



ISUP

PROMOTION 2009

Mémoire présenté devant

**L'Institut de Statistique
de l'Université Pierre et Marie Curie**

Pour l'obtention du

Diplôme de Statisticien Mention Actuariat

***A s s u r a n c e* □**

***F i n a n c e* □**

Par M. Antoine CHIMON

Sujet : Etude du comportement des investisseurs dans le cadre de la gestion du risque de liquidité

Lieu du stage : HSBC Global Asset Management

Responsable du stage : Mme Valérie HASSID

Invité(s) :

CONFIDENTIEL □

A ma famille, avec toute ma reconnaissance.

Remerciements

Je tiens à remercier en tout premier lieu mon maître de stage, Valérie Hassid, pour sa confiance, sa patience et les conseils prodigués tout au long du stage.

Ma reconnaissance s'adresse également à Mehdy Latreche qui, par sa collaboration et ses conseils, m'a permis de mener à bien les travaux qui m'avaient été confiés.

Je remercie également le personnel des différents services de la société HSBC Global Asset Management pour leur accueil et leur disponibilité.

Enfin, je remercie l'ensemble du corps professoral de l'ISUP, dont Jacques Chevalier, mon tuteur de mémoire, pour le savoir et les techniques qu'ils m'ont inculqués.

Résumé

La récente crise financière a multiplié les besoins d'expertise en analyse du risque dans tous les domaines de la banque.

Jadis plutôt réservé au domaine de la banque d'investissement, ces pratiques se sont étendues au secteur de la gestion d'actif il y a peu de temps.

Un des principaux objectifs du département risque d'une société de gestion d'actif est de faire face au **risque de liquidité**.

Le risque de liquidité est le risque porté par un actif qui ne peut être liquidé rapidement par manque d'acheteurs. Ce risque est d'autant plus important en gestion d'actif que le gérant peut être confronté au rachat d'encours de certains clients, ce qui lui imposerait de disposer de suffisamment d'actifs liquides pour être en mesure de les rembourser et ceci en tenant compte d'un principe fondamental de la gestion d'actif : **l'égalité des porteurs de part**.

En effet, c'est à ce principe que doit répondre toute société de gestion qui a dans son portefeuille de clients aussi bien des banques privées que des investisseurs institutionnels ou des particuliers. Ces clients doivent être traités indépendamment de leur statut financier ou de leurs poids dans le fonds en question (notamment dans le remboursement du rachat des clients qui ne doit donner ni priorité en temps ni en juste valeur du remboursement).

Dans ce contexte il est intéressant de se pencher sur le comportement des différents clients en terme de souscription et de rachat pour pouvoir les anticiper (dans une certaine mesure) et ainsi être en mesure de disposer de suffisamment d'actifs liquide pour faire face à ces rachats et respecter ainsi le principe d'égalité des porteurs de parts.

Ce mémoire a donc pour objectif principal d'étudier le comportement des différentes catégories de clients à travers l'amplitude de leurs souscriptions et de leurs rachats, pour pouvoir proposer in fine plusieurs scénarios de rachats qui pourront orienter la gestion du risque de liquidité.

Pour mettre en place ces différents scénarios de rachats, nous nous appuierons en particulier sur certains modèles statistiques tels que les modèles autorégressifs ou les régressions linéaires.

Nous tenterons également d'analyser, dans une dernière partie, dans quelle mesure la performance d'un fonds peut influencer sur le comportement des différents investisseurs.

Mots clés : risque de liquidité, stress test, modèles autorégressifs, modèles ARMA, comportement de l'investisseur, aversion au risque, scénarios de rachat, performance.

Abstract

The recent financial crisis has increased the need of a deep analysis in every field of the banking industry.

Even though these technics of risk watching were used to be developped in the CIB industry, it has recently emerged in asset management companies.

One of the main targets of the risk department of an asset management company is to cope with the **liquidity risk**.

The liquidity risk is the risk that a given security or asset cannot be traded quickly enough in the market to prevent a loss. This risk is major in the asset management business, because asset managers may face the repurchase of the customers parts of the fund in which they have signed in. In this case, he would have to possess enough liquid assets in order to reimburse every customer that want to be paid off.

This is a legal issue that asset managers have to respect at the risk of being sued by the regulator. Every customer must be treated in an equal way, whatever an individual, a bank or an institutional investor. The managers must consider customers independently of their weight in the given fund or their financial reputation, especially when it comes to the fair value of the repayment of customers' parts.

In this context, it is crucial to focus on customer's behavior in terms of investments and on the differences between each customer group. The study of behavior's investors in subscription or in repurchase of their fund aims at anticipating customer's repurchases (to a certain extent) and so to be in possession of enough liquid assets in order to cope with every customer's repurchase and to respect the equality of every customer.

This report aims at studying customer groups behaviors through the amplitude of their subscriptions and their repurchases in order to be able to propose repurchase scenarios that would direct liquidity risk management. We will also set stress tests in order to monitor the impact of huge paid off on the liquidity of the funds.

In order to set these scenarios, we will use statistical models and especially autoregressive models (such as Arma models: Autoregressive moving average) or linear regressions.

In the last part of our study, we will attempt to analyze how the performance of a fund can influence investor behavior in repurchases or subscriptions

Key words : Liquidity risk, stress test, autoregressive models, Arma models, investor behavior, risk aversion, repurchase scenarios.

Sommaire

Résumé	3
Abstract.....	5
Sommaire	7
Généralités sur la gestion d’actifs.....	10
I.1 Types de fonds	11
I.2 Impacts de la crise : focus sur le risk management	12
I.3 La gestion du risque de liquidité	13
I.3 a/ Détermination pratique de la liquidité d’un actif	13
I.3 b/ Lien avec le passif.....	14
I.3 c/ Les récentes évolutions réglementaires.....	15
Présentation des données.....	16
II.1 Présentation des différentes catégories de clients.....	17
II.2 Evolution de la répartition des encours des catégories de client par type de fonds.....	17
II.2 a/ Les entreprises bancaires (<i>GBM</i>).....	17
II.2 b/ Les investisseurs institutionnels (<i>GBM-Instit</i>).....	18
II.2 c/ Les particuliers (<i>PFS</i>)	19
II.2 d/ La banque privée (<i>GPB</i>)	20
II.2 e/ Synthèse	21
II.3 Théorie de la décision en environnement incertain	22
II.3 a/ L’axiomatique de von Neumann et Morgenstern	22
II.3 b/ Les comportements des agents face à l’incertain.....	24
II.4 Périmètre de l’étude	25
II.5 Détermination du comportement des investisseurs par rapport à leurs encours	26
Mise en place de scénarios de rachats	29
III.1 Mise en place de stress tests	30
III.1 a/ Stress test 1	30
III.1 b/ Stress test 2.....	31
III.2 Utilisation des processus autorégressifs pour modéliser les comportements des investisseurs	33
III.2 a/ Rappels : outils statistiques.....	33

III.2 b/ Introduction au concept de modélisation temporelle et processus ARMA	35
III.2 c/ Procédure	36
III.2 d/ Identification	37
III.2 d/ i. Estimation du paramètre d'intégration d.....	37
III.2 d/ii. Estimation des ordres p et q d'un modèle ARMA (p, q).....	42
III.2 e/ Détermination des paramètres	46
III.2 f/ Vérifications à posteriori : Tests d'adéquation de modèles.....	49
III.2 f/ i. Test de bruit blanc et de stationnarité	49
III.2 f/ ii. Normalité des résidus	52
III.2 g/ Etude de la stabilité temporelle des modèles utilisés	57
III.2 g/ i. Détermination graphique de la stabilité.....	57
III.2 g/ ii. Détermination de la stabilité des modèles à l'aide de tests statistiques	58
III. 2 g/ Prévisions à l'aide d'un modèle ARMA (p, q).....	61
III. 3 Synthèse	65
Etude de l'influence de la performance des fonds sur le comportement des investisseurs	66
IV. 1 Contexte	67
IV. 2 Outils de mesure de performance d'un OPCVM	67
IV. 2 a/ La performance nette d'un OPCVM	67
IV. 2 b/ La comparaison avec le benchmark	68
IV. 2 c/ La comparaison avec d'autres OPCVM.....	68
IV. 3 Présentations des données	69
IV.3 a/ Fonds retenus pour l'étude	69
IV.3 b/ Calcul des performances glissantes.....	69
IV.4 Etude de la corrélation entre les performances glissantes des fonds et le pourcentage de souscription/rachat	72
IV.5 Etude de la relation de causalité entre performance et souscription/rachat.....	74
IV.5 a/ Définition de la causalité au sens de Granger	74
IV.5 b/ Conclusion sur la causalité.....	77
IV.5 c/ Etude de la relation entre les comportements d'investissement et la différence de performance du fonds et du benchmark	77
Conclusion générale	79
Annexes	82
Annexe N°1	83

Annexe N° 2 : tables statistiques utilisées.....	85
Table de Shapiro-Wilk	85
Table de Fisher	87
Annexe N°3	88
Annexe N°4.....	90
Bibliographie	91

Première partie

Généralités sur la gestion d'actifs

L'Asset Management, en français la « Gestion d'Actifs » est sans aucun doute la seule partie des marchés financiers connue de tous les français.

Filiale d'une banque ou d'un assureur (le statut de filiale est obligatoire depuis 1999), la société de gestion d'actifs créée et gère au quotidien les produits de placements de ses diverses catégories de clients (particuliers, entreprises, institutionnels,...), d'où le nom de « gestion pour compte de tiers ».

La gestion pour compte de tiers se scinde en deux parties :

- **La Gestion Collective** : les produits proposés par la gestion collective sont des OPCVM. Ils sont destinés à de nombreux porteurs qui en achètent des parts.
- **La Gestion Sous Mandat** : un mandat de gestion n'a pour sa part pas de matérialisation, c'est simplement un contrat entre la société de gestion et un client : un mandat de gestion ne s'adresse qu'à un seul client.

La gestion sous mandat est donc un style de gestion particulier dans la mesure où le client a certains droits quant à la gestion du fonds qui lui est dédié.

En gestion collective, l'argent des clients est géré par les Asset Managers au sein de structures juridiques prédéfinies : les **OPCVM** (Organisme de Placement Collectif en Valeurs Mobilières). Ces OPCVM, qui peuvent être des FCP ou des SICAV, se composent de parts vendues par des distributeurs. SICAV et FCP opèrent de façon identique mais leur nature juridique diffère.

