au décalage k ;
   - m est le nombre de décalages considérés.

   Dans le contexte d'un modèle SARIMAX(p,d,q)(P,D,Q)[s] avec k variables exogènes, les degrés de liberté sont ajustés comme suit :

$$df = p + q + P + Q + k$$

   Cet ajustement permet de tenir compte de la complexité du modèle SARIMAX dans l'évaluation de l'auto-corrélation des résidus.

4. Interprétation :
   La statistique Q_{LB} suit une distribution du χ^2 sous l'hypothèse nulle H_0. Une p-valeur

[1] Statistique de Box-Pierce : $Q = T \sum_{k=1}^{m} r_k^2$

inférieure au seuil choisi (généralement 0.05) indique le rejet de l'hypothèse nulle, suggérant la présence d'auto-corrélation significative dans les résidus.

3.3.7.2 Tests de stationnarité

Comme pour les variables Y et X_i, la stationnarité des résidus est vérifiée à l'aide des deux tests complémentaires ADF et KPSS, une condition essentielle pour la validité du modèle (S)ARIMAX.

3.3.7.3 Analyses graphiques

En complément des tests statistiques, plusieurs méthodes graphiques sont employées pour évaluer visuellement les propriétés des résidus :

- **Évaluation de la normalité** : Deux représentations graphiques sont utilisées pour comparer la distribution des résidus à une distribution normale théorique :
  - Le QQ-plot, réalisé avec les fonctions R `qqnorm()` et `qqline()`. Dans ce graphique, une distribution normale des résidus est indiquée par des points alignés le long de la droite de référence.
  - Un histogramme des résidus, permettant de visualiser leur distribution et d'évaluer visuellement leur conformité à une distribution normale.
- **Fonction d'auto-corrélation (ACF)** : Ce graphique visualise l'auto-corrélation des résidus à différents décalages temporels. Il offre un aperçu de la structure de dépendance temporelle résiduelle et permet d'illustrer graphiquement les résultats du test de Ljung-Box effectué précédemment.

Ces analyses graphiques complètent les tests statistiques en fournissant une interprétation intuitive et graphique des propriétés des résidus, renforçant ainsi l'évaluation globale de l'adéquation du modèle (S)ARIMAX.

L'ensemble de ces tests et visualisations fournit une évaluation complète des résidus du modèle (S)ARIMAX, vérifiant les hypothèses clés du modèle : absence d'auto-corrélation, stationnarité et normalité des résidus. Cette approche combinée de tests statistiques formels et d'analyses graphiques offre une méthode robuste pour évaluer la performance du modèle SARIMAX et identifier d'éventuelles améliorations nécessaires dans sa spécification.

Si l'un de ces tests statistiques ou visuels n'est pas concluant, la combinaison (p,d,q)(P,D,Q) en question n'est pas retenue pour la suite de la méthodologie.

3.3.8 Projection, validation croisée et analyse vis à vis de la base test

Après avoir ajusté le modèle (S)ARIMAX, il est essentiel de valider sa performance et sa capacité à généraliser. Il faut donc s'assurer que le modèle ne soit pas simplement adapté aux données d'entraînement, mais qu'il puisse également fournir des prévisions fiables sur des données inédites.

La base test joue un rôle fondamental dans ce processus. Pour cette étude, comme évoqué *précédemment*, les trois dernières années de données sont choisies comme base test. Ce choix n'est pas arbitraire. Une période de trois ans offre une fenêtre suffisamment large pour évaluer la robustesse du modèle face à différentes conditions de marché. Cette durée permet également d'observer comment le modèle réagit à divers scénarios d'évolution des taux d'affaires nouvelles.

La séparation des données en bases d'entraînement et de test permet d'évaluer la performance du modèle sur des observations qui ne sont pas utilisées lors de l'apprentissage. Cela teste la capacité

du modèle à généraliser ses prévisions à des situations nouvelles.

Une validation croisée adaptée aux séries temporelles est alors mise en place. Cette approche est appliquée sur trois périodes distinctes : $2019 - 2021$, $2020 - 2022$ et $2021 - 2023$. Cette méthode respecte l'ordre chronologique des données, ce qui est essentiel dans l'analyse des séries temporelles. La *Figure 3.2* ci-dessous illustre l'approche de validation croisée, montrant comment les données sont divisées pour chacune des périodes de test.

FIGURE 3.2 – Validation croisée de la série temporelle

Cette validation croisée offre plusieurs avantages. Elle permet d'évaluer la performance du modèle sur différentes périodes, capturant ainsi diverses conditions de marché. Elle fournit également une estimation plus robuste de la capacité de généralisation du modèle et aide à identifier si sa performance est stable dans le temps.

Pour chacune des trois périodes de test, une projection du modèle est efefctuée. Cette projection permet de comparer visuellement les prévisions aux valeurs réelles. Des métriques de performance sont également calculées sur chaque période de test et les erreurs de prévision sont analysées pour identifier d'éventuels biais ou faiblesses du modèle.

L'utilisation de trois bases test distinctes permet d'évaluer la stabilité des prévisions dans le temps. Cela est particulièrement important pour les séries temporelles, où les comportements peuvent changer en fonction des cycles économiques ou saisonniers.

Dans la section suivante, les métriques utilisées pour mesurer la performance du modèle sur ces différentes bases test sont abordées en détail. Ces indicateurs fournissent une évaluation quantitative et qualitative des prévisions effectuées par le modèle SARIMAX, permettant ainsi une analyse approfondie de son efficacité et de sa robustesse à travers le temps.

3.3.9 Critères de sélection et métriques d'erreur (AIC, BIC, MSE, MAE, RMSE, MAPE)

Pour évaluer et comparer les performances des différents modèles (S)ARIMAX, il est nécessaire d'utiliser une combinaison de critères de sélection et de métriques d'erreur. Ces outils permettent de choisir le modèle le plus approprié et d'évaluer sa précision prédictive.

Critères de sélection de modèles

- Critère d'Information d'Akaike **(AIC - Akaike Information Criterion)** :

$$AIC = 2k - 2ln(\hat{L})$$

- Critère d'Information Bayésien **(BIC - Bayesian Information Criterion)** :

$$BIC = ln(n)k - 2ln(\hat{L})$$

Dans ces formules :

- k est le nombre de paramètres du modèle ;
- $\hat{L}$ est la valeur maximale de la fonction de vraisemblance pour le modèle ;
- n est le nombre d'observations.

L'AIC est un critère qui équilibre la qualité de l'ajustement du modèle avec sa complexité, tandis que le BIC, similaire à l'AIC, pénalise plus fortement la complexité du modèle, surtout pour les grands échantillons. Pour ces deux critères, un modèle avec une valeur plus faible est considéré comme meilleur.

Métriques d'erreur

- Erreur Quadratique Moyenne **(MSE - Mean Squared Error)** :

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

- Erreur Absolue Moyenne **(MAE - Mean Absolute Error)** :

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i|$$

- Racine de l'Erreur Quadratique Moyenne **(RMSE - Root Mean Squared Error)** :

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$$

- Erreur Absolue Moyenne en Pourcentage **(MAPE - Mean Percentage Error)** :

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right|$$

Dans ces formules :

- Y_i représente la valeur observée ;
- $\hat{Y}_i$ la valeur prédite ;
- n est le nombre d'observations.

Ces métriques offrent différentes perspectives sur la performance du modèle : le MSE et le RMSE pénalisent plus fortement les grandes erreurs mais sont sensibles aux valeurs aberrantes, le MAE est moins sensible aux valeurs aberrantes et donne une idée de l'erreur moyenne en unités originales, et enfin le MAPE fournit une mesure relative de l'erreur, utile pour comparer des séries à différentes échelles.

L'utilisation combinée de ces critères et métriques permet une évaluation complète et nuancée des modèles (S)ARIMAX, facilitant la sélection du modèle le plus performant et le plus approprié pour l'analyse des lois abordées.

3.3.10 Synthèse

En conclusion, cette section présente une méthodologie complète et structurée pour la modélisation des séries temporelles, couvrant toutes les étapes essentielles du processus. De la préparation initiale des données à l'évaluation finale du modèle, chaque étape est soigneusement détaillée pour assurer une approche rigoureuse et scientifique.

Cette méthodologie fournit un cadre solide pour aborder la modélisation de séries temporelles complexes, en tenant compte des spécificités telles que la non-stationnarité, la saisonnalité et l'influence de variables exogènes. L'accent mis sur les tests statistiques, la validation croisée et l'utilisation de diverses métriques d'erreur garantit une évaluation approfondie des modèles développés.

Dans les sections suivantes, cette méthodologie est appliquée concrètement à la modélisation du taux d'affaires nouvelles et des lois de rachat en assurance emprunteur. Ces applications pratiques permettent de démontrer l'efficacité de l'approche proposée et d'obtenir des informations précieuses sur les dynamiques de ces variables clés du secteur de l'assurance emprunteur.

La modélisation du taux d'affaires nouvelles est d'abord abordée dans le *Chapitre 4*, suivie de celle des lois de rachat dans le *Chapitre 5*. Chaque application met en lumière les particularités propres à ces séries temporelles et illustre comment la méthodologie peut être adaptée pour répondre aux défis spécifiques de chaque cas.

Chapitre 4

APPLICATION : Construction des taux d'Affaires Nouvelles

La modélisation du taux d'affaires nouvelles représente l'étape centrale dans l'analyse des *Flux mensuels de crédits nouveaux à l'habitat* en France. En s'appuyant sur les données historiques couvrant près de 21 ans, ce chapitre se concentre sur l'application des méthodes de séries temporelles pour capturer les dynamiques sous-jacentes du marché. L'objectif est d'identifier le meilleur modèle prédictif à l'aide de techniques telles que SARIMA et SARIMAX, tout en évaluant les performances à travers des métriques appropriées et une validation croisée. Cette approche permet non seulement de prévoir les taux d'affaires nouvelles, mais également d'analyser les sensibilités face aux variables exogènes clés, telles que les taux d'intérêt des crédits nouveaux à l'habitat et l'évolution de l'IPC. En intégrant ces éléments, cette modélisation vise à fournir des outils pertinents pour la gestion des risques et la prise de décision stratégique dans le secteur de l'assurance emprunteur.

4.1 Remise en contexte

Dans le cadre de cette étude, un jeu de données économiques est utilisé, couvrant la période de mars 2003 à décembre 2023. Il comprend une variable dépendante Y, représentant le rapport entre le *Flux mensuels des crédits nouveaux à l'habitat* et l'*Encours des crédits à l'habitat accordés aux particuliers résidents*, ainsi que plusieurs variables économiques explicatives X_i, comme indiqué dans le *Tableau 2.5*. Ces variables englobent divers indicateurs économiques et financiers susceptibles d'influencer le marché du crédit immobilier.

Pour analyser ces données et modéliser l'impact du contexte économique sur l'assurance emprunteur, un modèle SARIMAX est choisi. Ce modèle est particulièrement adapté pour capturer efficacement les caractéristiques complexes des séries temporelles disponibles, notamment :

- Les tendances à long terme observées dans le marché du crédit immobilier ;
- Les variations saisonnières inhérentes à ce secteur ;
- Les cycles économiques qui influencent l'activité de crédit ;
- L'impact des variables exogènes X_i sur la variable dépendante Y.

L'utilisation du modèle SARIMAX permet non seulement d'identifier les *patterns* intrinsèques de la série temporelle des flux de crédits, mais aussi d'intégrer l'influence des facteurs économiques externes. Cette approche vise à développer une compréhension approfondie des dynamiques du marché et à fournir des prévisions robustes pour le taux d'affaires nouvelles en assurance emprunteur.

4.2 Application de la méthodologie au taux d'affaires nouvelles

Cette section présente l'application détaillée de la méthodologie de modélisation des séries temporelles au taux d'affaires nouvelles en assurance emprunteur. La démarche suit rigoureusement les étapes établies dans la *Partie 3.3*, offrant ainsi une illustration concrète de l'approche théorique.

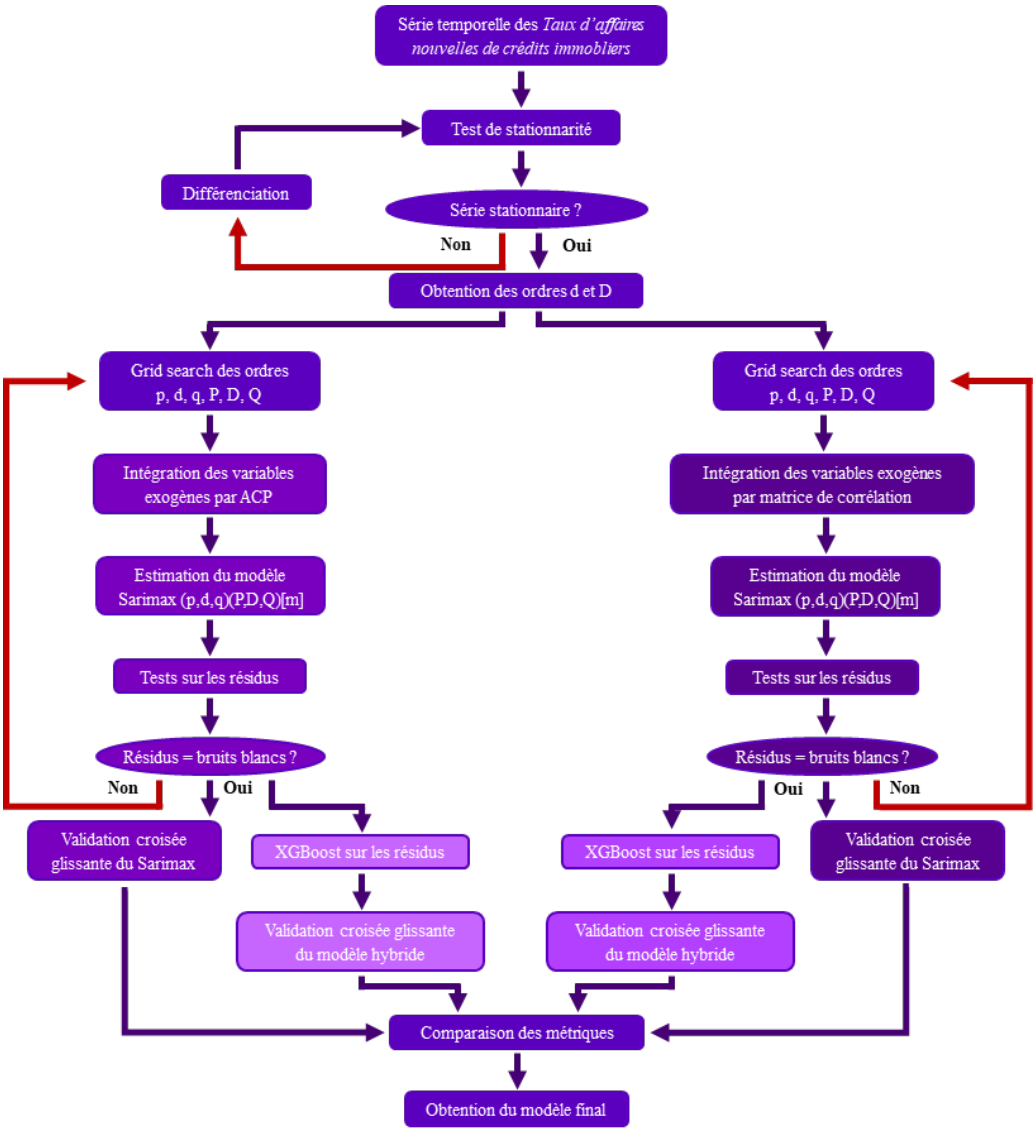


FIGURE 4.1 – Schéma du processus de modélisation du taux d'affaires nouvelles

Ce *Schéma 4.1* récapitulatif ci-dessus accompagne cette section, fournissant une représentation visuelle claire de chaque étape du processus de modélisation. Cette illustration permet de mieux appréhender la progression logique de l'analyse et les interconnexions entre les différentes phases de la modélisation.

Outre l'application de la méthodologie générale, cette section introduit également une approche de modélisation hybride. Cette technique consiste à modéliser les résidus du modèle SARIMAX initial, offrant ainsi une perspective complémentaire et potentiellement plus précise de la dynamique du taux d'affaires nouvelles.

Chaque sous-section détaille une étape spécifique du processus, de la préparation initiale des données à l'interprétation finale des résultats, en passant par les tests de stationnarité, l'identification des paramètres du modèle, et l'évaluation des performances. Cette approche structurée permet une compréhension approfondie de l'application pratique des concepts théoriques présentés précédemment.

4.2.1 Préparation des données

Pour l'application de la méthodologie au taux d'affaires nouvelles, les données sont préparées conformément à la procédure décrite dans la *Partie 3.3.1.* La variable dépendante Y, représentant les taux d'affaires nouvelles de crédits immobiliers, ainsi que les variables explicatives X_i, sont transformées en objets de série temporelle à l'aide de la fonction `ts()` de R. Cette transformation est essentielle pour capturer la structure temporelle inhérente aux données et permettre l'application des techniques de modélisation SARIMAX.

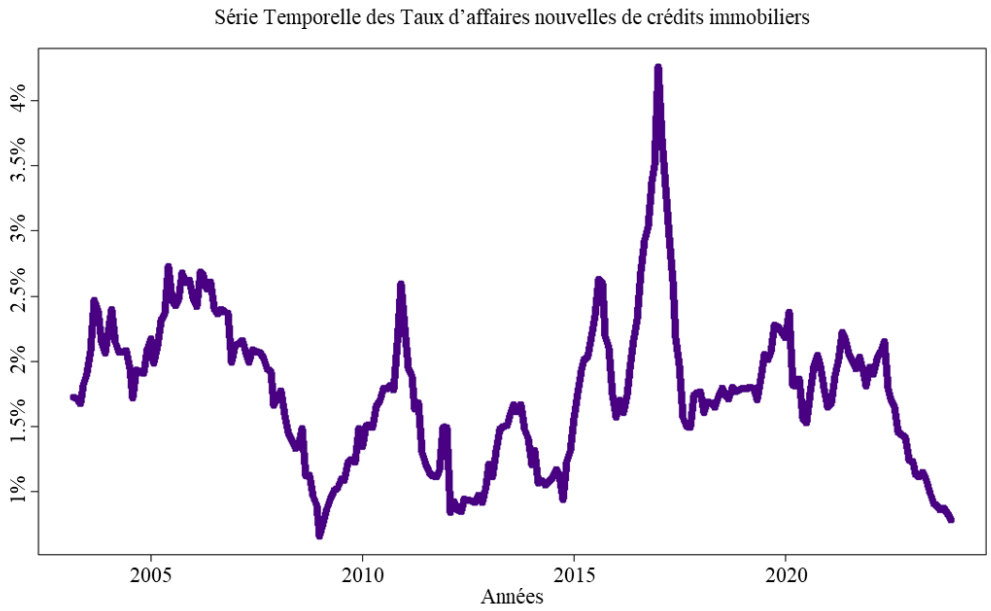


FIGURE 4.2 – Taux d'affaires nouvelles de crédits immobiliers (*Banque de France* [3])

Cette structuration des données en format de série temporelle, illustrée dans la *Figure 4.2*, permet de préserver l'ordre chronologique et la périodicité des observations, éléments obligatoires pour l'analyse subséquente.

4.2.2 Tests de stationnarité, Différenciation et transformation des données

Après la transformation de la variable Y en série temporelle, l'étape suivante est de vérifier sa stationnarité. Comme vu théoriquement en *Partie 3.3.2*, pour obtenir une évaluation robuste, deux tests complémentaires sont appliqués : le *test ADF* et le *test KPSS*. Ces tests sont choisis pour leur complémentarité, car pour rappel ils présentent des hypothèses nulles inverses :

— Le test ADF pose comme hypothèse nulle la non-stationnarité de la série ;

— Le test KPSS, à l'inverse, pose comme hypothèse nulle la stationnarité de la série.

Cette approche permet une analyse plus complète et robuste de la stationnarité de la série temporelle, en réduisant le risque d'erreur lié à l'utilisation d'un seul test.

Test ADF

Le test ADF est réalisé à l'aide de la fonction `ur.df()` sur R [50], il en ressort 4 éléments d'analyse :

— Le type de régression du test, ainsi il est obtenu "`Test regression none`", signifiant que la régression du test n'inclut ni constante (intercept), ni terme de tendance, soit un marché aléatoire pure ;

— Les valeurs critiques pour la taille de l'échantillon (250 mois) et ce type de régression (sans dérive ni tendance) sont donc :

T	1%	5%	10%
	marché (1)		
250	**-2.58**	**-1.95**	**-1.62**

TABLE 4.1 – Valeurs critiques du test ADF sur la variable Y

— La valeur de la statistique de test : "`Value of test-statistic is : -0.8218`" ;

— La p-valeur du test "`p-value = 0.08513`".

L'analyse des résultats du test ADF révèle que la statistique de test `-0.8218` est supérieure aux valeurs critiques à tous les seuils de significativité usuels (1%, 5% et 10%). De plus, la p-valeur associée `0.08513` excède le seuil conventionnel de 0.05. Ces deux observations convergent vers la même conclusion : l'hypothèse nulle de présence d'une racine unitaire ne peut être rejetée.

Test KPSS

Le test KPSS est réalisé à l'aide de la fonction `ur.kpss()` sur R [48], il en ressort 4 éléments d'analyse :

— Le type de test et le nombre de retards : "`Test is of type: mu with 5 lags`", indiquant que le test est effectué avec une constante μ et 5 retards ;

— Les valeurs critiques de ce test :

10%	5%	2.5%	1%
0.347	**0.463**	**0.574**	**0.739**

TABLE 4.2 – Valeurs critiques du test KPSS sur la variable Y

— La valeur de la statistique de test : "`Value of test-statistic is : 2.4722`" ;

— La p-valeur du test est "`p-value = 0.01`".

L'analyse des résultats du test KPSS révèle que la statistique de test `2.4722` est supérieure aux valeurs critiques à tous les seuils de significativité usuels (10%, 5%, 2.5% et 1%). De plus, la p-valeur associée `0.01` est inférieure au seuil conventionnel de 0.05. Ces deux observations convergent vers la même conclusion : l'hypothèse nulle de stationnarité est rejetée.

Ce non-rejet de l'hypothèse alternative dans le test KPSS, combinée aux résultats du test ADF, renforce la conclusion que la série temporelle est non-stationnaire. En d'autres termes, les *Taux d'affaires nouvelles* présente des caractéristiques qui indiquent la présence d'une tendance stochastique ou d'une racine unitaire, rendant la série non-stationnaire dans sa forme actuelle.

Différenciation des données

Étant donné la non-stationnarité de la série originale, une différenciation de la variable Y (observable sur la *Figure 4.3*) est effectuée à l'aide de la fonction `diff(Y, lag = 1, differences = 1)`. Cette fonction calcule les différences entre les valeurs consécutives de la série temporelle, ce qui peut éliminer les tendances et rendre la série stationnaire. La syntaxe de la fonction est telle que `Y` est la série temporelle, `lag` spécifie l'intervalle entre les observations à soustraire (ici un intervalle mensuel), et `differences` indique le nombre de fois que la différenciation est appliquée (ici une fois). On obtient alors $\forall t \in [[2, T]]$:

$$Z_t = \Delta Y_t = Y_t - Y_{t-1}$$

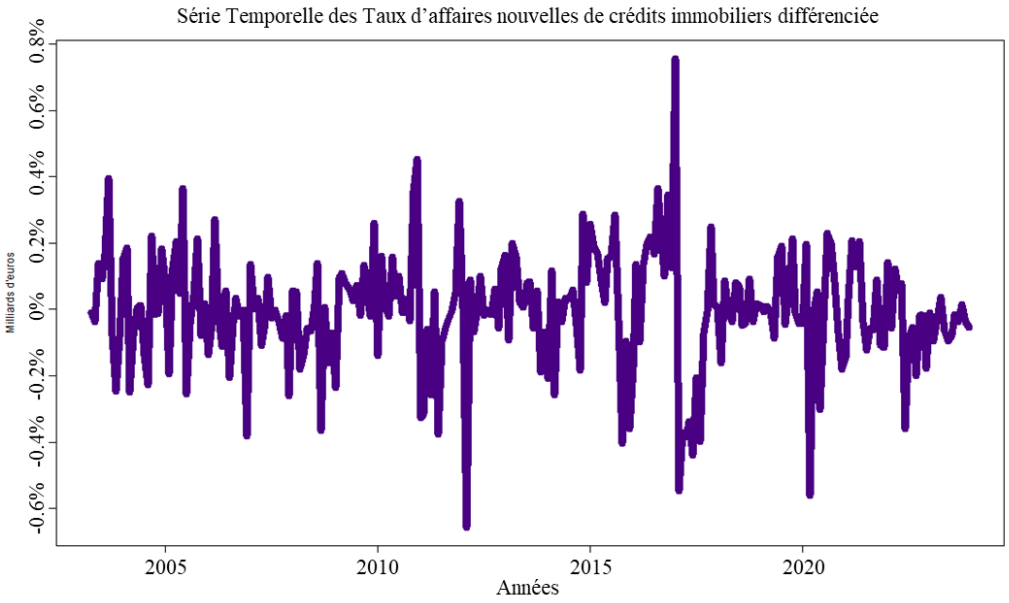


FIGURE 4.3 – Différenciation des Taux d'affaires nouvelles de crédits immobiliers (*Banque de France* [3])

Après cette transformation, de nouveaux tests *ADF* et *KPSS* sont réalisés sur la série différenciée pour vérifier si elle est devenue stationnaire. Les résultats sont les suivants :

Test ADF sur la série différenciée :

Le test ADF est réalisé à l'aide de la fonction `ur.df()`, il en ressort 4 éléments d'analyse :

— Le type de régression du test est `"Test regression none"`, indiquant une régression sans constante ni tendance ;

— Les valeurs critiques sont :

T	1%	5%	10%
	marché (1)		
250	**-2.58**	**-1.95**	**-1.62**

TABLE 4.3 – Valeurs critiques du test ADF sur la série différenciée Z

— La statistique de test obtenue est `"-8.8896"` ;

— La p-valeur du test est `"0.01"`.

L'analyse des résultats du test ADF révèle que la statistique de test `-8.8896` est inférieure aux valeurs critiques à tous les seuils de significativité usuels. De plus, la p-valeur associée `0.01` est inférieure au seuil conventionnel de 0.05. Ces observations indiquent que l'hypothèse nulle de présence d'une racine unitaire est rejetée, suggérant que la série différenciée est stationnaire.

Test KPSS sur la série différenciée :

Le test KPSS est réalisé à l'aide de la fonction `ur.kpss()`, il en ressort 4 éléments d'analyse :

— Le type de test et le nombre de retards : `"Test is of type: mu with 5 lags."`, indiquant que le test est effectué avec une constante μ et 5 retards ;

— Les valeurs critiques pour différents niveaux de significativité :

10%	5%	2.5%	1%
0.347	**0.463**	**0.574**	**0.739**

TABLE 4.4 – Valeurs critiques du test KPSS sur la série différenciée Z

— La valeur de la statistique de test : `"Value of test-statistic is: 0.08"` ;

— La p-valeur du test : `"p-value = 0.1"`.

L'analyse des résultats du test KPSS montre que la statistique de test `0.08` est inférieure aux valeurs critiques à tous les seuils de significativité usuels. De plus, la p-valeur associée `0.1` est supérieure au seuil conventionnel de 0.05. Ces observations indiquent que l'hypothèse nulle de stationnarité ne peut être rejetée.

En conclusion, les résultats des deux tests sur la série différenciée convergent vers la même conclusion : **la série devient stationnaire après différenciation**.

Suite à la confirmation de la stationnarité de la série différenciée, il convient de poursuivre l'analyse en utilisant cette série transformée comme base pour la modélisation SARIMAX. Pour rappel, dans ce cas, il est important de noter qu'une fois le modèle ajusté et les prévisions effectuées, il sera nécessaire de "dé-différencier" les résultats pour les interpréter dans l'échelle originale de la série temporelle.

4.2.3 Identification des ordres p, d, q (ARIMA) et P, D, Q (composante saisonnière)

Pour identifier les ordres optimaux du modèle SARIMAX(p,d,q)(P,D,Q)[m], la méthode de recherche par grille (*grid search*) est appliquée, conformément à la méthodologie décrite *Partie 3.3.3*. Cette approche systématique permet d'évaluer différentes combinaisons de paramètres pour les composantes non saisonnières (p, d, q) et saisonnières (P, D, Q), ainsi que pour la période saisonnière (m).

Pour cette recherche par grille, les plages de valeurs suivantes sont explorées, basées sur des considérations théoriques et l'expertise du domaine :

- p : 1 à 4
- d : 0 à 2
- q : 0 à 4

- P : 0 à 2
- D : 0 à 2
- Q : 0 à 2
- m : fixée à 12, correspondant à des données mensuelles

Ces plages sont choisies pour couvrir un large éventail de spécifications de modèles (soit 1620 combinaisons) tout en restant dans des limites raisonnables pour l'analyse de séries temporelles mensuelles.

Les résultats de cette recherche par grille servent de base pour une comparaison ultérieure des modèles candidats. La sélection finale des paramètres optimaux est effectuée en utilisant, l'analyse des résidus, les critères de sélection et les métriques d'erreur détaillés dans la section *Analyse comparative des modèles et résultats du modèle optimal.*

4.2.4 Intégration des variables exogènes

Les 29 variables exogènes sont normalisées à l'aide de la fonction `"scale"` de R pour assurer une échelle commune et faciliter leur intégration dans le modèle SARIMAX (*cf Annexe C.2*).

Pour l'intégration efficace des variables exogènes, deux approches sont appliquées :

— La *matrice de corrélation* ;

— L'*Analyse en Composantes Principales.*

Matrice de corrélation

La matrice de corrélation est utilisée pour identifier et éliminer les redondances dans les données, évitant ainsi les problèmes potentiels lors de l'inversion de la matrice des variables exogènes dans le modèle SARIMAX.

Pour cette étude, la matrice de corrélation est calculée pour les 29 variables exogènes et la variables Y. Les variables présentant une corrélation inférieure à 90% (en valeur absolue) entre elles sont conservées, en privilégiant celles ayant la plus forte corrélation avec la variable dépendante (taux d'affaires nouvelles) lorsque deux variables sont fortement corrélées.

Cette approche permet de réduire le nombre de variables explicatives de 29 à 17, préservant l'information la plus pertinente pour le modèle.

	y	x1	x4	x8	x9	x10	x11	x12	x13	x14	x16	x21	x22	x23	x24	x26	x27	x28
y	100%	-86%	45%	-8%	67%	-26%	63%	19%	27%	36%	51%	-44%	33%	-43%	59%	52%	-35%	5%
x1	-86%	100%	-45%	-1%	-83%	24%	-74%	1%	-31%	-27%	-47%	64%	-24%	61%	-55%	-61%	50%	-13%
x4	45%	-45%	100%	-47%	71%	39%	83%	-8%	56%	27%	28%	-47%	1%	-55%	-3%	70%	6%	60%
x8	-8%	-1%	-47%	100%	-3%	-63%	-24%	-40%	-48%	-1%	1%	-16%	14%	7%	5%	-9%	6%	-10%
x9	67%	-83%	71%	-3%	100%	0%	86%	-33%	53%	10%	25%	-85%	-6%	-85%	12%	74%	-34%	51%
x10	-26%	24%	39%	-63%	0%	100%	21%	-14%	19%	-5%	-27%	5%	-24%	-8%	-44%	2%	-4%	31%
x11	63%	-74%	83%	-24%	86%	21%	100%	-17%	29%	22%	29%	-77%	2%	-76%	14%	90%	-7%	64%
x12	19%	1%	-8%	-40%	-33%	-14%	-17%	100%	-1%	29%	28%	50%	40%	43%	51%	-26%	-7%	-52%
x13	27%	-31%	56%	-48%	53%	19%	29%	-1%	100%	-8%	4%	-21%	-25%	-43%	-8%	13%	-38%	14%
x14	36%	-27%	27%	-1%	10%	-5%	22%	29%	-8%	100%	58%	17%	69%	23%	50%	14%	-8%	-18%
x16	51%	-47%	28%	1%	25%	-27%	29%	28%	4%	58%	100%	6%	60%		66%	24%	-13%	-20%
x21	-44%	64%	-47%	-16%	-85%	5%	-77%	50%	-21%	17%	6%	100%	30%	89%	9%	-77%	10%	-68%
x22	33%	-24%	1%	14%	-6%	-24%	2%	40%	-25%	69%	60%	30%	100%	46%	71%	0%	-11%	-43%
x23	-43%	61%	-55%	7%	-85%	-8%	-76%	43%	-43%	23%	5%	89%	46%	100%	16%	-69%	19%	-64%
x24	59%	-55%	-3%	5%	12%	-44%	14%	51%	-8%	50%	66%	9%	71%	16%	100%	6%	-32%	-52%
x26	52%	-61%	70%	-9%	74%	2%	90%	-26%	13%	14%	24%	-77%	0%	-69%	6%	100%	26%	78%
x27	-35%	50%	6%	6%	-34%	-4%	-7%	-7%	-38%	-8%	-13%	10%	-11%	19%	-32%	26%	100%	43%
x28	5%	-13%	60%	-10%	51%	31%	64%	-52%	14%	-18%	-20%	-68%	-43%	-64%	-52%	78%	43%	100%

FIGURE 4.4 – Matrice de corrélation des variables exogènes retenues

La *Figure 4.4* ci-dessus illustre la matrice de corrélation des variables retenues après l'application de cette méthode de sélection.

Les variables conservées après cette analyse sont présentées dans le *Tableau 4.5* suivant :

	Données retenues par la matrice de corrélation					
Y	**Taux d'affaires nouvelles de crédits immobiliers**					
X_1	**Taux d'intérêt annuel des crédits nouveaux à l'habitat des particuliers**					
X_4	**Indice des prix des logements**		Anciens		Maisons	
X_8	**Taux de chômage au sens du BIT**					
X_9	**Dette publique en % de PIB**					
X_{10}	**Indice des prix à la consommation**					
X_{11}	**Produit Intérieur Brut en Mds€ 2020**					
X_{12}	**Évolution annuelle relative du Pouvoir d'achat du revenu disponible brut**					
X_{13}	**Taux d'épargne des ménages en % du revenu brut**					
X_{14}	**Volume de construction**	Nombre de logements	Autorisés		Collectif et Résidence	
X_{15}			Commencés		Collectif et Résidence	
X_{21}		Surfaces de logements	Commencés		Individuel	
X_{22}	**Volume de commercialisation**	Nombre de logements	Mis en vente aux particuliers	Nouvelles constructions	Collectif et Résidence	
X_{23}					Individuel	
X_{24}			Réservés à la vente par des particuliers	Nouvelles constructions	Collectif et Résidence	
X_{26}		Encours de logements disponibles à la vente aux particuliers		Nouvelles constructions	Collectif et Résidence	
X_{27}					Individuel	
X_{28}		Délai moyen d'écoulement de l'encours de logements disponibles à la vente aux particuliers		Nouvelles constructions	Collectif et Résidence	

TABLE 4.5 – Variables économiques exogènes retenues pour la modélisation

Cette réduction du nombre de variables améliore la stabilité numérique du modèle SARIMAX et simplifie son interprétation en se concentrant sur les variables les plus pertinentes et les moins redondantes.

Comme pour la variable Y, il est nécessaire de différencier les 17 variables exogènes pour que celles-ci soient stationnaires. Pour prouver leur stationnarité, des tests ADF et KPSS sont effectués à l'aide des fonctions `ur.df()` et `ur.kpss()`. Il en ressort 4 éléments d'analyse chacun :

— Le type de régression du test pour le test ADF, il est obtenu `"Test regression none"`, signifiant que la régression du test n'inclut ni constante (intercept), ni terme de tendance, soit un marché aléatoire pure ;

— Le type de test et le nombre de retards pour le test KPSS : `"Test is of type: mu with 5 lags"`, indiquant que le test est effectué avec une constante μ et 5 retards ;

— Les valeurs critiques pour les deux tests sont donc :

T	1%	5%	10%
	marché (1)		
250	**-2.58**	**-1.95**	**-1.62**

10%	5%	2.5%	1%
0.347	**0.463**	**0.574**	**0.739**

TABLE 4.6 – Valeurs critiques des tests ADF (à gauche) et KPSS (à droite) sur les variables exogènes

— Les valeurs de la statistique de test et les p-valeurs pour les 17 variables exogènes sont :

	Test ADF		Test KPSS	
	Value of test-statistic	p-value	Value of test-statistic	p-value
X_1	**-4.6077**	**0.01**	**0.24755**	**0.1**
X_4	**-17.261**	**0.01**	**0.0537**	**0.1**
X_8	**-16.9864**	**0.01**	**0.0685**	**0.1**
X_9	**-16.1048**	**0.01**	**0.0701**	**0.1**
X_{10}	**-2.6154**	**0.01**	**0.34491**	**0.1**
X_{11}	**-10.7025**	**0.01**	**0.1589**	**0.1**
X_{12}	**-9.8517**	**0.01**	**0.1116**	**0.1**
X_{13}	**-4.3497**	**0.01**	**0.1162**	**0.1**
X_{24}	**-5.0465**	**0.01**	**0.1524**	**0.1**
X_{16}	**-10.1554**	**0.01**	**0.1217**	**0.1**
X_{21}	**-4.022**	**0.01**	**0.1883**	**0.1**
X_{22}	**-9.0626**	**0.01**	**0.2691**	**0.1**
X_{23}	**-6.2331**	**0.01**	**0.1146**	**0.1**
X_{25}	**-12.3247**	**0.01**	**0.3419**	**0.1**
X_{26}	**-12.4356**	**0.01**	**0.0594**	**0.1**
X_{27}	**-20.7697**	**0.01**	**0.048**	**0.1**
X_{28}	**-10.8475**	**0.01**	**0.1094**	**0.1**

TABLE 4.7 – Valeurs critiques et p-valeurs des tests ADF (à gauche) et KPSS (à droite) sur les variables exogènes différenciées

L'analyse des résultats du test ADF révèle que les statistiques de test sont inférieures aux valeurs critiques à tous les seuils de significativité usuels (1%, 5% et 10%). De plus, les p-valeurs associées sont inférieures au seuil conventionnel de 0.05. Ces observations indiquent que l'hypothèse nulle de présence d'une racine unitaire est rejetée, suggérant que les CP différenciées sont stationnaires.