1.1 Types de fonds

Pour séduire un maximum de clients, les Asset Managers ont développé des gammes de fonds de plus en plus larges correspondant aux différents niveaux de risques (et donc espérances de gains) et horizons de placement souhaités par les clients. Des milliers de fonds existent non seulement en gestion traditionnelle (actions, obligations et monétaire) mais également de plus en plus dans les nouveaux types de gestion proposés par la gestion alternative (fonds garantis, fonds de fonds,...) ou la gestion ISR (Investissement Socialement Responsable).

HSBC Global Asset Management, de par ses filières de gestion spécialisées propose également une large gamme de fonds adaptée aux besoins de sa clientèle diversifiée.

1.2 Impacts de la crise : focus sur le risk management

Il semble que la récente crise économique ait affecté les sociétés de gestion sous de multiples aspects. Mais au-delà de ses nombreux aspects négatifs (faibles performances, décollectes des clients), cette crise a permis de mettre en lumière l'importance de la gestion des risques.

Les gérants ainsi que les autorités de contrôle et les clients ont désormais conscience du caractère primordial de la gestion de ces risques.

Dans le secteur de la gestion d'actifs, on retrouve principalement quatre grandes classes de risques :

- **Le risque de contrepartie**, c'est le risque que la partie avec laquelle un contrat a été conclu ne tienne pas ses engagements. (livraison, paiement, remboursement, etc.)
- **Le risque de taux**, c'est le risque des prêts-emprunts. C'est le risque que les taux de crédit évoluent défavorablement.
- **Le risque de change**, c'est le risque sur les variations des cours des monnaies entre elles.
- **Le risque de liquidité**, c'est le risque sur la facilité à acheter ou à revendre un actif. Si un marché n'est pas liquide, on risque de ne pas trouver d'acheteur quand on le souhaite ou de ne pas trouver de vendeur lorsqu'on en a besoin. C'est un risque lié à la nature du sous-jacent mais aussi à la crédibilité de l'acheteur-Vendeur. En effet, il est facile d'acheter ou de vendre un produit courant à une contrepartie de confiance, mais plus difficile avec un produit très spécialisé. C'est la liquidité de ce produit. De plus, si l'acheteur/vendeur n'est pas crédible, le risque de contrepartie pour les éventuels fournisseurs/clients, les dissuade de traiter.

Il y a d'autres risques mais la plupart se rapprochent de ceux-là. Par exemple, on parle de **risque pays**. Si un pays connaît une crise très grave (guerre, révolution, faillite en cascade, etc.) alors même les entreprises de confiance, malgré leur crédibilité vont se retrouver en difficulté. C'est un risque de contrepartie lié à l'environnement de la contrepartie.

La récente crise financière a été marquée par une crise de confiance relativement importante. En effet, les banques et autres acteurs du monde financier refusent de se prêter mutuellement de l'argent (c'est-à-dire des liquidités). La raison est une perte de confiance : les autres établissements sont susceptibles de détenir des créances douteuses ou de faire faillite et donc

de ne pas pouvoir faire face à leurs engagements. Ce manque de confiance entraîne une baisse significative des transactions et donc par un phénomène de cercle vicieux, la crise est auto-entretenu. Pour enrayer ce phénomène, les banques centrales injectent des liquidités sur les marchés financiers.

En ce sens, la gestion du risque de liquidité est au centre des départements de risk management des différents acteurs du secteur financier et plus particulièrement de la gestion d'actifs.

1.3 La gestion du risque de liquidité

La récente crise de liquidité a impacté les sociétés de gestion par des sorties massives de capitaux.

Nous allons rendre compte dans cette partie de la gestion du risque de liquidité à l'actif du bilan d'une société de gestion.

I.3 a/ Détermination pratique de la liquidité d'un actif

Il est important de noter que l'on mesure la liquidité d'un fonds au travers de la composition de ce dernier. Un fonds peut en effet être constitué d'actifs très liquides et d'autres actifs moins liquides.

Pour mesurer la liquidité d'un actif, on utilise divers outils suivant l'actif considéré.

On retiendra principalement :

- _ L'écart entre le cours acheteur et le cours vendeur (la fourchette ou bid-ask spread) : plus la fourchette est petite, plus l'actif en question est liquide.
- _ La moyenne périodique du nombre de transactions quotidiennes : à comparer à l'historique ou à d'autres actifs comparables.

On distingue alors trois classes dans la liquidité des actifs, indépendamment de leurs catégories (action, obligation,...) : les actifs très liquides, les actifs liquides et les actifs peu liquides.

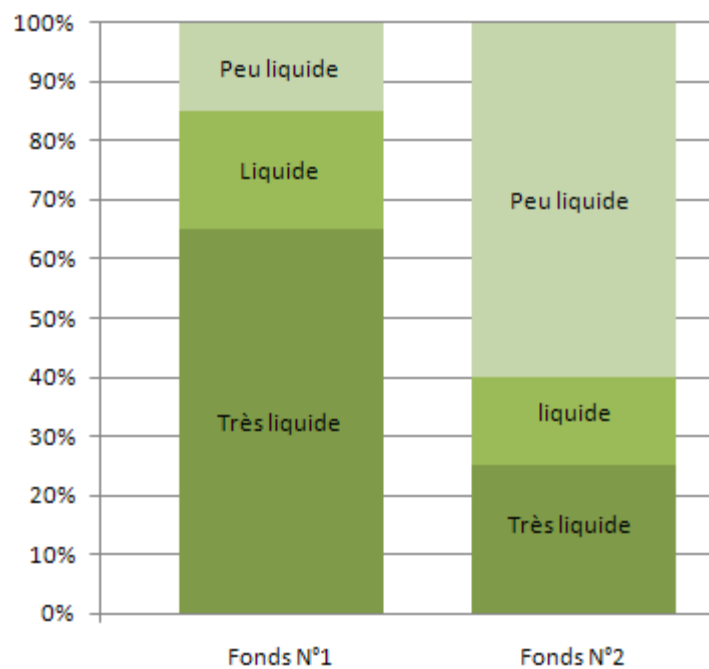
A partir de la classification des actifs d'un fonds en terme de liquidité, on peut déterminer la répartition de la liquidité d'un fonds.

Prenons l'exemple des deux fonds (resp. N°1 et N°2) composés comme suit :

- _ 40% d'actions A
- _ 25% d'actions B
- _ 20% d'obligations C
- _ 15% d'obligations D

- _ 10% d'actions A
- _ 15% d'actions B
- _ 15% d'obligations C
- _ 60% d'obligations D

S'il s'avère que les actions A et les obligations C sont jugées très liquides, que les actions B sont jugées liquides et que les obligations D sont jugées peu liquides, alors le fonds N°1 sera globalement liquide, alors que le fonds N°2 le sera nettement moins.



Exemple de répartition de deux fonds en terme de liquidité des actifs

I.3 b/ Lien avec le passif

Dès lors on peut faire le lien avec le passif de la société de gestion. En effet, les clients sont en droit, dans la grande majorité des contrats, de souscrire ou de racheter tout ou partie de leurs encours d'un fonds.

Un rachat trop important de la part des clients peut alors poser des problèmes en terme de délai ou de juste valeur de l'actif. On doit en effet garder en tête le principe d'égalité des porteurs de parts (article 322-24 du RGAMF) lorsque l'on reverse l'encours d'un client.

L'égalité des porteurs de parts est un des principes fondamentaux de la gestion collective : les opérations (achats/ventes, comptabilisation,...) ne doivent porter préjudice ni aux anciens porteurs, présents depuis un certain temps, ni aux jeunes porteurs qui viennent de souscrire.

Nous allons donc étudier le risque de liquidité à travers l'analyse du comportement des différents investisseurs pour tenter d'anticiper les éventuels rachats des clients et ainsi guider la gestion de ce risque.

I.3 c/ Les récentes évolutions réglementaires

L'article 156 de la loi du 4 août 2008, dite loi de modernisation de l'économie, a rappelé aux établissements financiers qu'ils étaient tenus de respecter des normes de gestion destinées à garantir leurs liquidités et leur solvabilité à l'égard des déposants et, plus généralement, des tiers, et l'équilibre de leur structure financière. Ils doivent en particulier respecter des ratios de couverture et de division de risques.

Les établissements financiers doivent également disposer d'un système adéquat de contrôle interne leur permettant notamment de mesurer les risques et la rentabilité de leurs activités. Lorsque la surveillance est exercée sur la base de la situation financière consolidée, les groupes financiers ou mixtes doivent adopter des procédures de contrôle interne adéquates pour la production des informations et renseignements utiles aux fins de l'exercice de cette surveillance.

Au sein des établissements de crédit, les conditions d'information des organes de direction, d'administration et de surveillance concernant l'efficacité des systèmes de contrôle interne, d'audit interne et de gestion des risques et le suivi des incidents révélés, notamment par ces systèmes, sont réglementées.

En réponse à la récente crise de liquidité, un arrêté du 5 mai 2009 vient de préciser les conditions dans lesquelles doivent s'organiser l'identification, la mesure, la gestion et le contrôle du risque de liquidité par les établissements de crédit. Il est complété par un second arrêté du même jour, portant modification du règlement (n°97-02) du 21 février 1997 relatif au contrôle interne des établissements de crédit et des entreprises d'investissement.

(Voir les recommandations de l'AMF quant à la gestion du risque de liquidité en annexe)

Deuxième partie

Présentation des données

II.1 Présentation des différentes catégories de clients

Au passif, nous avons une classification en terme de clients. On dénombre cinq catégories de clients :

- Les entreprises bancaires
- La clientèle institutionnelle (instituts de prévoyance, mutuelles, caisses de retraites,...)
- Les particuliers
- Les clients de la banque privée

Ces catégories de clients ou Customer group (CG) sont par nature très différents. En particulier, les horizons de placements et les profils d'investissements sont naturellement différents.

Chacune de ces catégories est composée d'un certain nombre de clients. Nous verrons l'importance que peut avoir le nombre de clients au sein d'une catégorie pour un gérant dans une optique de diversification du risque client.

Chaque client est investi dans un ou plusieurs fonds. Ces fonds, tout comme les clients, sont classés en quatre catégories classiques : actions, obligations, garantis et alternatifs.

II.2 Evolution de la répartition des encours des catégories de client par type de fonds

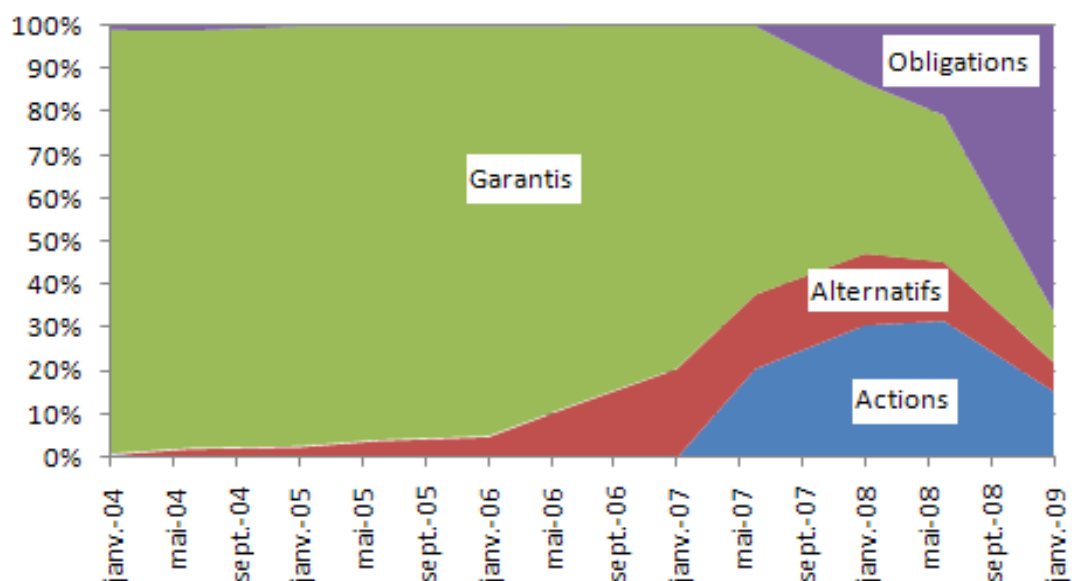
On étudie l'évolution de la répartition des encours de chaque catégorie de client par classe d'actifs au cours du temps. On présente par la même occasion ces différentes catégories de clients.

Pour cela, on dispose d'un historique mensuel sur les cinq dernières années.

Voici les graphes représentant les évolutions des différentes catégories de clients :

II.2 a/ Les entreprises bancaires (GBM)

Les entreprises bancaires, françaises ou étrangères, par le biais de souscription de prêts ou de placement de leur trésorerie ont souvent recours aux sociétés de gestion. Ce sont dans la plupart des cas, les conseils d'administration qui prennent *in fine* les décisions d'investissements.



Evolution de la répartition des encours des entreprises bancaires

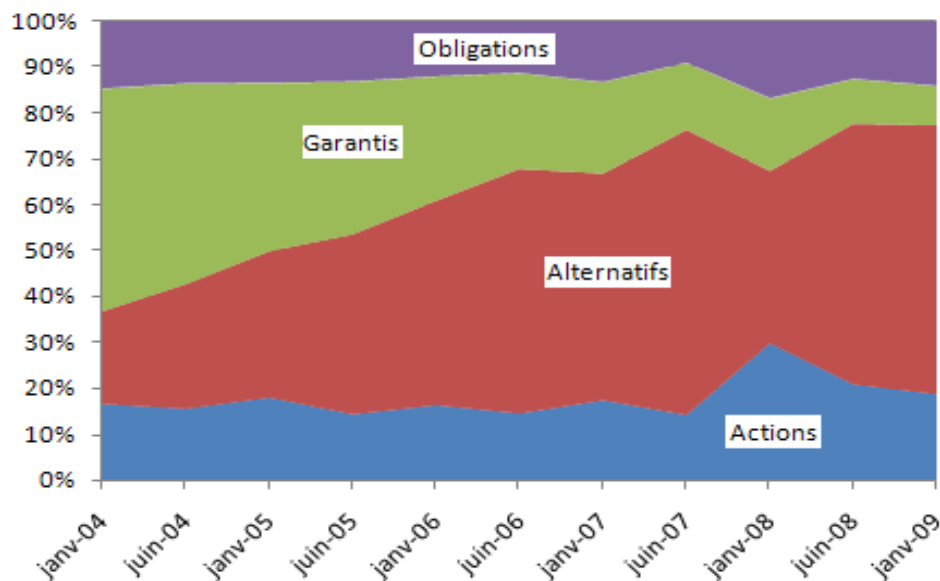
On constate un profond changement dans la répartition des actifs de cette catégorie. En effet, les banques ont privilégiées les obligations ou les actions au détriment des fonds de type garantis qui représentaient l'intégralité des fonds détenus par cette clientèle au début de l'historique. Il ne faut pas voir l'apparition des fonds actions dans le profil de gestion des banques comme la volonté d'apporter de la performance supplémentaire, mais comme la nécessité de diversifier les fonds détenus par ces investisseurs.

On note également que cette catégorie est composée de 55 à 89 clients selon les périodes de l'historique.

II.2 b/ Les investisseurs institutionnels (GBM-Instit)

Ces investisseurs sont pour la plupart des sociétés d'assurances, des mutuelles ou des fonds de pensions. Ils ont donc pour objectif de placer les encours de leurs clients sur les marchés avec des horizons de placements relativement élevés.

Autre particularité de cette catégorie de clients, le rôle de la réglementation est particulièrement influent sur les décisions d'investissement de la clientèle institutionnelle qui doit respecter un certain nombre de critères et de ratios.



Evolution de la répartition des encours de la clientèle institutionnelle

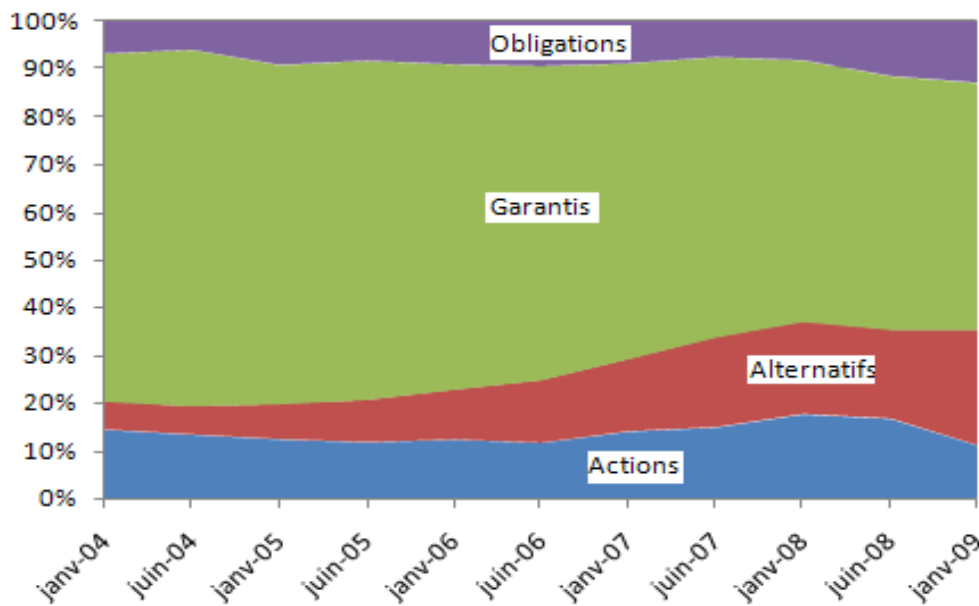
Pour la clientèle institutionnelle, on retrouve un changement important, notamment une augmentation des encours des fonds alternatifs au dépit des fonds de type garantis. Ces fonds qui constituaient 50% des encours de cette catégorie au début de l'historique ne pèsent plus que 10% 5 ans après.

En revanche on constate une certaine stabilité des proportions des fonds obligataires et de type actions.

II.2 c/ Les particuliers (PFS)

Les clients particuliers souscrivent le plus souvent leurs contrats sur les différents points de vente des établissements bancaires. Pour ce type de clients, le rôle du chargé de clientèle est souvent très influent, car les particuliers, contrairement aux autres types de clients ont pour la plupart une mauvaise connaissance du domaine financier et des placements en général.

Il faudra garder à l'esprit cette caractéristique lors de la modélisation du comportement de cette catégorie d'investisseurs.



Evolution de la répartition des encours des particuliers

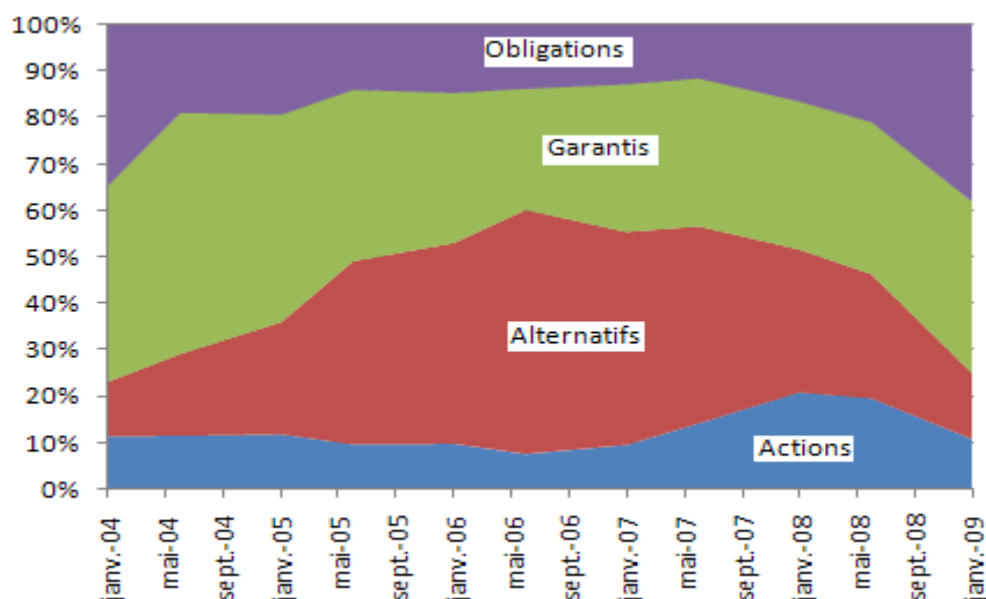
On remarque que la proportion des encours des fonds alternatifs passe de 5% à 25% en cinq ans, au détriment des fonds de type garantis. Les proportions des encours des fonds obligataires et de type actions restent quant à eux relativement stables au cours de l'historique.

Cette stabilité peut s'expliquer par le nombre relativement important de particuliers, comparativement aux autres groupes, ce qui crée une certaine homogénéité. D'autre part, l'horizon d'investissement étant relativement long, on peut penser que l'on n'observera pas d'évolution brutale sur une période de cinq années.

II.2 d/ La banque privée (GPB)

La clientèle de la banque privée du groupe est composée d'une quinzaine de regroupements de clients en 2009. Ces clients étant des acteurs particuliers du monde de la finance, ils ont souvent un accès privilégié à l'information et aux données de marchés.

D'autre part, on note que ces clients doivent respecter certains critères d'investissements qui sont dictés par l'AMF (notamment concernant les délits d'initiés), ce qui peut influencer les décisions des investisseurs.