De même, l'analyse des résultats du test KPSS montre que les statistiques de test sont inférieures aux valeurs critiques à tous les seuils de significativité usuels. De plus, les p-valeurs associées sont supérieures au seuil conventionnel de 0.05. Ces observations indiquent que l'hypothèse nulle de stationnarité ne peut être rejetée.

En conclusion, les résultats des deux tests sur les séries différenciées convergent vers la même conclusion : **les 17 variables exogènes deviennent stationnaires après différenciation**.

Analyse en Composantes Principales (ACP)

Comme expliqué dans la *méthodologie générale*, l'ACP est utilisée comme méthode alternative à la matrice de corrélation pour l'intégration des variables exogènes dans le modèle SARIMAX. De plus, étant donné que l'ACP est indépendante de la variable Y, celle-ci est déjà calculée lors de l'étape de préparation des données. Pour rappel, la méthodologie détaillée de l'ACP, incluant le processus de sélection des composantes principales, les tests de stationnarité ainsi que les différenciations, est consultable dans la section précédente.

Dans le cadre de cette application spécifique au taux d'affaires nouvelles, les cinq premières composantes principales sont retenues. Elles expliquent cumulativement 90% de la variance totale des données. Ces composantes, déjà différenciées et testées pour leur stationnarité, servent de variables exogènes dans le modèle SARIMAX.

Cette approche permet de bénéficier des avantages de l'ACP, notamment la réduction de la dimensionnalité des données et l'élimination de la multicolinéarité, tout en conservant l'essentiel de l'information contenue dans l'ensemble original des variables explicatives.

4.2.5 Base train et test

Le jeu de données est divisé en bases d'entraînement (*train*) et de test. Suivant le principe établi, les trois dernières années de données sont réservées comme base de test, tandis que le reste constitue la base d'entraînement.

Cette séparation chronologique permet d'évaluer la performance et la capacité de généralisation du modèle SARIMAX sur une période suffisamment longue pour englober divers scénarios d'évolution des taux d'affaires nouvelles.

4.2.6 Mise en place du modèle SARIMAX(p,d,q)(P,D,Q)[m]

Conformément à la méthodologie détaillée précédemment, le modèle SARIMAX est ajusté en utilisant la fonction `Arima` de R. Les paramètres (p,d,q)(P,D,Q)[m] sont déterminés selon la *procédure de recherche par grille* décrite antérieurement. L'intégration des variables exogènes est réalisée en utilisant soit la méthode de la matrice de corrélation, soit l'ACP, comme expliqué dans les *Partie 4.2.4*. Cette étape d'ajustement du modèle SARIMAX constitue le cœur de l'analyse, intégrant les composantes auto-régressives, les moyennes mobiles, la saisonnalité, et l'influence des variables exogènes sélectionnées ou transformées.

4.2.7 Analyse et Tests statistiques des résidus

L'analyse des résidus est menée pour chaque combinaison de paramètres (p,d,q)(P,D,Q) explorée dans le *grid search*. Cette analyse comprend :

- Le test de Ljung-Box pour évaluer l'auto-corrélation des résidus ;
- Les tests de stationnarité ADF et KPSS ;
- Des analyses graphiques incluant le QQ-plot, l'histogramme des résidus, et la fonction d'auto-corrélation (ACF).

Ces tests et analyses visent à vérifier les hypothèses clés du modèle SARIMAX : absence d'auto-corrélation, stationnarité et normalité des résidus.

Il est important de noter que si l'un de ces tests statistiques ou visuels n'est pas concluant pour une combinaison de paramètres donnée, cette combinaison n'est pas retenue pour la suite de l'analyse. Cette approche permet de filtrer les modèles qui ne satisfont pas les conditions fondamentales de validité.

Les résultats détaillés de ces tests pour les différentes combinaisons de paramètres ne sont pas présentés dans cette section. Ils sont plutôt intégrés dans l'évaluation globale des modèles, qui est abordée dans la *Partie 4.2.10* sur la comparaison des résultats des métriques. Cette approche permet de se concentrer sur les modèles les plus performants et valides selon les critères établis.

4.2.8 Modélisation hybride

La modélisation hybride est une approche qui combine les forces de différentes techniques de modélisation pour améliorer la précision des prévisions. Dans le cas présent, la méthode utilisée

associe le modèle SARIMAX à un modèle de gradient boosting[1] (XGBoost) pour capturer les structures résiduelles non-linéaires que le SARIMAX ne pourrait pas pu saisir.

Le principe de cette approche est d'abord d'ajuster un modèle SARIMAX aux données pour capturer les composantes linéaires, saisonnières et les effets des variables exogènes comme effectué jusqu'à maintenant. Ensuite, les résidus de ce modèle sont modélisés à l'aide de l'algorithme XGBoost, qui construit un ensemble d'arbres de décision de manière séquentielle, chaque nouvel arbre étant entraîné pour corriger les erreurs des arbres précédents. Les prévisions finales sont obtenues en combinant les prédictions du modèle SARIMAX avec celles du modèle XGBoost sur les résidus.

Cette modélisation hybride mise en place utilise les résidus du modèle SARIMAX comme variable cible pour l'entraînement du modèle XGBoost, avec les variables exogènes et les dates comme variables explicatives. Une recherche par grille est effectuée pour optimiser les hyperparamètres du modèle XGBoost, notamment le taux d'apprentissage (`eta` : valeurs entre 0.01 et 0.30), la profondeur maximale des arbres (`max depth` : valeurs entre 3 et 6), et le nombre d'arbres (`nrounds` : 50, 100 ou 200).

Le code R utilisé pour cette modélisation hybride comprend la préparation des données, l'optimisation des hyperparamètres, l'entraînement du modèle XGBoost avec les meilleurs paramètres trouvés, et l'évaluation des performances en calculant le MSE et le RMSE sur l'ensemble de test. Cette approche hybride vise à tirer parti à la fois de la capacité du modèle SARIMAX à capturer les tendances linéaires et saisonnières, et de la flexibilité du modèle XGBoost pour modéliser les relations non-linéaires complexes dans les résidus. L'objectif est d'obtenir des prévisions plus précises et robustes que celles obtenues avec le modèle SARIMAX seul.

4.2.9 Projection, validation croisée et analyse vis à vis de la base test

Conformément à la *méthodologie générale*, une approche rigoureuse de projection, de validation croisée et d'analyse sur la base test est appliquée, pour évaluer la performance et la capacité de généralisation des modèles de l'étude. Cette méthodologie est systématiquement appliquée aux configurations suivantes :

- Modèle SARIMAX avec variables exogènes sélectionnées par matrice de corrélation ;
- Modèle SARIMAX avec variables exogènes issues de l'ACP ;
- Modèle hybride combinant SARIMAX (matrice de corrélation) et XGBoost ;
- Modèle hybride combinant SARIMAX (ACP) et XGBoost.

Pour chaque configuration, la validation croisée est utilisée sur trois périodes distinctes (2019-2021, 2020-2022, 2021-2023), respectant l'ordre chronologique des données. Cette approche permet d'évaluer la stabilité et la robustesse des modèles à travers différentes conditions de marché.

Les projections sont effectuées pour chaque période de test, permettant une comparaison visuelle et quantitative des prévisions avec les valeurs réelles. Cette méthodologie fournit une base solide pour l'analyse comparative des performances des différents modèles.

Les résultats détaillés de cette analyse, incluant l'évaluation des résidus, les critères de sélection et les métriques de performance, sont présentés et discutés dans la *Partie 4.2.10*. Cette approche

[1] Il est à noter que la partie théorique détaillée du gradient boosting, incluant ses fondements mathématiques et son fonctionnement algorithmique, est présentée en *Annexe D.1.*

permet de comparer efficacement les différentes configurations de modèles et d'identifier la plus performante pour ce cas d'étude.

4.2.10 Analyse comparative des modèles et résultats du modèle optimal

Cette section présente une analyse détaillée des résultats pour les quatre modèles développés (SARIMAX avec matrice de corrélation, SARIMAX avec ACP, modèle hybride avec matrice de corrélation, et modèle hybride avec ACP), suivie d'une présentation du modèle optimal choisi.

4.2.10.1 Tableau comparatif des critères de sélection et des métriques d'erreur

Suite à la recherche par grille effectuée pour chacun des quatre modèles testés, les configurations optimales suivantes sont identifiées :

— SARIMAX avec ACP : SARIMAX (4,0,1)(0,1,1)[12]

— SARIMAX avec matrice de corrélation : SARIMAX (4,0,1)(0,0,1)[12]

— Modèle hybride avec ACP : SARIMAX (4,0,4)(1,1,1)[12] + XGBoost (eta[1] = 0.08, depth[2] = 4, nrounds[3] = 100)

— Modèle hybride avec matrice de corrélation : SARIMAX (4,0,2)(1,0,1)[12] + XGBoost (eta = 0.16, depth = 5, nrounds = 100).

Le *Tableau 4.8* ci-dessous présente les montants des critères de sélection, les métriques et les p-valeurs des différents tests effectués pour l'analyse des résidus :

		Base Test	AIC	BIC	MSE	MAE	RMSE	MAPE	Ljung.Box p.value	ADF.test p.value	KPSS.test p.value
ACP	SARIMAX (4,0,1)(0,1,1)[12]	**Moyenne CV**	-1807.93	-1769.05	**0.00045%**	**0.172%**	**0.210%**	10.75%	0.16	0.01	0.10
		2019-2021	-1925.81	-1886.17	0.00065%	0.231%	0.255%	17.08%	0.19	0.01	0.10
		2020-2022	-1811.91	-1773.01	0.00040%	0.154%	0.200%	8.39%	0.11	0.01	0.10
		2021-2023	-1686.07	-1647.96	0.00031%	0.131%	0.176%	6.79%	0.19	0.01	0.10
Matrice de corrélation	SARIMAX (4,0,1)(0,0,1)[12]	**Moyenne CV**	**-1826.09**	**-1855.16**	0.00102%	0.246%	0.309%	13.83%	0.13	0.01	0.10
		2019-2021	-1943.56	-1980.51	0.00120%	0.298%	0.346%	22.66%	0.16	0.01	0.10
		2020-2022	-1826.72	-1857.68	0.00215%	0.395%	0.464%	20.18%	0.09	0.01	0.10
		2021-2023	-1707.01	-1727.68	0.00029%	0.107%	0.169%	4.58%	0.15	0.01	0.10
Hybride - ACP	SARIMAX (4,0,4)(1,1,1)[12] + XGBoost (eta = 0.08, depth = 4, nrounds = 100)	**Moyenne CV**	-1801.73	-1761.09	0.00051%	0.174%	0.224%	**10.05%**	0.19	0.01	0.10
		2019-2021	-1927.39	-1884.79	0.00118%	0.338%	0.344%	24.82%	0.17	0.01	0.10
		2020-2022	-1813.12	-1771.75	0.00052%	0.175%	0.227%	9.31%	0.10	0.01	0.10
		2021-2023	-1668.03	-1629.62	0.00019%	0.081%	0.139%	3.78%	0.96	0.01	0.10
Hybride - Matrice de corrélation	SARIMAX (4,0,2)(1,0,1)[12] + XGBoost (eta = 0.16, depth = 5, nrounds = 100)	**Moyenne CV**	-1817.57	-1842.47	0.00072%	0.275%	0.266%	13.62%	0.15	0.01	0.10
		2019-2021	-1931.21	-1970.40	0.00139%	0.371%	0.373%	30.97%	0.16	0.01	0.10
		2020-2022	-1824.62	-1845.35	0.00112%	0.326%	0.335%	14.71%	0.08	0.01	0.10
		2021-2023	-1696.13	-1712.97	0.00021%	0.166%	0.145%	4.83%	0.66	0.01	0.10

TABLE 4.8 – Résultats d'analyse des 4 modèles de l'étude

L'analyse des critères de sélection AIC et BIC, qui évaluent l'équilibre entre la qualité de l'ajustement et la complexité du modèle, montre que tous les modèles présentent des valeurs fortement négatives. Le modèle basé sur la matrice de corrélation **SARIMAX (4,0,1)(0,0,1)[12]** affiche les valeurs les plus faibles pour l'AIC (−1826.09) et le BIC (−1855.16). Cependant, étant donné que tous les modèles ont des AIC et BIC très négatifs, ces critères ne sont pas déterminants pour le choix final.

Dans ce cas, l'objectif principal est de minimiser l'erreur de prévision. À cet égard, le modèle le plus performant s'avère être : **SARIMAX (4,0,1)(0,1,1)[12]** avec ACP. Ce modèle affiche les meilleures performances avec un MSE de 0.00045%, un MAE de 0.172% et un RMSE de 0.210%.

[1] `eta` : Taux d'apprentissage

[2] `depth` : Profondeur de l'arbre

[3] `nrounds` : Nombre d'arbres

Ces métriques d'erreur, qui mesurent la précision des prévisions, sont essentielles pour cette étude.

Il est intéressant de noter que le modèle hybride utilisant l'ACP (**SARIMAX (4,0,4)(1,1,1)[12] + XGBoost**) présente le meilleur MAPE (Mean Absolute Percentage Error) de 10.05%, indiquant une bonne performance en termes d'erreur relative.

Pour la suite de cette étude sur le taux d'affaires nouvelles, le modèle **SARIMAX (4,0,1)(0,1,1)[12]** avec ACP est retenu. Ce choix est motivé par ses meilleures performances en termes de métriques d'erreur (MSE, MAE, RMSE). Bien que ce modèle n'ait pas les meilleurs AIC et BIC, la priorité donnée à la minimisation de l'erreur de prévision justifie cette sélection.

Ce modèle offre le meilleur équilibre entre précision des prévisions et complexité, ce qui est important pour cette étude où l'accent est mis sur la qualité des prédictions. La capacité du modèle à minimiser les erreurs de prévision est particulièrement importante dans le contexte de la modélisation des affaires nouvelles, où une prédiction précise peut avoir un impact significatif sur les décisions stratégiques et opérationnelles.

4.2.10.2 Analyse des résidus du modèle optimal

L'analyse des résidus du modèle SARIMAX (4,0,1)(0,1,1)[12] avec ACP est effectuée pour vérifier les hypothèses clés du modèle : absence d'auto-corrélation, stationnarité et normalité des résidus. Les résultats des tests sont les suivants :

1. **Test de Ljung-Box** : Le test de Ljung-Box est appliqué sur les trois périodes de validation croisée pour évaluer l'auto-corrélation des résidus. Les résultats sont les suivants :
   - 2019-2021 : p-valeur = 0.19
   - 2020-2022 : p-valeur = 0.11
   - 2021-2023 : p-valeur = 0.19

   Toutes les p-valeurs sont supérieures au seuil de significativité de 0.05, ce qui indique qu'il n'est pas possible rejeter l'hypothèse nulle d'absence d'auto-corrélation. Cela suggère que le modèle capture efficacement les dynamiques de la série temporelle, ne laissant pas de structure d'auto-corrélation significative dans les résidus.

2. **Tests de stationnarité** : Les tests ADF et KPSS sont effectués pour vérifier la stationnarité des résidus. Pour prouver leur stationnarité, des tests ADF et KPSS sont effectués à l'aide des fonctions `ur.df()` et `ur.kpss()`. Il en ressort 4 éléments d'analyse chacun :
   - Le type de régression du test pour le test ADF, il est obtenu `"Test regression none"`, signifiant que la régression du test n'inclut ni constante (intercept), ni terme de tendance, soit un marché aléatoire pure ;
   - Le type de test et le nombre de retards pour le test KPSS : `"Test is of type: mu with 5 lags"`, indiquant que le test est effectué avec une constante μ et 5 retards ;
   - Les valeurs critiques pour les deux tests sont donc :

T	1%	5%	10%
	marché (1)		
250	**-2.58**	**-1.95**	**-1.62**

10%	5%	2.5%	1%
0.347	**0.463**	**0.574**	**0.739**

TABLE 4.9 – Valeurs critiques des tests ADF (à gauche) et KPSS (à droite) sur les variables exogènes

— Les valeurs de la statistique de test et les p-valeurs pour les résidus du modèle sont :

	Test ADF		Test KPSS	
	Value of test-statistic	p-value	Value of test-statistic	p-value
`Base train 2003-18`	**-9.6529**	**0.01**	**0.0932**	**0.1**
`Base train 2003-19`	**-9.9544**	**0.01**	**0.0837**	**0.1**
`Base train 2003-20`	**-10.1519**	**0.01**	**0.0286**	**0.1**

TABLE 4.10 – Valeurs critiques et p-valeurs des tests ADF (à gauche) et KPSS (à droite) sur les résidus

L'analyse des résultats du test ADF révèle que les statistiques de test sont inférieures aux valeurs critiques à tous les seuils de significativité usuels (1%, 5% et 10%). De plus, les p-valeurs associées sont inférieures au seuil conventionnel de 0.05. Ces observations indiquent que l'hypothèse nulle de présence d'une racine unitaire est rejetée, suggérant que les résidus sont stationnaires.

De même, l'analyse des résultats du test KPSS montre que les statistiques de test sont inférieures aux valeurs critiques à tous les seuils de significativité usuels. De plus, les p-valeurs associées sont supérieures au seuil conventionnel de 0.05. Ces observations indiquent que l'hypothèse nulle de stationnarité ne peut être rejetée.

En conclusion, les résultats des deux tests sur les séries différenciées convergent vers la même conclusion : **les résidus sont bien stationnaires**.

3. **Analyses graphiques** :

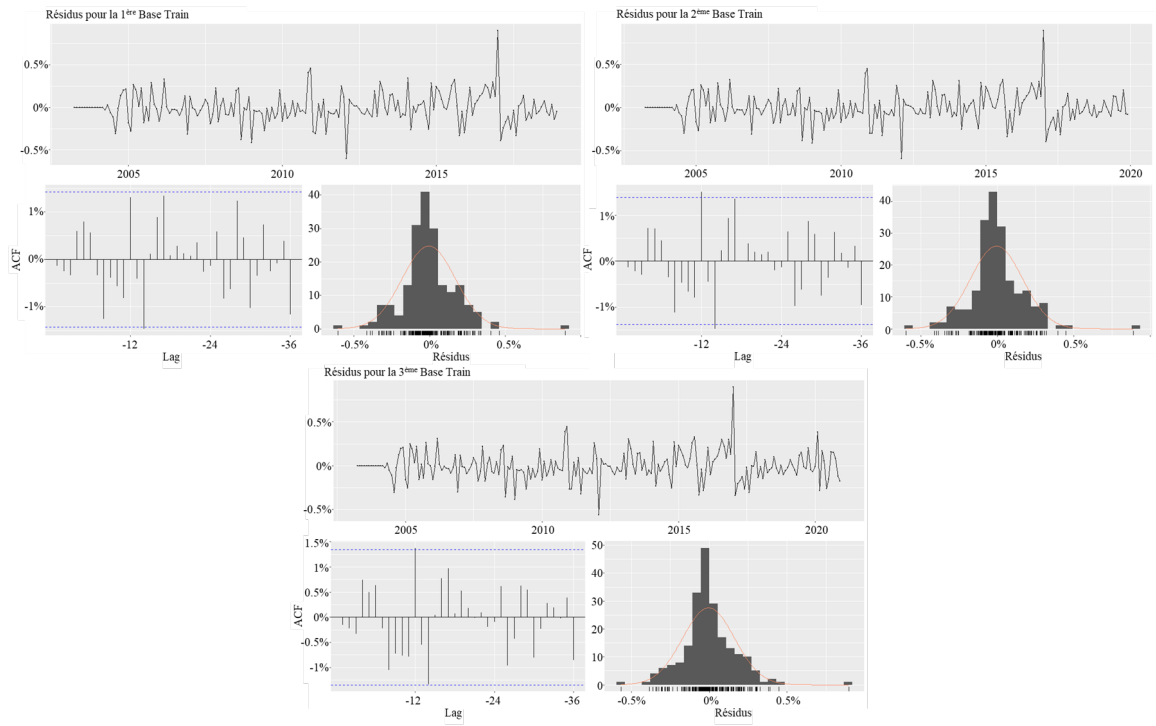


FIGURE 4.5 – Tendance, ACF et histogramme des résidus du modèle SARIMAX (4,0,1)(0,1,1)[12] sur les trois Bases Trains de la validation croisée

— **Tendance des résidus** : La tendance visuelle des résidus, comme illustrée dans la *Figure 4.5*, corrobore les résultats théoriques sur la stationnarité. Il est possible d'identifier une fluctuation des résidus autour de zéro sans tendance apparente, ce qui est caractéristique d'un processus stationnaire.

— **Fonction d'auto-corrélation (ACF)** : L'analyse du graphique ACF dans la *Figure 4.5* confirme les résultats du test de Ljung-Box, c'est-à-dire que les auto-corrélations sont majoritairement contenues dans les limites de confiance, indiquant une absence d'auto-corrélation significative dans les résidus.

— **Histogramme des résidus** : Les histogrammes présentés dans la *Figure 4.5* pour les trois bases d'entraînement montrent une distribution qui s'apparente à une distribution normale. Cette observation suggère que l'hypothèse de normalité des résidus, importante pour la validité du modèle SARIMAX, est raisonnablement satisfaite.

— **QQ-plot** :

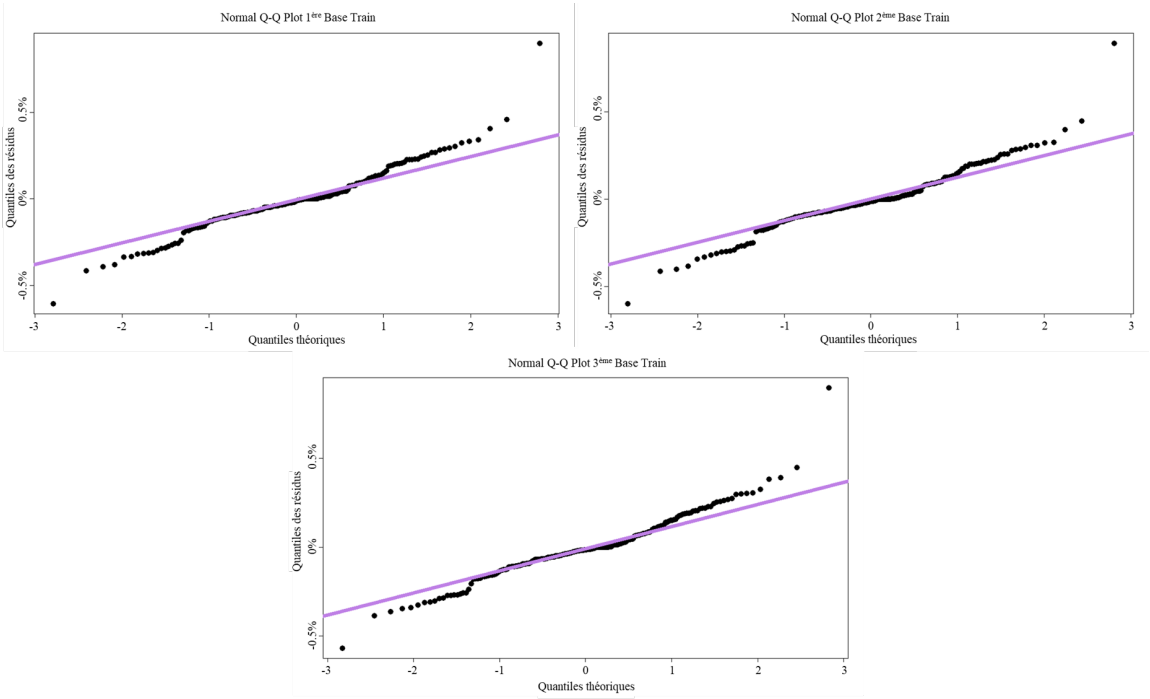


FIGURE 4.6 – QQ-Plot des résidus du modèle SARIMAX (4,0,1)(0,1,1)[12] sur les trois Bases Trains de la validation croisée

Les QQ-plots illustrés dans la Figure 4.6 renforcent l'hypothèse de normalité des résidus. On observe que les points suivent relativement bien la ligne théorique de normalité pour les trois bases d'entraînement, avec seulement quelques légères déviations aux extrémités, ce qui est courant dans la pratique.

4.2.10.3 Conclusion de l'analyse des résidus

L'analyse des résidus du modèle SARIMAX (4,0,1)(0,1,1)[12] avec ACP fournit des preuves solides de sa validité et de sa robustesse : les tests de Ljung-Box démontrent l'absence d'auto-corrélation significative dans les résidus, indiquant que le modèle capture efficacement la structure temporelle des données.

Les tests de stationnarité (ADF et KPSS) confirment que les résidus sont stationnaires, une

condition essentielle pour la validité du modèle SARIMAX. Les analyses graphiques, notamment la tendance des résidus, l'ACF, les histogrammes et les QQ-plots, corroborent les résultats statistiques et fournissent une confirmation visuelle de la stationnarité et de la normalité des résidus.

Ces résultats convergent tous vers la même conclusion : le modèle SARIMAX (4,0,1)(0,1,1)[12] avec ACP satisfait les hypothèses clés nécessaires à sa validité. La stationnarité des résidus, l'absence d'auto-corrélation significative et la distribution normale des résidus indiquent que le modèle réussit à capturer efficacement les dynamiques sous-jacentes de la série temporelle étudiée.

Cette analyse renforce la confiance dans la capacité du modèle à produire des prévisions fiables et robustes pour le taux d'affaires nouvelles. La cohérence des résultats à travers les trois périodes de validation croisée souligne également la stabilité du modèle dans le temps, un aspect capital pour son application pratique dans la prévision des taux d'affaires nouvelles de crédits immobiliers.

4.2.10.4 Projections du modèle optimal

Les résultats visuels des projections du modèle SARIMAX (4,0,1)(0,1,1)[12] choisi sont présentés ci-dessous, illustrant sa performance sur les différentes périodes de validation croisée temporelle.

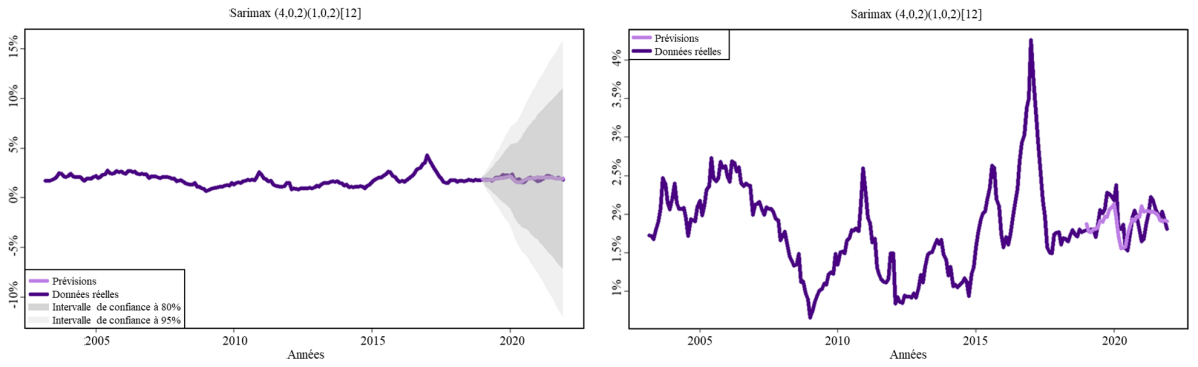


FIGURE 4.7 – Projections du modèle SARIMAX (4,0,1)(0,1,1)[12] sur la période 2019-2021

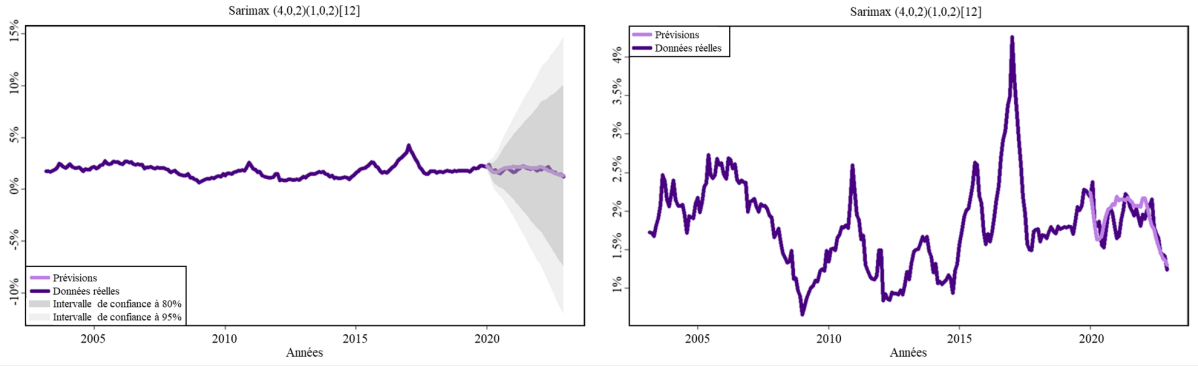


FIGURE 4.8 – Projections du modèle SARIMAX (4,0,1)(0,1,1)[12] sur la période 2020-2022

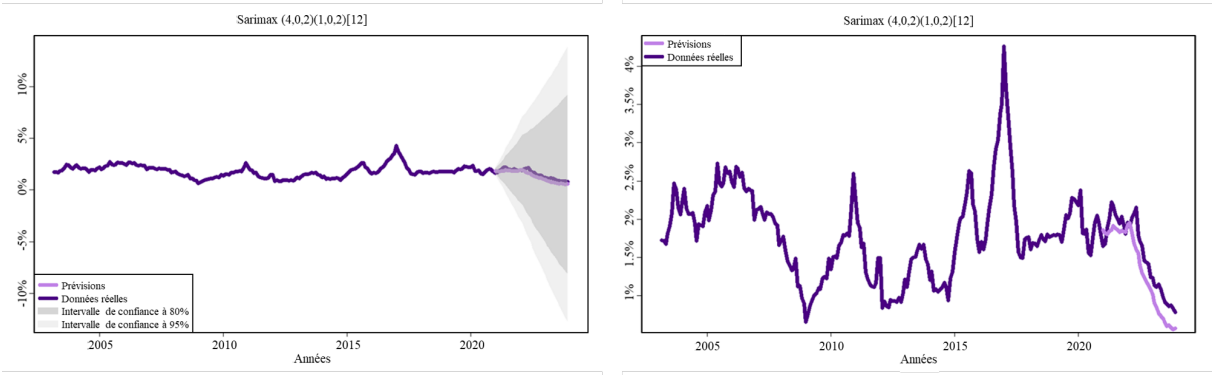


FIGURE 4.9 – Projections du modèle SARIMAX (4,0,1)(0,1,1)[12] sur la période 2021-2023

L'analyse des résultats de projections de modélisation du taux d'affaires nouvelles en assurance emprunteur révèle que le modèle SARIMAX (4,0,1)(0,1,1)[12] avec ACP se distingue comme la solution optimale parmi les configurations testées. Ce modèle démontre une bonne capacité à capturer la tendance générale de l'évolution des taux d'affaires nouvelles, offrant ainsi une base solide pour les projections à moyen et long terme.

La performance du modèle est particulièrement notable dans sa capacité à saisir les mouvements globaux du marché, reflétant avec précision les changements de dynamique sur l'ensemble de la période étudiée. Cette aptitude à suivre la tendance générale est intéressante pour comprendre les évolutions structurelles du marché de l'assurance emprunteur et pour anticiper les besoins futurs en termes de gestion des risques et de stratégies commerciales.

Concernant les variations mensuelles, le modèle SARIMAX montre une performance satisfaisante, bien que légèrement moins précise que pour la tendance globale. Il parvient à capturer une partie significative des fluctuations à court terme, offrant ainsi des informations précieuses sur les dynamiques saisonnières et les variations ponctuelles du marché. Cette capacité, bien que non parfaite, représente un atout pour les prévisions à court terme et la planification opérationnelle.

La robustesse du modèle est également mise en évidence par sa performance constante sur les différentes périodes de validation croisée. Cette stabilité renforce la confiance dans la fiabilité des projections, même face à des conditions de marché changeantes. Néanmoins, il est important de reconnaître que le modèle présente certaines limites, notamment dans la capture précise de tous les pics et creux à court terme, particulièrement visible sur les graphiques zoomés à droite des 3 *Figures 4.7, 4.8 et 4.9*. Ces écarts mineurs entre les prévisions et les valeurs réelles pour certaines variations mensuelles soulignent la complexité inhérente au marché de l'assurance emprunteur et la difficulté à prévoir avec une exactitude parfaite tous les mouvements du marché.

4.2.10.5 Conclusion

Le modèle SARIMAX (4,0,1)(0,1,1)[12] avec ACP offre un équilibre optimal entre la capture des tendances à long terme et la sensibilité aux variations à court terme. Sa capacité à fournir des projections fiables, tout en s'adaptant aux changements structurels du marché, en fait un outil précieux pour la prise de décision stratégique et la gestion des risques dans le secteur de l'assurance emprunteur. Les résultats obtenus démontrent l'efficacité de l'approche choisie et ouvrent la voie à des applications pratiques prometteuses dans la modélisation et la prévision des taux d'affaires nouvelles.

4.3 Limites et perspectives d'amélioration

L'une des principales limitations de cette étude est la quantité relativement restreinte de données historiques disponibles. Avoir un historique plus étendu permettrait au modèle d'apprendre à partir d'un plus grand nombre de scénarios économiques différents, améliorant ainsi sa capacité à généraliser et à produire des prévisions plus robustes.

Concernant le modèle hybride, bien qu'il soit envisagé comme une alternative prometteuse, il ne donne pas les résultats escomptés. Le modèle SARIMAX capture efficacement la tendance générale des taux d'affaires nouvelles, mais il rencontre des difficultés avec la volatilité mensuelle. Bien que le modèle hybride aurait pu potentiellement combler cette lacune, ses performances se sont révélées inférieures. Cela pourrait être attribué à l'utilisation de XGBoost ; une exploration de réseaux de neurones pourrait offrir une meilleure capacité à modéliser les relations non-linéaires complexes dans les résidus. Malheureusement, en raison de contraintes de temps mais également de ressources informatiques, cette option n'a pas pu être explorée.

De plus, l'exploration d'autres types de modèles de séries temporelles, tels que les modèles à espace d'états ou les modèles de lissage exponentiel, pourrait également être bénéfique pour comparer leurs performances et améliorer la précision des prévisions. Une analyse plus approfondie des variables exogènes pourrait également identifier des indicateurs économiques supplémentaires avec pour but d'améliorer la capacité prédictive du modèle.

Ces limites identifiées ouvrent la voie à de futures recherches et améliorations. En particulier, l'exploration de techniques d'apprentissage profond pour la modélisation hybride représente une perspective prometteuse pour capturer plus efficacement la volatilité à court terme tout en préservant la capacité du modèle SARIMAX à modéliser les tendances à long terme.

4.4 Conclusion sur la modélisation du taux d'affaires nouvelles

Au terme de cette analyse, le modèle SARIMAX (4,0,1)(0,1,1)[12] utilisant l'ACP pour la sélection des variables exogènes se révèle être le plus performant pour modéliser le taux d'affaires nouvelles. Ce modèle démontre sa capacité à capturer efficacement les tendances et les dynamiques complexes de la série temporelle, tout en offrant une bonne précision prédictive.

Bien que le modèle hybride ait été envisagé comme une alternative prometteuse, notamment pour mieux saisir la volatilité à court terme, ses performances ne surpassent pas celles du modèle SARIMAX retenu. Cela souligne la robustesse du modèle SARIMAX dans ce contexte spécifique, tout en ouvrant des perspectives pour de futures améliorations, notamment dans la modélisation des composantes non-linéaires.

L'utilisation de l'ACP pour la sélection des variables exogènes s'avère particulièrement efficace, permettant de réduire la dimensionnalité des données tout en préservant l'information pertinente. Cette approche contribue à améliorer la stabilité et l'interprétabilité du modèle.