Evolution de la répartition des encours des entreprises bancaires.

Concernant la répartition des encours de ce type de clients, on remarque un transfert (en proportion) des encours des fonds obligataires vers les fonds alternatifs en début d'année 2004, cette tendance s'inverse dès le mois de mai 2007. La proportion des fonds actions quant à elle, reste stable (environ 10%) jusqu'à ce qu'elle augmente en janvier 2007 pour atteindre les 20% un an après et ensuite revenir à son niveau précédent. Il semblerait que ces clients aient profité de la baisse de la valeur des actions lors de la crise financière (le marché action a été particulièrement affecté) pour investir dans cette catégorie de fonds. En effet la crise financière a pu être considérée par cette catégorie de clients (comme par d'autres) comme un point d'entrée dans des fonds essentiellement composés par des actions.

II.2 e/ Synthèse

On peut conclure de cette présentation, que les différentes catégories de clients semblent avoir des profils d'investissements totalement différents.

En réalité, les différents clients ont des profils de risques différents, et ces différences se retrouvent dans la répartition de leurs encours par catégories de fonds. Par exemple, un investisseur institutionnel ne cherchera pas à obtenir une performance très élevée, mais aura pour soucis de limiter la volatilité de ses actifs, contrairement aux sociétés bancaires qui sont prêtes à prendre plus de risques pour obtenir éventuellement une performance plus élevée.

Autre facteur non négligeable, la législation peut avoir une influence sur le comportement des investisseurs. En effet, les banques, les instituts de prévoyances ou les sociétés d'assurances n'ont pas les mêmes contraintes en terme d'investissement de leurs liquidités, dans la mesure où le législateur est différent. Par exemple, les sociétés d'assurance ne peuvent excéder x% de leurs actifs investis en actions (selon la législation en vigueur).

Ces différences en terme de répartition des encours se retrouvent également dans la gestion des souscriptions et des rachats des clients. En effet, après avoir analysé statistiquement les collectes et les décollectes des différentes catégories de clients, on remarque des différences significatives chez ces groupes. Ces différences semblent liées à de multiples facteurs explicatifs comme l'horizon de gestion et les profils de risque. Ce dernier élément semble en effet primordial pour expliquer l'étendue des disparités au sein des différentes catégories en terme de souscription ou de rachat.

Dès lors, il semble intéressant de procéder à une étude sur le comportement des investisseurs en considérant un regroupement par catégorie de client, ce qui permettra de dégager un comportement général de la catégorie étudiée.

Il existe de nombreuses théories qui étudient le comportement des agents et leur rapport au risque, on peut notamment citer les travaux de Bernoulli ou de Von Neumann et Morgenstern. Nous allons en exposer les bases dans la partie suivante.

II.3 Théorie de la décision en environnement incertain

II.3 a/ L'axiomatique de von Neumann et Morgenstern

Le but de cette théorie est de modéliser le comportement des investisseurs et d'analyser leur rapport au risque.

Le point de départ de cette théorie est constitué de deux éléments : un ensemble de consommation X que l'on supposera fini par commodité et d'une relation de préférence sur C , notée $\succsim$.

- La notion de loteries : Soit $X = (x_1, \dots, x_s)$ l'ensemble des résultats possibles, qui peut s'interpréter comme l'ensemble des prix ou des lots que les loteries permettent d'obtenir. Une loterie $L \in \mathcal{L}$ est une mesure probabiliste $(p_1, \dots, p_s)$ associée à X telle que $\sum_{p=1}^s p_i = 1$, avec s le nombre d'états de la nature. $\mathcal{L}$ est l'espace de toutes les loteries possibles.

On note encore $L = \sum_{p=1}^s p_i \cdot x_i$.

Deux loteries distinctes diffèrent par les probabilités associées à chaque lot.

- La notion de préférences sur des loteries : Un individu rationnel a des préférences sur les loteries, formalisées par la relation de préférence $\succsim$ (ou $\approx$ pour la relation d'équivalence). C'est un préordre total (axiome 1) et continu (axiome 2), autrement dit, qui vérifie les axiomes suivants :

➤ **Axiome 1** : $\succsim$ est un préordre total si elle vérifie les propriétés suivantes :

- Deux loteries L et $L' \in \mathcal{L}$ sont toujours comparables, autrement dit : soit $L \succsim L'$, soit $L' \succsim L$, soit $L \approx L'$ (Ordre total).
- Pour toute loterie $L \in \mathcal{L}$, $L \succsim L$ (Réflexivité).

- Pour trois loteries $L, L' \text{ et } L'' \in \mathcal{L}$, si $L \succcurlyeq L'$ et $L' \succcurlyeq L''$ alors $L \succcurlyeq L''$ (Transitivité).

- **Axiome 2** : $\succcurlyeq$ est dite continue en probabilité si pour les trois loteries $L, L', L'' \in \mathcal{L}$, telles que $L \succcurlyeq L' \succcurlyeq L''$, il existe $\lambda \in [0,1]$ tel que : $\lambda \circ L + (1 - \lambda) \circ L'' \approx L'$.

Avant d'énoncer le principal résultat de Von Neumann et Morgenstern, il est nécessaire de définir encore deux axiomes :

- **Axiome 3** (Axiome d'Indépendance) : Soient L et L' , deux loteries telles que $L \succcurlyeq L'$ alors quelle que soit une troisième loterie L'' , on a :

$$\forall \lambda \in [0, 1], \lambda L + (1 - \lambda)L'' \succcurlyeq \lambda L' + (1 - \lambda)L''$$

- **Axiome 4** (Axiome de Composition) : Le choix entre loteries ne dépend que des conséquences finales et des probabilités.

Dès lors, le résultat fondamental de la théorie de VNM est énoncé dans le théorème de représentation des préférences dans le modèle d'espérance d'utilité (Von Neumann-Morgenstern, 1944, Fishburn 1970) :

Si $\succcurlyeq$ est une relation de préférence sur les loteries qui vérifie les axiomes 1, 2, 3 et 4, alors il existe une fonction continue $U : X \rightarrow \mathbb{R}$ qui représente $\succcurlyeq$ au sens suivant :

$\forall$ couple de loteries aléatoires $(L, L') \in \mathcal{L}^2$, définies par : $L = (p_1, \dots, p_s)$ et $L' = (q_1, \dots, q_s)$, on a :

$$L \succcurlyeq L' \Leftrightarrow \sum_{i=1}^s p_i \cdot U(x_i) \geq \sum_{i=1}^s q_i \cdot U(x_i) \Leftrightarrow U(L) \geq U(L')$$

avec : U définie sur les loteries telle que $U(L) = \sum_{i=1}^s p_i \cdot U(x_i)$.

$U(\cdot)$ est appelée fonction d'utilité de Bernouilli et $E[U(\cdot)]$ fonction d'utilité espérée de VNM.

Ce théorème peut évidemment s'appliquer à des problèmes de choix en environnement certain. La fonction d'utilité est alors définie, non plus sur un espace de loteries, mais sur un espace de biens (X) et le résultat précédent devient :

$$\forall \text{ couple de biens } (x, y) \in X^2 : x \succcurlyeq y \Leftrightarrow U(x) \geq U(y)$$

II.3 b/ Les comportements des agents face à l'incertain

La fonction d'utilité U définie comme précédemment permet d'introduire la diversité des comportements des agents économiques face à l'incertitude :

1. Un individu présente une aversion pour le risque s'il préfère une loterie donnant un gain certain à une autre loterie risquée ayant la même espérance de gain. Autrement dit, un individu averse au risque affectera une fraction positive de sa richesse à des placements en actif risqué si et seulement si leur espérance mathématique de rendement est supérieure au rendement de l'actif certain. Le comportement d'aversion pour le risque se formalise ainsi : $\{L, L \preceq E[L]\}$.
2. Un individu a une propension au risque s'il préfère une loterie risquée à une loterie donnant un gain certain, égal à l'espérance de gain de la loterie risquée, autrement dit, si et seulement si : $\{L, L \succeq E[L]\}$.
3. Un individu est neutre au risque si il est indifférent au choix entre une loterie donnant un gain certain et une loterie risquée ayant la même espérance de gain, i.e. si et seulement si : $\{L, L \sim E[L]\}$.

Ces trois comportements peuvent être reformulés à partir de la fonction d'utilité espérée de VNM :

Soit x une variable aléatoire continue représentant un gain et $U(x)$ une fonction continue et croissante ($U : \mathbb{R}_+ \rightarrow \mathbb{R}$),

- * Si $E[U(x)] < U[E(x)]$, U concave, l'agent est averse au risque.
- * Si $E[U(x)] = U[E(x)]$, U linéaire, l'agent est neutre au risque.
- * Si $E[U(x)] > U[E(x)]$, U convexe, l'agent est risquophile.

S'il est relativement simple d'analyser les comportements d'investisseurs particuliers grâce à cette théorie, il paraît plus compliqué de l'adapter aux autres types d'investisseurs. En effet, concernant les banques ou les investisseurs institutionnels, d'autres critères entrent en jeu dans le processus de décision relatif à l'investissement. Ces derniers sont soumis à une législation particulière, ils doivent rendre des comptes aux autorités de marché et à leurs actionnaires (pour les sociétés cotées).

Ces caractéristiques rendent cette théorie inadaptée pour ces catégories de clients. Cependant, dans le cadre de notre étude, nous supposons que tous les clients ont un comportement rationnel face au risque.

II.4 Périmètre de l'étude

Dans ce mémoire, il est important de signaler que nous n'étudierons pas les fonds dédiés ou les mandats, dans la mesure où la gestion de ces derniers est relativement particulière. D'autre part, avec ce style de gestion, on ne bénéficie pas des propriétés de la loi des grands nombres, qui nous est indispensable dans le cadre d'études statistiques. On n'étudiera donc uniquement les fonds de gestion collective.

On n'étudie également que les fonds européens émis en euros (qui représentent l'essentiel des fonds gérés) pour des raisons d'homogénéité et pour éviter les complications liées aux fluctuations des taux de changes.

II.5 Détermination du comportement des investisseurs par rapport à leurs encours

Le but est ici de déterminer un indice qui nous permettra de suivre l'évolution du comportement des clients en terme de souscription, mais surtout (dans le cadre de l'étude du suivi du risque de liquidité), de rachat.

Pour cet indice, nous avons à notre disposition l'évolution mensuelle de l'encours ainsi que du montant de la collecte/décollecte de chaque client pour chaque fonds.

Pour autant, suivre l'évolution de l'encours du fonds d'un client comporte un biais pour rendre compte du comportement de ce dernier. En effet, en se rapportant à la définition de l'encours d'un fonds, on réalise que deux facteurs influencent directement celui-ci.

L'encours d'un fonds d'un client au mois n est donné par la formule :

$$\text{Encours}(n) = \text{VL}(n) * \text{Nombre de parts}(n)$$

L'évolution de l'encours est donc lié à la fois à l'évolution de la valeur liquidative (VL) qui est propre au fonds et dépend de l'évolution des valeurs de marchés des actifs qui composent ce dernier et du nombre de parts du fonds que le client détient.

Il est donc possible que l'encours d'un client double d'un mois à un autre, sans pour autant que le client ait réinvesti dans le fonds qu'il détient, dû à une évolution favorable des marchés. On ne pourra donc pas se baser uniquement sur l'encours pour suivre l'évolution des souscriptions ou des rachats d'un client.

D'autre part, étudier le seul nombre de parts d'un client dans un fonds n'a pas réellement de sens dans la mesure où le client raisonnera toujours en euros et non en nombre de parts.

En effet, lorsqu'il récupère de la liquidité en rachetant une partie des parts qu'il détient dans un fonds, un client aura souvent pour habitude de réinvestir cette somme dans un autre fonds ou dans d'autres actifs, il ne se basera donc pas sur les parts détenues dans un fonds, mais raisonnera toujours en terme de cash.

L'indice que nous avons utilisé pour étudier les souscriptions et les rachats des clients utilise à la fois la notion de monnaie tout en minimisant l'effet marché dû à l'évolution de la valeur liquidative.

Définition de l'indice retenu pour le suivi des investissements des clients

L'indice I_n retenu pour étudier le comportement des investisseurs en terme de souscriptions et de rachats est basé sur le montant des collectes ou des décollectes mensuelles. On le définit à partir du premier encours de notre historique comme suit :

$$I_n = \text{Encours}(1) + \sum_{i=1}^n \text{Collecte/décollecte}(i)$$

Où $\text{Encours}(1)$ est le premier encours connu du client considéré pour le fonds considéré. Cet encours sert en réalité de référence pour cet indice. On ajoute ensuite toutes les collectes et toutes les décollectes effectuées par ce client dans son fonds mois après mois. L'unité de cet indice est donc l'euro.

Cet indice ne fait que minimiser l'effet marché, car on le retrouve dans les collectes ou les décollectes qui sont valorisées avec la VL de chaque mois i :

$$\text{Collecte/décollecte}(i) = \text{VL}(i) * \text{Nombre de parts}(i)$$

Cet indice est calculé pour chaque client de chaque fonds. Cependant, il peut être agrégé par type de client ou par type de fonds pour constater des tendances qui pourraient se dégager. En agrégeant ainsi les données, on compense ainsi l'effet marché entre les souscriptions et les rachats.

A partir de cet indice, nous définissons le pourcentage de souscription/rachat que nous étudierons par la suite dans la mise en place de nos scénarios de rachats :

$$\%SR_n = \frac{I_n - I_{n-1}}{I_{n-1}}$$

En revanche dans la modélisation du comportement des clients par des modèles statistiques, nous prendrons le soin d'agréger ces pourcentages par catégories de clients pour pouvoir dégager des tendances grâce à la théorie de la loi des grands nombres :

$$\%SR_n(A) = \frac{\sum_{j \in A} I_n(j) - \sum_{j \in A} I_{n-1}(j)}{\sum_{j \in A} I_{n-1}(j)}$$

On remarque dans un premier temps que ce ratio ne pourra, par définition, évoluer de façon symétrique. En effet, d'un mois à un autre, un client pourra choisir d'investir autant qu'il le souhaite (sauf contraintes propres à la législation), son pourcentage de souscription peut donc être infini. En revanche, il ne pourra jamais décollecter une somme supérieure à son encours, son pourcentage de rachat est donc quant à lui minoré par -100%.

Nous verrons par la suite l'influence que peut avoir cette dissymétrie sur l'étude statistique de nos données.

Le but final de l'étude statistique du comportement des investisseurs est de pouvoir proposer des scénarios de type prévisionnels en complément des stress tests qui sont basés sur diverses hypothèses plus ou moins fortes.

Dans un premier temps, nous présenterons les scénarios de rachats, que l'on peut appeler « scénarios catastrophes » ou « stress tests » car ils sont basés sur des hypothèses relativement fortes. Dans un deuxième temps, nous introduirons les processus autorégressifs (modèles ARMA) pour modéliser les séries de pourcentages de souscription/rachat des différents Customer group et ainsi proposer des scénarios de type prévisionnels basés sur ces modèles.

Troisième partie

Mise en place de scénarios de rachats

Le but de ce mémoire est d'analyser l'impact que peuvent avoir les comportements des différentes catégories d'investisseurs sur la liquidité future des fonds.

Pour cela, nous mettons en place divers techniques projectives dont l'objectif n'est pas d'obtenir les prévisions les plus précises possibles, mais d'avoir un échantillon de scénario le plus complet possible pour être en mesure d'analyser au final le passif des fonds avec le plus d'informations concernant les différents investisseurs.

Cette étude sera donc scindée en deux parties. Nous étudierons dans un premier temps des scénarios de rachat dont on décidera des hypothèses retenues, puis on complètera cette étude par la mise en place de scénarios basés sur des études statistiques et plus particulièrement sur des modèles autorégressifs.

III.1 Mise en place de stress tests

III.1 a/ Stress test 1

Pour ce scénario, nous faisons l'hypothèse que le client le plus important (au moment du dernier mois connu) de chaque Customer group rachète l'intégralité de son encours.

Ce scénario repose donc uniquement sur le dernier mois connu et non sur l'historique complet ce qui a pour avantage de paraître plus concret.

Le pourcentage de rachat que l'on obtient pour ce stress test est en réalité le pourcentage de détention du client le plus important de chaque Customer group.

On note que ce scénario n'est pas adapté pour toutes les catégories de clients. En particulier, ce scénario présente peu de sens pour les clients particuliers. En effet, cette catégorie est composée de plus de 65 millions de clients au total, donc même dans l'éventualité où le client le plus important décide de retirer l'intégralité de ses encours, cette situation ne sera pas considérée comme catastrophique, contrairement à l'impact de cette hypothèse aux autres catégories de clients.

Voici les résultats que nous obtenons pour le dernier mois connu de notre historique d'étude :

Customer Group	Scénario 1
Banque privée	-16,5%
Institutionnels	-29,9%
Banques	-22,5%

Pourcentage de souscription/rachat mensuel obtenu par le scénario N°1

On constate alors que l'on doit faire face à un rachat équivalent à 30% de l'encours total des investisseurs institutionnels dans le cas où le client le plus important viendrait à racheter tous ses encours.

Il en est de même pour la catégorie des banques et compagnies financières, où un rachat du client le plus important représenterait un rachat de 22,5% du total des encours de cette catégorie.

Il est donc intéressant de suivre l'évolution de ces clients et plus particulièrement, les fonds dans lesquels ces clients ont investis.

Ce scénario met en lumière l'importance de la diversification du portefeuille de clients, pour limiter les risques de rachat.

A partir de ce scénario, il est aisé d'obtenir d'autres scénarios avec des degrés de stress plus importants, on peut en effet faire l'hypothèse que les deux ou trois plus importants clients décollectent l'intégralité de leurs encours.

III.1 b/ Stress test 2

Le deuxième scénario que nous proposons est un scénario basé sur le minimum historique du pourcentage de souscription/rachat de chaque Customer group.

Ce scénario peut être considéré comme un scénario catastrophe, car il reproduit les rachats les plus importants sur les cinq dernières années pour chaque catégorie de clients.

On peut effectivement supposer qu'en temps de crise, chaque client souhaite obtenir des réserves de liquidités et décollecte ainsi un fort pourcentage de ses encours placés sur les marchés financiers (bien qu'après avoir analysé l'influence de la récente crise financière, on s'aperçoit que ce n'est pas exactement ce qui s'est produit).

Nous obtenons ces résultats pour le mois dernier mois connu de l'historique :

Customer Group	Scénario 2
Banque privée	-17,0%
Investisseurs institutionnels	-16,5%
Banques	-15,7%
Particuliers	-12,4%

Pourcentage de souscription/rachat mensuel obtenu par le scénario N°2

La principale caractéristique de ce scénario (comme celui de nombreux scénarios basés sur les extremums historiques) est d'être relativement constant d'un mois sur l'autre. Cela provient également du fait que notre historique étant assez complet, il est peu probable qu'une nouvelle donnée remplace le minimum (ou le maximum) établi.

Cependant ce scénario à l'intérêt de se baser sur des événements concrets, qui se sont déjà produits par le passé.

Nous avons mis en place bien d'autres scénarios de ce type basés sur des hypothèses plus ou moins susceptibles de se produire. Le but n'est pas ici de tous les passer en revue, mais d'analyser comment les hypothèses retenues impactent les différentes catégories de clients et comment la répartition de ces catégories au sein d'un même fonds peuvent impacter à leur tour ce dernier en terme de risque de liquidité.

On remarque principalement que ces scénarios catastrophe sont extrêmement prudents dans leur approche face au risque de rachat. En effet, tous les pourcentages de rachat envisagés par ces scénarios sont à deux chiffres, ce qui est relativement exigeant dans le cadre du risque de liquidité.

En confrontant ces résultats de rachats à la liquidité des fonds gérés, on se rend compte que ces stress tests mettent en lumière bien trop de fonds. En effet, ces scénarios bien qu'intéressant en terme d'analyse de la liquidité des fonds, ne peuvent être utilisés seuls. Il paraît très pertinent d'utiliser ces scénarios « chocs » en complément d'autres scénarios, basés sur des modèles statistiques plus fiables et au caractère plus objectif.

III.2 Utilisation des processus autorégressifs pour modéliser les comportements des investisseurs

Pour procéder à l'étude des modèles autorégressifs, nous avons utilisé les logiciels statistiques R et Eviews, nous allons expliquer la démarche que nous avons utilisée pour modéliser chaque série de données relatives aux différents groupes de clients.

Bien que nous présentions les résultats obtenus pour tous les Customer group, nous ne les détaillerons que pour deux des quatre catégories de clients, ce qui aboutira à une étude comparative entre ces deux groupes : les particuliers et les banques et institutions financières.

L'étude des séries temporelles, ou séries chronologiques, correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. Elles ont été utilisées en astronomie, en météorologie, en théorie du signal ou en économie à différentes époques.

Dans notre cas, nous étudions l'évolution mensuelle sur les cinq dernières années du pourcentage de souscription/rachat des différentes catégories de clients.