Avec la validation de ce modèle pour le taux d'affaires nouvelles, l'étude dispose d'un outil fiable pour la prévision de cette composante essentielle de l'assurance emprunteur. Cette étape franchie, il est à présent temps de se concentrer sur la modélisation de la loi de rachat, un autre élément clé dans la compréhension et la prévision des dynamiques du portefeuille d'assurance emprunteur. La transition vers la modélisation de la loi de rachat permet d'obtenir une vision plus complète et d'intégrer des flux entrants et sortants dans un portefeuille, contribuant ainsi à une gestion plus précise et proactive des risques associés à l'assurance emprunteur.

Chapitre 5

APPLICATION : Construction de la loi de Rachat

Après avoir modélisé le taux d'affaires nouvelles, qui représente l'entrée des assurés dans le portefeuille d'assurance emprunteur, ce chapitre se focalise sur un aspect tout aussi important : la modélisation de la loi de rachat. Cette loi caractérise la sortie des assurés du portefeuille et joue un rôle fondamental dans la gestion des risques et la prévision des flux financiers.

Contrairement au taux d'affaires nouvelles, la loi de rachat présente des caractéristiques uniques qui nécessitent une approche de modélisation différente. L'analyse ici, repose sur une matrice complexe comprenant 20 lois de rachat distinctes, chacune correspondant à une année d'ancienneté de prêt, s'étendant sur une période de 25 ans. Ces lois sont exprimées mensuellement, couvrant la période de décembre 2010 à décembre 2023.

Il est important de noter que les lois des 5 dernières années d'ancienneté ne sont pas modélisées en raison de leur valeur nulle ou quasi-nulle, reflétant la rareté des rachats pour les prêts très anciens. Cette particularité souligne la nécessité d'une approche sur mesure pour capturer efficacement les dynamiques de rachat au fil du temps et de l'ancienneté des prêts.

5.1 Remise en contexte

Le contexte de la modélisation de la loi de rachat diffère significativement de celui du taux d'affaires nouvelles. Le travail ici se fait avec un ensemble de données couvrant la période de décembre 2010 à décembre 2023, soit environ 13 ans d'historique. La structure de ces données se présente sous forme d'une matrice où :

- Chaque ligne représente un mois spécifique de décembre 2010 à décembre 2023.
- Chaque colonne correspond à une année d'ancienneté du prêt, allant de 1 à 20 ans.
- Chaque cellule contient le taux de rachat mensuel pour une combinaison donnée de mois et d'ancienneté de prêt.

Ce jeu de données comprend une variable dépendante R_j, représentant les *Taux de rachat mensuels* pour chaque année d'ancienneté j du prêt, ainsi que plusieurs variables économiques explicatives X_i, comme indiqué dans le *Tableau 2.5*. Cette structure matricielle permet d'observer l'évolution des taux de rachat non seulement dans le temps, mais aussi en fonction de l'ancienneté des prêts. Il est important de noter que les données originales sont annuelles et sont mensualisées pour cette analyse. Cette mensualisation explique l'absence de saisonnalité marquée dans les taux de rachat, contrairement à ce qui est observé dans le taux d'affaires nouvelles. Malgré l'absence de variations saisonnières, il est possible d'anticiper des tendances potentielles liées à l'évolution

du marché immobilier et des conditions économiques générales. Ces tendances pourraient se manifester à travers des changements graduels dans les taux de rachat au fil du temps, reflétant les dynamiques à long terme du marché du crédit immobilier et de l'environnement économique global.

Pour cette modélisation, l'approche abordée tient compte de la nature spécifique de ces données. Plutôt que d'utiliser un modèle SARIMAX comme pour le taux d'affaires nouvelles, des techniques adaptées aux séries temporelles sans composante saisonnière sont explorées, tout en prenant en compte la dimension de l'ancienneté du prêt. Cette approche permet de développer un modèle capable de capturer les tendances à long terme et les variations spécifiques à chaque année d'ancienneté, fournissant ainsi une base solide pour la prévision des taux de rachat futurs et la gestion des risques associés.

5.2 Application de la méthodologie à la loi de rachat

Cette section présente l'application détaillée de la méthodologie de modélisation des séries temporelles à la loi de rachat en assurance emprunteur. Contrairement à la modélisation du taux d'affaires nouvelles, il s'agit ici non pas d'une seule modélisation, mais de 20 modélisations distinctes, correspondant aux 20 années d'ancienneté du prêt considérées. La démarche suit les étapes établies dans le *Chapitre 3*, tout en intégrant des ajustements spécifiques à cette modélisation particulière.

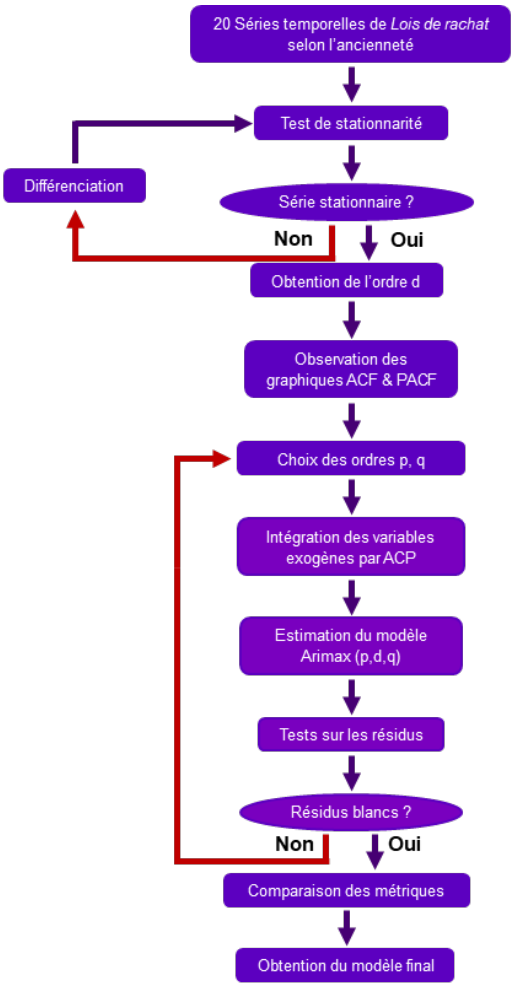


FIGURE 5.1 – Schéma du processus de modélisation de la loi de rachat

Le *Schéma 5.1* ci-dessus synthétise le processus de modélisation adapté à la loi de rachat. Cette illustration met en évidence les particularités de cette approche, notamment la modélisation séparée pour chaque année d'ancienneté et les ajustements méthodologiques nécessaires. Ces spécificités, visibles dans le schéma, sont explicitées plus en détail dans les sections suivantes.

La méthodologie appliquée ici, bien que basée sur le cadre général présenté précédemment, intègre des modifications importantes pour s'adapter à la nature unique des données de rachat. Ces ajustements concernent notamment la gestion de l'absence de saisonnalité, la prise en compte de l'ancienneté du prêt comme dimension supplémentaire, et l'adaptation des techniques de modélisation à des séries temporelles mensualisées à partir de données annuelles.

Chaque sous-section qui suit détaille une étape spécifique du processus, en mettant l'accent sur les adaptations réalisées pour la modélisation de la loi de rachat. Cette approche permet de comprendre comment les concepts théoriques sont appliqués et ajustés dans le contexte spécifique de la modélisation des taux de rachat en fonction de l'ancienneté du prêt.

5.2.1 Préparation des données

Pour l'application de la méthodologie à la loi de rachat, la préparation des données diffère légèrement de celle utilisée pour le taux d'affaires nouvelles. Dans ce cas, le travail est effectué sur une matrice de taux de rachat R_j, où j représente l'année d'ancienneté du prêt, allant de 1 à 20 ans. Chaque série R_j est transformée en objet de série temporelle à l'aide de la fonction `ts()` de R, avec une fréquence mensuelle.

Les variables explicatives X_i, représentant divers indicateurs économiques, sont également préparées comme des séries temporelles mensuelles sur la même période que les taux de rachat (de décembre 2010 à décembre 2023).

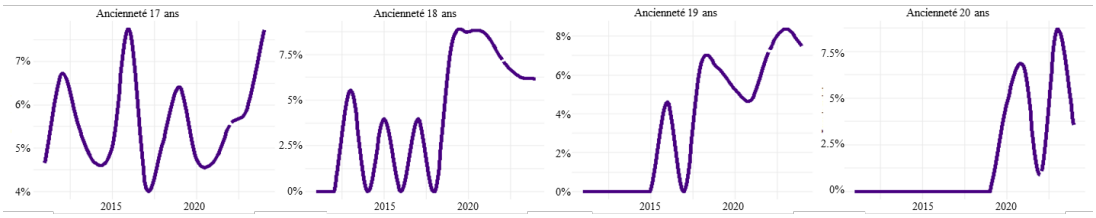


FIGURE 5.2 – Taux de rachat mensuels en % pour les 20 premières années d'ancienneté du prêt

La *Figure 5.2* illustre la structure des données de taux de rachat pour les 20 premières années d'ancienneté. Cette visualisation met en évidence la nature multidimensionnelle de l'analyse, où chaque ligne représente l'évolution du taux de rachat pour une année d'ancienneté spécifique au fil du temps.

Cette structuration des données en format de série temporelle permet de préserver l'ordre chronologique et la périodicité des observations pour chaque année d'ancienneté, tout en facilitant l'analyse comparative entre les différentes durées de prêt. Cette approche permet de capturer non seulement les tendances temporelles générales, mais aussi les variations spécifiques liées à l'ancienneté des prêts.

La préparation inclut également une vérification de la qualité des données, notamment la gestion des valeurs manquantes ou aberrantes, et l'alignement temporel de toutes les séries (taux de rachat et variables explicatives) pour assurer une cohérence dans l'analyse subséquente.

5.2.2 Tests de stationnarité et transformation des données

Après la transformation des variables R_j en séries temporelles, où j représente l'année d'ancienneté du prêt, l'étape suivante consiste à vérifier leur stationnarité. Pour obtenir une évaluation robuste, deux tests complémentaires sont appliqués : le *test ADF* et le *test KPSS*. Ces tests sont choisis pour leur complémentarité, présentant des hypothèses nulles inverses :

— Le test ADF pose comme hypothèse nulle la non-stationnarité de la série ;

— Le test KPSS, à l'inverse, pose comme hypothèse nulle la stationnarité de la série.

Cette approche permet une analyse plus complète et robuste de la stationnarité des séries temporelles, réduisant le risque d'erreur lié à l'utilisation d'un seul test.

Dans un souci de concision et pour alléger la partie, seuls les résultats des tests de stationnarité sur les séries différenciées ΔR_j sont présentés ci-dessous. Les tests préliminaires montrant que les séries originales R_j ne sont pas stationnaires, la section se concentre ici sur l'analyse des séries différenciées ΔR_j, où :

$$\Delta R_{j,t} = R_{j,t} - R_{j,t-1}$$

Les tests ADF et KPSS sont effectués à l'aide des fonctions `ur.df()` et `ur.kpss()` respectivement. Les valeurs critiques pour ces tests sont les suivantes :

T	1%	5%	10%
marché (1)			
250	**-2.58**	**-1.95**	**-1.62**

10%	5%	2.5%	1%
0.347	**0.463**	**0.574**	**0.739**

TABLE 5.1 – Valeurs critiques des tests ADF (à gauche) et KPSS (à droite) sur les variables exogènes

Les résultats des tests pour chaque année d'ancienneté sont présentés dans le tableau suivant :

Ancienneté	Test ADF		Test KPSS	
	Stat. de test	p-value	Stat. de test	p-value
1	-16.1	0.0165	0.0737	0.1
2	-12.2	0.0332	0.137	0.1
3	-13.2	0.0143	0.218	0.1
4	-18.2	0.01	0.263	0.1
5	-19.3	0.0139	0.132	0.1
6	-12.6	0.0745	0.117	0.1
7	-18.1	0.0489	0.0819	0.1
8	-16.2	0.0417	0.109	0.1
9	-16.5	0.0505	0.122	0.1
10	-10.6	0.0238	0.27	0.1
11	-12	0.0169	0.157	0.1
12	-7.03	0.0499	0.324	0.1
13	-14.3	0.0579	0.207	0.1
14	-10.7	0.0219	0.127	0.1
15	-8.05	0.0482	0.222	0.1
16	-8.26	0.0135	0.122	0.1
17	-11.1	0.01	0.123	0.1
18	-6.58	0.01	0.0779	0.1
19	-6.4	0.01	0.055	0.1
20	-7.25	0.01	0.0672	0.1

TABLE 5.2 – Résultats des tests ADF et KPSS pour chaque année d'ancienneté

L'analyse des résultats montre que pour l'ensemble des séries différenciées ΔR_j, les tests ADF et KPSS convergent vers la conclusion de stationnarité :

— Pour le test ADF, l'ensemble des statistiques de test sont inférieures aux valeurs critiques, avec des p-valeurs inférieures à 0.05, indiquant le rejet de l'hypothèse nulle de non-stationnarité ;

— Pour le test KPSS, l'ensemble des statistiques de test sont inférieures aux valeurs critiques, avec des p-valeurs égales à 0.1, indiquant que l'hypothèse nulle de stationnarité ne peut être rejetée.

En conclusion, l'utilisation des séries différenciées ΔR_j comme base pour la modélisation ultérieure est justifiée. A l'instar de la modélisation des affaires nouvelles, ces modèles et prévisions sont basés sur des séries différenciées, nécessitant une "dé-différenciation" des résultats pour les interpréter dans l'échelle originale des taux de rachat.

5.2.3 Identification des ordres p, d, q (ARIMA)

Pour identifier les ordres optimaux du modèle ARIMAX(p,d,q) pour chaque année d'ancienneté du prêt, la méthode de recherche par grille (*grid search*) est appliquée, conformément à la méthodologie décrite *Partie 3.3.3*. Cette approche systématique permet d'évaluer différentes combinaisons de paramètres pour les composantes du modèle (p, d, q). Contrairement à la modélisation du taux d'affaires nouvelles, la composante saisonnière n'est pas incluse ici, en raison de la mensualisation des données de rachat qui a éliminé les *patterns* saisonniers.

Pour cette recherche par grille, les plages de valeurs suivantes sont explorées, basées sur des considérations théoriques et l'expertise du domaine :

- p : 0 à 4 (ordre autorégressif)
- d : 0 à 2 (ordre de différenciation)
- q : 0 à 4 (ordre de moyenne mobile)

Ces plages sont choisies pour couvrir un large éventail de spécifications de modèles (soit 75 combinaisons par modèle) tout en restant dans des limites raisonnables pour l'analyse de séries temporelles mensuelles non saisonnières. La recherche par grille est effectuée séparément pour chacune des 20 séries temporelles correspondant aux différentes années d'ancienneté du prêt.

Il est important de noter que, bien que les séries soient différenciées pour les tests de stationnarité, l'ordre de différenciation d est tout de même inclus dans la recherche par grille. Cela permet au processus de sélection de confirmer ou d'ajuster le niveau de différenciation optimal pour chaque série.

Les résultats de cette recherche par grille servent de base pour une comparaison ultérieure des modèles candidats pour chaque année d'ancienneté. La sélection finale des paramètres optimaux est effectuée, comme pour les affaires nouvelles, en utilisant l'analyse des résidus, les critères de sélection et les métriques d'erreur détaillés dans la section *Analyse comparative des modèles et résultats du modèle optimal*.

Cette approche individualisée pour chaque année d'ancienneté permet de capturer les dynamiques spécifiques qui peuvent varier selon la maturité du prêt, offrant ainsi une modélisation plus précise et adaptée de la loi de rachat dans son ensemble.

5.2.4 Intégration des variables exogènes

Fort de l'expérience acquise lors de la modélisation des affaires nouvelles, l'intégration des variables exogènes pour la loi de rachat est exclusivement faite par la méthode de l'Analyse en Composantes Principales (ACP). Cette décision est motivée par plusieurs facteurs :

1. **Supériorité démontrée de l'ACP :** Les résultats précédents montre clairement la supériorité de l'ACP par rapport à la méthode de la matrice de corrélation, tant pour les modèles SARIMAX de base que pour les modèles hybrides. Cette performance supérieure justifie l'adoption exclusive de l'ACP pour cette nouvelle modélisation ;
2. **Historique de données plus court :** La loi de rachat dispose d'un historique de données plus limité que celui des affaires nouvelles. Dans ce contexte, l'ACP présente un avantage significatif car elle permet de réduire la dimensionnalité des variables exogènes. Cette réduction est particulièrement bénéfique dans le cadre d'un ensemble de données plus restreint ;
3. **Meilleure adéquation au modèle :** La dimension réduite résultant de l'ACP est mieux adaptée à la structure du modèle, surtout compte tenu de l'historique plus court. Cela permet d'éviter les problèmes de sur-paramétrisation qui pourraient survenir avec un trop grand nombre de variables explicatives par rapport au nombre d'observations disponibles ;
4. **Cohérence méthodologique :** L'utilisation exclusive de l'ACP assure une cohérence méthodologique avec l'approche qui s'est révélée la plus efficace dans la modélisation précédente. Cette cohérence facilite la comparaison des résultats et l'interprétation des modèles entre les différentes composantes de l'étude.

La procédure d'intégration des variables exogènes par ACP suit la même méthodologie que celle détaillée précédemment dans la section *Analyse en Composantes Principales (ACP)*. Cette

approche systématique assure une cohérence dans le traitement des variables exogènes entre la modélisation des affaires nouvelles et celle de la loi de rachat. Elle permet de capturer efficacement l'information pertinente tout en maintenant un modèle parcimonieux, particulièrement adapté à la modélisation des 20 séries temporelles distinctes avec un historique de données limité.

5.2.5 Base train et test

Pour la modélisation de la loi de rachat, la division du jeu de données en bases d'entraînement (*train*) et de test est ajustée pour tenir compte de l'historique plus court. Dans ce cas, les deux dernières années de données sont réservées comme base de test, tandis que le reste constitue la base d'entraînement.

Cette séparation chronologique, bien que plus courte que pour la modélisation des affaires nouvelles, permet néanmoins d'évaluer la performance et la capacité de généralisation du modèle ARIMAX sur une période suffisante pour capturer les tendances récentes des taux de rachat. Cette approche équilibre la nécessité d'avoir une base d'entraînement substantielle avec le besoin d'une période de test représentative pour chacune des 20 séries temporelles.

5.2.6 Mise en place du modèle ARIMAX(p,d,q)

Conformément à la méthodologie détaillée précédemment, le modèle ARIMAX est ajusté en utilisant la fonction `Arima` de R. Les paramètres (p,d,q) sont déterminés selon la *procédure de recherche par grille* décrite antérieurement.

L'intégration des variables exogènes est réalisée par ACP, comme expliqué dans la *Partie 5.2.4*. Cette étape d'ajustement du modèle ARIMAX constitue le cœur de l'analyse, intégrant les composantes auto-régressives, les moyennes mobiles, et l'influence des variables exogènes transformées.

5.2.7 Analyse et Tests statistiques des résidus

L'analyse des résidus est menée pour chaque combinaison de paramètres (p,d,q) explorée dans le *grid search*. Cette analyse comprend :

- Le test de Ljung-Box pour évaluer l'auto-corrélation des résidus ;
- Les tests de stationnarité ADF et KPSS ;
- Des analyses graphiques incluant le QQ-plot, l'histogramme des résidus, et la fonction d'auto-corrélation (ACF).

Ces tests et analyses visent à vérifier les hypothèses clés du modèle ARIMAX : absence d'auto-corrélation, stationnarité et normalité des résidus.

Il est important de noter que si l'un de ces tests statistiques ou visuels n'est pas concluant pour une combinaison de paramètres donnée, cette combinaison n'est pas retenue pour la suite de l'analyse. Cette approche permet de filtrer les modèles qui ne satisfont pas les conditions fondamentales de validité.

Les résultats détaillés de ces tests pour les différentes combinaisons de paramètres ne sont pas présentés dans cette section. Ils sont plutôt intégrés dans l'évaluation globale des modèles, qui est abordée dans la *Partie 5.2.10* sur la comparaison des résultats des métriques. Cette approche permet de se concentrer sur les modèles les plus performants et valides selon les critères établis.

5.2.8 Modélisation hybride

Dans le contexte de la modélisation de la loi de rachat, l'approche hybride combinant ARIMAX et XGBoost n'est pas mise mise en œuvre. Cette décision s'appuie sur plusieurs considérations importantes. Premièrement, l'expérience acquise lors de la modélisation du taux d'affaires nouvelles montre que le modèle hybride n'apporte pas systématiquement d'amélioration significative par rapport au modèle ARIMAX seul. De plus, la complexité inhérente à la modélisation de 20 lois de rachat distinctes rend l'application d'une approche hybride particulièrement chronophage et computationnellement intensive.

Étant donné les contraintes de temps et la nécessité de produire des résultats robustes pour chacune des 20 séries temporelles, il est jugé plus prudent et efficace de se concentrer sur l'optimisation des modèles ARIMAX individuels. Cette approche permet une analyse et une interprétation claire des résultats pour chaque année d'ancienneté, sans introduire la complexité supplémentaire d'un modèle hybride.

Néanmoins, il est important de noter que l'exploration d'approches hybrides pour la modélisation de la loi de rachat reste une piste intéressante pour de futures recherches. Une telle approche pourrait potentiellement capturer des dynamiques non linéaires subtiles dans les taux de rachat, particulièrement pour les années d'ancienneté où les modèles ARIMAX montrent certaines limites. Dans un contexte où les ressources en temps et en calcul seraient moins contraintes, l'application d'une méthodologie hybride pourrait offrir des perspectives enrichissantes sur la dynamique complexe des rachats de prêts immobiliers.

5.2.9 Projection, validation croisée et analyse vis à vis de la base test

Pour la modélisation de la loi de rachat, l'approche de projection, de validation croisée et d'analyse sur la base test est adaptée pour tenir compte de la période d'historique plus courte. Conformément à la *méthodologie générale*, cette approche est appliquée pour évaluer la performance et la capacité de généralisation des modèles ARIMAX développés pour chaque année d'ancienneté du prêt.

Dans ce contexte spécifique, la validation croisée est réalisée sur deux périodes distinctes : 2020-2022 et 2021-2023. Cette adaptation est nécessaire en raison de l'historique de données plus limité pour la loi de rachat. L'utilisation d'une troisième période de test plus ancienne pourrait compromettre la qualité des prédictions en raison du manque de données suffisantes pour l'entraînement du modèle.

Pour chaque configuration et chaque année d'ancienneté, les projections sont effectuées sur ces deux périodes de test. Cette approche permet une comparaison visuelle et quantitative des prévisions avec les valeurs réelles, tout en maintenant la robustesse de l'analyse malgré la contrainte d'un historique plus court.

Les résultats détaillés de cette analyse, incluant l'évaluation des résidus, les critères de sélection et les métriques de performance, sont présentés et discutés dans la *Partie 5.2.10*. Cette méthodologie, bien qu'adaptée, fournit une base solide pour évaluer la performance des modèles ARIMAX pour chaque année d'ancienneté du prêt, permettant ainsi une compréhension approfondie de la dynamique des rachats à travers différentes maturités de prêt.

5.2.10 Analyse comparative des modèles et résultats du modèle optimal

Cette section présente une analyse détaillée des résultats pour les modèles ARIMAX développés pour chaque année d'ancienneté du prêt, suivie d'une présentation des modèles optimaux choisis.

5.2.10.1 Tableau comparatif des critères de sélection et des métriques d'erreur

Suite à la recherche par grille effectuée pour chacune des 20 années d'ancienneté, les configurations optimales sont identifiées et sont présentées dans le *Tableau 5.3*. Ce tableau résume les moyennes des résultats d'analyse de la validation croisée temporelle pour chaque année d'ancienneté, incluant le meilleur modèle ARIMAX, les critères de sélection (AIC et BIC), les métriques d'erreur (RMSE et MAE), ainsi que les p-valeurs des tests de Ljung-Box, ADF et KPSS pour l'analyse des résidus.

Ancienneté	Best Model	AIC	BIC	RMSE	MAE	Ljung.Box p.value	ADF p.value	KPSS p.value
1	ARIMA(4,1,2)	-1175.65	-1141.73	0.16%	0.14%	0.21	0.01	0.10
2	ARIMA(2,1,2)	-1113.54	-1085.27	0.11%	0.09%	0.27	0.01	0.10
3	ARIMA(4,1,2)	-979.87	-945.94	0.32%	0.24%	0.09	0.01	0.10
4	ARIMA(2,1,1)	-905.26	-879.81	0.45%	0.33%	0.13	0.01	0.10
5	ARIMA(3,1,2)	-961.03	-929.94	0.69%	0.51%	0.23	0.01	0.10
6	ARIMA(2,1,3)	-1058.49	-1027.39	0.34%	0.29%	0.09	0.01	0.10
7	ARIMA(3,0,1)	-1035.31	-1004.12	0.29%	0.21%	0.12	0.03	0.10
8	ARIMA(2,1,1)	-1056.79	-1031.34	0.17%	0.14%	0.39	0.01	0.10
9	ARIMA(2,1,1)	-971.54	-946.09	0.46%	0.38%	0.24	0.01	0.10
10	ARIMA(4,1,1)	-1063.18	-1032.08	0.42%	0.36%	0.25	0.01	0.10
11	ARIMA(3,0,2)	-1104.06	-1070.03	0.37%	0.28%	0.37	0.03	0.10
12	ARIMA(1,1,3)	-1169.63	-1141.36	0.12%	0.10%	0.12	0.02	0.10
13	ARIMA(4,1,3)	-979.28	-942.53	0.42%	0.33%	0.06	0.01	0.10
14	ARIMA(2,0,1)	-1057.99	-1029.64	0.66%	0.51%	0.12	0.01	0.10
15	ARIMA(2,1,2)	-1000.70	-972.43	0.46%	0.31%	0.75	0.01	0.10
16	ARIMA(2,0,2)	-920.52	-889.33	0.48%	0.36%	0.29	0.02	0.10
17	ARIMA(4,1,1)	-925.98	-894.88	0.59%	0.45%	0.46	0.01	0.10
18	ARIMA(1,1,3)	-427.85	-399.58	0.39%	0.33%	0.13	0.03	0.10
19	ARIMA(1,1,3)	-508.12	-479.85	0.40%	0.32%	0.14	0.02	0.10
20	ARIMA(4,0,1)	-559.80	-525.78	1.72%	1.49%	0.75	0.02	0.10

TABLE 5.3 – Moyenne des résultats d'analyse de la validation croisée pour chaque année d'ancienneté

La diversité des modèles ARIMA optimaux, variant de ARIMA(1,1,3) à ARIMA(4,1,3) selon l'année d'ancienneté, souligne la complexité et la variabilité des dynamiques de rachat en fonction de la maturité du prêt. Cette hétérogénéité des modèles reflète la nécessité d'une approche adaptative pour capturer efficacement les comportements de rachat à différents stades du cycle de vie du prêt.

Les critères de performance AIC et BIC présentent des variations significatives entre les années d'ancienneté, indiquant des niveaux de complexité et d'ajustement différents selon la période considérée. Les modèles pour les années d'ancienneté 1, 2 et 12 se distinguent par leurs valeurs AIC et BIC particulièrement basses, suggérant un ajustement optimal pour ces périodes spécifiques.

En termes de précision des prévisions, les métriques RMSE et MAE affichent généralement des valeurs faibles, témoignant d'une bonne précision globale des modèles. Les années d'ancienneté 2

et 12 se démarquent avec des RMSE particulièrement bas, de 0.11% et 0.12% respectivement, indiquant une bonne performance prédictive pour ces périodes.

Une particularité notable concerne le modèle pour l'année d'ancienneté 20, qui présente des critères de performance et des métriques d'erreur nettement plus élevées, avec notamment un RMSE de 1.72%. Cette observation suggère une difficulté accrue dans la modélisation des rachats pour les prêts les plus anciens, potentiellement due à une plus grande variabilité ou à des facteurs spécifiques influençant les décisions de rachat à ce stade avancé du prêt.

Il est important de noter que les résultats détaillés pour les deux bases de validation croisée (2020-2022 et 2021-2023) sont disponibles en *Annexes E.1 et E.2*, permettant une analyse plus approfondie de la stabilité des modèles dans le temps.

Pour la suite de cette étude sur la loi de rachat, les modèles ARIMA spécifiques à chaque année d'ancienneté sont retenus, comme présentés dans le *Tableau 5.3*. Ce choix est motivé par la nécessité de capturer les dynamiques uniques de rachat pour chaque période de maturité du prêt, offrant ainsi une modélisation plus précise et adaptée de la loi de rachat dans son ensemble.

Cette approche individualisée permet de tenir compte des variations significatives dans les comportements de rachat selon l'ancienneté du prêt, tout en maintenant une bonne performance globale en termes de critères de sélection et de métriques d'erreur.

5.2.10.2 Analyse des résidus du modèle optimal

L'analyse des résidus des modèles ARIMAX optimaux pour chaque année d'ancienneté est effectuée pour vérifier les hypothèses clés des modèles : absence d'auto-corrélation, stationnarité et normalité des résidus. Les résultats des tests sont les suivants :

1. **Test de Ljung-Box** : Le test de Ljung-Box est appliqué sur les deux périodes de validation croisée pour évaluer l'auto-corrélation des résidus. Les résultats sont présentés dans le *Tableau 5.3*, avec des détails supplémentaires disponibles en *Annexes E.1 et E.2* pour les données précises des deux bases.

   Toutes les p-valeurs sont supérieures au seuil de significativité de 0.05, ce qui indique qu'il n'est pas possible rejeter l'hypothèse nulle d'absence d'auto-corrélation. Cela suggère que les modèles capturent efficacement les dynamiques des séries temporelles pour la plupart des années d'ancienneté, ne laissant pas de structure d'auto-corrélation significative dans les résidus.

2. **Tests de stationnarité** : Les tests ADF et KPSS sont effectués pour vérifier la stationnarité des résidus. Pour prouver leur stationnarité, des tests ADF et KPSS sont effectués à l'aide des fonctions `ur.df()` et `ur.kpss()`. Il en ressort 4 éléments d'analyse chacun :

   — Le type de régression du test pour le test ADF, il est obtenu `"Test regression none"`, signifiant que la régression du test n'inclut ni constante (intercept), ni terme de tendance, soit un marché aléatoire pure ;
   — Le type de test et le nombre de retards pour le test KPSS : `"Test is of type: mu with 5 lags"`, indiquant que le test est effectué avec une constante μ et 5 retards ;
   — Les valeurs critiques pour les deux tests sont donc :

T	1%	5%	10%
marché (1)			
250	**-2.58**	**-1.95**	**-1.62**

10%	5%	2.5%	1%
0.347	**0.463**	**0.574**	**0.739**

TABLE 5.4 – Valeurs critiques des tests ADF (à gauche) et KPSS (à droite) sur les variables exogènes

— Les valeurs de la statistique de test et les p-valeurs pour les résidus des modèles sont :

	Test ADF				Test KPSS			
	Value of test-statistic		p-value		Value of test-statistic		p-value	
	2021-22	2022-23	2021-22	2022-23	2021-22	2022-23	2021-22	2022-23
1	**-8.88**	**-9.22**	**1.4E-15**	**753.9E-18**	**0.11**	**0.04**	**0.10**	**0.10**
2	**-8.50**	**-7.88**	**1.3E-12**	**1.2E-12**	**0.07**	**0.13**	**0.10**	**0.10**
3	**-8.82**	**-9.17**	**1.7E-15**	**1.0E-15**	**0.11**	**0.04**	**0.10**	**0.10**
4	**-10.91**	**-10.13**	**-695.2E-18**	**4.5E-18**	**0.12**	**0.03**	**0.10**	**0.10**
5	**-9.39**	**-9.76**	**736.7E-18**	**37.8E-18**	**0.10**	**0.03**	**0.10**	**0.10**
6	**-7.89**	**-7.30**	**28.1E-12**	**26.3E-12**	**0.03**	**0.10**	**0.10**	**0.10**
7	**-7.46**	**-7.76**	**2.2E-12**	**2.3E-12**	**0.31**	**0.24**	**0.10**	**0.10**
8	**-10.60**	**-9.84**	**-675.4E-18**	**23.0E-18**	**0.22**	**0.04**	**0.10**	**0.10**
9	**-9.64**	**-10.01**	**708.5E-18**	**8.7E-18**	**0.10**	**0.03**	**0.10**	**0.10**
10	**-9.66**	**-8.97**	**2.7E-15**	**3.2E-15**	**0.28**	**0.05**	**0.10**	**0.10**
11	**-7.15**	**-7.44**	**12.5E-12**	**12.9E-12**	**0.34**	**0.27**	**0.10**	**0.10**
12	**-8.10**	**-7.51**	**9.7E-12**	**9.1E-12**	**0.19**	**0.04**	**0.10**	**0.10**
13	**-7.29**	**-7.59**	**5.6E-12**	**5.8E-12**	**0.15**	**0.08**	**0.10**	**0.10**
14	**-7.40**	**-6.85**	**299.9E-12**	**280.3E-12**	**0.09**	**0.15**	**0.10**	**0.10**
15	**-7.21**	**-7.51**	**8.8E-12**	**9.1E-12**	**0.23**	**0.16**	**0.10**	**0.10**
16	**-7.38**	**-6.83**	**322.9E-12**	**301.7E-12**	**0.01**	**0.08**	**0.10**	**0.10**
17	**-8.43**	**-8.76**	**10.3E-15**	**9.9E-15**	**0.12**	**0.05**	**0.10**	**0.10**
18	**-7.96**	**-7.37**	**19.8E-12**	**18.5E-12**	**0.11**	**0.03**	**0.10**	**0.10**
19	**-7.23**	**-7.53**	**7.9E-12**	**8.1E-12**	**0.10**	**0.03**	**0.10**	**0.10**
20	**-9.13**	**-8.46**	**54.2E-15**	**51.3E-15**	**0.11**	**0.16**	**0.10**	**0.10**

TABLE 5.5 – Valeurs critiques et p-valeurs des tests ADF (à gauche) et KPSS (à droite) sur les résidus par ancienneté

L'analyse des résultats du test ADF révèle que pour toutes les années d'ancienneté, les statistiques de test sont inférieures aux valeurs critiques à tous les seuils de significativité usuels (1%, 5% et 10%). De plus, les p-valeurs associées sont extrêmement faibles (inférieures à 1E-12 dans la plupart des cas), bien en-dessous du seuil conventionnel de 0.05. Ces observations indiquent que l'hypothèse nulle de présence d'une racine unitaire est fortement rejetée pour tous les modèles, suggérant que les résidus sont stationnaires.

De même, l'analyse des résultats du test KPSS montre que pour toutes les années d'ancienneté, les statistiques de test sont inférieures aux valeurs critiques à tous les seuils de significativité usuels. Les p-valeurs associées sont généralement égales à 0.10, ce qui est supérieur au seuil conventionnel de 0.05. Ces observations indiquent que l'hypothèse nulle de stationnarité ne peut être rejetée pour l'ensemble des modèles.

En conclusion, les résultats des deux tests convergent vers la même conclusion pour l'en-

semble des modèles : **les résidus sont bien stationnaires pour toutes les années d'ancienneté**. Cette stationnarité des résidus confirme la validité des modèles ARIMAX sélectionnés pour chaque année d'ancienneté et renforce la fiabilité des prévisions qui en découlent.

3. **Analyses graphiques** :
   Les analyses graphiques des résidus pour chaque modèle ARIMAX sont effectuées. En raison du grand nombre de graphiques générés (40 au total, correspondant aux 20 années d'ancienneté et aux 2 bases de test), ces visualisations sont placées en *Annexes E.1*. Néanmoins, une analyse synthétique de ces graphiques est présentée ci-dessous :

   — **Tendance des résidus** : Pour la majorité des modèles, la tendance visuelle des résidus corrobore les résultats théoriques sur la stationnarité. Une fluctuation des résidus est autour de zéro sans tendance apparente, ce qui est caractéristique d'un processus stationnaire.
   — **Fonction d'auto-corrélation (ACF)** : L'analyse des graphiques ACF confirme globalement les résultats du test de Ljung-Box. Pour la plupart des modèles, les auto-corrélations sont majoritairement contenues dans les limites de confiance, indiquant une absence d'auto-corrélation significative dans les résidus
   — **Histogramme des résidus** : Les histogrammes des résidus pour les différents modèles et bases de test montrent des distributions qui s'apparentent à une distribution normale. Cette observation suggère que l'hypothèse de normalité des résidus, importante pour la validité des modèles ARIMAX, est raisonnablement satisfaite pour la plupart des années d'ancienneté.
   — **QQ-plot** : Les QQ-plots (*Annexes E.2 et E.3*) viennent renforcer l'hypothèse de normalité des résidus pour la majorité des modèles. Les points suivent relativement bien la ligne théorique de normalité, avec seulement quelques légères déviations aux extrémités, ce qui est courant dans la pratique.

Ces analyses visuelles, combinées aux tests statistiques précédents, renforcent la validité globale des modèles ARIMAX sélectionnés pour chaque année d'ancienneté dans la modélisation de la loi de rachat. Elles confirment que les hypothèses fondamentales de stationnarité, d'absence d'auto-corrélation et de normalité des résidus sont généralement respectées, soutenant ainsi la fiabilité des prévisions issues de ces modèles.

5.2.10.3 Projections du modèle optimal

Les résultats visuels des projections des modèles ARIMAX choisis pour chaque année d'ancienneté sont présentés ci-dessous. Ces graphiques comparent les valeurs prédites aux valeurs réelles pour les deux périodes de validation croisée (2021-2022 et 2022-2023).

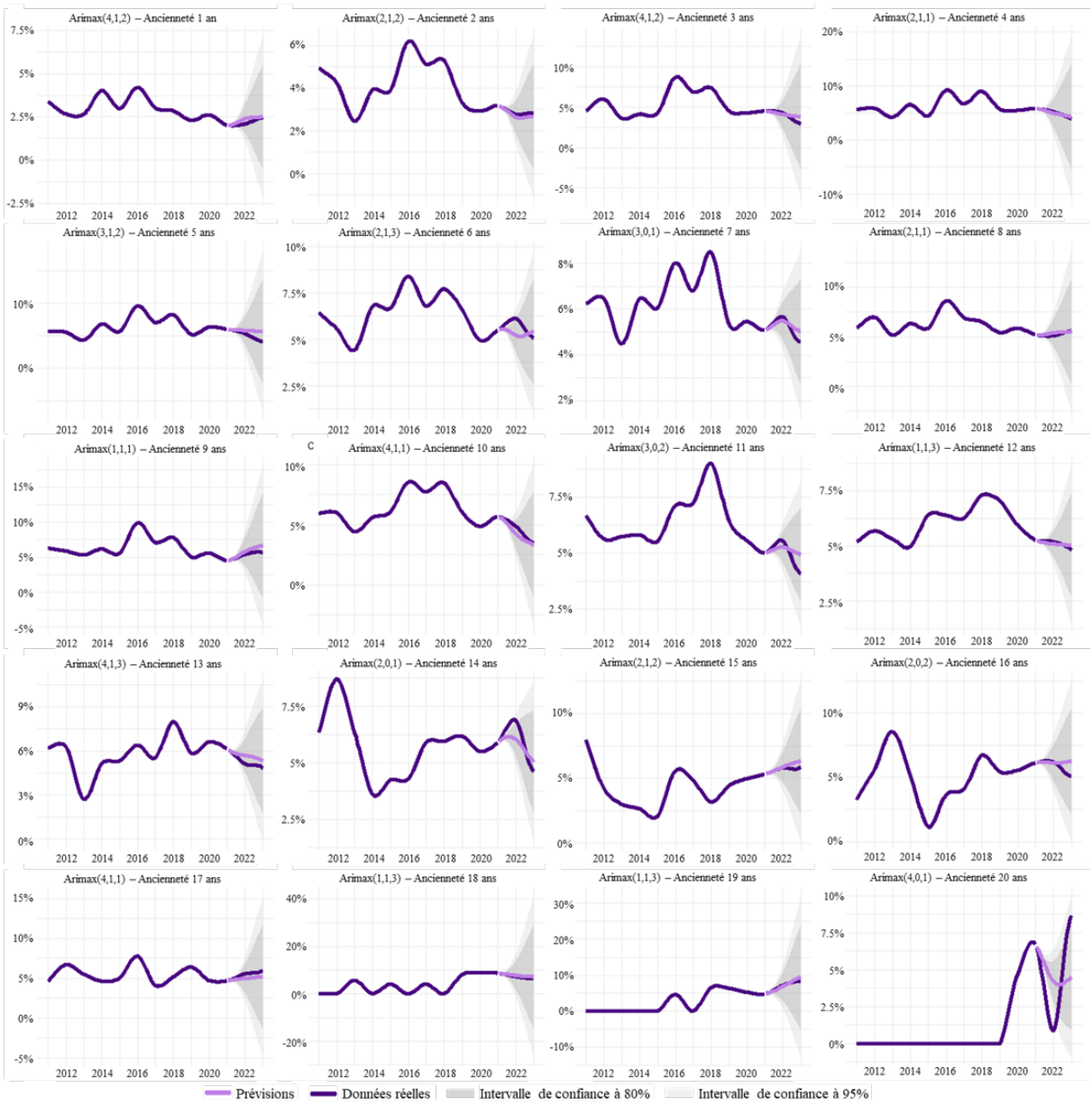


FIGURE 5.3 – Projections des modèles ARIMAX (4,0,2)(1,0,2)[12] sur la période 2021-22

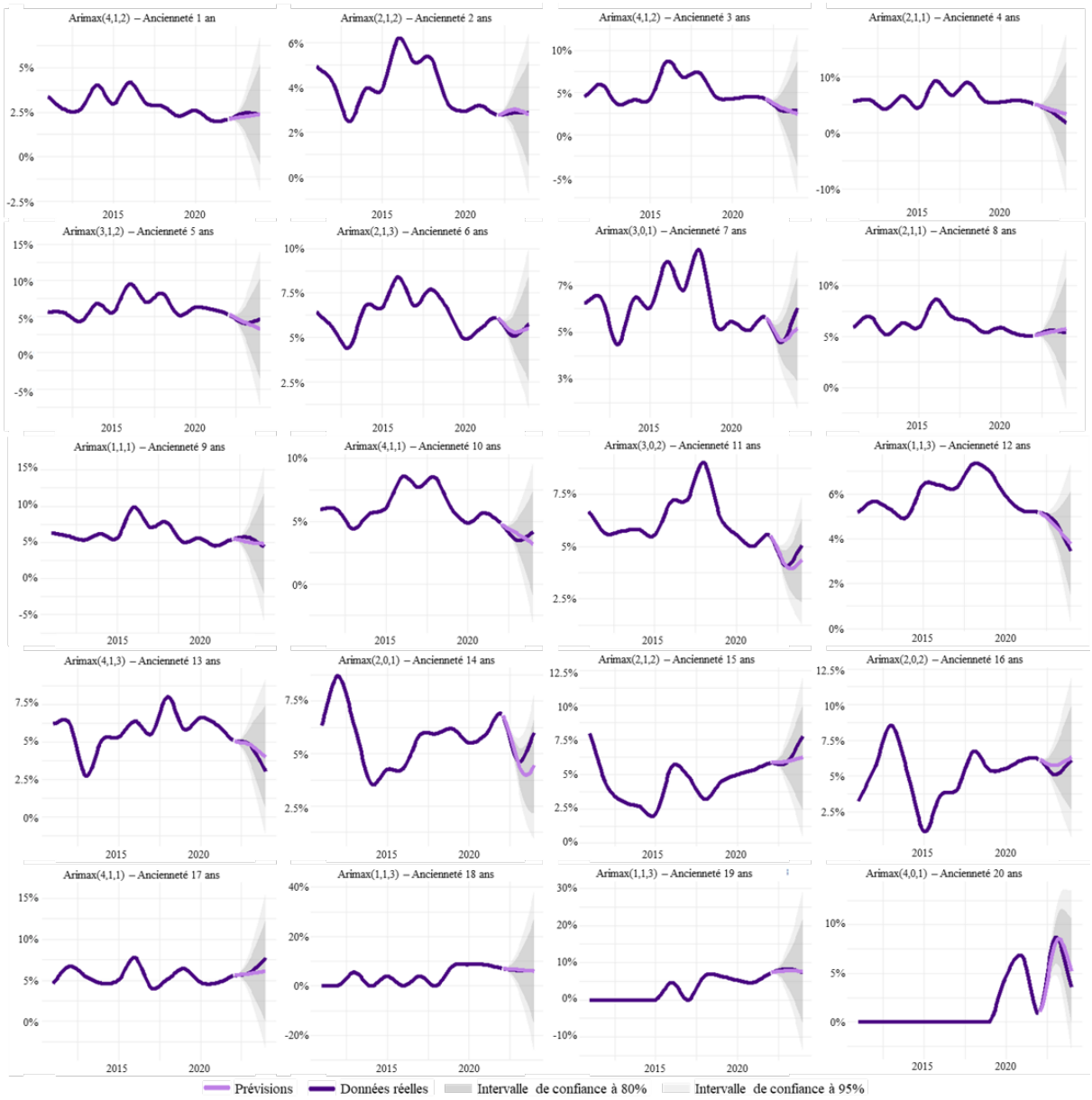


FIGURE 5.4 – Projections des modèles ARIMAX (4,0,2)(1,0,2)[12] sur la période 2022-23

Ces graphiques illustrent la performance des modèles ARIMAX pour chacune des 20 années d'ancienneté du prêt. Il est possible observer la capacité des modèles à capturer les tendances et les variations des taux de rachat pour différentes maturités de prêt.

5.2.10.4 Conclusion

L'analyse de la modélisation de la loi de rachat en assurance emprunteur révèle que les modèles ARIMAX optimisés pour chaque année d'ancienneté du prêt offrent une solution robuste et efficace. Cette approche individualisée permet de capturer les dynamiques spécifiques des taux de rachat à différents stades de la vie du prêt, offrant ainsi une compréhension nuancée du comportement des assurés.

Les modèles ARIMAX sélectionnés démontrent une capacité à saisir les tendances générales des taux de rachat pour chaque année d'ancienneté. Cette performance est particulièrement notable

dans la capture des mouvements à moyen et long terme, reflétant avec précision les changements structurels dans les comportements de rachat au fil du temps. Cette aptitude est intéressante pour comprendre l'évolution des risques associés aux différentes cohortes de prêts et pour anticiper les besoins futurs en termes de gestion de portefeuille.

Concernant les variations à court terme, les modèles ARIMAX montrent une performance satisfaisante, bien que légèrement moins précise que pour les tendances globales. Ils parviennent à capturer une partie significative des fluctuations mensuelles, offrant des informations précieuses sur les dynamiques et les variations ponctuelles des taux de rachat. Cette capacité, bien que non parfaite, représente un atout pour les prévisions à court terme et la planification opérationnelle.

La robustesse des modèles est mise en évidence par leur performance constante sur les différentes périodes de validation croisée et à travers les différentes années d'ancienneté. Cette stabilité renforce la confiance dans la fiabilité des projections, même face à des conditions de marché changeantes et pour des prêts à différents stades de maturité.

5.3 Limites et perspectives d'amélioration

Malgré les performances satisfaisantes des modèles ARIMAX, plusieurs limitations et pistes d'amélioration sont identifiées :

La quantité limitée de données historiques constitue une première contrainte. L'historique relativement court des données de rachat restreint la capacité des modèles à apprendre de scénarios économiques variés. Un historique plus étendu permettrait d'améliorer la robustesse et la capacité de généralisation des modèles.

La modélisation de la volatilité à court terme représente un défi. Bien que les modèles ARIMAX capturent efficacement les tendances générales, ils peuvent parfois sous-estimer la volatilité mensuelle des taux de rachat. L'exploration de modèles plus sophistiqués, tels que les modèles GARCH ou les modèles à volatilité stochastique, pourrait améliorer la modélisation de cette volatilité.

La modélisation conjointe des différentes années d'ancienneté constitue une piste de recherche future. Une approche qui modéliserait simultanément les taux de rachat pour toutes les années d'ancienneté pourrait potentiellement capturer des interactions entre ces différentes cohortes et améliorer la cohérence globale des prévisions.

Ces limitations identifiées ouvrent la voie à de futures recherches et améliorations, soulignant le potentiel d'évolution de la modélisation des lois de rachat en assurance emprunteur, avec pour objectif d'affiner davantage la précision et la robustesse des prévisions dans ce domaine complexe et dynamique.

5.4 Conclusion sur la modélisation de la loi de rachat

Au terme de cette analyse, les modèles ARIMAX optimisés pour chaque année d'ancienneté se sont révélés être les plus performants pour modéliser la loi de rachat en assurance emprunteur. Cette approche individualisée démontre sa capacité à capturer efficacement les tendances et les dynamiques complexes propres à chaque stade de la vie du prêt, tout en offrant une bonne précision prédictive.

L'utilisation de modèles ARIMAX spécifiques pour chaque année d'ancienneté, plutôt qu'un modèle unique, permet de saisir les nuances et les particularités des comportements de rachat à différents moments du cycle de vie du prêt. Cette approche granulaire offre une compréhension plus fine et précise des dynamiques de rachat.

Bien que des modèles plus complexes, tels que des modèles hybrides, aient été envisagés, la performance des modèles ARIMAX souligne leur pertinence dans ce contexte spécifique. Cela ouvre néanmoins des perspectives pour de futures améliorations, notamment dans la modélisation des composantes non linéaires et des interactions entre les différentes années d'ancienneté.

L'analyse des résidus et les tests de validation confirment la qualité des modèles sélectionnés, démontrant leur capacité à capturer les tendances principales tout en minimisant les erreurs de prédiction. Cette robustesse permet de fournir des estimations fiables des taux de rachat futurs.

Avec la validation de ces modèles pour la loi de rachat, l'étude dispose maintenant d'un outil pour la prévision de cette composante essentielle de l'assurance emprunteur. Cette modélisation, combinée à celle du taux d'affaires nouvelles précédemment réalisée, offre une vision complète et intégrée des flux entrants et sortants dans un portefeuille d'assurance emprunteur.

Cette approche globale permet une gestion précise et proactive des risques associés à l'assurance emprunteur. Elle fournit aux assureurs des outils pour anticiper les comportements de rachat, ajuster leurs stratégies de tarification et de provisionnement, et optimiser la gestion de leur portefeuille à long terme.

En conclusion, cette modélisation de la loi de rachat constitue une avancée significative dans la compréhension et la prévision des dynamiques du portefeuille d'assurance emprunteur. Elle ouvre la voie à des pratiques de gestion des risques plus sophistiquées et mieux adaptées aux réalités complexes du marché de l'assurance emprunteur.

Le chapitre suivant explore l'impact de l'inflation sur un portefeuille fictif d'assurance emprunteur en calculant et analysant son Best Estimate (BE). Cette étude démontre l'application pratique des modèles développés précédemment pour anticiper et gérer les effets des variations macroéconomiques sur la valorisation et la performance financière des portefeuilles. L'analyse met en lumière l'interaction entre les dynamiques de marché modélisées et les exigences de reporting financier, contribuant à une gestion plus intégrée et prospective des risques dans le secteur.

Chapitre 6

Impact du contexte économique sur un portefeuille d'assurance emprunteur

Le chapitre final de ce mémoire évalue l'impact concret du contexte économique sur un portefeuille d'assurance emprunteur. Dans un environnement caractérisé par des fluctuations de l'inflation et des conditions de marché instables, il devient important pour les assureurs de comprendre comment ces variations affectent la performance et la valeur de leurs portefeuilles. En s'appuyant sur les modèles développés pour le taux d'affaires nouvelles et la loi de rachat, cette étude simule divers scénarios économiques pour prédire les taux correspondants. L'impact du contexte économique sur le marché de l'assurance est ainsi quantifié en utilisant le Best Estimate (BE) comme métrique principale.

Ce chapitre s'ouvre par la présentation d'un portefeuille fictif, construit à partir de données réalistes reflétant les spécificités du marché de l'assurance emprunteur. Cette section inclut une description détaillée des sources, des hypothèses utilisées et des caractéristiques du portefeuille. Ensuite, la méthodologie de projection est développée, en détaillant les scénarios économiques basés sur l'inflation, les hypothèses sous-jacentes (notamment les tables BCAC et les lois d'entrée, de décès et de maintien), ainsi que le fonctionnement du modèle de projection.

La deuxième partie du chapitre est consacrée au calcul des BE, en précisant leur définition dans le contexte de l'étude, les courbes de taux utilisées, et les méthodes spécifiques pour le calcul des BE de primes, de sinistres et globaux. Une analyse comparative des résultats met en évidence les différences entre les scénarios d'inflation et les interactions complexes entre les variables économiques et les comportements des assurés.

Enfin, ce chapitre conclut sur une synthèse des principaux résultats et sur les implications pratiques pour la gestion des risques et la stratégie des assureurs dans un contexte économique incertain.

6.1 Présentation du portefeuille fictif

Cette section détaille la construction et les caractéristiques du portefeuille fictif d'assurance emprunteur utilisé pour l'étude de l'impact du contexte économique. Ce portefeuille, conçu pour refléter fidèlement la réalité du marché français, se compose de 15 000 models points, offrant ainsi une large gamme de profils d'assurés. Le portefeuille est créé à la date du 31/12/2023, fournissant une base de référence précise pour les projections et analyses ultérieures.

6.1.1 Sources et méthodologie de construction du portefeuille

La création du portefeuille fictif s'appuie sur diverses sources et actuelles :

— Le rapport 2023 du Comité consultatif du secteur financier (CCSF) sur le bilan de l'assurance emprunteur (*CCSF* [6]) ;

— Les données de l'Observatoire français des drogues et des tendances addictives (OFDT) sur le tabagisme en France (*OFDT* [39]) ;

— Les statistiques de l'Observatoire Crédit Logement sur la durée des prêts immobiliers (*L'Observatoire Crédit Logement* [32]) ;

— Les données de la Banque de France sur les encours de crédit (*Banque de France* [3]) ;

— Le site Empruntis pour l'historique des taux de crédit immobilier (*Empruntis* [11]) ;

— Des benchmarks internes ;

— Des études de marché sur les primes d'assurance (*Simulassur* [52]).

6.1.2 Caractéristiques détaillées du portefeuille

Le portefeuille se caractérise par une grande diversité de profils, reflétant la variété des assurés sur le marché français de l'assurance emprunteur :

- **Âge de l'assuré** : Entre 25 et 65 ans durant toute la vie de l'emprunt, avec une répartition selon le *CCSF* : 23% de 25 à 30 ans, 36% de 30 à 40 ans, 28% de 40 à 50 ans, 13% de 50 à 65 ans ;
- **Sexe** : Répartition équilibrée (50% hommes, 50% femmes) ;
- **Statut tabagique** : Basé sur les données de l'*OFDT*, variant selon l'âge et le sexe ;
- **Catégorie socio-professionnelle (CSP)** : Répartie en trois catégories selon la *Figure 1.7* ;
- **Caractéristiques du prêt** :
  - **Date de début** : Remonte au maximum à 25 ans avant le 31/12/2023 ;
  - **Date d'entrée en portefeuille** : Entre la date de début de prêt et 31/12/2023 ;
  - **Durée** : Repartie selon les proportions données par l'*Observatoire Crédit Logement* : 10 ans (5.2%), 15 ans (10.3%), 20 ans (20.4%), 25 ans (64.1%) ;
  - **Montant** : Dépend de la durée du prêt et de la CSP de l'assuré, selon le *CCSF* ;
  - **Taux d'intérêt** : Dépend de la date de début et de la durée du prêt et est basé sur les données de la *Banque de France* et d'*Empruntis* ;
  - **Profil Lemoine** : Applicable si le début du prêt est plus récent que 01/06/2022 ou que l'entrée en portefeuille est plus récente que 01/09/2022 avec un montant du prêt inférieur à 200k€ par emprunteur et un âge de fin de prêt inférieur à 60 ans ;
  - **Prime d'assurance** : Calculée selon un benchmark interne et une *étude de marché*, dépend de la durée du prêt, l'âge de début de prêt, le statut tabagique, la CSP, le montant du prêt et est augmentée de 15 à 20% pour les profils Lemoine ;
  - **Mensualité** : Fixe et basée sur l'équation suivante :

$$montant.pret \cdot \frac{taux.annuel}{12} \cdot \frac{(1+\frac{taux.annuel}{12})^{dure.pret.en.mois}}{(1+\frac{taux.annuel}{12})^{duree.pret.en.mois}-1} \quad ;$$

  - **Capital Restant Dû (CRD)** : Calculé en fonction de l'ancienneté du prêt :

$$montant.pret \cdot \left(1 - \frac{(1+\frac{taux.annuel}{12})^{anciennete.pret.en.mois}-1}{(1+\frac{taux.annuel}{12})^{duree.pret.en.mois}-1}\right) \quad ;$$

- **Capital total remboursé** : Différence entre le CRD et le montant du prêt initial.

Cette structure détaillée du portefeuille permet une analyse de l'impact des variations économiques sur différents segments de l'assurance emprunteur. À la date du 31/12/2023, le portefeuille présente un CRD total de 2.683 milliards d'euros. Il est important de noter que tous les assurés sont considérés comme étant en état de bonne santé, plus précisément en "état valide" ou "état sain", ce qui signifie qu'aucun n'est en situation d'incapacité, d'invalidité ou de sinistre à la date de référence.

6.2 Méthodologie de projection

Cette section détaille la méthodologie employée pour projeter le portefeuille d'assurance emprunteur sur une période de cinq ans, du 31/12/2023 au 31/12/2028. L'approche s'appuie sur les modèles de taux de rachat et d'affaires nouvelles développés dans les *Chapitres 4 et 5*, tout en intégrant divers scénarios économiques et tables d'incidence.

La projection vise à simuler l'évolution du portefeuille en tenant compte des caractéristiques spécifiques de chaque contrat et des conditions macroéconomiques anticipées. Cette méthodologie permet d'évaluer la sensibilité du portefeuille aux variations économiques, offrant ainsi une vision prospective de sa performance et de sa valeur dans différents contextes.

Dans ce cadre, six états sont considérés pour chaque contrat du portefeuille :

— La validité ;
— La fin de prêt ;
— Le rachat ;
— L'incapacité ;
— L'invalidité ;
— Le décès.

Les mouvements possibles entre états sont récapitulés dans la *Figure 6.1* avec une différenciation colorée :

- Violet foncé : États en portefeuille ;
- Violet clair : États de sortie du portefeuille.

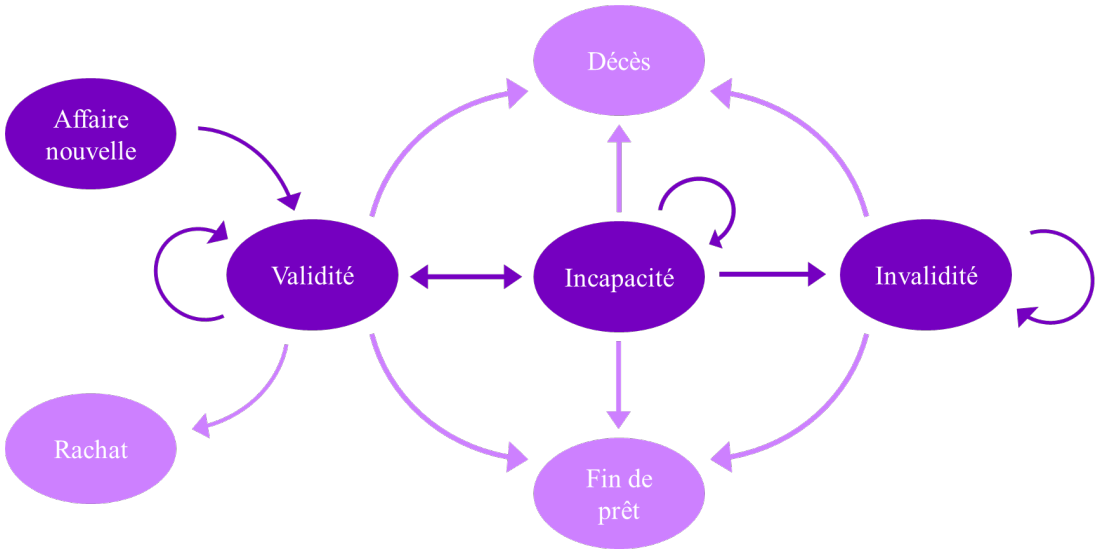


FIGURE 6.1 – Modèle à états

A noter qu'un individu valide peut soit être maintenu en validité, soit passer en incapacité soit mourir. Un individu en incapacité peut revenir en état de validité, être maintenu en état d'incapacité, passer en invalidité ou décéder. Enfin, un individu en invalidité peut uniquement être maintenu en état d'invalidité ou mourir.

Dans ce modèle, une hypothèse importante est adoptée : seul un client en état de validité peut racheter son prêt. Cette simplification permet de mieux refléter la réalité du marché, où les rachats sont généralement effectués par des emprunteurs en bonne santé et capables de renégocier leurs conditions de prêt.

Les sous-sections suivantes détaillent les tables d'incidence utilisées, les scénarios économiques basés sur l'inflation, la méthode de projection des variables économiques dont dépendent les lois de rachat et d'affaires nouvelles, et le fonctionnement du modèle de projection du portefeuille.

6.2.1 Tables d'incidence

Cette section se concentre sur les tables d'incidence utilisées pour projeter les models points :

— **Loi de rachat** : Déterminée dans le *Chapitre 5* et adaptée selon les scénarios d'inflation présentés dans la *Partie 6.2.3* sur la projection des variables économiques ;

— **Loi d'entrée** :
  - En incapacité : Benchmark interne ;
  - En invalidité : Table du BCAC de 2013 ;

— **Loi de décès** :
  - Pour les individus en validité : Tables TH02 et TF02 (*Source [54]*) ;
  - Pour les individus en incapacité : Table du BCAC de 2002 ;
  - Pour les individus en invalidité : Table du BCAC de 2013 ;

— **Loi de maintien** :
  - En validité : Déduite des autres lois ;
  - En incapacité et invalidité : Tables du BCAC de 2013 ;

— **Loi d'affaires nouvelles** : Déterminée dans le *Chapitre 4* et adaptée selon les scénarios d'inflation présentés dans la *Partie 6.2.3* sur la projection des variables économiques.

6.2.1.1 Les tables du BCAC

Trois tables réglementaires sont construites (maintien en invalidité, passage d'incapacité à invalidité et maintien en invalidité) par le Bureau Commun des Assurances Collectives (BCAC) en 1993 à partir de données des portefeuilles d'assurances collectives des principales compagnies françaises d'assurances de l'époque : AGF, AXA, GAN et UAP. Celles-ci sont homologuées et rendues obligatoires en 1996 pour tous les organismes assureurs. Ensuite, dans le cadre de la réforme des retraites de 2010, l'arrêté du 24 décembre 2010 a modifié les tables à utiliser pour le provisionnement à compter de 2011. Enfin, de nouvelles tables sont construites en 2013 mais ne sont pas homologuées.

Cependant, au lieu d'utiliser ces tables réglementaires, l'entreprise d'assurances peut construire ses propres tables d'expérience, basées sur l'historique de ses portefeuilles à condition qu'elles soient certifiées par un actuaire indépendant.

Pour cet exemple, les tables du BCAC datant de 2013 étant les plus récentes, ce sont elles qui sont détaillées et utilisées.

6.2.1.2 Lois d'entrée

Loi d'entrée en incapacité

La table d'entrée en incapacité est unisexe et se présente comme en *Table 6.1* :

Age	Tx annuel	Tx mensuel
25	1.00%	0.08%
26	1.00%	0.08%
27	1.00%	0.08%
28	1.00%	0.08%
29	1.00%	0.08%
30	1.00%	0.08%

TABLE 6.1 – Loi d'entrée en incapacité

Dans le cadre de ce mémoire, les taux d'entrée en incapacité sont constants et sont mensualisés pour pouvoir adapter les calculs aux données du mémoire qui possèdent une granularité mensuelle. Cela est réalisé grâce à la formule actuarielle suivante :

$$1 + i_{annuel} = (1 + i_{mensuel})^{12}$$
$$i_{mensuel} = (1 + i_{annuel})^{\frac{1}{12}} - 1$$

Avec i le taux d'entrée.

Loi de passage d'un état incapable à invalide

La loi de passage de l'état d'incapacité à l'état d'invalidité est présentée dans la *Table 6.2* ci-dessous :

Age \ Ancienneté	0	1	2	3	4
25	0.007%	0.012%	0.016%	0.019%	0.019%
26	0.007%	0.012%	0.016%	0.018%	0.019%
27	0.007%	0.012%	0.016%	0.019%	0.019%
28	0.008%	0.014%	0.020%	0.024%	0.026%
29	0.009%	0.016%	0.023%	0.029%	0.032%
30	0.010%	0.018%	0.027%	0.034%	0.039%

TABLE 6.2 – Table 2013 du BCAC : Loi de passage d'incapacité à invalidité (*BCAC* [5])

La *Table 6.2* présente les probabilités de transition d'incapacité à invalidité établies par le BCAC. Chaque ligne correspond à l'âge de l'individu au moment de l'entrée en incapacité, variant de 20 à 65 ans. Les colonnes représentent le nombre de mois écoulés depuis l'arrêt de travail, allant de 0 à 35 mois, indiquant l'ancienneté de l'individu dans l'état d'incapacité. Étant déjà mensualisée et en probabilité de passage, aucun traitement n'est fait dessus.

6.2.1.3 Lois de décès

Loi de décès des valides

Les tables de décès utilisées pour les valides sont les tables TH00-02 pour les hommes et TF00-02 pour les femmes. La *Table 6.3* représente la TH00-02 :

Age	L_x	Q_x	Q_x mensuel
25	98 406	0.10%	0.009%
26	98 303	0.11%	0.009%
27	98 198	0.11%	0.009%
28	98 091	0.11%	0.009%
29	97 982	0.11%	0.010%
30	97 870	0.12%	0.010%

TABLE 6.3 – Loi de décès TH00-02 (*Source [54]*)

Ici, L_x représente le nombre de personnes encore en vie pour l'âge x et Q_x la probabilité de décéder pour ce même âge. Q_x est calculé de la manière suivante : $Q_x = 1 - \frac{L_{x+1}}{L_x}$. Enfin, le Q_x mensuel est obtenu via le même calcul que la *mensualisation* des taux d'entrée en incapacité.

Loi de décès en incapacité

La loi de décès en incapacité est présentée dans la *Table 6.4*. Celle-ci est construite en 2002 :

Age \ Ancienneté	0	1	2	3	4
25	10000	9985	9974	9965	9957
26	10000	9985	9974	9965	9957
27	10000	9985	9974	9965	9957
28	10000	9985	9974	9965	9956
29	10000	9984	9973	9964	9955
30	10000	9984	9972	9963	9953

TABLE 6.4 – Table 2002 du BCAC : Loi de décès en incapacité (*BCAC* [5])

Chaque ligne correspond à l'âge de l'individu au moment de leur décès, sachant qu'ils étaient dans un état d'incapacité. Il est exprimé en années et varie de 25 ans à 65 ans dans la table établie par le BCAC. Les colonnes représentent le nombre de mois écoulés depuis l'entrée en incapacité et varient de 0 à 35 mois. Il s'agit de l'ancienneté d'un individu dans cet état d'incapacité au moment du décès.

Cette table en nombre permet de déduire les probabilités de décès des individus en incapacité par mois. Elle n'a pas besoin d'être mensualisée car elle l'est déjà. Ainsi, pour passer en probabilité plutôt qu'en nombre et appliquer les bons taux de décès en incapacité, la formule suivante est appliquée :

$$T_{x,N} = 1 - \frac{N_{x,N}}{N_{x,N-1}}$$

Où :

— $T_{x,N}$ le taux de décès en incapacité d'un individu d'âge x pour le mois N ;

— $N_{x,N}$ le nombre d'individu d'âge x en incapacité pour le mois N ;

— $N_{x,N-1}$ le nombre d'individu d'âge x en incapacité pour le mois $N-1$.

Loi de décès en invalidité

La seule sortie possible de l'état d'invalidité étant le décès, un individu qui n'est pas maintenu en invalidité est considéré comme décédé. Ainsi, la table de décès en invalidité est calculée comme suit : $P_{DC,x,N} = 1 - P_{maintien,x,N}$, avec :

— $P_{DC,x,N}$ la probabilité de décès en invalidité d'un individu d'âge x pour le mois N ;

— $P_{maintien,x,N}$ la probabilité de maintien en invalidité d'un individu d'âge x pour le mois N.

6.2.1.4 Lois de maintien

Loi de maintien en validité

Aucune loi de maintien en validité n'est retenue dans ce mémoire. En effet, le maintien en validité se fait si un valide n'est pas parti dans un état autre que celui de validité. Le maintien est donc défini par défaut par rapport aux autres lois impactant la population valide des portefeuilles.

Loi de maintien en incapacité

La loi de maintien en incapacité présentée dans la *Table 6.5*, est similaire à la loi de décès en incapacité, avec les âges d'entrée dans l'état en ligne et l'ancienneté de l'individu dans l'état sur chaque colonne.

Age \ Ancienneté	0	1	2	3	4
25	10 000	5 273	3 395	2 456	1 886
26	10 000	5 357	3 478	2 528	1 944
27	10 000	5 427	3 550	2 590	1 996
28	10 000	5 477	3 604	2 641	2 043
29	10 000	5 526	3 654	2 687	2 087
30	10 000	5 567	3 699	2 731	2 129

TABLE 6.5 – Table 2013 du BCAC : Loi de maintien en incapacité (*BCAC* [5])

Comme pour la loi de décès en incapacité, cette table est déjà mensualisée et permet de déduire les probabilités de maintien en incapacité par mois. La même formule est utilisée pour convertir les nombres en probabilités :

$$T_{x,N} = 1 - \frac{N_{x,N}}{N_{x,N-1}}$$

Où :

— $T_{x,N}$ le taux de maintien en incapacité d'un individu d'âge x pour le mois N ;

— $N_{x,N}$ le nombre d'individu d'âge x en incapacité pour le mois N ;

— $N_{x,N-1}$ le nombre d'individu d'âge x en incapacité pour le mois $N-1$.

Cette approche cohérente permet une interprétation et une application uniformes des lois de

décès et de maintien en incapacité dans le modèle de projection.

Loi de maintien en invalidité

La loi de maintien en invalidité est présentée dans la *Table 6.6* :

Age \ Ancienneté	0	1	2	3	4
25	10000	9557	9263	9045	8889
26	10000	9568	9281	9070	8922
27	10000	9581	9303	9100	8958
28	10000	9596	9328	9131	8994
29	10000	9613	9354	9163	9028
30	10000	9630	9379	9191	9056

TABLE 6.6 – Table 2013 du BCAC : Loi de maintien en invalidité (*BCAC* [5])

Chaque ligne correspond à l'âge de l'individu au moment de son entrée en invalidité. Il est exprimé en années et varie de 20 ans à 65 ans. Ici, les colonnes représentent le nombre d'années écoulées depuis l'entrée en invalidité et varie de 0 à 45 ans. Il s'agit de l'ancienneté d'un individu dans cet état d'invalidité.

Elle est exprimée en année et doit donc être mensualisée. La mensualisation est faite directement sur les effectifs de maintien en invalidité. Les probabilités de passage sont ensuite déduites de la table mensualisée.
La table en nombre est mensualisée de la façon suivante :

1. Calcul de la différence entre le nombre d'individus maintenus en invalidité en année N +1 et N ;
2. Déduction du montant mensuel d'individus à maintenir en divisant le nombre obtenu précédemment par 12 ;
3. Soustraction de ce montant au nombre d'individus du mois 0 de l'année N et ce jusqu'à l'année N + 1.

Un exemple de la mensualisation est présenté dans la *Table 6.7* :

Age \ Ancienneté	0	1	2	3	4	5	6	7	8	9	10	11	12
25	10 000	9 963.06	9 926.13	9 889.19	9 852.26	9 815.32	9 778.38	9 741.45	9 704.51	9 667.57	9 630.64	9 593.70	9 556.77
Impact mensualisation	36.94	36.94	36.94	36.94	36.94	36.94	36.94	36.94	36.94	36.94	36.94	36.94	
26	10 000	9 963.99	9 927.97	9 891.96	9 855.95	9 819.93	9 783.92	9 747.91	9 711.89	9 675.88	9 639.87	9 603.85	9 567.84
Impact mensualisation	36.01	36.01	36.01	36.01	36.01	36.01	36.01	36.01	36.01	36.01	36.01	36.01	

TABLE 6.7 – Exemple de mensualisation des maintiens en invalidité en nombre

Comme énoncé précédemment, les 36.94 sont obtenus via la différence entre le nombre d'invalides en année 0 et en année 1, divisée par 12 : $(10000 - 9556.77)\frac{1}{12} = 36.94$. Ce montant est ensuite retranché au mois 0 et ce, jusqu'au mois 11. Ce processus est répété pour chaque âge (de 25 à 65 ans) et pour chaque mois (de 0 à 480 mois).

Une fois la mensualisation faite, le taux de maintien en invalidité en probabilité est calculé de la même façon que pour la loi de maintien en incapacité. Soit :

$$T_{x,N} = 1 - \frac{N_{x,N}}{N_{x,N-1}}$$

Où :

— $T_{x,N}$ le taux de maintien en invalidité d'un individu d'âge x pour le mois N ;

— $N_{x,N}$ le nombre d'individu d'âge x en invalidité pour le mois N ;

— $N_{x,N-1}$ le nombre d'individu d'âge x en invalidité pour le mois $N-1$.