Nous commençons par rappeler les principaux concepts théoriques gravitant autour des processus autorégressifs car ils seront utilisés à de nombreuses reprises par la suite.

Pour modéliser les différents pourcentages de souscription/rachat, nous avons choisis les processus ARIMA. Pour cela, on suppose que la variance conditionnelle de la série est constante à travers le temps. Cette hypothèse est renforcée par la faible longueur des séries étudiées. Ces séries sont bien moins sujettes au phénomène de survolatilité (même en temps de crise économique) que des données financières (indices boursiers, taux obligataires,...) qui seraient modélisées par des processus de type ARCH ou GARCH (modèles qui prennent en compte le caractère stochastique de la volatilité).

III.2 a/ Rappels : outils statistiques

On rappelle ici les différentes définitions de concepts qui seront utilisés dans la suite de notre étude.

Stationnarité : Soit un processus temporel à valeurs réelles et en temps discret $Z_1, Z_2, \dots, Z_t$.

- Il est dit stationnaire au sens fort si pour toute fonction f mesurable:
$$f(Z_1, Z_2, \dots, Z_t) = f(Z_{1+k}, Z_{2+k}, \dots, Z_{t+k}).$$

- Il est dit stationnaire au sens faible (ou "de second ordre", ou "en covariance") si
 $\forall i, k = 1 \dots t,$
 $E(Z_i) = \mu$
 $Var(Z_i) = \sigma^2 < \infty$
 $Cov(Z_i, Z_{i-k}) = f(k)$

Autocovariance : L'autocovariance d'un processus stochastique est la covariance de ce processus avec une version décalée de lui-même. Si le processus a une espérance μ alors son autocovariance d'ordre k , notée souvent γ_k , est donnée par :

$$\gamma_k = Cov(X_t, X_{t-k}) = E[(X_t - \mu)(X_{t-k} - \mu)]$$

Fonction d'autocorrélation : À partir de la notion d'autocovariance définie plus haut, on définit la fonction d'autocorrélation $\rho_X(h)$ pour le processus (X_t) comme étant :

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_0(h)} = \frac{Cov(X_t, X_{t+h})}{V(X_t)}$$

Le graphe de l'application qui à h , associe $\rho_X(h)$ est l'autocorrélogramme.

Bruit blanc : Un processus ε_t est qualifié de bruit blanc si :

$$E(\varepsilon_t) = 0$$

$$E(\varepsilon_t^2) = \sigma^2$$

$$E(\varepsilon_s \varepsilon_t) = 0, \text{ avec } s \neq t.$$

Un processus de bruit blanc est donc par définition stationnaire de second ordre. La troisième condition, $E(\varepsilon_s \varepsilon_t) = 0$ (ou $Cov(\varepsilon_s, \varepsilon_t) = 0$), signifie que l'autocovariance est nulle.

De plus, si les ε_t suivent une loi normale centrée, on parle de bruit blanc gaussien.

III.2 b/ Introduction au concept de modélisation temporelle et processus ARMA

On appelle processus ARMA(p, q), un processus stationnaire (X_t) vérifiant une relation du type :

$$X_t + \sum_{i=1}^p \phi_i X_{t-i} = \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad , \quad \text{pour tout } t \in \mathbb{Z}$$

où les θ_j sont des réels et (ε_t) est un bruit blanc de variance σ^2 . Cela est équivalent à l'écriture :

$$\Phi(L)X_t = \Theta(L)\varepsilon_t$$

avec $\Phi(L) = 1 + \phi_1 L + \dots + \phi_p L^p$ et $\Theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$ et L opérateur retard : $L(X_t) = X_{t-1}$.

Remarque : Dans le cas d'un processus non centré, $\Phi(L)X_t = \mu + \Theta(L)\varepsilon_t$, on peut noter que $E(X_t) = \mu / \Phi(1) = m$, et que $\Phi(L)(X_t - m) = \Theta(L)\varepsilon_t$. Il est donc toujours possible, en translatant le processus, de se ramener à un processus centré.

On supposera de plus que les polynômes Φ et Θ n'ont pas de racines en module strictement supérieures à 1 (écriture sous forme canonique), et n'ont pas de racine commune. On supposera de plus que les degrés de Φ et Θ sont respectivement q et p, au sens où $\phi_p \neq 0$ et $\theta_q \neq 0$.

III.2 c/ Procédure

L'ouvrage de Box et Jenkins (*Time series analysis, forecasting and control*), publié en 1970 a proposé une démarche de prévision pour les séries temporelles, fondée sur l'utilisation de processus ARIMA.

Nous allons procéder selon les étapes suivantes pour mener à bien l'étude de nos différentes séries et ainsi effectuer des prévisions :

(1) Identification

- (i) choix de d : combien de fois faut-il différencier pour obtenir une série stationnaire ?
- (ii) choix de p et q : ordres respectifs des composantes AR et MA

(2) Estimation des paramètres

Estimation des ϕ_i et des θ_j : paramètres respectifs des composantes AR et MA

(3) Vérification a posteriori

- (i) significativité des paramètres
- (ii) validation de l'hypothèse de bruit blanc des résidus

Cette procédure est résumée par le schéma suivant :

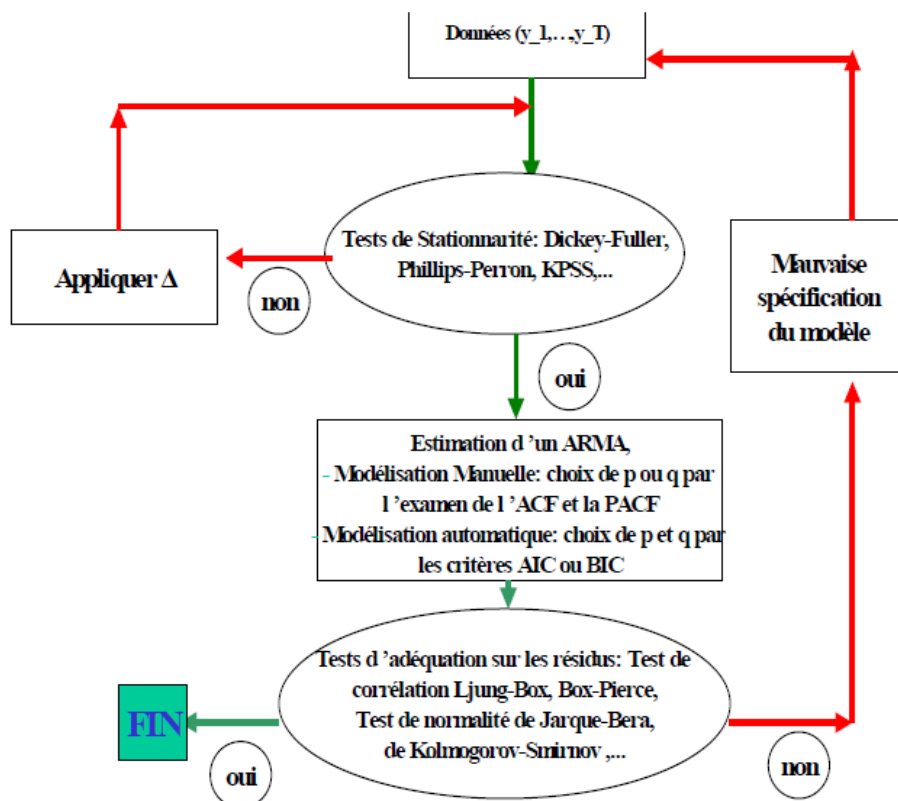


Schéma général de la modélisation d'une série temporelle par un modèle ARIMA.

Etant entendu qu'il convient de retenir le modèle le plus parcimonieux, utilisant le moins de paramètres, et ayant le meilleur comportement en prévision.

Il faut également garder à l'esprit le but de nos modélisations. En effet, l'intérêt de modéliser nos séries temporelles est de pouvoir fournir des prévisions basées sur ces modèles. Ces prévisions seront alors utilisées dans le cadre de la gestion du risque de liquidité.

III.2 d/ Identification

III.2 d/ i. Estimation du paramètre d'intégration d

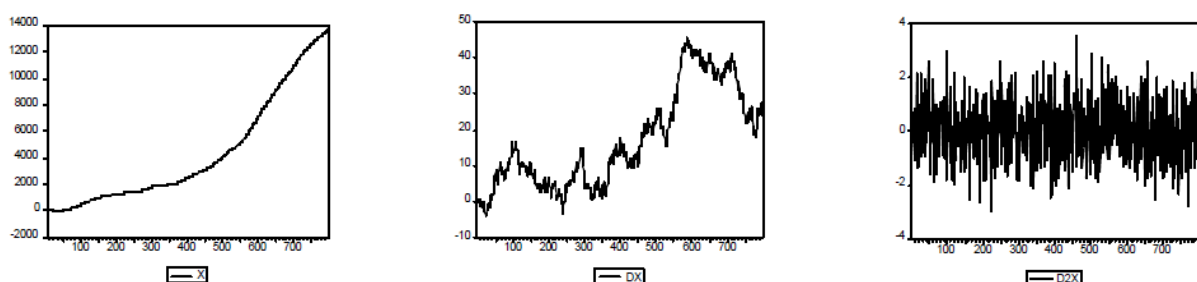
Le paramètre d'intégration d est le nombre de fois qu'il faut différencier la série de départ afin qu'elle devienne stationnaire (dans le cas où elle ne l'est pas à l'origine). On entend par stationnarité le fait que la structure du processus sous-jacent supposé évolue ou non avec le temps. Si la structure reste la même, le processus est dit alors stationnaire.

On peut donc chercher à n'étudier que les modèles aléatoires dont le comportement dans le temps est assez stable.

Pour déterminer le paramètre d , il suffit la plupart du temps de différencier la série de départ jusqu'à obtenir une série stationnaire, en effet, avec des données réelles on n'excède jamais un d supérieur à 3 (Il existe d'autres techniques pour modifier la stationnarité d'une série temporelle, mais nous utiliserons celle-ci tout au long de notre étude).