De cette façon, la *Table 6.8* de maintien en invalidité mensuelle est obtenue ci-après :

Age \ Ancienneté	0	1	2	3	4	5	6	7	8	9	10	11	12
25	100.00%	99.63%	99.63%	99.63%	99.63%	99.63%	99.62%	99.62%	99.62%	99.62%	99.62%	99.62%	99.61%
26	100.00%	99.64%	99.64%	99.64%	99.64%	99.63%	99.63%	99.63%	99.63%	99.63%	99.63%	99.63%	99.63%
27	100.00%	99.65%	99.65%	99.65%	99.65%	99.65%	99.64%	99.64%	99.64%	99.64%	99.64%	99.64%	99.64%
28	100.00%	99.66%	99.66%	99.66%	99.66%	99.66%	99.66%	99.66%	99.66%	99.65%	99.65%	99.65%	99.65%
29	100.00%	99.68%	99.68%	99.68%	99.67%	99.67%	99.67%	99.67%	99.67%	99.67%	99.67%	99.67%	99.67%
30	100.00%	99.69%	99.69%	99.69%	99.69%	99.69%	99.69%	99.69%	99.69%	99.68%	99.68%	99.68%	99.68%

TABLE 6.8 – Loi de maintien en invalidité

6.2.2 Explication des 7 scénarios économiques basés sur l'inflation

L'étude se concentre sur l'inflation comme variable économique clé afin de refléter son impact significatif sur le marché de l'assurance emprunteur, tel que démontré dans le *Chapitre 1*. L'inflation, mesurée par l'évolution annuelle de l'Indice des Prix à la Consommation (IPC), influence directement les taux d'intérêt, le coût du crédit, et par conséquent, la demande de crédit, la demande d'assurance emprunteur et les comportements de rachat.

Sept scénarios d'inflation sont élaborés pour la période 2023-2028 :

1. **Scénario standard** : Basé sur les prévisions de l'EIOPA [10] et de la BCE [9], qui anticipent une inflation à moyen terme de 2% pour la zone euro. Étant donné que la France fait partie de l'Union européenne, il est raisonnable de supposer que le marché français se calquera sur cette tendance ;
2. **Trois scénarios haussiers** : Ceux-ci sont envisagés, avec des augmentations croissantes par rapport au scénario standard. Ces scénarios ne dépassent pas 5% d'inflation afin de rester dans des limites réalistes. Cela permet d'évaluer les impacts d'une inflation modérée à élevée sur le portefeuille ;
3. **Trois scénarios baissiers** : Ceux-ci sont définis avec une diminution plus ou moins prononcée par rapport au scénario standard tout en restant positifs. Cette approche prend en compte l'historique économique récent, qui ne montre pas de périodes prolongées de désinflation depuis plusieurs décennies. Ainsi le scénario le plus extrême envisage une inflation nulle à long terme.

Ces scénarios partent d'une inflation de 4.9% au 31/12/2023 pour aboutir aux valeurs suivantes au 31/12/2028 :

1. **Scénario standard** :
   - 2.31% (scénario long terme à 2%).
2. **Scénarios à la baisse** :
   - 1.85% (scénario long terme à 1.3%) ;
   - 1.58% (scénario long terme à 0.65%) ;
   - 1.05% (scénario long terme à 0%).
3. **Scénarios à la hausse** :
   - 2.76% (scénario long terme à 3%) ;
   - 3.15% (scénario long terme à 4%) ;

- 3.37% (scénario long terme à 5%).

Cette approche permet d'évaluer l'impact des différentes trajectoires inflationnistes sur les lois de rachat et le taux d'affaires nouvelles, offrant ainsi une vision complète des risques et opportunités pour le marché de l'assurance emprunteur dans divers contextes économiques. En intégrant ces scénarios, il devient possible d'anticiper les tendances du marché et d'adapter les stratégies de tarification ainsi que la gestion des risques en conséquence.

Le *Graphique 6.2* ci-dessous illustre visuellement ces différents scénarios d'inflation, permettant une meilleure compréhension des trajectoires envisagées pour la période 2023-2028.

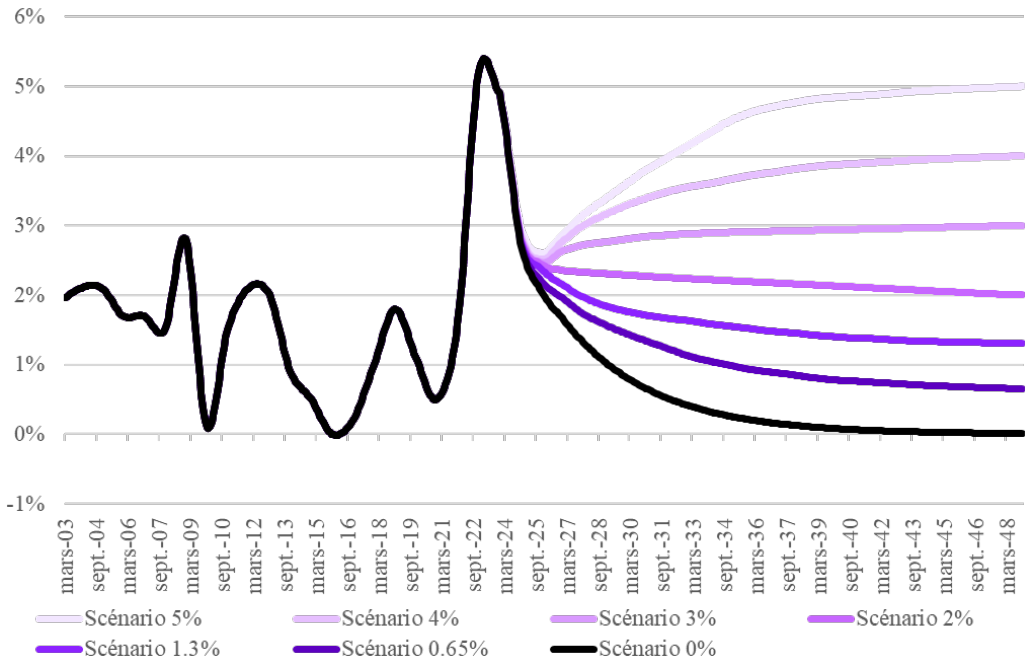


FIGURE 6.2 – Scénarios d'inflation

Cette représentation graphique, combinée à l'analyse détaillée des scénarios, fournit une base solide pour évaluer l'impact potentiel des variations de l'inflation sur le marché de l'assurance emprunteur. Elle sert de fondement pour les projections et analyses ultérieures dans cette étude.

6.2.3 Projection des variables économiques, des taux d'affaires nouvelles et des lois de rachat

Pour projeter le portefeuille jusqu'à fin 2028 selon les sept scénarios d'inflation établis, il est nécessaire de disposer de données économiques sur cette période future. Ces données n'existant pas encore, une modélisation est requise pour générer des scénarios économiques cohérents.

La méthodologie employée pour cette projection s'inspire directement de celle utilisée dans le *Chapitre 3* pour modéliser les taux d'affaires nouvelles et les taux de rachat. L'approche consiste à créer un générateur de scénarios économiques (GSE) basé sur les sept scénarios d'inflation précédemment définis. Le Schéma récapitulatif des étapes du processus de modélisation du GSE est disponible en *Annexe Figure F.1*

Pour chacune des variables économiques présentées dans le *Tableau 2.5*, une série temporelle est modélisée en fonction de l'inflation. Cette modélisation s'appuie sur la méthode SARIMAX, per-

mettant de capturer les relations complexes entre l'inflation et les autres indicateurs économiques.

La sélection du modèle SARIMAX le plus approprié pour chaque variable suit un processus identique à celui détaillé dans le *Chapitre 3*. Les paramètres p, d, q, P, D, Q sont déterminés pour chaque variable, optimisant ainsi la précision des projections. Un tableau récapitulatif des modèles SARIMAX retenus pour chaque variable, ainsi que les métriques ayant guidé ces choix, est disponible en *Annexe Tableau F.1*.

Cette approche permet de générer des projections économiques du 01/01/2024 au 31/12/2028, fournissant ainsi les données nécessaires pour modéliser les taux d'affaires nouvelles et les taux de rachat sur cette période future.

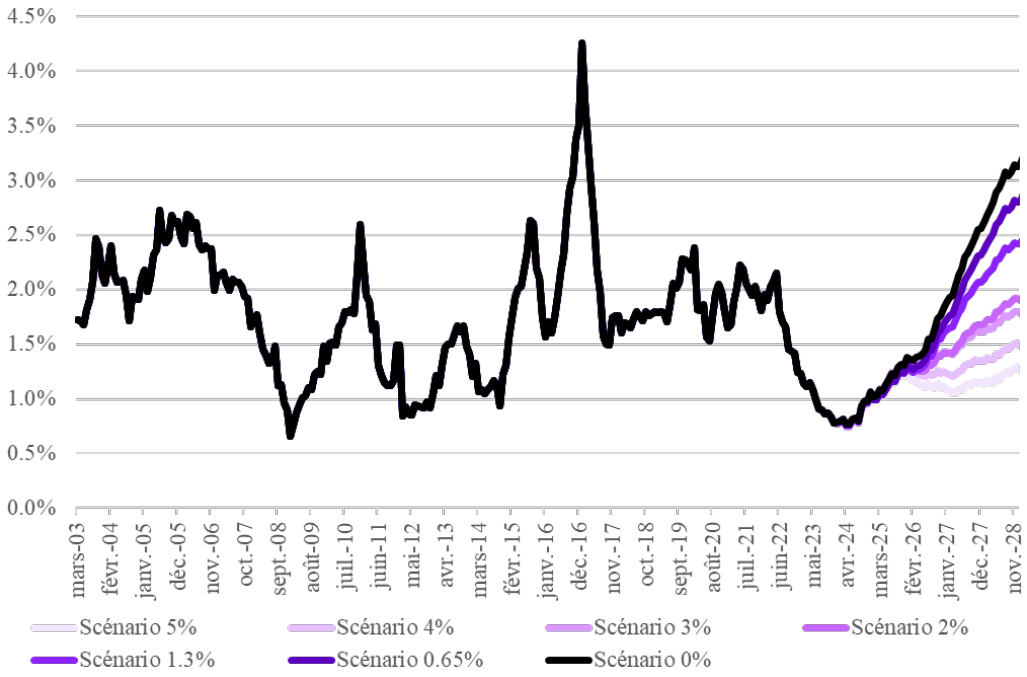


FIGURE 6.3 – Projection des taux d'affaires nouvelles jusqu'en 2028

La *Figure 6.3* illustre ces projections économiques, offrant une visualisation claire des tendances anticipées pour chaque variable selon les différents scénarios d'inflation.

Cette figure confirme la tendance négative observée précédemment entre les taux d'affaires nouvelles et l'inflation. Il est possible de clairement distinguer que les scénarios avec une inflation plus élevée correspondent à des taux d'affaires nouvelles plus faibles, et inversement.

Cependant, une observation intéressante émerge de ces projections : malgré les variations liées aux différents scénarios d'inflation, une tendance générale à la hausse des taux d'affaires nouvelles est visible. Cette tendance contraste avec le déclin observé depuis 2022, suggérant une possible inversion de la dynamique du marché.

Cette évolution souligne que l'inflation, bien qu'étant un facteur important, n'est pas le seul déterminant de la croissance ou de la décroissance des taux d'affaires nouvelles. L'utilisation de l'inflation comme variable influençant le marché, qui à son tour impacte les taux d'affaires nouvelles, met en lumière la complexité des interactions économiques en jeu.

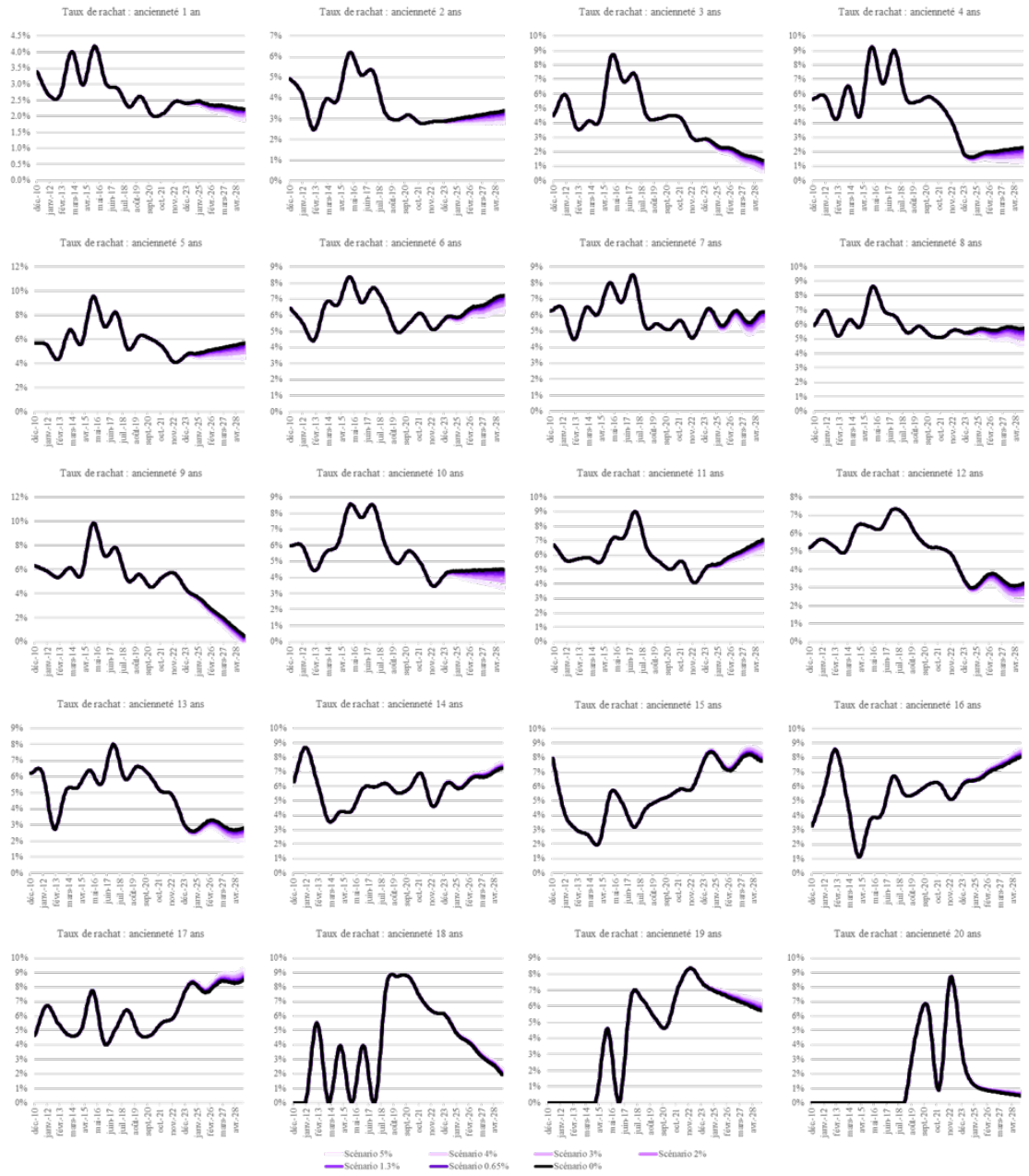


FIGURE 6.4 – Projection des taux de rachat jusqu'en 2028

La *Figure 6.4* illustre les projections des taux de rachat jusqu'en 2028 en fonction de l'ancienneté du prêt, révélant des impacts variés de l'inflation selon la durée écoulée depuis le début du prêt.

Pour les prêts récents (ancienneté faible), une relation inverse entre l'inflation et les taux de rachat est observable. Lorsque l'inflation diminue, les taux d'intérêt ont tendance à baisser également, comme observé dans le *Chapitre 1*. Cette baisse des taux d'intérêt entraîne une augmentation des taux de rachat. L'ampleur de cet effet varie selon les années d'ancienneté du prêt, mais reste cohérente pour les prêts les plus récents.

À l'inverse, pour les prêts plus anciens (ancienneté élevée), la relation s'inverse. Une baisse de l'inflation et des taux d'intérêt correspond à une diminution des taux de rachat. Cette inversion de tendance pourrait s'expliquer par des facteurs tels que la maturité du prêt, les conditions initiales du contrat, ou les comportements différents des emprunteurs à long terme.

6.2.4 Fonctionnement du modèle de projection du portefeuille

Le modèle de projection du portefeuille permet de suivre l'évolution des assurés en fonction des transitions entre différents états et des lois économiques applicables. La projection couvre la période du 31/01/2024 au 31/12/2028 et se déroule de la manière suivante :

1. Vieillissement mensuel du portefeuille initial : les 15 000 model points sont vieillis, et leur état est actualisé en fonction des tables de transition (valide, incapable, invalide, décédé ou fin de prêt) ;
2. Application des lois économiques : après vieillissement, des ajustements spécifiques sont appliqués :
   - Les assurés valides sont soumis à la loi de rachat selon les probabilités définies par le scénario économique considéré ;
   - De nouveaux assurés sont intégrés au portefeuille chaque mois, en fonction des taux d'affaires nouvelles projetés ;
3. Vieillissement des nouveaux assurés : les nouveaux entrants suivent ensuite le même processus de vieillissement et de transition que les assurés du portefeuille initial.

Pour appliquer les lois, le modèle utilise la fonction `runif` dans R qui génère un nombre aléatoire entre 0 et 1. Si la valeur générée est inférieure à la probabilité de la loi étudiée au moment T, alors le model point passe dans l'état en question. Les transitions possibles entre états sont illustrées dans le *Schéma 6.1.*

Le code R est optimisé grâce à la vectorisatoin et l'utilisation de `data.table`, qui permet des opérations rapides et efficaces sur les ensembles de données volumineux, tout en offrant une syntaxe concise et une gestion efficace de la mémoire.

Les affaires nouvelles sont ajoutées mensuellement après le traitement de tous les models points existants, évitant ainsi que leur situation ne change immédiatement après leur entrée en portefeuille. Ces nouvelles affaires sont basées sur un profil de client "moyen" du portefeuille au 31/12/2023, dont les caractéristiques sont détaillées dans le *Tableau 6.9.* Cette approche permet de maintenir la physionomie générale du portefeuille. Le mois suivant leur entrée, ces affaires nouvelles sont intégrées au stock classique.

Sexe	Fumeur	Lemoine	CSP	Age	Date en début de prêt	Ancienneté en mois / année	Durée du prêt en mois / année	Montant du prêt	Taux d'intérêt	Mensualité	Prime	Capital total remboursé	CRD
M	Non	Non	CSP2	36.8	30/11/2017	72 / 6	240 / 20	236 000 €	1.63%	-1 153.39 €	35.16 €	-62 898.82 €	173 101.18 €

TABLE 6.9 – Client "moyen" du portefeuille

6.2.5 Résultats de la projection du portefeuille

Le *Tableau 6.10* ci-dessous présente l'évolution des effectifs du portefeuille pour chaque année de projection, dans le cadre d'un scénario d'inflation à 2%. Il détaille les transitions entre différents états (en cours, décès, incapacité, invalidité, etc...) ainsi que les affaires nouvelles. Les totaux sont également indiqués pour chaque année.

	Effectif		2023	2024	2025	2026	2027	2028
Projection d'inflation 2%	**Stock**	**En cours**	15 000	14 266	13 669	13 108	12 516	11 811
		Décès	-	30	25	32	35	17
		Incapacité	-	39	40	44	43	39
		Décès Incap	-	-	-	3	1	2
		Invalidité	-	1	2	3	5	10
		Décès Inval	-	-	-	-	-	-
		Racheté	-	487	525	522	551	537
		Terminé	-	177	162	153	175	342
		Total	**15 000**	**15 000**	**14 423**	**13 865**	**13 326**	**12 758**
	Affaires nouvelles annuelles	**En cours**	-	116	151	166	185	219
		Décès	-	-	-	1	-	-
		Incapacité	-	-	-	1	-	-
		Décès Incap	-	-	-	-	-	-
		Invalidité	-	-	-	-	-	-
		Décès Inval	-	-	-	-	-	-
		Racheté	-	1	3	3	9	8
		Terminé	-	-	-	-	-	-
		Total	**-**	**117**	**154**	**171**	**194**	**227**
	Total	**Total**	**15 000**	**15 117**	**14 577**	**14 036**	**13 520**	**12 985**

TABLE 6.10 – Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 2%

Les tableaux correspondant aux autres scénarios d'inflation sont disponibles en *Annexes F.2 à F.7* pour permettre une comparaison complète des projections selon les hypothèses économiques retenues.

Le *Tableau 6.11* ci-dessous illustre l'impact de différents scénarios d'inflation sur un portefeuille d'assurances. Pour rappel, le scénario d'inflation à 2%, prévu par l'EIOPA, sert de référence, tandis que les autres scénarios (0%, 0.65%, 1.3%, 3%, 4%, et 5%) représentent des chocs par rapport à cette prévision. Il faut donc en retenir :

1. Une diminution générale de l'effectif :
   - Dans tous les scénarios d'inflation, l'effectif total du portefeuille diminue au fil des années, passant d'environ 15 000 en 2023 à environ 12 100-12 300 en 2028, selon le scénario ;
   - Cette baisse est constante, quelle que soit l'inflation, bien que son ampleur varie légèrement.
2. Des affaires nouvelles fortement impactées :
   - L'inflation a un effet direct sur les affaires nouvelles : plus elle est élevée, moins il y a d'entrées dans le portefeuille. Cela reflète probablement une diminution de l'attractivité des produits pour les nouveaux assurés dans un contexte inflationniste ;
   - À l'inverse, dans les scénarios de baisse d'inflation (comme à 0%), les affaires nouvelles augmentent en proportion et atteignent des niveaux plus élevés (jusqu'à 2.95% en 2028), ce qui contribue à stabiliser davantage le portefeuille.
3. Des rachats en diminution relative :
   - Les rachats diminuent légèrement à mesure que l'inflation augmente, mais cette diminution est moins marquée que celle des affaires nouvelles ;
   - Par exemple, dans le scénario à 2%, les rachats représentent environ 4.28% de l'effectif total en 2028, tandis qu'ils ne représentent que 3.79% dans le scénario à 5%. Cette différence reste modérée ;

- Dans les scénarios de faible inflation (comme à 0%), les rachats restent stables ou légèrement plus élevés (jusqu'à 4.57% en 2028), mais leur impact est compensé par une augmentation significative des affaires nouvelles.

4. Un effet combiné inflation-affaires nouvelles-rachat dont découlent 3 tendances :
   - Pour les scénarios d'inflation faibles (0% à 1.3%) : Ces scénarios entraînent une diminution plus rapide de l'effectif total au cours des premières années, mais conduisent à un meilleur maintien des assurés en fin de période (2028) par rapport au scénario standard à 2%. Cette dynamique s'explique par une augmentation des taux de rachat, dont l'impact domine celui des affaires nouvelles dans les premières années, avant de s'inverser en dernière année de projection ;
   - Pour les scénarios d'inflation à la hausse (3%, 4% et 5%) : Les effets sont variés selon le niveau d'inflation :
     - Pour un scénario à 3%, l'effectif fluctue autour de celui du scénario standard. Cela est dû à un équilibre entre des affaires nouvelles modérées et des taux de rachat légèrement plus élevés, compensant les différences observées en fin de projection ;
     - Pour un scénario à 4%, l'effectif total est constamment inférieur à celui du scénario standard. Cette tendance s'explique par une augmentation marquée des taux de rachat, combinée à un volume limité d'affaires nouvelles, qui ne permet pas de compenser les sorties du portefeuille ;
     - Pour un scénario extrême à 5%, l'effectif final est supérieur à celui du scénario standard. Contrairement aux scénarios d'inflation faibles, cette tendance s'explique par une diminution des taux de rachat dont l'impact est plus marqué que la baisse des affaires nouvelles, particulièrement en fin de période.

Ce *Tableau 6.11* met en évidence que l'inflation impacte davantage les entrées (affaires nouvelles) que les sorties (rachats), ce qui influence la dynamique de l'effectif total du portefeuille. Bien que l'effectif diminue dans tous les scénarios d'inflation, cette diminution est moins prononcée dans les scénarios de faible inflation. En effet, dans ces derniers, la hausse des affaires nouvelles contribue à ralentir la décroissance du portefeuille, principalement grâce à un meilleur maintien des nouvelles entrées. À l'inverse, lorsque l'inflation augmente, l'effectif total du portefeuille diminue progressivement. Il est important de noter que l'inflation basse ne stabilise pas ou ne renforce pas le portefeuille en termes absolus, mais elle permet de modérer sa décroissance, malgré un niveau stable ou légèrement accru des rachats. Ces observations soulignent l'importance d'une gestion proactive face aux variations économiques pour maintenir la viabilité du portefeuille d'assurances.

	Effectif	2023	2024	2025	2026	2027	2028
Projection d'inflation à 0%	**Portefeuille**	15 000	14 411	13 890	13 292	12 829	12 287
	Dont Affaires nouvelles	-	118	157	196	282	363
		0.00%	0.82%	1.13%	1.47%	2.20%	2.95%
	Rachats	-	531	524	632	575	589
		0.00%	3.55%	3.63%	4.54%	4.29%	4.57%
	Rachats d'Affaires Nouvelles	-	1	4	5	9	15
		0.00%	0.01%	0.03%	0.04%	0.07%	0.12%
Projection d'inflation à 0.65%	**Portefeuille**	15 000	14 414	13 895	13 380	12 840	12 264
	Dont Affaires nouvelles	-	115	152	184	255	326
		0.00%	0.80%	1.09%	1.38%	1.99%	2.66%
	Rachats	-	526	511	539	614	575
		0.00%	3.52%	3.55%	3.87%	4.56%	4.47%
	Rachats d'Affaires Nouvelles	-	3	2	3	8	11
		0.00%	0.02%	0.01%	0.02%	0.06%	0.09%
Projection d'inflation à 1.3%	**Portefeuille**	15 000	14 389	13 903	13 374	12 914	12 280
	Dont Affaires nouvelles	-	117	150	175	231	278
		0.00%	0.81%	1.08%	1.31%	1.79%	2.26%
	Rachats	-	557	474	545	504	600
		0.00%	3.73%	3.30%	3.91%	3.75%	4.65%
	Rachats d'Affaires Nouvelles	-	-	4	6	8	15
		0.00%	0.00%	0.03%	0.04%	0.06%	0.12%
Projection d'inflation à 2%	**Portefeuille**	15 000	14 451	13 912	13 402	12 852	12 184
	Dont Affaires nouvelles	-	116	151	168	185	219
		0.00%	0.80%	1.09%	1.25%	1.44%	1.80%
	Rachats	-	488	528	525	560	545
		0.00%	3.27%	3.66%	3.77%	4.17%	4.28%
	Rachats d'Affaires Nouvelles	-	1	3	3	9	8
		0.00%	0.01%	0.02%	0.02%	0.07%	0.06%
Projection d'inflation à 3%	**Portefeuille**	15 000	14 391	13 930	13 389	12 859	12 217
	Dont Affaires nouvelles	-	113	156	169	180	206
		0.00%	0.79%	1.12%	1.26%	1.40%	1.69%
	Rachats	-	546	458	552	535	524
		0.00%	3.65%	3.18%	3.96%	3.99%	4.11%
	Rachats d'Affaires Nouvelles	-	4	6	5	12	9
		0.00%	0.03%	0.04%	0.04%	0.09%	0.07%
Projection d'inflation à 4%	**Portefeuille**	15 000	14 409	13 911	13 352	12 848	12 174
	Dont Affaires nouvelles	-	113	154	154	157	168
		0.00%	0.78%	1.11%	1.15%	1.22%	1.38%
	Rachats	-	526	494	553	492	530
		0.00%	3.52%	3.43%	3.97%	3.69%	4.17%
	Rachats d'Affaires Nouvelles	-	3	2	7	5	5
		0.00%	0.02%	0.01%	0.05%	0.04%	0.04%
Projection d'inflation à 5%	**Portefeuille**	15 000	14 440	13 929	13 435	12 879	12 220
	Dont Affaires nouvelles	-	113	149	145	138	145
		0.00%	0.78%	1.07%	1.08%	1.07%	1.19%
	Rachats	-	502	492	484	517	481
		0.00%	3.36%	3.41%	3.48%	3.86%	3.79%
	Rachats d'Affaires Nouvelles	-	3	3	3	1	2
		0.00%	0.02%	0.02%	0.02%	0.01%	0.02%

TABLE 6.11 – Focus sur le rachat et les affaires nouvelles des tableaux d'effectif selon les années et scénarios d'inflation

6.3 Calcul du Best Estimate (BE)

Cette section aborde le calcul du Best Estimate (BE), un élément clé dans l'évaluation des engagements d'assurance. Le BE est calculé pour chaque année et pour les différents scénarios d'inflation afin de refléter les variations possibles des hypothèses économiques. En tenant compte des dynamiques du portefeuille (affaires nouvelles, rachats, et transitions d'états), ce calcul permet de projeter les flux futurs et d'évaluer leur impact sur la valeur des engagements.

Les résultats obtenus offrent une vision claire de l'évolution du portefeuille sous différents chocs d'inflation, avec comme scénario central celui prévu par l'EIOPA à 2%. Ces analyses permettent non seulement de quantifier les impacts financiers, mais également de dégager des conclusions stratégiques pour la gestion du portefeuille face aux incertitudes économiques.

6.3.1 Définition du BE dans le contexte de l'étude

Dans le cadre de cette étude, le BE représente la meilleure estimation de la valeur actuelle des flux de trésorerie futurs liés aux engagements d'assurance du portefeuille analysé. Cette évaluation prend en compte les spécificités du produit d'assurance emprunteur et les différents scénarios d'inflation envisagés.

Le BE se décompose en deux composantes principales :

1. **BE de primes** : Il correspond à la valeur actualisée des flux futurs liés aux contrats en cours, incluant :
   - Les primes futures attendues ;
   - Les commissions futures à verser ;
   - Les frais de gestion générale ;
   - Les rachats anticipés.

   Cette composante reflète l'engagement net de l'assureur lié à la perception des primes et aux dépenses associées sur la durée restante des contrats ;
2. **BE de sinistres** : Il représente la valeur actualisée des paiements futurs attendus en cas de sinistres, incluant :
   - Le Capital Restant Dû (CRD) à verser en cas de décès des assurés ;
   - Les prestations liées à l'Incapacité Temporaire Totale (ITT) ;
   - Les prestations liées à l'Invalidité Permanente Totale (IPT) ;
   - Les prestations en cas de décès suite à une incapacité ou une invalidité ;
   - Les frais de gestion des sinistres anticipés.

Il est important de noter que, dans cette étude, le BE de rachat n'est pas calculé séparément, car son impact financier est déjà intégré dans le BE de primes. En effet, les rachats entraînent une cessation du paiement des primes sans générer d'obligation supplémentaire pour l'assureur, contrairement aux décès ou à l'arrêt de travail.

6.3.2 Méthode de calcul pour chaque année et scénario

Cette sous-section détaille la méthodologie employée pour calculer le Best Estimate (BE) pour chaque année de projection et pour les différents scénarios d'inflation envisagés. Le calcul du BE s'appuie sur trois éléments fondamentaux :

1. Les courbes des taux EIOPA, qui servent de base pour l'actualisation des flux futurs ;
2. Le calcul du BE de primes, qui reflète les engagements liés aux primes futures ;

3. Le calcul du BE de sinistres, qui englobe les engagements relatifs aux prestations futures.

Ces calculs prennent en compte plusieurs facteurs clés :

— Les projections d'effectifs du portefeuille, incluant les transitions entre les différents états (en cours, incapacité, invalidité) ;

— Les taux de mortalité, de rachat, d'entrée et maintien en incapacité et en invalidité ;

— Les différents scénarios d'inflation et leurs impacts sur les flux financiers ;

— La structure temporelle des courbes des taux EIOPA pour l'actualisation des flux futurs ;

— Les frais de gestion, de gestion des sinistres et les commissions.

Il est important de noter que si un assuré se trouve en état d'incapacité, cet individu ne peut pas revenir à un état valide dans le futur lors du calcul du BE. Cette hypothèse vise à simplifier le processus de calcul du BE et à le rendre le plus économique possible.

Cette approche permet d'obtenir une évaluation réaliste et prudente des engagements de l'assureur, tout en reflétant la sensibilité du portefeuille aux variations des conditions économiques, notamment l'inflation. Elle offre également une vision dynamique de l'évolution du portefeuille sous différents chocs d'inflation, permettant ainsi d'anticiper les impacts financiers potentiels et d'ajuster les stratégies de gestion en conséquence.

6.3.2.1 Courbes des taux EIOPA

Le modèle développé en R permet d'adapter les courbes des taux EIOPA annuelles du 31/12/2023 pour les besoins spécifiques de l'évaluation du portefeuille d'assurances. Cette adaptation se fait en trois étapes principales : la mensualisation des courbes des taux, leur passage en courbes des taux forwards et le calcul des déflateurs. Ces transformations sont essentielles pour actualiser avec précision les différents flux associés au portefeuille d'assurés, en tenant compte de la temporalité mensuelle des projections et de l'évolution future des taux. Cette approche garantit une évaluation fine et réaliste des engagements financiers sur l'horizon de projection considéré.

Mensualisation des courbes des taux

Le modèle d'adaptation des courbes des taux importe d'abord une courbe des taux courante, puis la mensualise. La mensualisation de la courbe des taux annuelle se fait via la formule actuarielle suivante :

$$t_{i,0} = 1 + (t_{\lfloor \frac{i}{12} \rfloor + 1} - t_{\lfloor \frac{i}{12} \rfloor})^{\frac{i\%12}{12}} - 1 + t_{\lfloor \frac{i}{12} \rfloor}$$

Où :

— $t_{i,0}$, le taux mensuel à i mois en année 0 ;

— $t_{\lfloor \frac{i}{12} \rfloor + 1}$, le taux à $\lfloor \frac{i}{12} \rfloor + 1$ ans en année 0, $\lfloor \frac{i}{12} \rfloor + 1$ représentant le quotient de i par 12 ;

— $t_{\lfloor \frac{i}{12} \rfloor}$, le taux à $\lfloor \frac{i}{12} \rfloor$ ans en année 0, $\lfloor \frac{i}{12} \rfloor$ représentant le quotient de i par 12 ;

— $i\%12$, le reste de la division euclidienne de i par 12.

Un exemple est donné en *Tableau 6.12* :

Mois	0	1	...	11	12
Taux	$t_{0,0} = t_0$	$t_{1,0} = 1 + (t_{12,0} - t_{0,0})^{\frac{1}{12}} - 1 + t_{0,0}$	...	$t_{11,0} = 1 + (t_{12,0} - t_{0,0})^{\frac{11}{12}} - 1 + t_{0,0}$	$t_{12,0} = t_1$

TABLE 6.12 – Mensualisation de la courbe des taux EIOPA annuelle

Ici, $t_{i,0}$ représente le taux mensuel à i mois pour l'année de référence 0, t_0 représente le taux annuel en 0 et t_1 le taux annuel à 1 an. Le modèle adapte donc la courbe des taux annuelle en une unique courbe des taux mensuelle.

Passage en courbe des taux forwards

Une fois la mensualisation faite, il convient de "forwarder" les taux mensuels obtenus sur la durée de l'étude, ce qui permet d'avoir une mensualisation aussi bien sur les lignes que sur les colonnes. Pour se faire, la formule suivante est utilisée :

$$t_{i,j} = \begin{cases} (\frac{(1+t_{j,0})^j}{(1+t_{i,0})^i})^{\frac{1}{j-i}} - 1 & \text{si } i \leq n \\ 0 & \text{sinon} \end{cases}$$

Avec le *Tableau 6.13* associé :

Mois \ Passage dans le temps (Mois)	0	1	2	3	4
0	0	0	0	0	0
1	$t_{1,0}$	0	0	0	0
2	$t_{2,0}$	$t_{2,1}$	0	0	0
3	$t_{3,0}$	$t_{3,1}$	$t_{3,2}$	0	0
4	$t_{4,0}$	$t_{4,1}$	$t_{4,2}$	$t_{4,3}$	0
5	$t_{5,0}$	$t_{5,1}$	$t_{5,2}$	$t_{5,3}$	$t_{5,4}$

TABLE 6.13 – Table des taux "forwardés"

Comme précédemment, $t_{i,j}$ représente le taux mensuel à i mois en mois j.

Calculs des déflateurs

Suite au calcul des taux forwards, le modèle va calculer les déflateurs associés à chaque courbe des taux "forwardée". Les déflateurs sont calculés comme montré sur le *Tableau 6.14* :

Mois \ Passage dans le temps (Mois)	0	1	2	3	4
0	0	0	0	0	0
1	$(\frac{1}{1+t_{1,0}})^{\frac{1}{12}}$	0	0	0	0
2	$(\frac{1}{1+t_{2,0}})^{\frac{2}{12}}$	$(\frac{1}{1+t_{2,1}})^{\frac{1}{12}}$	0	0	0
3	$(\frac{1}{1+t_{3,0}})^{\frac{3}{12}}$	$(\frac{1}{1+t_{3,1}})^{\frac{2}{12}}$	$(\frac{1}{1+t_{3,2}})^{\frac{1}{12}}$	0	0
4	$(\frac{1}{1+t_{4,0}})^{\frac{4}{12}}$	$(\frac{1}{1+t_{4,1}})^{\frac{3}{12}}$	$(\frac{1}{1+t_{4,2}})^{\frac{2}{12}}$	$(\frac{1}{1+t_{4,3}})^{\frac{1}{12}}$	0
5	$(\frac{1}{1+t_{5,0}})^{\frac{5}{12}}$	$(\frac{1}{1+t_{5,1}})^{\frac{4}{12}}$	$(\frac{1}{1+t_{5,2}})^{\frac{3}{12}}$	$(\frac{1}{1+t_{5,3}})^{\frac{2}{12}}$	$(\frac{1}{1+t_{5,4}})^{\frac{1}{12}}$

TABLE 6.14 – Table des déflateurs

La formule généralisée pour calculer les déflateurs et qui est utilisée dans le *Tableau 6.14* est la suivante :

$$def_{i,j} = (\frac{1}{1+t_{i,j}})^{\frac{j-i}{12}}$$

Ici, i représente les colonnes tandis que j représente les lignes.