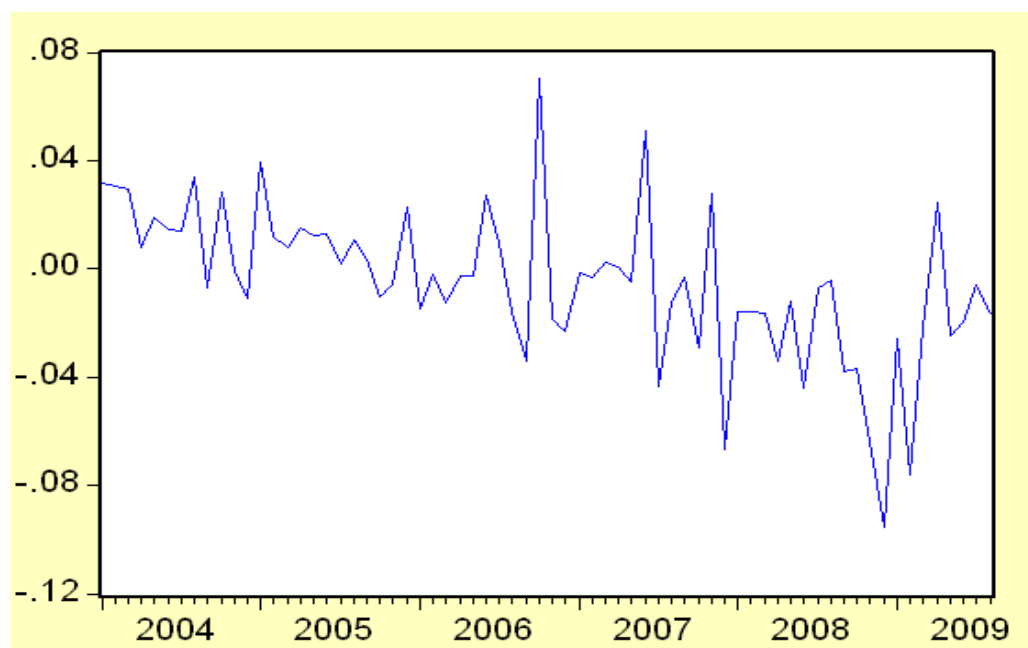
On notera $\Delta X_t = X_t - X_{t-1}$ la série différenciée et $\Delta^k X_t = \Delta^{k-1}(X_t - X_{t-1})$ la série k -fois différenciée.

Même si on peut déterminer graphiquement la stationnarité d'une série (voir exemples ci-dessous), il existe de nombreux tests statistiques pour obtenir des résultats plus fiables, tels que les tests KPSS, de Dickey-Fuller ou de Dickey-Fuller augmenté.



Graphiques d'une série et de ses deux premières différenciées ΔX_t et $\Delta^2 X_t$. Il apparaît que seule la deuxième différenciée est stationnaire.

Nous présentons dans un premier temps les graphiques associées aux séries de pourcentages de souscription/rachat pour les particuliers et les clients institutionnels.



Graphique de l'évolution temporelle de %SR des particuliers

A partir de ce graphique, il ne paraît pas illusoire de penser que cette série est stationnaire. Cette intuition est confirmée par le test de Dickey-Fuller augmenté :

Null Hypothesis: PFS has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic based on SIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-9.144008	0.0000
Test critical values: 1% level	-4.100935	
5% level	-3.478305	
10% level	-3.166788	

*MacKinnon (1996) one-sided p-values.

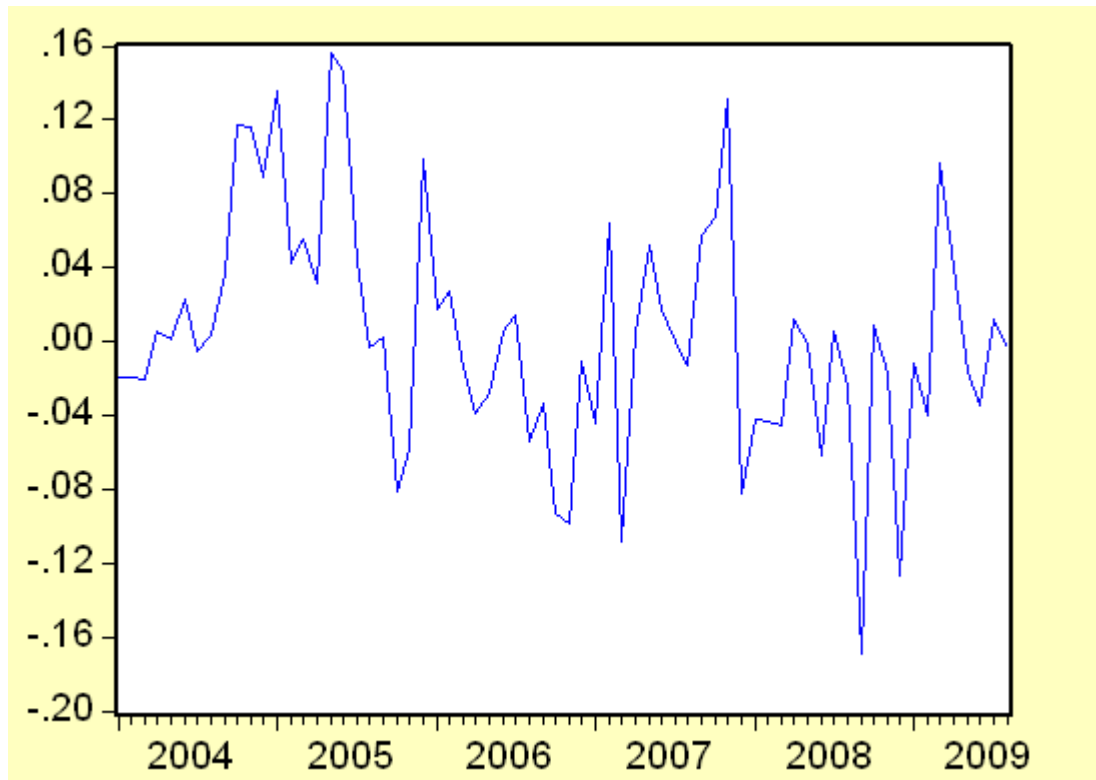
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(PFS)
Method: Least Squares
Date: 11/14/09 Time: 16:29
Sample (adjusted): 2004M02 2009M08
Included observations: 67 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PFS(-1)	-1.135365	0.124165	-9.144008	0.0000
C	0.026101	0.006811	3.832103	0.0003
@TREND(2004M01)	-0.000940	0.000188	-4.985055	0.0000
R-squared	0.566564	Mean dependent var	-0.000720	
Adjusted R-squared	0.553019	S.D. dependent var	0.036848	
S.E. of regression	0.024635	Akaike info criterion	-4.525524	
Sum squared resid	0.038842	Schwarz criterion	-4.426806	
Log likelihood	154.6050	F-statistic	41.82868	
Durbin-Watson stat	1.987686	Prob(F-statistic)	0.000000	

En effet, la statistique du test de Dickey-Fuller est supérieure à toutes les valeurs critiques du test.

Pour cette série, l'ordre d'intégration est donc $d=0$.

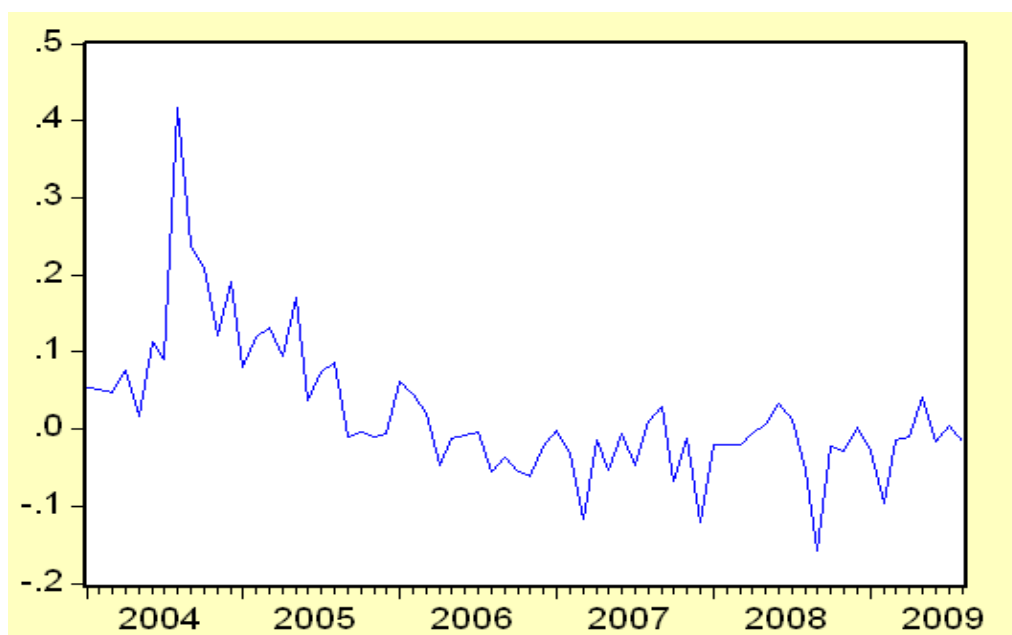
L'autre série que nous étudions, la série des clients institutionnels, l'ordre d'intégration est également de 0, ce qui est décelable graphiquement.



Graphique de l'évolution temporelle de %SR des investisseurs institutionnels

Les deux séries que nous allons étudier sont donc stationnaires.

Prenons l'exemple de la série des pourcentages de souscription/rachat des clients de la banque privée. Cette série ne semble en effet pas stationnaire à première vue :



Evolution temporelle du %SR des clients de la banque privée

En effet, il semble se dégager une non-stationnarité due à une tendance baissière s'étendant de septembre 2004 à février 2007. Cette interprétation graphique semble être confirmée par les résultats du test de Dickey-Fuller augmenté :

Null Hypothesis: GPB has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic based on SIC, MAXLAG=10)