Le modèle des courbes des taux donne donc en sortie une courbe des taux mensualisée et "forwardée" de la courbe des taux courant, ainsi que les déflateurs associés.

6.3.2.2 Calcul du BE de primes

Dans le cadre de ce portefeuille d'assurance emprunteur, le calcul du Best Estimate (BE) de primes est effectué individuellement et mensuellement pour chaque assuré aux 5 dates de projections (du 31/12/2024 au 31/12/2028) et il se poursuit jusqu'à la fin de chaque contrat T. Le montant de la prime est ajusté en fonction de l'état de l'assuré au mois de projection t, ce qui influence directement les flux de trésorerie anticipés.

Le BE de primes du portefeuille correspond à la somme des BE de primes individuels :

$$BE_{prime} = \sum_{i=1}^{NBassurs} BE_{prime,i}$$

3 calculs de BE de primes individuels sont donc calculés en fonction de l'état de l'assuré :

1. Si l'assuré est valide :

$$BE_{prime,i} = \sum_{t=1}^{T} (P_{t,i} \cdot (1 - Tx_{deces.valide,t,i}) \cdot (1 - Tx_{rachat,t,i}) - C_{t,i} - F_{t,i}) \cdot def_{t,i}$$

2. Si l'assuré est incapable :

$$BE_{prime,i} = \sum_{t=1}^{T} (P_{t,i} \cdot (1 - Tx_{deces.incap,t,i}) - C_{t,i} - F_{t,i}) \cdot def_{t,i}$$

3. Si l'assuré est invalide :

$$BE_{prime,i} = \sum_{t=1}^{T} (P_{t,i} \cdot (1 - Tx_{deces.inval,t,i}) - C_{t,i} - F_{t,i}) \cdot def_{t,i}$$

Où :

— $P_{t,i}$ sont les primes de l'individu i au mois de projection t ;
— $Tx_{deces,t,i}$ sont les taux de décès de l'individu i au mois de projection t en fonction de son état ;
— $C_{t,i}$ sont les commissions anticipées de l'individu i au mois de projection t, qui s'élèvent à 10% des primes anticipées ;
— $F_{t,i}$ sont les frais de gestions anticipées de l'individu i au mois de projection t, qui s'élèvent à 3% des primes anticipées ;
— $def_{t,i}$ sont les déflateurs de l'individu i au mois de projection t calculés *précédémment*. Cela permet d'obtenir une valeur actuelle des flux futurs, tenant compte du temps et des taux d'intérêt.

Il est important de noter que le taux de rachat modélisé dans la *Partie 6.2.3* dépend du temps et n'est modélisé que jusqu'au 31/12/2028. A partir de cette date, une hypothèse est faite pour considérer que le taux de rachat devient constant (flat), car le modèle ne peut pas prédire avec précision les taux à un horizon aussi éloigné en raison des limitations des données historiques disponibles.
Ce processus de calcul du BE de primes permet d'obtenir une estimation précise et dynamique des primes futures, en tenant compte des différents états possibles des assurés et des coûts associés à leur gestion.

6.3.2.3 Calcul du BE de sinistres

Dans le cadre de ce portefeuille d'assurance emprunteur, les prestations en cas de sinistre sont structurées comme suit :

1. **Incapacité Temporaire Totale (ITT)** : L'assureur prend en charge le remboursement des mensualités du prêt à partir du 91e jour d'incapacité. Cette période de 90 jours est appelée "délai de franchise" ;
2. **Invalidité Permanente Totale (IPT)** : De manière similaire à l'ITT, l'assureur assume le paiement des mensualités du prêt à compter du 91e jour d'invalidité ;
3. **Décès** : En cas de décès de l'assuré, l'assureur verse le Capital Restant Dû (CRD) du prêt.

Il est important de noter que, pour des raisons de simplification du modèle, aucun délai de traitement administratif n'est pris en compte dans la gestion des sinistres. Tous les événements sont traités instantanément dans le cadre de cette modélisation. Également, le modèle ne distingue pas les situations où un assuré entre et sort d'un état d'ITT pour y retourner dans la même année. Chaque nouvel épisode d'incapacité est traité indépendamment, et le délai de franchise est systématiquement appliqué dans tous les cas. Cette hypothèse simplificatrice permet une gestion uniforme des événements dans le cadre de cette modélisation.

A l'instar du calcul de BE de primes, le calcul du BE des sinistres est effectué individuellement et mensuellement pour chaque assuré aux 5 dates de projections (du 31/12/2024 au 31/12/2028) et se poursuit jusqu'à la fin de chaque contrat T.
Le BE des sinistres du portefeuille correspond à la somme des BE des sinistres individuels :

$$BE_{sinistre} = \sum_{i=1}^{NBassurs} BE_{sinistre,i}$$

3 calculs de BE de sinistres individuels sont donc calculés en fonction de l'état de l'assuré :

1. Si l'assuré est valide :

$$\begin{aligned}
BE_{sinistre,i} &= Incap_i + Inval_i + DC_{valide,i} + DC_{incap,i} + DC_{inval,i} + FGS_i \\
FGS_i &= (Incap_i + Inval_i + DC_{valide,i} + DC_{incap,i} + DC_{inval,i}) \cdot 2.7\% \\
Incap_i &= -\sum_{t=1}^{T} \left(def_{t,i} \cdot \frac{M_{t,i}}{CRD_{t,i}} \cdot \begin{cases} \sum_{c=4}^{min(t-4,35)} Maintien.incap_{t,i,c} & \text{si } 4 \leq t-c \\ 0 & \text{sinon} \end{cases} \right) \\
Inval_i &= -\sum_{t=1}^{T} \left(def_{t,i} \cdot \frac{M_{t,i}}{CRD_{t,i}} \cdot \begin{cases} \sum_{v=4}^{min(t-4,T)} Maintien.inval_{t,i,v} & \text{si } 4 \leq t-v \\ 0 & \text{sinon} \end{cases} \right) \\
DC_{valide,i} &= -\sum_{t=1}^{T} def_{t,i} \cdot CRD_{t,i} \cdot Tx_{deces.valide,i_t} \\
DC_{incap,i} &= -\sum_{t=1}^{T} \left(def_{t,i} \cdot \sum_{c=1}^{min(t-1,35)} Maintien.incap_{t,i,c} \cdot Tx_{deces.incap,i_t,c} \right) \\
DC_{inval,i} &= -\sum_{t=1}^{T} \left(def_{t,i} \cdot \sum_{v=1}^{min(t-1,T)} Maintien.inval_{t,i,v} \cdot (1 - Tx_{maintien.inval,i_t,v}) \right)
\end{aligned}$$

Où :

$$
\begin{aligned}
Entree.incap_{t,i} &= CRD_{t,i} \cdot Tx_{entree.incap,i_t} \\
Maintien.incap_{t,i,c} &= Entree.incap_{t-c,i} \cdot \prod_{u=1}^{c} Tx_{maintien.incap,i_t,u} \\
Pass.incap.inval_{t,i} &= Entree.incap_{t,i} \cdot Tx_{pass.incap.inval,i_t,0} + \\
&\qquad \sum_{c=1}^{min(t-1,35)} Maintien.incap_{t,i,c} \cdot Tx_{pass.incap.inval,i_t,c} \\
Maintien.inval_{t,i,v} &= Pass.incap.inval_{t-v,i} \cdot \prod_{u=1}^{v} Tx_{maintien.inval,i_t,u}
\end{aligned}
$$

2. Si l'assuré est incapable avec une ancienneté c dans cet état :

$$
\begin{aligned}
BE_{sinistre,i} &= Incap_i + Inval_i + DC_{incap,i} + DC_{inval,i} + FGS_i \\
FGS_i &= (Incap_i + Inval_i + DC_{incap,i} + DC_{inval,i}) \cdot 2.7\% \\
Incap_i &= -\sum_{t=1}^{T} \left(def_{t,i} \cdot \frac{M_{t,i}}{CRD_{t,i}} \cdot \begin{cases} Maintien.incap_{t,i,c} & \text{si } 4 \leq c \leq 35 \\ 0 & \text{sinon} \end{cases} \right) \\
Inval_i &= -\sum_{t=1}^{T} \left(def_{t,i} \cdot \frac{M_{t,i}}{CRD_{t,i}} \cdot \begin{cases} \sum_{v=4}^{min(t-4,T)} Maintien.inval_{t,i,v} & \text{si } 4 \leq t-v \\ 0 & \text{sinon} \end{cases} \right) \\
DC_{incap,i} &= -\sum_{t=1}^{T} def_{t,i} \cdot \begin{cases} Maintien.incap_{t,i,c} \cdot Tx_{deces.incap,i_t,c} & \text{si } c \leq 35 \\ 0 & \text{sinon} \end{cases} \\
DC_{inval,i} &= -\sum_{t=1}^{T} \left(def_{t,i} \cdot \sum_{v=1}^{min(t-1,T)} Maintien.inval_{t,i,v} \cdot (1 - Tx_{maintien.inval,i_t,v}) \right)
\end{aligned}
$$

Où :

$$
\begin{aligned}
Maintien.incap_{t,i,c} &= CRD_{t,i} \cdot Tx_{maintien.incap,i_t,c} \\
Pass.incap.inval_{t,i} &= \begin{cases} Maintien.incap_{t,i,c} \cdot Tx_{pass.incap.inval,i_t,c} & \text{si } c \leq 35 \\ Maintien.incap_{t,i,c} & \text{si } c > 35 \end{cases} \\
Maintien.inval_{t,i,v} &= Pass.incap.inval_{t-v,i} \cdot \prod_{u=1}^{v} Tx_{maintien.inval,i_t,u}
\end{aligned}
$$

3. Si l'assuré est invalide avec une ancienneté v dans cet état :

$$
\begin{aligned}
BE_{sinistre,i} &= Inval_i + DC_{inval,i} + FGS_i \\
FGS_i &= (Inval_i + DC_{inval,i}) \cdot 2.7\% \\
Inval_i &= -\sum_{t=1}^{T} \left(def_{t,i} \cdot \frac{M_{t,i}}{CRD_{t,i}} \cdot \begin{cases} Maintien.inval_{t,i,v} & \text{si } 4 \leq v \\ 0 & \text{sinon} \end{cases} \right) \\
DC_{inval,i} &= -\sum_{t=1}^{T} (def_{t,i} \cdot Maintien.inval_{t,i,v} \cdot (1 - Tx_{maintien.inval,i_t,v}))
\end{aligned}
$$

Où :

$$Maintien.inval_{t,i,v} = CRD_{t,i} \cdot Tx_{maintien.inval,i_t,v}$$

Avec :

- $M_{t,i}$: le montant des mensualités du prêt pour l'individu i au mois de projection t ;
- $CRD_{t,i}$: le CRD pour l'individu i au mois de projection t ;
- $Tx_{entree.incap,i_t}$: le taux d'entrée en incapacité pour l'individu i d'âge i_t ;
- $Tx_{maintien.incap,i_t,c}$: le taux de maintien en incapacité pour l'individu i d'âge i_t et d'ancienneté en état d'incapacité c ;
- $Tx_{pass.incap.inval,i_t,c}$: le taux de passage d'incapacité en invalidité pour l'individu i d'âge i_t et d'ancienneté en état d'incapacité c ;
- $Tx_{maintien.inval,i_t,v}$: le taux de maintien en invalidité pour l'individu i d'âge i_t et d'ancienneté en état d'invalidité v ;
- $Tx_{deces.valide,i_t}$: le taux de décès pour l'individu valide i d'âge i_t ;
- $Tx_{deces.incap,i_t,c}$: le taux de décès en incapacité pour l'individu i d'âge i_t et d'ancienneté en état d'incapacité c ;
- $Tx_{deces.inval,i_t,c}$: le taux de décès en invalidité pour l'individu i d'âge i_t et d'ancienneté en état d'invalidité c ;
- FGS_i : les frais de gestions de sinistres anticipés de l'individu i, qui s'élèvent à 2.7% de l'ensemble des sinistres anticipés ;
- $def_{t,i}$: le déflateur pour l'individu i au mois de projection t, permettant d'obtenir une valeur actuelle des flux futurs.

Ce processus de calcul du BE des sinistres permet d'obtenir une estimation précise et dynamique des engagements liés aux sinistres, en tenant compte des différents types d'événements et des coûts associés à leur gestion.

6.3.3 Analyse comparative des résultats

Cette section se concentre sur l'analyse comparative des résultats obtenus pour le BE de primes, le BE de sinistres, et enfin le BE global. En examinant ces différentes composantes, il est possible de mieux comprendre l'impact des hypothèses économiques et des scénarios d'inflation sur la performance du portefeuille d'assurance emprunteur. Cette analyse permet également d'identifier les tendances et les variations significatives entre les différents états des assurés, ainsi que d'évaluer la viabilité financière du portefeuille dans son ensemble.

6.3.3.1 Analyse des BE de Primes

Les BE de primes sont calculés pour chaque assuré, pour chacune des 5 années de projection et chaque scénario d'inflation, comme expliqué dans la *Partie 6.3.2.2.* Les résultats obtenus sont disponibles dans le *Tableau 6.15.* Il est possible d'y constater que les BE de primes deviennent moins négatifs au fil des années, quel que soit le scénario d'inflation. Par exemple, dans le scénario d'inflation à 2%, les BE passent de -132.77 M€ au T4 2024 à -82.21 M€ au T4 2028. Cela reflète la diminution progressive de l'effectif total du portefeuille due à une baisse combinée des affaires nouvelles et à des rachats modérés (cf *Tableau 6.11*). Cependant, contrairement à ce qui pourrait être attendu suite au constat sur les effectifs, les scénarios d'inflation plus élevée ne montrent pas systématiquement des BE de primes plus négatifs.

en M€	BE de primes Déflatés				
	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028
Inflation 0%	-132.27	-118.76	-105.54	-93.89	-83.02
Inflation 0.65%	-132.53	-118.92	-105.56	-94.15	-83.53
Inflation 1.3%	-132.37	-118.71	-105.63	-94.00	-82.88
Inflation 2%	-132.77	-119.31	-105.93	-93.88	-82.21
Inflation 3%	-132.25	-118.86	-105.49	-94.33	-83.44
Inflation 4%	-132.53	-119.29	-106.00	-94.25	-82.89
Inflation 5%	-132.56	-119.82	-106.23	-94.37	-82.79

TABLE 6.15 – BE de primes du portefeuille selon la date de projection et le scénario d'inflation

Le *Tableau 6.16* présente les variations des BE de primes selon différents scénarios d'inflation. Bien que ces écarts restent relativement faibles, ils sont significatifs et illustrent la sensibilité des primes au taux d'inflation. Pour les scénarios d'inflation modérément faibles (0%, 0.65%, et 1.3%), les BE de primes sont initialement inférieurs (en valeur absolue) que ceux du scénario de référence à 2%. Cependant, cette tendance s'inverse vers la fin de la période de projection, où les primes commencent à se redresser.

Dans les scénarios d'inflation plus élevés (3%, 4%, et 5%), la situation devient plus complexe. À mesure que l'inflation augmente, les variations des BE de primes montrent une dynamique similaire à celle observée dans les scénarios de faible inflation, avec des montants qui peuvent dépasser (en valeur absolue) que ceux du scénario de référence à 2% dans certaines années. Cela souligne que, même si les BE de primes sont affectés par l'inflation, les tendances peuvent varier considérablement selon le niveau d'inflation et la période considérée.

en K€	BE de primes Déflatés				
	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028
Inflation 0%	500.3 -0.38%	555.7 -0.47%	386.3 -0.36%	-10.2 0.01%	-811.2 0.99%
Inflation 0.65%	249.4 -0.19%	395.2 -0.33%	370.2 -0.35%	-268.7 0.29%	-1 321.5 1.61%
Inflation 1.3%	404.3 -0.30%	607.1 -0.51%	299.1 -0.28%	-116.1 0.12%	-673.0 0.82%
Inflation 2%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%
Inflation 3%	521.2 -0.39%	453.4 -0.38%	439.9 -0.42%	-449.5 0.48%	-1 230.4 1.50%
Inflation 4%	248.7 -0.19%	22.0 -0.02%	-69.9 0.07%	-365.8 0.39%	-686.9 0.84%
Inflation 5%	211.9 -0.16%	-507.3 0.43%	-303.2 0.29%	-487.4 0.52%	-582.9 0.71%

TABLE 6.16 – Comparaison des BE de primes du portefeuille selon la date de projection et le scénario d'inflation à 2%

Les divergences observées entre les tendances des BE de primes et celles des effectifs (cf *Tableau 6.11*) peuvent s'expliquer par plusieurs hypothèses :

Le profil des assurés rachetant leurs contrats : Dans les scénarios d'inflation élevée, il est possible que les assurés qui rachètent leurs contrats aient un profil de risque différent de la moyenne du portefeuille. Si les assurés avec des primes plus faibles et un risque moins élevé sont plus enclins à racheter leurs contrats, cela pourrait expliquer pourquoi la diminution des BE de primes n'est pas proportionnelle à la diminution des effectifs.

L'impact des affaires nouvelles : Bien que basées sur un profil de client "moyen", l'impact des affaires nouvelles sur le BE de primes pourrait être plus complexe qu'anticipé. Dans les scénarios de faible inflation, où les affaires nouvelles sont plus nombreuses, elles pourraient contribuer à maintenir un niveau de BE de primes plus élevé à long terme, malgré un effectif total plus faible.

L'effet de l'actualisation : Les BE de primes étant déflatés, l'impact de l'inflation sur l'actualisation des flux futurs pourrait expliquer certaines variations contre-intuitives. Dans les scénarios d'inflation élevée, la valeur actuelle des primes futures pourrait être plus élevée, compensant partiellement la diminution de l'effectif.

Les dynamique des taux de rachat : La modélisation des taux de rachat en fonction de l'inflation pourrait avoir des effets non linéaires sur le BE de primes. Si les rachats concernent principalement des contrats moins rentables dans les scénarios d'inflation élevée, cela pourrait expliquer pourquoi le BE de primes ne diminue pas autant qu'attendu.

Les interaction complexe entre inflation et comportement des assurés : L'inflation pourrait influencer non seulement les rachats et les affaires nouvelles, mais aussi d'autres comportements des assurés, comme la propension à déclarer des sinistres, qui affectent indirectement le BE de primes.

Ces divergences soulignent la complexité des interactions entre l'inflation, le comportement des assurés, et la valorisation actuarielle des engagements.

6.3.3.2 Analyse des BE de Sinistres

Avant de commencer l'analyse des BE de sinistres, il convient d'examiner les effectifs d'incapables et d'invalides dans le portefeuille, afin d'anticiper les potentiels impacts de proportion d'effectifs.

L'examen des données du *Tableau 6.17* révèle plusieurs tendances intéressantes. Tout d'abord, il est possible de constater que l'inflation n'impacte pas directement les états des assurés, mais plutôt les effectifs globaux du portefeuille. Cependant, les proportions d'effectifs selon les divers états varient en fonction des scénarios d'inflation. Pour les incapacités, une tendance à l'augmentation de la proportion avec l'inflation est visible, passant d'environ 0.3% dans les scénarios à faible inflation à plus de 0.7% dans les scénarios à forte inflation en 2028. Les invalidités, quant à elles, montrent une évolution plus stable, avec des proportions généralement comprises entre 0.07% et 0.09% en 2028, quel que soit le scénario d'inflation. Il est également notable que les effectifs d'invalides augmentent progressivement au fil des années dans tous les scénarios, reflétant probablement une accumulation des cas au cours du temps. Ces variations dans la composition du portefeuille selon les états des assurés pourraient avoir des implications significatives sur les BE de sinistres, qu'il convient d'analyser en détail.

	Effectif	2023	2024	2025	2026	2027	2028
Projection d'inflation à 0%	**Portefeuille**	15 000	14 411	13 890	13 292	12 829	12 287
	Incapacité	- 0.00%	42 0.29%	39 0.28%	49 0.37%	50 0.39%	41 0.33%
	Invalidité	- 0.00%	- 0.00%	- 0.00%	3 0.02%	7 0.05%	10 0.08%
Projection d'inflation à 0.65%	**Portefeuille**	15 000	14 414	13 895	13 380	12 840	12 264
	Incapacité	- 0.00%	44 0.31%	49 0.35%	53 0.40%	57 0.44%	40 0.33%
	Invalidité	- 0.00%	- 0.00%	1 0.01%	2 0.01%	8 0.06%	11 0.09%
Projection d'inflation à 1.3%	**Portefeuille**	15 000	14 389	13 903	13 374	12 914	12 280
	Incapacité	- 0.00%	52 0.36%	49 0.35%	51 0.38%	51 0.39%	43 0.35%
	Invalidité	- 0.00%	1 0.01%	1 0.01%	3 0.02%	5 0.04%	8 0.07%
Projection d'inflation à 2%	**Portefeuille**	15 000	14 451	13 912	13 402	12 852	12 184
	Incapacité	- 0.00%	39 0.27%	40 0.29%	45 0.34%	43 0.33%	39 0.32%
	Invalidité	- 0.00%	1 0.01%	2 0.01%	3 0.02%	5 0.04%	10 0.08%
Projection d'inflation à 3%	**Portefeuille**	15 000	14 391	13 930	13 389	12 859	12 217
	Incapacité	- 0.00%	89 0.62%	101 0.73%	104 0.78%	101 0.79%	81 0.66%
	Invalidité	- 0.00%	1 0.01%	1 0.01%	3 0.02%	5 0.04%	8 0.07%
Projection d'inflation à 4%	**Portefeuille**	15 000	14 409	13 911	13 352	12 848	12 174
	Incapacité	- 0.00%	88 0.61%	96 0.69%	106 0.79%	109 0.85%	78 0.64%
	Invalidité	- 0.00%	- 0.00%	1 0.01%	2 0.01%	8 0.06%	11 0.09%
Projection d'inflation à 5%	**Portefeuille**	15 000	14 440	13 929	13 435	12 879	12 220
	Incapacité	- 0.00%	78 0.54%	73 0.52%	85 0.63%	84 0.65%	80 0.65%
	Invalidité	- 0.00%	- 0.00%	- 0.00%	3 0.02%	7 0.05%	10 0.08%

TABLE 6.17 – Focus sur les effectifs d'incapables et d'invalides selon les années et scénarios d'inflation

Les BE de sinistres sont calculés pour chaque assuré, pour chacune des 5 années de projection et chaque scénario d'inflation, comme expliqué dans la *Partie 6.3.2.3*. Les résultats obtenus sont disponibles dans le *Tableau 6.18*. À l'instar des BE de primes, les BE de sinistres diminuent constamment pour tous les scénarios d'inflation au fil du temps, passant d'environ 95 M€ au T4 2024 à environ 63 M€ au T4 2028. Cette baisse générale pourrait s'expliquer par la diminution de l'effectif du portefeuille observée *précédemment*, entraînant une réduction des engagements liés aux sinistres.

en M€	BE de sinistres Déflatés				
	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028
Inflation 0%	94.48	86.92	77.77	70.26	63.65
Inflation 0.65%	95.10	87.58	78.57	72.61	65.05
Inflation 1.3%	95.07	87.20	78.59	71.82	64.02
Inflation 2%	95.13	87.64	78.64	70.90	63.13
Inflation 3%	94.28	88.14	78.59	72.06	64.08
Inflation 4%	94.93	87.13	78.34	71.34	63.50
Inflation 5%	94.42	87.29	78.25	71.24	63.03

TABLE 6.18 – BE de sinistres du portefeuille selon la date de projection et le scénario d'inflation

Il est notable que l'impact de l'inflation sur les BE de sinistres semble moins prononcé que sur les BE de primes. Les écarts entre les différents scénarios d'inflation sont relativement faibles, généralement inférieurs à 2 M€ pour une année donnée. Cette observation suggère une certaine résilience des engagements liés aux sinistres face aux variations d'inflation.

La comparaison des résultats avec le scénario d'inflation à 2% (cf *Tableau 6.19*) montre que les écarts sont généralement inférieurs à 1% en valeur relative. Une tendance intéressante se dégage : les scénarios d'inflation plus faibles montrent des BE de sinistres légèrement plus élevés en fin de période (T4 2028). Cela pourrait s'expliquer par une plus grande persistance des contrats dans ces scénarios de faible inflation.

L'évolution des BE de sinistres au fil du temps et selon les scénarios d'inflation reflète probablement des interactions complexes entre plusieurs facteurs. Les BE incluent les sinistres d'incapacité, d'invalidité, de décès, de décès en incapacité et de décès en invalidité. Chacun de ces types de sinistres peut réagir différemment aux variations d'inflation et à l'évolution du portefeuille.

Il est important de noter que les scénarios d'inflation plus élevée (3%, 4%, 5%) montrent des variations plus importantes par rapport au scénario de référence, particulièrement vers la fin de la période de projection. Cela suggère que l'impact de l'inflation sur les BE de sinistres pourrait se manifester plus fortement à long terme.

En conclusion, bien que l'inflation ait un impact sur les BE de sinistres, cet impact semble moins prononcé et moins linéaire que pour les BE de primes. Les variations observées soulignent la complexité des interactions entre l'inflation, la composition du portefeuille et les différents types de sinistres couverts.

en K€	BE de sinistres Déflatés				
	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028
Inflation 0%	-648.4 -0.68%	-722.9 -0.82%	-868.8 -1.10%	-639.9 -0.90%	517.1 0.82%
Inflation 0.65%	-34.6 -0.04%	-59.1 -0.07%	-67.1 -0.09%	1 709.3 2.41%	1 914.8 3.03%
Inflation 1.3%	-64.3 -0.07%	-441.0 -0.50%	-49.4 -0.06%	919.0 1.30%	886.2 1.40%
Inflation 2%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%
Inflation 3%	-854.8 -0.90%	493.5 0.56%	-50.7 -0.06%	1 155.3 1.63%	950.2 1.51%
Inflation 4%	-200.7 -0.21%	-515.5 -0.59%	-302.2 -0.38%	433.8 0.61%	372.5 0.59%
Inflation 5%	-709.5 -0.75%	-346.6 -0.40%	-391.9 -0.50%	336.5 0.47%	-101.2 -0.16%

TABLE 6.19 – Comparaison des BE de sinistres du portefeuille selon la date de projection et le scénario d'inflation à 2%

6.3.3.3 Analyse des BE globaux

L'analyse des BE globaux, qui combinent les BE de primes et de sinistres, révèle des tendances intéressantes qui reflètent l'interaction complexe entre l'inflation, la composition du portefeuille, et les différents types d'engagements.

Tout d'abord, une diminution générale des BE globaux (en valeur absolue) est observable au fil du temps pour tous les scénarios d'inflation, passant d'environ -38 à -37 M€ au T4 2024 à environ -20 à -18 M€ au T4 2028. Cette tendance à la baisse est cohérente avec les analyses effectuées précédemment et souligne que, même dans un contexte d'inflation modérée, la performance du portefeuille est affectée par la diminution du nombre d'assurés.

L'impact de l'inflation sur les BE globaux semble plus prononcé que sur les BE de sinistres seuls, mais moins marqué que sur les BE de primes. Les écarts entre les différents scénarios d'inflation sont relativement faibles en début de période, mais s'accentuent vers la fin, atteignant jusqu'à 1.28 M€ de différence entre les scénarios extrêmes au T4 2028 (cf. *Tableau 6.21*).

en M€	BE Déflatés				
	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028
Inflation 0%	-37.79	-31.84	-27.78	-23.63	-19.37
Inflation 0.65%	-37.43	-31.33	-26.99	-21.54	-18.48
Inflation 1.3%	-37.30	-31.50	-27.04	-22.18	-18.86
Inflation 2%	-37.64	-31.67	-27.29	-22.98	-19.07
Inflation 3%	-37.98	-30.72	-26.90	-22.28	-19.35
Inflation 4%	-37.60	-32.16	-27.66	-22.91	-19.39
Inflation 5%	-38.14	-32.52	-27.99	-23.13	-19.76

TABLE 6.20 – BE du portefeuille selon la date de projection et le scénario d'inflation

La comparaison des résultats avec le scénario d'inflation à 2% laisse apparaître des variations

intéressantes :

— Les scénarios de faible inflation (0,65% et 1,3%) présentent des BE globaux systématiquement plus faibles (en valeur absolue) que ceux du scénario de référence à 2%. Cette tendance s'aligne avec les effectifs plus faibles observés dans ces scénarios, entraînant une diminution des BE de primes tout en maintenant une sinistralité comparable. Il est intéressant de noter que, malgré des effectifs légèrement supérieurs à la dernière date de projection par rapport au scénario de référence, cela ne se traduit pas par une augmentation des BE ;

— Les scénarios d'inflation élevée (3%, et 4%) montrent des variations des BE globaux par rapport au scénario à 2%, cohérentes avec les observations sur les effectifs.

— Les scénarios d'inflation extrême (0% et 5%) présentent les BE globaux les plus élevés (en valeur absolue). Pour le scénario d'inflation élevée, cela s'accorde avec l'augmentation de l'effectif au cours des projections, comme constaté *précédemment*. Cependant, le scénario d'inflation nulle, bien qu'ayant un comportement des effectifs similaire aux scénarios d'inflation faible, ne suit pas la même tendance en termes de BE. Cette divergence s'explique principalement par une sinistralité moins importante dans ce scénario, soulignant l'impact complexe de l'inflation sur différents aspects du portefeuille.

en K€	BE Déflatés				
	T4 2024	T4 2025	T4 2026	T4 2027	T4 2028
Inflation 0%	-148.1 0.39%	-167.2 0.53%	-482.6 1.77%	-650.2 2.83%	-294.1 1.54%
Inflation 0.65%	214.7 -0.57%	336.1 -1.06%	303.2 -1.11%	1 440.6 -6.27%	593.4 -3.11%
Inflation 1.3%	340.0 -0.90%	166.1 -0.52%	249.7 -0.92%	802.9 -3.49%	213.2 -1.12%
Inflation 2%	0 0.00%	0 0.00%	0 0.00%	0 0.00%	0 0.00%
Inflation 3%	-333.5 0.89%	946.9 -2.99%	389.2 -1.43%	705.9 -3.07%	-280.1 1.47%
Inflation 4%	48.0 -0.13%	-493.5 1.56%	-372.1 1.36%	68.0 -0.30%	-314.4 1.65%
Inflation 5%	-497.6 1.32%	-853.9 2.70%	-695.1 2.55%	-150.9 0.66%	-684.1 3.59%

TABLE 6.21 – Comparaison des BE du portefeuille selon la date de projection et le scénario d'inflation à 2%

Il est important de noter que ces variations ne sont pas linéaires avec l'augmentation de l'inflation, ce qui souligne la complexité des interactions entre les différents facteurs économiques et actuariels.

En conclusion, l'analyse des BE globaux souligne l'importance d'une gestion proactive et adaptable du portefeuille face aux fluctuations d'inflation. Bien que tous les scénarios présentent une baisse des BE (en valeur absolue), les résultats suggèrent que des stratégies spécifiques doivent être mises en place pour naviguer efficacement dans des environnements inflationnistes variés. La gestion doit donc prendre en compte non seulement le niveau d'inflation, mais aussi son impact sur les affaires nouvelles et les rachats pour assurer la viabilité financière à long terme du portefeuille.

6.3.3.4 Synthèse

Les BE de primes, calculés pour chaque assuré et chaque scénario d'inflation, montrent une diminution progressive (en valeur absolue) au fil des années dans tous les scénarios d'inflation, reflétant la baisse des effectifs due à la réduction des affaires nouvelles et aux rachats. Toutefois, les scénarios d'inflation plus élevée ne produisent pas systématiquement des BE de primes moins élevés (en valeur absolue). Cela peut s'expliquer par des interactions complexes, telles que la sélection adverse parmi les assurés qui rachètent leurs contrats et l'impact de l'inflation sur l'actualisation des flux futurs. Les BE de sinistres, en revanche, présentent une sensibilité moindre à l'inflation, affichant une résilience face aux variations économiques tout en montrant également une diminution constante (en valeur absolue) liée à la réduction des effectifs.

Les BE globaux, combinant primes et sinistres, révèlent une tendance logique à la baisse (en valeur absolue) au fil du temps, tout en mettant en évidence des variations significatives entre les scénarios d'inflation. Les scénarios de faible inflation tendent à réduire les BE globaux (en valeur absolue) en raison de la stabilité des effectifs et d'une sinistralité comparable. À l'inverse, les scénarios d'inflation nulle ou élevée présentent des BE globaux plus élevés (en valeur absolue) en raison de dynamiques spécifiques, comme une meilleure stabilité du portefeuille ou une compensation partielle de la diminution des effectifs par des primes actualisées plus élevées.

Ces résultats soulignent l'importance d'une gestion proactive des portefeuilles face aux fluctuations d'inflation, en anticipant ses effets non linéaires sur les affaires nouvelles, les rachats et les sinistres pour préserver la viabilité financière à long terme.

6.4 Conclusion du chapitre

Ce chapitre a présenté la modélisation d'un portefeuille fictif d'assurance emprunteur, avec une projection sur cinq ans selon plusieurs scénarios d'inflation. Les BE de primes et de sinistres sont calculés pour chaque assuré, permettant d'analyser l'impact du contexte économique sur la performance du portefeuille.

Les résultats montrent que les BE de primes comme de sinistres diminuent progressivement (en valeur absolue) dans tous les scénarios d'inflation, en raison de la baisse des effectifs due à la réduction des affaires nouvelles et aux rachats. Cependant, les scénarios d'inflation plus élevée ne conduisent pas systématiquement à des BE de primes plus faibles (en valeur absolue), ce qui souligne la complexité des interactions entre l'inflation et le comportement des assurés. En parallèle, les BE de sinistres, bien qu'affectés par les variations économiques, présentent des écarts limités entre les différents scénarios d'inflation.

Les BE globaux combinent ces dynamiques, révélant des variations significatives selon les scénarios. Si une faible inflation tend à réduire les BE globaux (en valeur absolue), l'inflation nulle ou élevée peut, dans certains cas, offrir de meilleures performances grâce à des facteurs compensatoires comme la stabilité du portefeuille ou l'effet de l'actualisation.

Ainsi, cette étude met en lumière non seulement les effets directs de l'inflation sur les primes et sinistres, mais également les interactions plus subtiles qui influencent la performance globale du portefeuille d'assurance emprunteur.

Conclusion

Ce mémoire s'est attaché à explorer l'impact du contexte économique sur les dynamiques clés d'un portefeuille d'assurance emprunteur, en modélisant le taux d'affaires nouvelles, les lois de rachat, et en analysant la performance du portefeuille sous divers scénarios d'inflation.

L'objectif principal était de développer une approche complète et rigoureuse pour mieux comprendre et anticiper les impacts des variations économiques sur un portefeuille d'assurance emprunteur. Cette démarche s'inscrit dans un contexte marqué par l'incertitude macroéconomique, où des facteurs tels que l'inflation et les taux d'intérêt jouent un rôle central dans les comportements des emprunteurs.

L'étude s'est structurée autour de plusieurs axes complémentaires. Dans un premier temps, une analyse approfondie du contexte économique, mettant en évidence le rôle clé des politiques monétaires et fiscales dans l'évolution du marché de l'assurance emprunteur. Cette contextualisation a permis de souligner les interactions complexes entre les facteurs macroéconomiques et les décisions des emprunteurs en matière de souscription ou de rachat de crédits. Les chapitres suivants ont successivement abordé la modélisation des taux d'affaires nouvelles et des lois de rachat. Ces travaux ont mis en lumière l'importance des taux d'intérêt, de l'inflation, ainsi que des effets saisonniers et des spécificités des prêts selon leur ancienneté. L'approche méthodologique employée, notamment les modèles SARIMAX et ARIMAX, a permis de capturer efficacement ces dynamiques, offrant des outils robustes et adaptés pour la prévision. Enfin, l'application des modèles développés à un portefeuille fictif d'assurance emprunteur a permis de projeter l'évolution des BE de primes, de sinistres et globaux sur une période de cinq ans selon différents scénarios d'inflation. Les résultats obtenus ont révélé la complexité des interactions entre l'inflation et les comportements des emprunteurs, ainsi que leur impact différencié sur les primes et les sinistres.

Malgré les avancées méthodologiques et les résultats obtenus, certaines limites méritent d'être soulignées. La disponibilité et la qualité des données, notamment en ce qui concerne les comportements des emprunteurs en période de chocs économiques, constituent un enjeu majeur. De même, l'intégration de modèles plus complexes, tels que des réseaux de neurones, pourrait enrichir l'analyse en capturant des interactions non linéaires et des dynamiques plus subtiles.

Les perspectives offertes par ce mémoire sont nombreuses. L'approche développée pourrait être étendue à d'autres segments d'assurance ou à des portefeuilles réels pour affiner les stratégies de tarification, de provisionnement et de gestion des risques. Par ailleurs, l'analyse des impacts économiques pourrait être enrichie par l'intégration de scénarios macroéconomiques prospectifs, incluant des variables telles que le chômage ou les politiques de logement.

En conclusion, ce mémoire contribue à une meilleure compréhension des dynamiques du marché de l'assurance emprunteur dans un contexte économique en mutation. Les outils développés offrent aux assureurs une base solide pour anticiper les effets des variations macroéconomiques, ajuster leurs stratégies et renforcer leur résilience face aux incertitudes futures.

Annexe A

CHAPITRE 1 : Cadre de l'étude

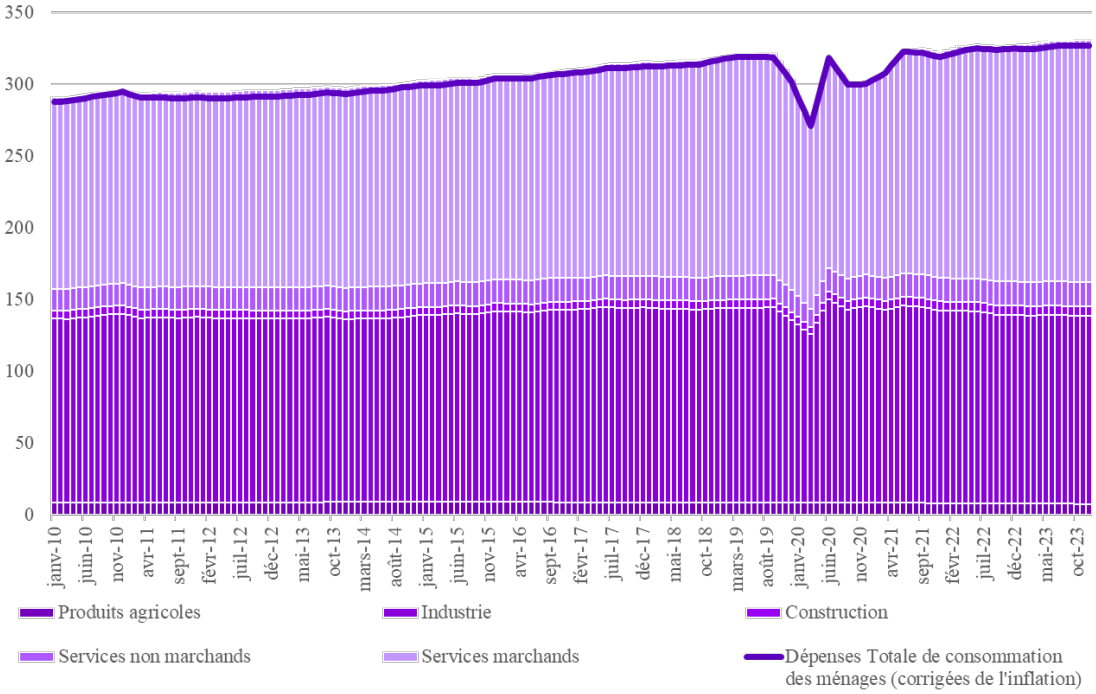


FIGURE A.1 – Répartition de la dépense totale de consommation des ménages (corrigés de l'inflation) en milliards d'euros(*INSEE* [19])

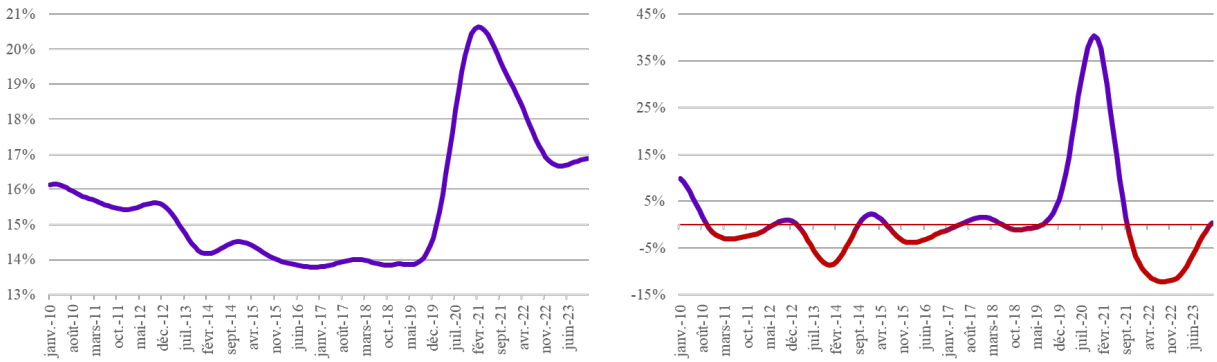


FIGURE A.2 – Taux d'épargne des ménages en % du revenu brut (gauche) et son glissement annuel en % (*INSEE* [28])

Annexe B

CHAPITRE 2 : Présentation des données économiques et de la démarche

B.1 Méthode d'interpolation cubique d'Hermite

L'interpolation cubique d'Hermite fonctionne en construisant un polynôme cubique entre chaque paire de points de données connus. Ces polynômes sont définis de manière à ce que la courbe résultante soit continue et ait une dérivée continue aux points de jonction. Mathématiquement, pour deux points connus (x_k, y_k) et (x_{k+1}, y_{k+1}), le polynôme d'interpolation $p(x)$ est défini comme :

$$p(x) = h_{00}(t)y_k + h_{10}(t)(x_{k+1} - x_k)m_k + h_{01}(t)y_{k+1} + h_{11}(t)(x_{k+1} - x_k)m_{k+1}$$

où m_k et m_{k+1} sont les dérivées aux points (x_k, y_k) et (x_{k+1}, y_{k+1}) respectivement, et h_{ij} sont les polynômes d'Hermite :

$$\begin{aligned} h_{00} &= 2t^3 - 3t^2 + 1; \\ h_{10} &= t^3 - 2t^2 + 1; \\ h_{01} &= -2t^3 + 3t^2; \\ h_{11} &= t^3 - t^2; \end{aligned}$$

tel que :

$$t = \frac{x - x_k}{x_{k+1} - x_k}.$$

	Données				
y	**Flux mensuels des crédits nouveaux à l'habitat aux particuliers résidents**				
r_j	**Taux de rachat des crédits immobiliers**				
x_1	**Taux d'intérêt annuel des crédits nouveaux à l'habitat des particuliers**				
x_2	**Encours de prêts à taux zéro accordés à la clientèle non financière résidente**				
x_3	**Indice des prix des logements**		Neufs	Tous logements	
x_4			Anciens	Maisons	
x_5				Appartements	
x_6	**Indice du coût de production dans la construction (ICP-F)**				
x_7	**Indice du coût de la construction (ICC)**				
x_8	**Taux de chômage au sens du BIT**				
x_9	**Dette publique en % de PIB**				
x_{10}	**Évolution annuelle de l'IPC**				
x_{11}	**Produit Intérieur Brut en Mds€ 2020**				
x_{12}	**Évolution annuelle relative du Pouvoir d'achat du revenu disponible brut**				
x_{13}	**Taux d'épargne des ménages en % du revenu brut**				
x_{14}	**Volume de construction**	Nombre de logements	Autorisés	Collectif et Résidence	
x_{15}				Individuel	
x_{16}			Commencés	Collectif et Résidence	
x_{17}				Individuel	
x_{18}		Surfaces de logements	Autorisés	Collectif et Résidence	
x_{19}				Individuel	
x_{20}			Commencés	Collectif et Résidence	
x_{21}				Individuel	
x_{22}	**Volume de commercialisation**	Nombre de logements	Mis en vente aux particuliers	Nouvelles constructions	Collectif et Résidence
x_{23}					Individuel
x_{24}			Réservés à la vente par des particuliers	Nouvelles constructions	Collectif et Résidence
x_{25}					Individuel
x_{26}		Encours de logements disponibles à la vente aux particuliers		Nouvelles constructions	Collectif et Résidence
x_{27}					Individuel
x_{28}		Délai moyen d'écoulement de l'encours de logements disponibles à la vente aux particuliers		Nouvelles constructions	Collectif et Résidence
x_{29}					Individuel

TABLE B.1 – Lexique des variables

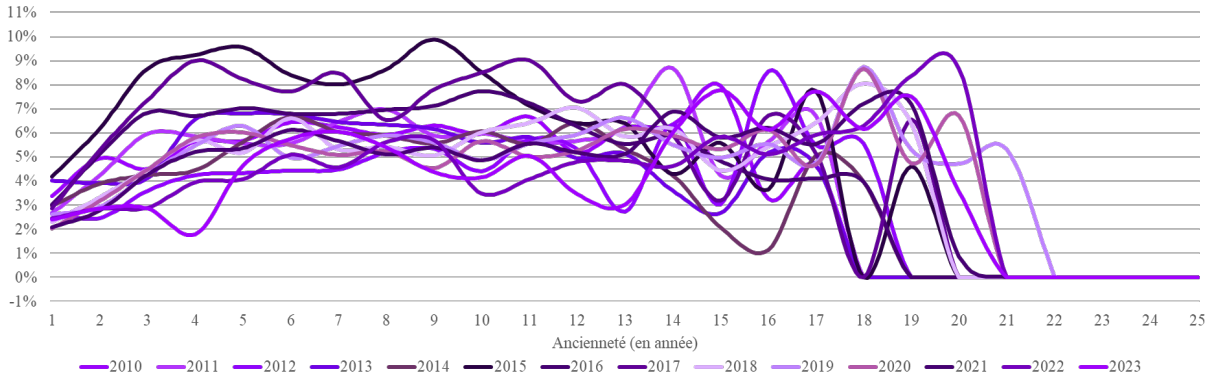


FIGURE B.1 – Taux de rachat de crédits immobiliers selon l'ancienneté du crédit

	Ancienneté																		
	1 an	2 ans	3 ans	4 ans	5 ans	6 ans	7 ans	8 ans	9 ans	10 ans	11 ans	12 ans	13 ans	14 ans	15 ans	16 ans	17 ans	18 ans	19 ans	20 ans
x1	37%	4%	-15%	-29%	-27%	-19%	-2%	17%	-4%	-19%	-18%	-38%	-46%	17%	-17%	-1%	32%	-59%	-62%	-33%
x2	-80%	-72%	-60%	-55%	-57%	-55%	-60%	-70%	-62%	-66%	-54%	-47%	-19%	17%	50%	48%	7%	77%	81%	80%
x3	-68%	-55%	-48%	-48%	-44%	-37%	-46%	-53%	-44%	-55%	-50%	-45%	-13%	2%	62%	25%	12%	69%	87%	83%
x4	-66%	-63%	-61%	-63%	-62%	-51%	-57%	-58%	-52%	-71%	-65%	-65%	-32%	13%	63%	30%	18%	57%	72%	79%
x5	-74%	-69%	-62%	-59%	-59%	-54%	-61%	-65%	-58%	-70%	-65%	-59%	-24%	11%	58%	35%	9%	70%	76%	85%
x6	-60%	-61%	-61%	-63%	-58%	-43%	-52%	-55%	-48%	-67%	-60%	-58%	-34%	-3%	56%	25%	22%	59%	78%	79%
x7	-60%	-59%	-58%	-62%	-54%	-42%	-50%	-53%	-46%	-64%	-58%	-58%	-34%	-3%	57%	25%	24%	61%	78%	79%
x8	76%	58%	51%	49%	51%	49%	50%	58%	53%	59%	50%	43%	3%	-24%	-68%	-32%	-8%	-68%	-80%	-81%
x9	-58%	-43%	-32%	-28%	-21%	-19%	-35%	-46%	-33%	-35%	-41%	-34%	-6%	-14%	46%	15%	-6%	68%	76%	76%
x10	-43%	-55%	-64%	-72%	-74%	-54%	-52%	-45%	-45%	-72%	-56%	-61%	-50%	15%	43%	31%	40%	26%	47%	55%
x11	-49%	-38%	-32%	-31%	-25%	-15%	-25%	-35%	-22%	-35%	-23%	-12%	-1%	-14%	45%	14%	19%	62%	89%	62%
x12	-36%	8%	23%	18%	19%	19%	17%	-4%	10%	19%	22%	33%	48%	22%	37%	-15%	-10%	34%	47%	6%
x13	-65%	-57%	-46%	-44%	-45%	-51%	-55%	-58%	-56%	-54%	-61%	-60%	-22%	26%	47%	32%	-8%	54%	37%	63%
x14	-14%	20%	33%	28%	13%	13%	24%	15%	22%	25%	33%	27%	26%	40%	10%	18%	4%	-11%	12%	-18%
x16	-26%	17%	35%	42%	25%	18%	21%	2%	12%	35%	40%	47%	49%	27%	-3%	17%	-23%	11%	19%	-6%
x18	-16%	16%	29%	25%	7%	8%	20%	12%	18%	20%	31%	25%	23%	47%	7%	24%	5%	-13%	7%	-21%
x20	-26%	14%	32%	39%	20%	12%	18%	0%	7%	32%	38%	46%	47%	36%	-7%	21%	-23%	7%	10%	-12%
x15	-18%	4%	17%	15%	-6%	-14%	1%	-4%	-2%	9%	18%	8%	9%	66%	7%	30%	-12%	-13%	-27%	-33%
x17	-12%	8%	13%	16%	-11%	-16%	3%	-3%	-2%	6%	19%	6%	16%	69%	4%	30%	-16%	-28%	-33%	-34%
x19	-19%	3%	16%	14%	-8%	-15%	-1%	-6%	-3%	7%	17%	8%	9%	66%	6%	30%	-13%	-12%	-27%	-33%
x21	-13%	4%	11%	14%	-14%	-19%	-1%	-5%	-4%	3%	15%	4%	13%	69%	2%	32%	-17%	-27%	-35%	-33%
x22	-3%	40%	56%	55%	33%	30%	38%	30%	39%	52%	59%	62%	46%	37%	-12%	11%	-10%	-18%	-4%	-38%
x24	-21%	27%	48%	56%	42%	35%	31%	11%	25%	50%	54%	67%	64%	19%	4%	5%	-24%	20%	29%	-13%
x26	-16%	-15%	-9%	-3%	9%	17%	-1%	-16%	0%	1%	7%	17%	-3%	-52%	-7%	5%	3%	45%	62%	33%
x28	7%	-32%	-47%	-55%	-30%	-17%	-24%	-20%	-24%	-41%	-42%	-52%	-62%	-48%	-3%	-2%	28%	11%	15%	30%
x23	37%	45%	49%	44%	25%	21%	34%	44%	41%	45%	49%	38%	8%	35%	-34%	7%	-2%	-58%	-63%	-71%
x25	29%	61%	70%	70%	53%	42%	52%	48%	49%	71%	73%	68%	49%	31%	-26%	-5%	-25%	-41%	-47%	-67%
x27	72%	29%	12%	12%	17%	28%	25%	35%	33%	22%	26%	17%	-32%	-43%	-68%	-21%	4%	-58%	-63%	-62%
x29	6%	-33%	-53%	-62%	-39%	-20%	-27%	-20%	-26%	-49%	-48%	-58%	-63%	-43%	8%	-8%	36%	7%	17%	32%

FIGURE B.2 – Corrélations des variables économiques exogènes avec les taux de rachat des crédits immobiliers selon l'année d'ancienneté du crédit

Annexe C

CHAPITRE 3 : DÉMARCHE THÉORIQUE : Construction du volume d'Affaires Nouvelles et de la loi de Rachat

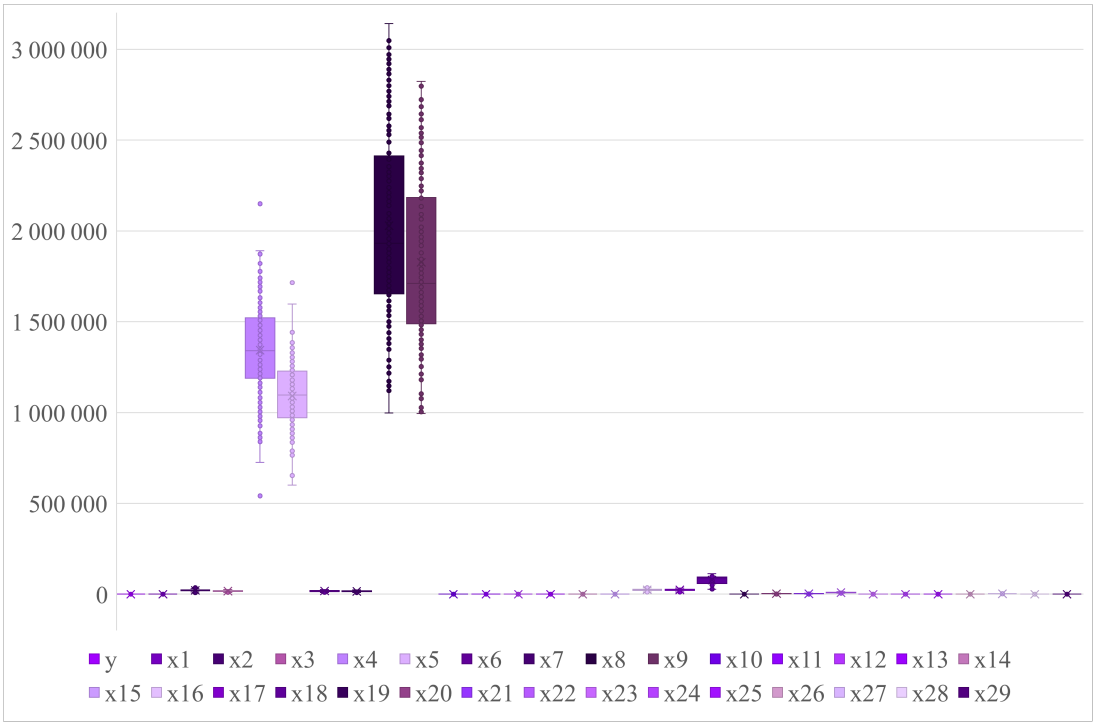


FIGURE C.1 – Boîte à moustache de l'ensemble du jeu de données

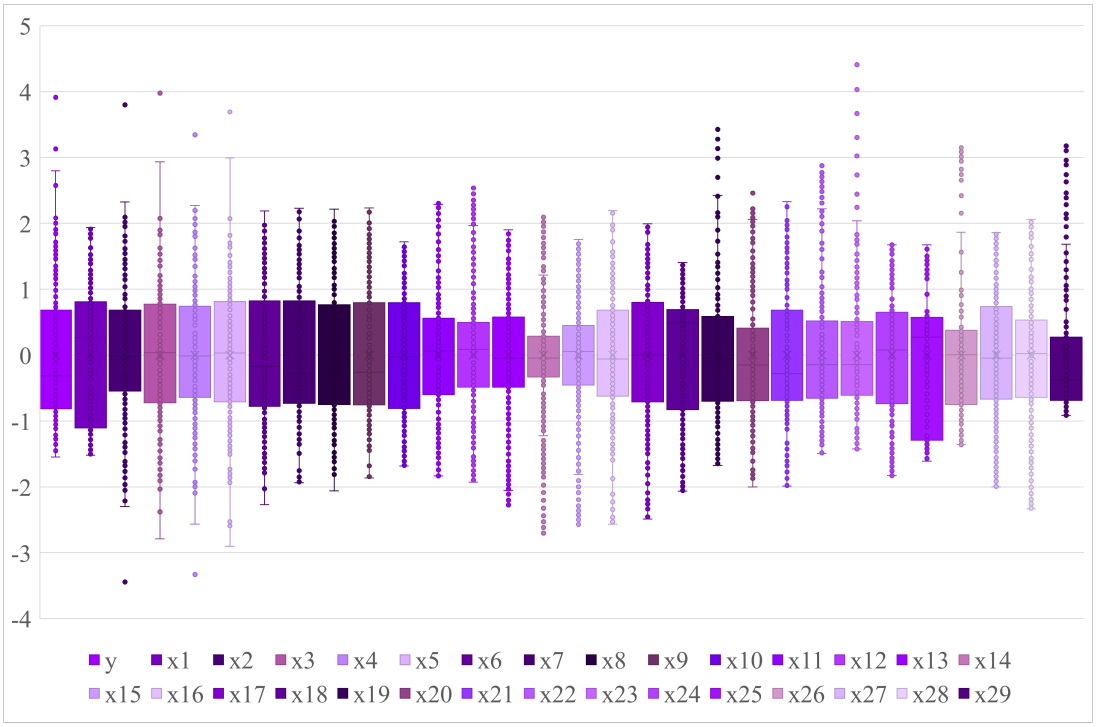


FIGURE C.2 – Boîte à moustache de l'ensemble du jeu de données normalisé

Annexe D

CHAPITRE 4 : APPLICATION : Construction de la loi d'Affaires Nouvelles

D.1 Méthode de Gradient Boosting

D'après le cours de l'*Université Paris-Dauphine [53]*

D.1.1 Généralités

Gradient Boosting est une méthode de machine learning pour des problèmes de régression et de classification. Elle est en quelque sorte une descente de gradient appliquée à une méthode de Boosting. Les méthodes de Boosting sont des techniques ensemblistes. Elles consistent à agréger des classifieurs de manière séquentielle sur un échantillon d'apprentissage dont les poids des individus sont améliorés au fur et à mesure. A chaque étape, ces classifieurs sont modifiés selon leurs performances. L'idée est d'optimiser une fonction de coût global sur les prédictions faites. Cette méthode tente d'ajuster chaque nouveau prédicteur aux erreurs résiduelles faites par le prédicteur précédent.

Pour expliquer cette méthode plus précisément mettons-nous dans le cadre d'une régression des moindres carrés. Dans ce cas, le but est d'apprendre à un modèle F de prédire la valeur de la variable expliquée, noté $\hat{y}$, connaissant les valeurs des variables explicatives x (i.e $\hat{y} = F(x)$) en minimisant l'erreur quadratique moyenne définie par :

$$\frac{1}{n}\sum_{i=0}^{n}(Y_i - \hat{Y}_i)^2$$

A chaque étape m, $1 \leq m \leq M$, de la méthode de gradient boosting, on peut supposer que le modèle F_m est imparfait. La méthode va donc améliorer F_m en construisant un nouveau modèle qui ajoute à F_m un estimateur e basé sur l'erreur de F_m : $F_{m+1}(x) = F_m(x) + e(x)$.

Pour trouver e, la méthode utilise les données de l'échantillon d'apprentissage de telle sorte que trouver un e parfait impliquerait que $F_m(x) + e(x) = y$ i.e $e(x) = y - F_m(x)$. Donc la méthode de gradient boosting prendra e comme le résidu de $y - F_m(x)$.

Ainsi comme vu précédemment, chaque F_{m+1} tente de corriger les erreurs de son prédécesseur F_m. De plus, en prenant comme fonction de coût $\frac{1}{2}(y - F(x))^2$, le résidu $y - F(x)$ n'est autre que l'opposé du gradient de la fonction de coût (en dérivant par rapport à $F(x)$). Le gradient boosting revient donc à un algorithme de descente de gradient à pas constant égal à 1 et permettant de

minimiser la fonction de coût.

D.1.2 Algorithme

Dans de nombreux problèmes d'apprentissage supervisés, il est retrouvé une variable de sortie est y et un vecteur de variables d'entrée x décrit par une distribution de probabilité $P(x, y)$. En partant d'un ensemble d'apprentissage (x_1, y_1), ...,(x_n, y_n), le but est de trouver une approximation $\hat{F}(x)$ qui minimise la valeur espérée de la fonction de coût, noté $L(y, F(x))$:

$$\hat{F} = \underset{F}{argmin}\mathbb{E}_{x,y}[L(y, F(x))]$$

La méthode du gradient boosting suppose y réel et cherche une approximation de $\hat{F}$ sous la forme d'une somme pondérée de $h_i(x)$ appartenant à la classe $\mathcal{H}$ appelée weak learner :

$$\hat{F}(x) = \sum_{i=0}^{n}(p_i h_i(x)) + cste$$

D'après le principe de minimisation du risque empirique (ERM), cette méthode essaie de trouver une approximation de $\hat{F}(x)$ qui minimise la valeur moyenne de la fonction de coût sur l'ensemble d'apprentissage c'est-à-dire le risque empirique.

Il peut alors écrire un premier algorithme sous la forme d'un problème d'optimisation :

$$\begin{cases} F_0(x) = \arg\min\limits_{\gamma} \sum\limits_{i=1}^{n} L(y_i, \gamma), \\ \\ F_m(x) = F_{m-1}(x) + \arg\min\limits_{h_m \in \mathcal{H}} \left[\sum\limits_{i=1}^{n} L(y_i, F_{m-1}(x_i) + h_m(x_i))\right]. \end{cases}$$

Trouver les h_m optimaux semble être un problème très compliqué à résoudre. Il faut plutôt se restreindre à une version simplifiée du problème d'optimisation en faisant une étape de steepest descent. Si le cas continu est considéré c'est-à-dire si $\mathcal{H}$ est l'ensemble des fonctions différentiables de $\mathbb{R}$, l'algorithme précédent peut être modifié en utilisant les deux équations suivantes :

$$\begin{cases} F_m(x) = F_{m-1}(x) - \gamma_m \sum\limits_{i=1}^{n} \nabla_{F_{m-1}} L(y_i, F_{m-1}(x_i)), \\ \\ \gamma_m = \arg\min\limits_{\gamma} \sum\limits_{i=1}^{n} L(y_i, F_{m-1}(x_i) - \gamma \nabla_{F_{m-1}} L(y_i, F_{m-1}(x_i))). \end{cases}$$

Ainsi l'algorithme de la méthode de gradient boosting est le suivant :

Entrée : ensemble d'apprentissage $(x_i, y_i)_{i=1}^{n}$, une fonction différentiable $L(y, F(x))$, nombre d'itération M

1. Initialisation :
$$\hat{F}_0(x) = \underset{\gamma}{argmin} \sum_{i=1}^{n}[L(y_i, \gamma)]$$
2. Pour $m = 1$ jusqu'à M :
   (a) Calculer le résidu : pour $i = 1, ..., n$
$$r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)}$$

(b) Ajuster un weak learner $h_m(x)$ au résidu, i.e. l'entrainer avec l'ensemble d'apprentissage $(x_i, r_{im})_{i=1}^n$;

(c) Calculer γ_m en résolvant le problème d'optimisation suivant :

$$\gamma_m = \arg\min_{\gamma} \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i))$$

(d) Mettre à jour le modèle :

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$$

3. Sortie de l'algorithme : $F_M(x)$

D.1.3 Gradient Boosted Regression Trees (GBRT)

Un choix de weak learners peut être de prendre des arbres de régressions de taille fixe. Dans ce cas, la méthode est appelée Gradient Boosted Regression Trees (GBRT).

En général à l'étape m, le gradient boosting va vouloir ajuster les arbres de régression $h_m(x)$ au résidu de l'étape précédente. Pour simplifier l'explication qu'à chaque étape la méthode construit un seul arbre de régression $h_m(x)$. Soit J_m son nombre de feuilles.

L'arbre hm$h_m(x)$ va alors partitionner l'ensemble des observations $(x_i)_i \in \mathbb{R}^d$ en J_m régions disjointes $R_{1,m}, ..., R_{J_m,m}$ et associer à chaque région une prédiction de la variable expliquée $\hat{y}_{.,m}$. En utilisant des indicatrices, sachant les variables explicatives x, l'expression de $h_m(x)$ est telle que :

$$h_m(x) = \sum_{i=1}^{J_m} (\hat{y}_{i,m} \mathbf{1}_{R_{i,m}}(x))$$

avec $\hat{y}_{i,m}$ la valeur de la prédiction de la région $R_{i,m}$.

Ainsi le modèle qui minimise la fonction de coût vu dans la partie précédente peut être modifié tel que :

$$\begin{cases} F_m(x) = F_{m-1}(x) - \gamma_m h_m(x), \\ \quad \gamma_m = \arg\min_{\gamma} \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i))). \end{cases}$$

Il est possible de choisir dans ce modèle un γ_m optimal pour l'ensemble de l'arbre. Une idée d'amélioration est de trouver J_m γ_m optimaux pour chaque prédiction $\hat{y}_{.,m}$ de l'arbre. Déterminant alors le nouveau modèle :

$$\begin{cases} F_m(x) = F_{m-1}(x) + \sum_{i=1}^{J_m} \gamma_{J_m} \mathbf{1}_{R_{J_m}}(x), \\ \quad \gamma_m = \arg\min_{\gamma} \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma). \end{cases}$$

Pour conclure cette partie, une condition de l'utilisation d'arbres de régressions pour cette méthode est de choisir une taille fixe des arbres. En posant J le nombre de nœuds terminaux, prendre $J = 2$ n'entraîne aucune interaction entre les variables et est donc dans la plupart des cas insuffisant. En général il est conseillé pour des méthodes de boosting de choisir J entre 3 et 6.

D.1.4 Régularisation

En machine learning, la notion de régularisation est très importante. Elle consiste à augmenter l'information pour résoudre des problèmes mal posés ou d'éviter le sur-apprentissage.

Nombre d'itérations M : Un premier paramètre de régularisation assez naturel est le nombre d'itérations que réalise l'algorithme du gradient boosting. Si la valeur de M est augmentée, il faut alors construire plus d'arbres ce qui va avoir pour effet de réduire l'erreur sur l'ensemble d'apprentissage. Mais prendre une valeur de M trop élevée conduit souvent à un sur-apprentissage qu'il est souhaitable d'éviter. Une solution à ce problème est souvent d'utiliser un ensemble de validation, séparé de l'ensemble d'apprentissage, et d'optimiser l'erreur sur cette ensemble en fonction des valeurs de M. Ainsi il en ressort une sorte de double optimisation : l'une sur l'ensemble d'apprentissage et l'autre sur celui de validation.

Shrinkage : Une autre méthode importante de régularisation de la méthode de gradient boosting est par shrinkage (taux d'apprentissage). Le modèle est modifiable comme suit :

$$F_m(x) = F_{m-1}(x) + r.\gamma_m h_m(x)$$

où $0 < r \leq 1$ est le taux d'apprentissage.

De façon empirique, une valeur basse du shrinkage ($< 0, 1$) demande plus d'arbres pour ajuster le modèle mais permet souvent d'obtenir des prédictions meilleures.

Nombre d'observations dans les feuilles : Une dernière méthode de régularisation est le choix du nombre d'observations minimal dans chaque feuille. Cela permet d'éviter une complexité trop importante des arbres et d'obtenir ainsi des prédictions par région peu significatives par rapport à l'ensemble d'apprentissage. Ceci permet alors d'empêcher le surapprentissage.

D.1.5 Avantages et inconvénients

Les GBM permettent la construction séquentielle d'arbres de décision peu profonds, chaque arbre essayant de corriger là où son prédécesseur a fait les plus grosses erreurs.

La méthode est très utilisée en machine learning car elle possède de nombreux avantages :

- Elle permet d'implémenter des problèmes non linéaires ;
- L'utilisateur peut entrer de nombreux paramètres (nombre d'arbres, taux d'apprentissage, profondeur des arbres..), ce qui permet d'adapter plus facilement la méthode au problème voulu ;
- La méthode fonctionne avec particulièrement bien les variables quantitatives, mais elle gère également les variables catégorielles ;
- Elle gère les données manquantes ;
- Elle renvoie de nombreuses analyses (l'influence relative des variables, les graphes de dépendance partielle...).

Cependant, cette méthode présente également quelques inconvénients qu'il faut savoir gérer :

- La méthode est lourde car le coût de calcul est très élevé. Il est souvent demandé aux GBM de construire un grand nombre d'arbres (généralement plus de 1000) ;
- Il peut être compliqué pour de gérer tous les paramètres : pour rappel, l'utilisateur doit en entrer un grand nombre (shrinkage, bagging, nombre d'arbres...) ;
- Le GBM est très sensible aux valeurs aberrantes puisque chaque arbre est obligé de corriger les erreurs dans les arbres qui le précèdent.

Annexe E

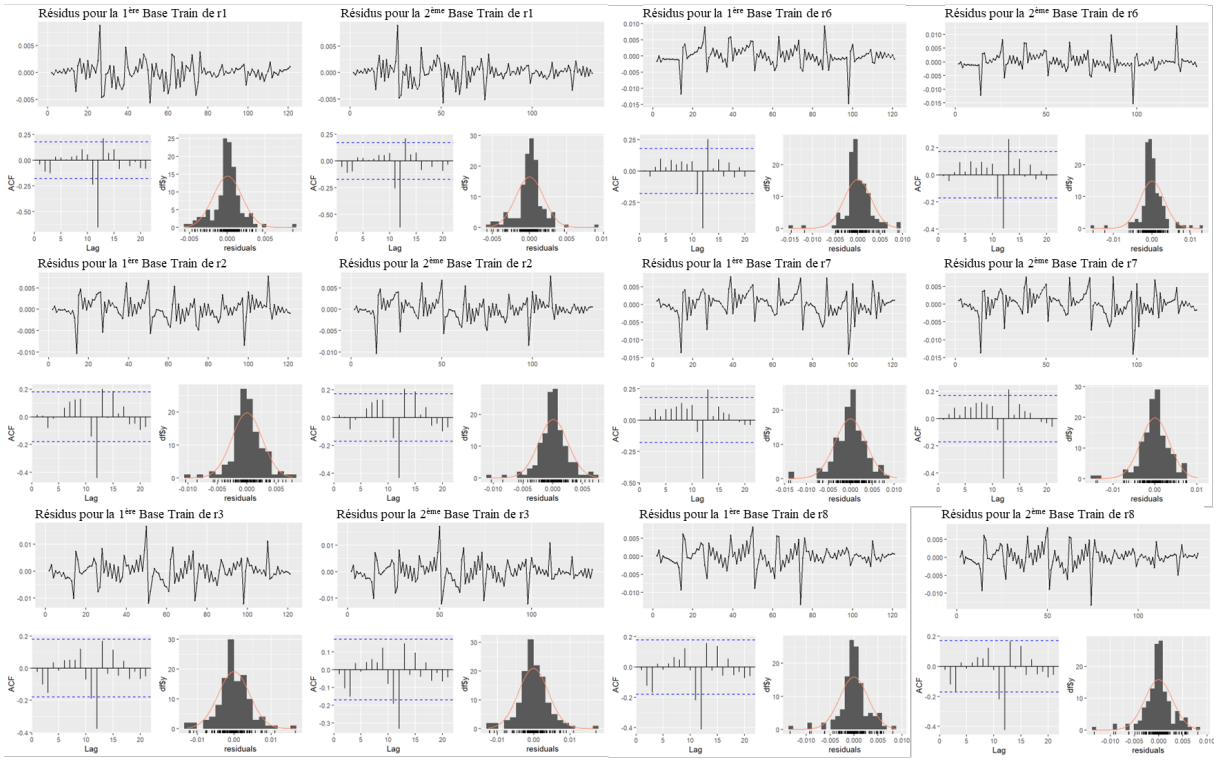
CHAPITRE 5 : APPLICATION : Construction de la loi de Rachat

Ancienneté	Best Model	AIC	BIC	RMSE	MAE	Ljung.Box p.value	ADF p.value	KPSS p.value
1	ARIMA(4,1,2)	-1115.07	-1081.72	0.19%	0.16%	0.18	0.01	0.10
2	ARIMA(2,1,2)	-1053.18	-1025.39	0.13%	0.11%	0.30	0.01	0.10
3	ARIMA(4,1,2)	-926.72	-893.38	0.40%	0.29%	0.12	0.01	0.10
4	ARIMA(2,1,1)	-855.21	-830.20	0.17%	0.14%	0.14	0.01	0.10
5	ARIMA(3,1,2)	-906.09	-875.52	0.83%	0.64%	0.39	0.01	0.10
6	ARIMA(2,1,3)	-1010.43	-979.86	0.54%	0.45%	0.07	0.01	0.10
7	ARIMA(3,0,1)	-979.95	-949.29	0.23%	0.18%	0.12	0.01	0.10
8	ARIMA(2,1,1)	-1001.25	-976.24	0.18%	0.15%	0.43	0.01	0.10
9	ARIMA(2,1,1)	-918.94	-893.93	0.52%	0.42%	0.27	0.01	0.10
10	ARIMA(4,1,1)	-1009.44	-978.87	0.40%	0.35%	0.26	0.01	0.10
11	ARIMA(3,0,2)	-1056.62	-1023.17	0.40%	0.31%	0.58	0.01	0.10
12	ARIMA(1,1,3)	-1107.86	-1080.07	0.09%	0.08%	0.14	0.01	0.10
13	ARIMA(4,1,3)	-925.79	-889.66	0.45%	0.40%	0.07	0.01	0.10
14	ARIMA(2,0,1)	-1014.74	-986.87	0.50%	0.42%	0.07	0.01	0.10
15	ARIMA(2,1,2)	-944.43	-916.64	0.28%	0.19%	0.94	0.01	0.10
16	ARIMA(2,0,2)	-870.63	-839.97	0.52%	0.34%	0.33	0.02	0.10
17	ARIMA(4,1,1)	-875.71	-845.14	0.50%	0.45%	0.47	0.01	0.10
18	ARIMA(1,1,3)	-401.04	-373.25	0.56%	0.48%	0.14	0.02	0.10
19	ARIMA(1,1,3)	-477.06	-449.27	0.46%	0.34%	0.16	0.02	0.10
20	ARIMA(4,0,1)	-543.53	-510.08	2.48%	2.15%	0.91	0.01	0.10

TABLE E.1 – Résultats d'analyse de la Base Train 2021-22 pour chaque année d'ancienneté

Ancienneté	Best Model	AIC	BIC	RMSE	MAE	Ljung.Box p.value	ADF p.value	KPSS p.value
1	ARIMA(4,1,2)	-1236.24	-1201.74	0.13%	0.11%	0.24	0.01	0.10
2	ARIMA(2,1,2)	-1173.90	-1145.15	0.10%	0.08%	0.24	0.01	0.10
3	ARIMA(4,1,2)	-1033.01	-998.51	0.23%	0.20%	0.06	0.01	0.10
4	ARIMA(2,1,1)	-955.30	-929.42	0.73%	0.52%	0.12	0.01	0.10
5	ARIMA(3,1,2)	-1015.98	-984.36	0.55%	0.37%	0.06	0.01	0.10
6	ARIMA(2,1,3)	-1106.55	-1074.93	0.15%	0.13%	0.11	0.01	0.10
7	ARIMA(3,0,1)	-1090.66	-1058.95	0.35%	0.24%	0.12	0.04	0.10
8	ARIMA(2,1,1)	-1112.33	-1086.45	0.16%	0.13%	0.35	0.01	0.10
9	ARIMA(2,1,1)	-1024.13	-998.26	0.41%	0.35%	0.22	0.01	0.10
10	ARIMA(4,1,1)	-1116.91	-1085.28	0.44%	0.37%	0.25	0.01	0.10
11	ARIMA(3,0,2)	-1151.49	-1116.90	0.34%	0.24%	0.15	0.05	0.10
12	ARIMA(1,1,3)	-1231.39	-1202.64	0.15%	0.13%	0.10	0.03	0.10
13	ARIMA(4,1,3)	-1032.78	-995.40	0.40%	0.26%	0.06	0.01	0.10
14	ARIMA(2,0,1)	-1101.24	-1072.41	0.82%	0.59%	0.18	0.02	0.10
15	ARIMA(2,1,2)	-1056.97	-1028.22	0.64%	0.42%	0.56	0.01	0.10
16	ARIMA(2,0,2)	-970.40	-938.69	0.43%	0.38%	0.24	0.02	0.10
17	ARIMA(4,1,1)	-976.25	-944.62	0.69%	0.45%	0.45	0.01	0.10
18	ARIMA(1,1,3)	-454.66	-425.91	0.21%	0.18%	0.11	0.03	0.10
19	ARIMA(1,1,3)	-539.18	-510.42	0.35%	0.29%	0.13	0.02	0.10
20	ARIMA(4,0,1)	-576.08	-541.48	0.96%	0.82%	0.59	0.02	0.10

TABLE E.2 – Résultats d'analyse de la Base Train 2022-23 pour chaque année d'ancienneté



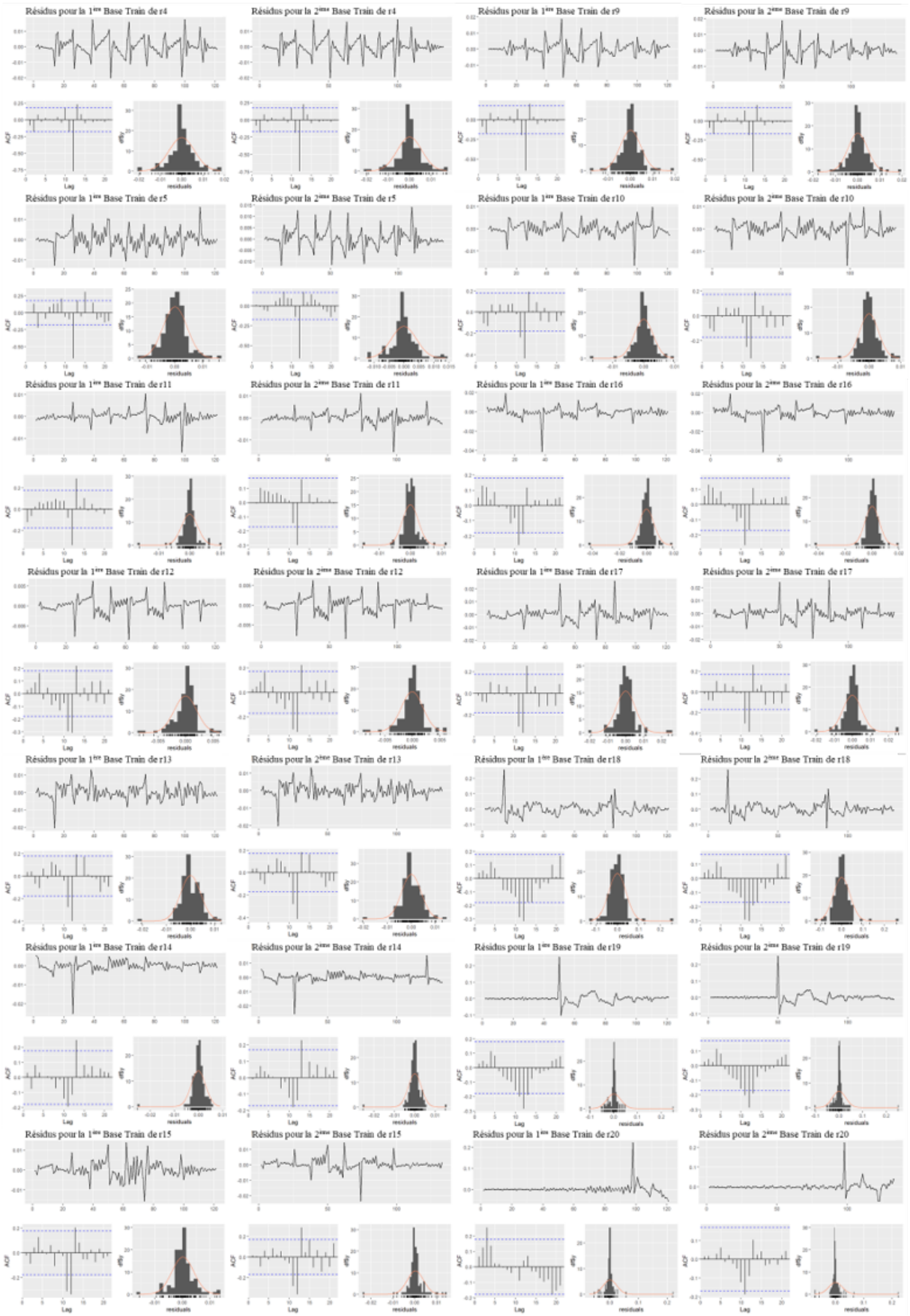


FIGURE E.1 – Tendance, ACF et histogramme des résidus des modèles ARIMAX des 1 à 20 ans d'ancienneté de prêt sur les deux Bases Trains de la validation croisée

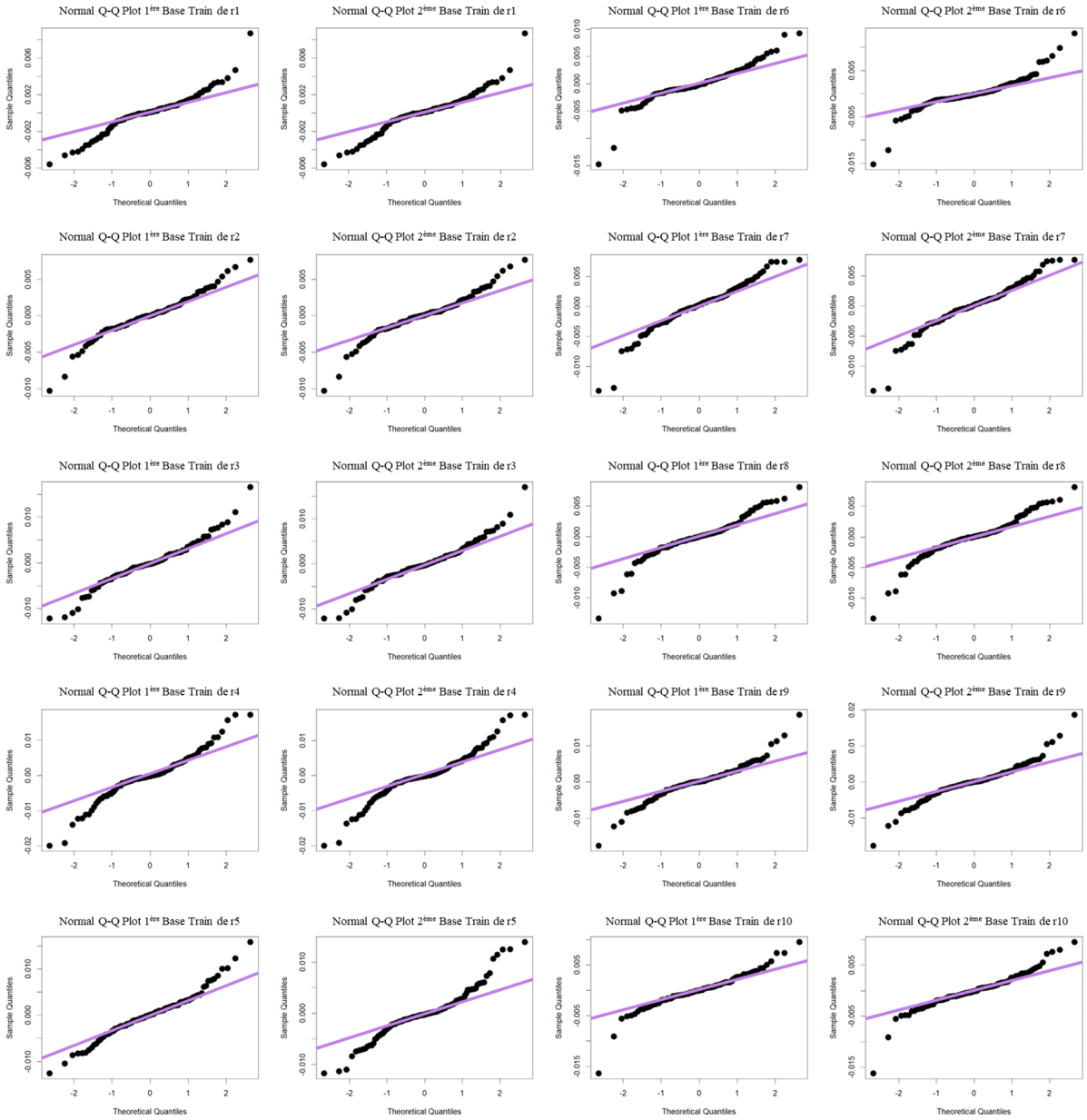


FIGURE E.2 – QQ-Plot des résidus des modèles ARIMAX des 1 à 10 ans d'ancienneté de prêt sur les deux Bases Trains de la validation croisée

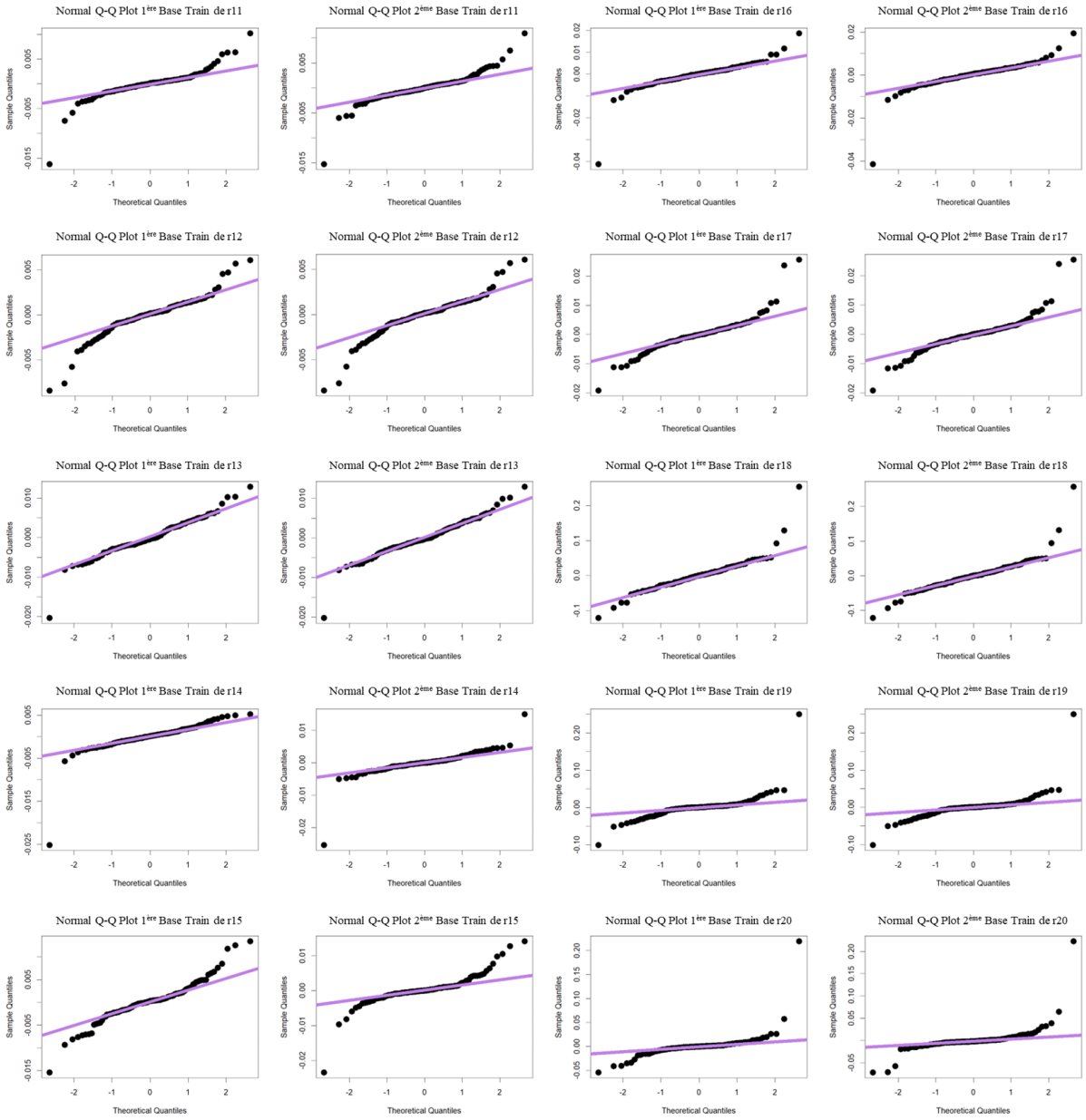


FIGURE E.3 – QQ-Plot des résidus des modèles ARIMAX des 11 à 20 ans d'ancienneté de prêt sur les deux Bases Trains de la validation croisée

Annexe F

CHAPITRE 6 : Impact du contexte économique sur un portefeuille d'assurance emprunteur

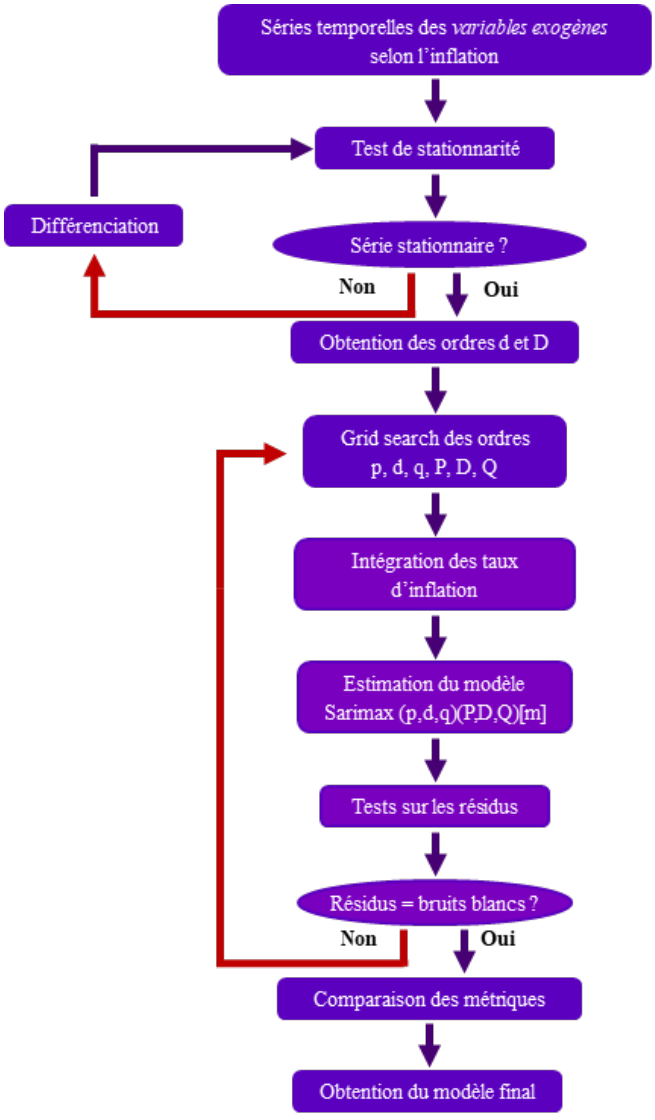


FIGURE F.1 – Schéma du processus de modélisation du GSE

	Model	AIC	BIC	MSE	MAE	RMSE	MAPE	Ljung.Box p.value	ADF.test p.value	KPSS.test p.value
x_1	Sarimax(4,0,3)(2,0,2)[12]	-2 439.73	-2 393.50	0.001%	0.242%	0.308%	14.77%	0.71	0.01	0.10
x_2	Sarimax(4,0,3)(0,0,2)[12]	3 664.65	3 704.28	21 479 002	3 227	4 524	17.69%	0.62	0.01	0.10
x_3	Sarimax(4,0,1)(0,1,2)[12]	3 464.29	3 493.45	5 117 605	1 647	2 260	10.12%	0.86	0.01	0.10
x_4	Sarimax(2,1,4)(0,0,0)[12]	5 298.10	5 324.48	80 567 006 411	197 127	278 781	17.26%	0.56	0.01	0.10
x_5	Sarimax(2,0,4)(0,1,1)[12]	5 008.23	5 037.39	17 161 132 694	93 660	130 937	9.25%	0.87	0.01	0.10
x_6	Sarimax(1,2,3)(0,2,2)[12]	3 129.67	3 154.97	7 135 352	2 045	2 666	13.81%	0.23	0.01	0.10
x_7	Sarimax(3,1,2)(0,0,2)[12]	3 229.67	3 259.34	2 173 193	1 130	1 464	9.05%	0.67	0.01	0.10
x_8	Sarimax(2,1,4)(2,1,1)[12]	5 041.44	5 077.03	98 750 822 922	235 106	313 464	14.11%	0.67	0.01	0.10
x_9	Sarimax(4,1,0)(2,0,2)[12]	5 158.15	5 191.12	39 895 608 645	166 160	195 172	11.34%	0.11	0.01	0.10
x_{10}	Sarimax(3,0,2)(2,1,2)[12]	- 131.22	- 95.57	3.48	1.35	1.77	3.30%	0.89	0.01	0.10
x_{11}	Sarimax(4,1,4)(2,2,0)[12]	9.68	47.71	7.80	2.25	2.75	1.93%	0.22	0.01	0.10
x_{12}	Sarimax(2,0,4)(0,0,1)[12]	47.15	8 0.17	6.15	1.93	2.35	1.66%	0.99	0.01	0.09
x_{13}	Sarimax(3,2,2)(2,1,2)[12]	- 180.36	- 144.83	2.75	1.28	1.56	1.03%	0.61	0.01	0.10
x_{14}	Sarimax(4,0,3)(1,0,0)[12]	- 208.75	- 172.42	38.05	4.43	5.83	3.51%	0.62	0.01	0.10
x_{15}	Sarimax(1,0,0)(2,0,2)[12]	- 280.45	- 254.03	25.32	3.74	4.67	2.96%	0.13	0.01	0.10
x_{16}	Sarimax(4,0,0)(0,1,1)[12]	3 031.83	3 054.51	31 709 592	4 424	5 571	20.09%	0.07	0.01	0.10
x_{17}	Sarimax(4,0,4)(0,0,1)[12]	3 169.48	3 209.11	27 770 965	4 215	5 263	17.60%	0.25	0.01	0.10
x_{18}	Sarimax(3,2,4)(2,0,0)[12]	3 119.17	3 155.38	41 111 621	4 882	5 965	5.15%	0.98	0.01	0.10
x_{19}	Sarimax(3,0,4)(2,0,0)[12]	- 377.24	- 337.61	3.48	1.44	1.79	36.76%	0.06	0.01	0.10
x_{20}	Sarimax(1,0,3)(1,0,2)[12]	2 450.01	2 483.03	199 517	367.58	443.27	22.69%	0.22	0.01	0.10
x_{21}	Sarimax(3,0,4)(2,0,1)[12]	2 313.75	2 356.68	391 174	552.49	621.17	35.45%	0.84	0.01	0.10
x_{22}	Sarimax(2,0,2)(1,2,2)[12]	2 166.13	2 194.70	657 746	615.62	734.78	8.27%	0.88	0.01	0.10
x_{23}	Sarimax(4,0,4)(2,2,2)[12]	- 291.86	- 247.41	1.72	1.07	1.28	24.65%	0.48	0.01	0.10
x_{24}	Sarimax(3,1,3)(1,1,2)[12]	-2 227.25	-2 191.67	0.003%	0.436%	0.517%	5.70%	1.00	0.01	0.10
x_{25}	Sarimax(1,0,2)(2,1,2)[12]	-1 718.62	-1 689.46	0.963%	8.013%	8.720%	7.08%	0.56	0.01	0.10
x_{27}	Sarimax(2,1,4)(1,0,1)[12]	- 514.10	- 481.13	7 136	63.18	76.80	2.60%	0.63	0.01	0.10
x_{28}	Sarimax(2,0,4)(2,2,2)[12]	-3 049.78	-3 011.68	0.011%	0.815%	1.015%	125.43%	0.28	0.01	0.09
x_{29}	Sarimax(4,0,4)(0,0,1)[12]	-3 837.94	-3 798.31	0.069%	1.921%	2.335%	10.29%	0.50	0.01	0.09

TABLE F.1 – Résultats d'analyse des validations croisées des Sarimax des variables économiques retenus pour le GSE

	Effectif		2023	2024	2025	2026	2027	2028
Projection d'inflation 0%	**Stock**	**En cours**	15 000	14 227	13 644	12 985	12 412	11 794
		Décès	-	24	30	16	34	26
		Incapacité	-	42	39	49	49	38
		Décès Incap	-	1	1	2	-	-
		Invalidité	-	-	-	3	7	10
		Décès Inval	-	-	-	-	-	1
		Racheté	-	530	520	627	566	574
		Terminé	-	176	154	162	170	316
		Total	**15 000**	**15 000**	**14 388**	**13 844**	**13 238**	**12 759**
	Affaires nouvelles annuelles	**En cours**	-	118	157	196	281	360
		Décès	-	-	-	-	-	-
		Incapacité	-	-	-	-	1	3
		Décès Incap	-	-	-	-	-	-
		Invalidité	-	-	-	-	-	-
		Décès Inval	-	-	-	-	-	-
		Racheté	-	1	4	5	9	15
		Terminé	-	-	-	-	-	-
		Total	**-**	**119**	**161**	**201**	**291**	**378**
	Total	**Total**	**15 000**	**15 119**	**14 549**	**14 045**	**13 529**	**13 137**

TABLE F.2 – Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 0%

	Effectif		2023	2024	2025	2026	2027	2028
Projection d'inflation 0.65%	Stock	En cours	15 000	14 231	13 648	13 078	12 435	11 797
		Décès	-	26	27	20	31	24
		Incapacité	-	44	48	52	57	39
		Décès Incap	-	1	-	3	1	-
		Invalidité	-	-	1	2	8	11
		Décès Inval	-	-	-	-	-	1
		Racheté	-	523	509	536	606	564
		Terminé	-	175	160	160	181	327
		Total	**15 000**	**15 000**	**14 393**	**13 851**	**13 319**	**12 763**
	Affaires nouvelles annuelles	En cours	-	115	151	183	254	325
		Décès	-	-	-	-	1	-
		Incapacité	-	-	1	1	-	1
		Décès Incap	-	-	-	-	-	-
		Invalidité	-	-	-	-	-	-
		Décès Inval	-	-	-	-	-	-
		Racheté	-	3	2	3	8	11
		Terminé	-	-	-	-	-	-
		Total	**-**	**118**	**154**	**187**	**263**	**337**
	Total	**Total**	**15 000**	**15 118**	**14 547**	**14 038**	**13 582**	**13 100**

TABLE F.3 – Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 0.65%

	Effectif		2023	2024	2025	2026	2027	2028
Projection d'inflation 1.3%	Stock	En cours	15 000	14 197	13 656	13 085	12 547	11 870
		Décès	-	20	30	19	32	24
		Incapacité	-	52	48	50	51	43
		Décès Incap	-	2	-	4	1	-
		Invalidité	-	1	1	3	5	8
		Décès Inval	-	-	-	-	-	-
		Racheté	-	557	470	539	496	585
		Terminé	-	171	162	159	187	312
		Total	**15 000**	**15 000**	**14 367**	**13 859**	**13 319**	**12 842**
	Affaires nouvelles annuelles	En cours	-	117	149	174	231	278
		Décès	-	-	-	-	-	-
		Incapacité	-	-	1	1	-	-
		Décès Incap	-	-	-	-	-	-
		Invalidité	-	-	-	-	-	-
		Décès Inval	-	-	-	-	-	-
		Racheté	-	-	4	6	8	15
		Terminé	-	-	-	-	-	-
		Total	**-**	**117**	**154**	**181**	**239**	**293**
	Total	**Total**	**15 000**	**15 117**	**14 521**	**14 040**	**13 558**	**13 135**

TABLE F.4 – Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 1.3%

	Effectif		2023	2024	2025	2026	2027	2028
Projection d'inflation 3%	Stock	En cours	15 000	14 216	13 681	13 109	12 553	11 891
		Décès	-	28	24	22	31	23
		Incapacité	-	37	53	54	50	38
		Décès Incap	-	1	-	4	1	-
		Invalidité	-	-	1	3	7	10
		Décès Inval	-	-	-	-	-	1
		Racheté	-	542	452	547	523	515
		Terminé	-	176	159	158	175	324
		Total	**15 000**	**15 000**	**14 370**	**13 897**	**13 340**	**12 802**
	Affaires nouvelles annuelles	En cours	-	113	156	168	180	205
		Décès	-	-	-	1	-	1
		Incapacité	-	-	-	-	-	-
		Décès Incap	-	-	-	-	-	-
		Invalidité	-	-	-	-	-	-
		Décès Inval	-	-	-	-	-	-
		Racheté	-	4	6	5	12	9
		Terminé	-	-	-	-	-	-
		Total	**-**	**117**	**162**	**174**	**192**	**215**
	Total	**Total**	**15 000**	**15 117**	**14 532**	**14 071**	**13 532**	**13 017**

TABLE F.5 – Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 3%

	Effectif		2023	2024	2025	2026	2027	2028
Projection d'inflation 4%	Stock	En cours	15 000	14 228	13 663	13 081	12 553	11 860
		Décès	-	26	26	21	30	26
		Incapacité	-	44	48	54	52	39
		Décès Incap	-	1	-	4	1	-
		Invalidité	-	-	1	2	6	10
		Décès Inval	-	-	-	-	-	1
		Racheté	-	523	492	546	487	525
		Terminé	-	178	158	160	169	312
		Total	**15 000**	**15 000**	**14 388**	**13 868**	**13 298**	**12 773**
	Affaires nouvelles annuelles	En cours	-	113	154	154	157	168
		Décès	-	-	-	-	-	-
		Incapacité	-	-	-	-	-	-
		Décès Incap	-	-	-	-	-	-
		Invalidité	-	-	-	-	-	-
		Décès Inval	-	-	-	-	-	-
		Racheté	-	3	2	7	5	5
		Terminé	-	-	-	-	-	-
		Total	**-**	**116**	**156**	**161**	**162**	**173**
	Total	**Total**	**15 000**	**15 116**	**14 544**	**14 029**	**13 460**	**12 946**

TABLE F.6 – Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 4%

	Effectif		2023	2024	2025	2026	2027	2028
Projection d'inflation 5%	Stock	En cours	15 000	14 266	13 704	13 191	12 615	11 914
		Décès	-	28	19	26	28	23
		Incapacité	-	36	34	36	35	42
		Décès Incap	-	-	4	1	-	1
		Invalidité	-	-	-	-	4	10
		Décès Inval	-	-	-	-	-	1
		Racheté	-	499	489	481	516	479
		Terminé	-	171	168	155	177	323
		Total	**15 000**	**15 000**	**14 418**	**13 890**	**13 375**	**12 793**
	Affaires nouvelles annuelles	En cours	-	113	148	145	138	145
		Décès	-	-	1	-	-	-
		Incapacité	-	-	-	-	-	-
		Décès Incap	-	-	-	-	-	-
		Invalidité	-	-	-	-	-	-
		Décès Inval	-	-	-	-	-	-
		Racheté	-	3	3	3	1	2
		Terminé	-	-	-	-	-	-
		Total	**-**	**116**	**152**	**148**	**139**	**147**
	Total	**Total**	**15 000**	**15 116**	**14 570**	**14 038**	**13 514**	**12 940**

TABLE F.7 – Tableau d'effectif du portefeuille pour chaque année de projection, pour le scénario d'inflation à 5%

Bibliographie

[1] Assurance emprunteur, le classement 2024. *L'Argus de l'assurance*, 2024. https://www.argusdelassurance.com.

[2] Aurélia Cadain Alexis André : Les contrats d'assurance de groupe. *L'Argus de l'assurance*, 2014. https://www.argusdelassurance.com.

[3] Banque de France : Crédits aux particuliers avr 2024. https://www.banque-france.fr, 2024.

[4] Banque de France : Panorama des prêts à l'habitat des ménages – avril 2024. https://www.banque-france.fr, 2024.

[5] BCAC : Arrêt de travail : lois de provisionnement du bcac. https://www.ressources-actuarielles.net.

[6] CCSF : Bilan de l'assurance emprunteur - 2023. https://www.ccsfin.fr, 2024.

[7] Paul Hubert Fabien Labondance Christophe Blot, Jérôme Creel : Les effets redistributifs des politiques monétaires de la bce. *In Revue de l'OFCE*, pages 165–180. OFCE, 2018. https://www.cairn.info.

[8] DG Trésor : Finances publiques : une inflation qui rapporte ? https://www.tresor.economie.gouv.fr, 2022.

[9] ECB : Monetary policy. https://www.ecb.europa.eu.

[10] EIOPA : Report on the calculation of the ufr for 2025. https://www.eiopa.europa.eu.

[11] Empruntis : Historique des taux d'intérêt du crédit immobilier. https://www.empruntis.com.

[12] European Central Bank : Main refinancing operations. https://data.ecb.europa.eu, 2024.

[13] Fonds monétaire international : Reprise en temps de pandémie. https://www.imf.org/fr, 2021.

[14] France Assureurs : Assurance française : données clés 2023. https://www.franceassureurs.fr, 2024.

[15] Imkiz : L'impact de la hausse des taux d'intérêt sur les prix immobiliers ? https://www.imkiz.com.

[16] INSEE : Guide pratique des séries non-stationnaires. https://www.bnsp.insee.fr, 1999.

[17] INSEE : Évolution de l'indice des prix à la consommation harmonisé dans l'union européenne et la zone euro. https://www.insee.fr/fr, 2023.

[18] INSEE : Commercialisation des logements neufs - vente aux particuliers au 1er trimestre 2024. https://www.insee.fr/fr, 2024.

[19] INSEE : Comptes nationaux trimestriels au deuxième trimestre 2024. https://www.insee.fr/fr, 2024.

[20] INSEE : Construction de logements : résultats à fin mars 2024 (france entière). https://www.insee.fr/fr, 2024.

[21] INSEE : Dette au sens de maastricht des administrations publiques en points de pib. https://www.insee.fr/fr, 2024.

[22] INSEE : Glissements annuels de l'ict – salaire horaire. https://www.insee.fr/fr, 2024.

[23] INSEE : Indices de coûts et de prix dans la construction. https://www.insee.fr/fr, 2024.

[24] INSEE : Indices des prix des logements neufs et indices notaires-insee des prix des logements anciens. https://www.insee.fr/fr, 2024.

[25] INSEE : Indices des prix à la consommation harmonisés. https://www.insee.fr/fr, 2024.

[26] INSEE : L'essentiel sur... le chômage. https://www.insee.fr/fr, 2024.

[27] INSEE : L'essentiel sur... l'inflation. https://www.insee.fr/fr, 2024.

[28] INSEE : Taux d'épargne des ménages. https://www.insee.fr/fr, 2024.

[29] INSEE : Évolution du produit intérieur brut et de ses composantes. https://www.insee.fr/fr, 2024.

[30] INSEE : Évolution du revenu disponible brut et du pouvoir d'achat. https://www.insee.fr/fr, 2024.

[31] Lambolez, G. : Construction d'un modèle à état prévoyance et application à la compréhension du résultat ifrs 17. Mémoire d'actuariat, Optimind, 2024.

[32] L'Observatoire Crédit Logement : Publication mensuelle. https://www.lobservatoirecreditlogement.fr.

[33] Légifrance : Loi n°2010-737 du 1er juillet 2010 portant réforme du crédit à la consommation. https://www.legifrance.gouv.fr, 2010.

[34] Légifrance : Loi n°2014-344 du 17 mars 2014 relative à la consommation. https://www.legifrance.gouv.fr, 2014.

[35] Légifrance : Loi n°2017-203 du 21 février 2017 ratifiant les ordonnances n°2016-301 du 14 mars 2016 relative à la partie législative du code de la consommation et n°2016-351 du 25 mars 2016 sur les contrats de crédit aux consommateurs relatifs aux biens immobiliers à usage d'habitation et simplifiant le dispositif de mise en œuvre des obligations en matière de conformité et de sécurité des produits et services. https://www.legifrance.gouv.fr, 2017.

[36] Légifrance : Loi n°2022-270 du 28 février 2022 pour un accès plus juste, plus simple et plus transparent au marché de l'assurance emprunteur. https://www.legifrance.gouv.fr, 2022.

[37] Ministère de l'Intérieur et de Outre-Mer : Cvs-cjo. https://mobile.interieur.gouv.fr, 2011.

[38] Thomas Moller-Nielsen : Bce : l'effet de la politique monétaire sur l'économie européenne a atteint son point culminant. *EurActiv*, 2024. https://www.euractiv.fr.

[39] OFDT : Tabagisme et arrêt du tabac en 2023. https://www.ofdt.fr.

[40] Organisation internationale du Travail : Rapport mondial sur les salaires 2022-23. https://webapps.ilo.org, 2023.

[41] Patrice Bertrand Denis Pasquignon : Analyse des données. *Université Paris-Dauphine*, 2022.

[42] Perplexity : Perplexity.ai. https://www.perplexity.ai.

[43] Pretto : Retour de l'inflation : une bonne nouvelle pour les emprunteurs ? https://www.pretto.fr, 2024.

[44] Quentin Lajaunie : Introduction à l'analyse des séries temporelles et applications sur le logiciel r. *Université Paris-Dauphine*, 2024.

[45] Rdocumentation : arima : Arima modelling of time series. https://www.rdocumentation.org.

[46] Rdocumentation : arima : Arima modelling of time series. https://www.rdocumentation.org.

[47] Rdocumentation : checkresiduals : Check that residuals from a time series model look like white noise. https://www.rdocumentation.org.

[48] Rdocumentation : kpss.test : Kpss test for stationarity. https://www.rdocumentation.org.

[49] Rdocumentation : prcomp : Principal components analysis. https://www.rdocumentation.org.

[50] Rdocumentation : ur.df : Augmented-dickey-fuller unit root test. https://www.rdocumentation.org.

[51] Quentin Romet : Analyse : comment les taux d'intérêt influencent le marché immobilier ? *Homunity.com : crowdfunding immobilier*, 2024. https://www.homunity.com/fr.

[52] Simulassur : Assurance emprunteur : quels sont les impacts de la loi lemoine 2022 ? https://www.simulassur.fr.

[53] Sixte JONGLEZ de LIGNE, Léa KARIM, Sophie PIGEON : Méthode de gradient boosting. *Université Paris-Dauphine*, 2020.

[54] SPAC Actuariat : Table de mortalité. https://www.spac-actuaires.fr.

[55] Yeva S. Nersisyan L. Randall Wray : Un excès de dette publique handicape-t-il réellement la croissance ? *In Revue de l'OFCE*, pages 173–190. OFCE, 2011. https://www.cairn.info.

[56] Zeste de savoir : Interpolation avec des splines cubiques d'hermite. https://zestedesavoir.com, 2022.