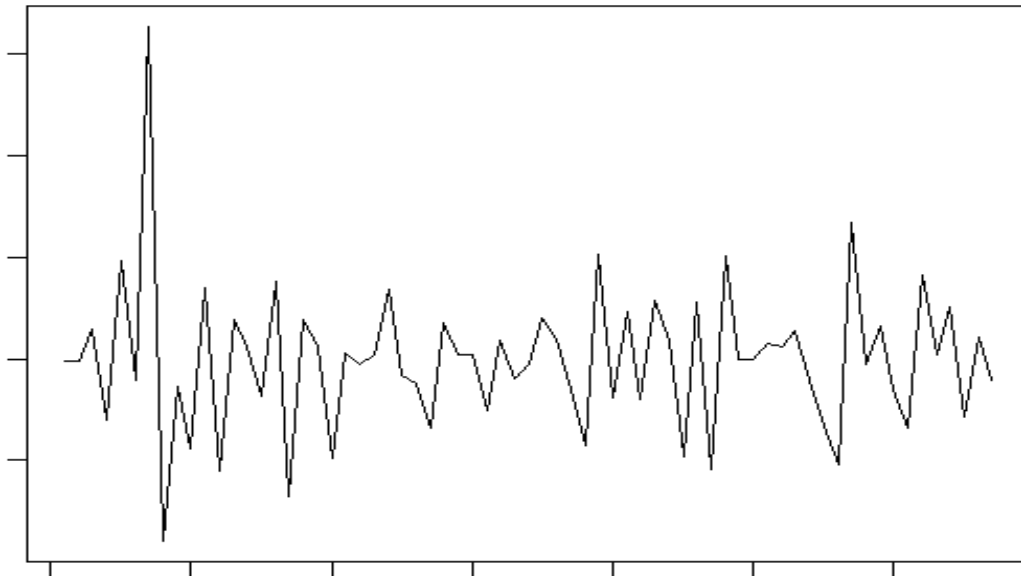
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.078196	0.1201
Test critical values: 1% level	-4.103198	
5% level	-3.479367	
10% level	-3.167404	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(GPB)
Method: Least Squares
Date: 11/14/09 Time: 18:08
Sample (adjusted): 2004M03 2009M08
Included observations: 66 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
GPB(-1)	-0.388586	0.126238	-3.078196	0.0031
D(GPB(-1))	-0.268079	0.122630	-2.186085	0.0326
C	0.045174	0.021895	2.063198	0.0433
@TREND(2004M01)	-0.001117	0.000541	-2.063050	0.0433
R-squared	0.318837	Mean dependent var	-0.001025	
Adjusted R-squared	0.285877	S.D. dependent var	0.073455	
S.E. of regression	0.062074	Akaike info criterion	-2.662293	
Sum squared resid	0.238896	Schwarz criterion	-2.529587	
Log likelihood	91.85567	F-statistic	9.673590	
Durbin-Watson stat	1.930099	Prob(F-statistic)	0.000025	

La statistique du test est effectivement supérieure à toutes les valeurs critiques, on en conclut que cette série n'est pas stationnaire, pour les différents seuils considérés (1%, 5% et 10%). On choisit alors de différencier cette série. Grâce au logiciel R, nous obtenons alors les résultats suivant :



Graphique de la série différenciée des %SR de la clientèle Banque

Augmented Dickey-Fuller Test

```
data: diff(GPB)
Dickey-Fuller = -4.5051, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(diff(GPB)) : p-value smaller than printed p-value
```

Le test de Dickey-Fuller augmenté du logiciel R valide l'hypothèse selon laquelle la série différenciée des pourcentages de souscription/rachat des clients de la banque privée est stationnaire. Pour cette série, on a donc $d=1$.

III.2 d/ii. Estimation des ordres p et q d'un modèle ARMA (p, q)

Pour l'estimation des paramètres p et q , on utilise le fait que si (X_t) suit un ARIMA (p, d, q), alors $(1-L)^d X_t$ suit asymptotiquement un processus ARMA(p, q). En pratique, l'idée est d'appliquer la règle suivante : si $(X_t) \sim \text{ARIMA}(p; d; q)$ alors $(1-L)^d X_t \sim \text{ARMA}(p, q)$.

Dans ce cas, l'approche utilisée est différente et repose sur la constatation suivante : un peu comme dans le cas du R^2 , le choix des paramètres p et q doit se faire en prenant en considération l'ajustement du modèle (le fit) contre le nombre de paramètres estimés.

Un modèle qui contient beaucoup de paramètres aura nécessairement un très bon ajustement à l'intérieur de l'échantillon (faible variance des résidus, etc.) mais pourrait être moins performant au niveau prévisionnel. En effet, il a souvent été constaté qu'un modèle contenant beaucoup de paramètres était par construction plus enclin à expliquer tous les phénomènes observés durant la période étudiée même si ces mêmes phénomènes étaient le plus souvent très spécifiques et peu susceptibles de se reproduire dans le futur.

Méthode pratique d'estimation des ordres p et q :

L'idée est de construire une mesure de l'écart entre le vrai modèle supposé être le processus générateur des pourcentages de souscription/rachat observés et les modèles ARMA dont on postule qu'ils peuvent approximer le vrai modèle.

La mesure usuellement retenue est la quantité d'information de Kullback-Leibler.

Soit $f_0(x)$ la densité inconnue des observations x dont on dispose et $f_{p,q}(x)$ la densité associée au modèle ARMA(p, q).

L'écart entre f_0 et $f_{p,q}$ est mesuré par :

$$I_K(f_0, f_{p,q}) = E_{f_0} \left(\ln \frac{f_0(X)}{f_{p,q}(X)} \right) = \int \frac{f_0(x)}{f_{p,q}(x)} f_0(x) dx$$

Comme on ne connaît pas cette quantité, il convient de l'estimer. Pour cela, il existe différents estimateurs dont les propriétés diffèrent.

Ainsi, si on dispose de T observations (dans notre étude $T \approx 70$), ces estimations s'obtiennent à partir des critères informationnels :

$$AIC(p, q) = \ln \hat{\sigma}_\varepsilon(p, q) + 2 \frac{p+q}{T}$$

$$BIC(p, q) = \ln \hat{\sigma}_\varepsilon(p, q) + (p+q) \frac{\ln T}{T}$$

$$HQ(p, q) = \ln \hat{\sigma}_\varepsilon(p, q) + 2(p + q) \frac{\ln(\ln T)}{T}$$

où $\hat{\sigma}_\varepsilon(p, q)$ représente la variance résiduelle non corrigée lorsque l'on estime le modèle ARMA dont les ordres sont fixés à p et q .

A partir du choix de l'un des critères, le modèle à retenir est celui dont le critère est minimal (on souhaite minimiser l'écart entre f_0 et $f_{p,q}$).

Pour chaque groupement de clients on utilise alors l'algorithme qui passe en revue les différents modèles et sélectionne in fine celui qui minimise le critère d'information retenu (voir code R en annexe). Cet algorithme est utilisé conjointement avec l'analyse des fonctions d'autocorrélation et d'autocorrélation partielles.

Lorsque l'on étudie les autocorrélogrammes de la série des pourcentages de souscription/rachat des clients particuliers, il semble raisonnable de proposer l'étude des modèles suivants :

- AR(2)
- MA(1)
- ARMA(2,1)
- ARMA(1,2)

Date: 11/14/09 Time: 20:12
Sample: 2004M01 2009M08
Included observations: 68

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.237	0.338	3.9943	0.046
		2	0.396	0.298	12.218	0.002
		3	0.338	0.263	23.682	0.000
		4	0.185	-0.001	26.235	0.000
		5	0.317	0.130	33.829	0.000
		6	0.182	-0.023	36.383	0.000
		7	0.144	-0.033	37.997	0.000
		8	0.263	0.108	43.496	0.000
		9	0.020	-0.128	43.529	0.000
		10	0.132	-0.003	44.963	0.000
		11	0.144	0.050	46.685	0.000
		12	0.012	-0.030	46.698	0.000
		13	0.126	0.021	48.072	0.000
		14	-0.015	-0.076	48.092	0.000
		15	0.046	0.022	48.281	0.000
		16	0.093	0.043	49.077	0.000
		17	0.041	0.091	49.236	0.000
		18	0.043	-0.044	49.414	0.000
		19	-0.011	-0.090	49.426	0.000
		20	0.002	-0.009	49.427	0.000
		21	0.099	0.101	50.420	0.000
		22	0.003	0.027	50.421	0.001
		23	0.011	-0.055	50.433	0.001
		24	0.067	0.021	50.920	0.001
		25	-0.008	-0.010	50.927	0.002
		26	0.069	0.059	51.461	0.002
		27	0.035	0.016	51.604	0.003
		28	0.026	-0.027	51.687	0.004

Autocorrélogramme et autocorrélogramme partiel du %SR des particuliers

On obtient alors les coefficients suivants pour les quatre modèles étudiés :

Modèle	logLik	AIC
AR(2)	152.1670	-296.3340
MA(1)	148.9107	-291.8213
ARMA(2,1)	157.7213	-305.4426
ARMA(1,2)	158.0971	-306.1942

Au vue des critères d'information, il apparait que l'on peut rejeter les modèles AR(2) et MA(1).

Le modèle ARMA(1,2) semble le plus intéressant, c'est celui que nous étudierons par la suite.

En revanche, concernant la série des pourcentages de souscription/rachat des entreprises multinationales, on obtient un modèle avec davantage de paramètres, puisque après avoir étudié les autocorrélogrammes et après avoir minimisé les critères d'information, le modèle retenu est un ARMA(2,9).

On résume ici les modèles obtenus pour tous les Customer group :

Customer group	Modèle
Particuliers	ARMA(1,2)
Banques	ARMA(2,9)
Banque privée	ARIMA(1,1,1)
Investisseurs institutionnels	ARMA(5,2)

Tableau résumant les différentes modélisations obtenus pour modéliser le comportement des différentes catégories d'investisseurs

On remarque une palette de modèles assez large, ce qui confirme la nature relativement diversifiée des investisseurs en terme de souscription et de rachat.

Cependant, à ce stade on ne peut encore juger de la pertinence des modèles obtenus. En effet, bien qu'ayant choisis, dans chaque cas, le modèle qui minimise le critère d'information retenu, il n'est pas garanti que ce modèle soit statistiquement fiable.

Nous jugerons de la fiabilité des modèles obtenus grâce à la mise en place de divers tests dans la partie III.2 f/.

III.2 e/ Détermination des paramètres

A cette étape, les paramètres p, d et q ont été fixés. Il convient alors d'estimer les coefficients ϕ_i et θ_j du processus ARIMA (p, d, q) de la forme :

$$X_t + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

Ces coefficients sont calculés par le logiciel R par des méthodes du type maximum de vraisemblance.

Pour déterminer la vraisemblance d'un processus ARMA, il est nécessaire de supposer que la loi des erreurs est gaussienne. Les erreurs étant normalement distribuées et indépendantes (ε_t étant un bruit blanc), le vecteur $(\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_1)$ est un vecteur gaussien. Les composantes du vecteur $(X_1, X_2, \dots, X_t)$ étant obtenues par combinaisons linéaires des composantes du vecteur $(\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_1)$, $(X_1, X_2, \dots, X_t)$ est un vecteur gaussien.

On peut alors calculer la vraisemblance du modèle :

$$L(X, \phi, \theta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \frac{1}{\sqrt{\det(XX')}} \exp\left(-\frac{1}{2\sigma^2} X'X\right)$$

Il reste alors à maximiser cette vraisemblance pour les différents paramètres recherchés en passant par le logarithme et utilisant des méthodes du type moindres carrés. Le logiciel R procède ainsi pour estimer les différents coefficients ϕ_i et θ_j .

On obtient alors les paramètres suivants pour le processus ARMA(1,2) adapté à la série des pourcentages de souscription/rachat des investisseurs particuliers :

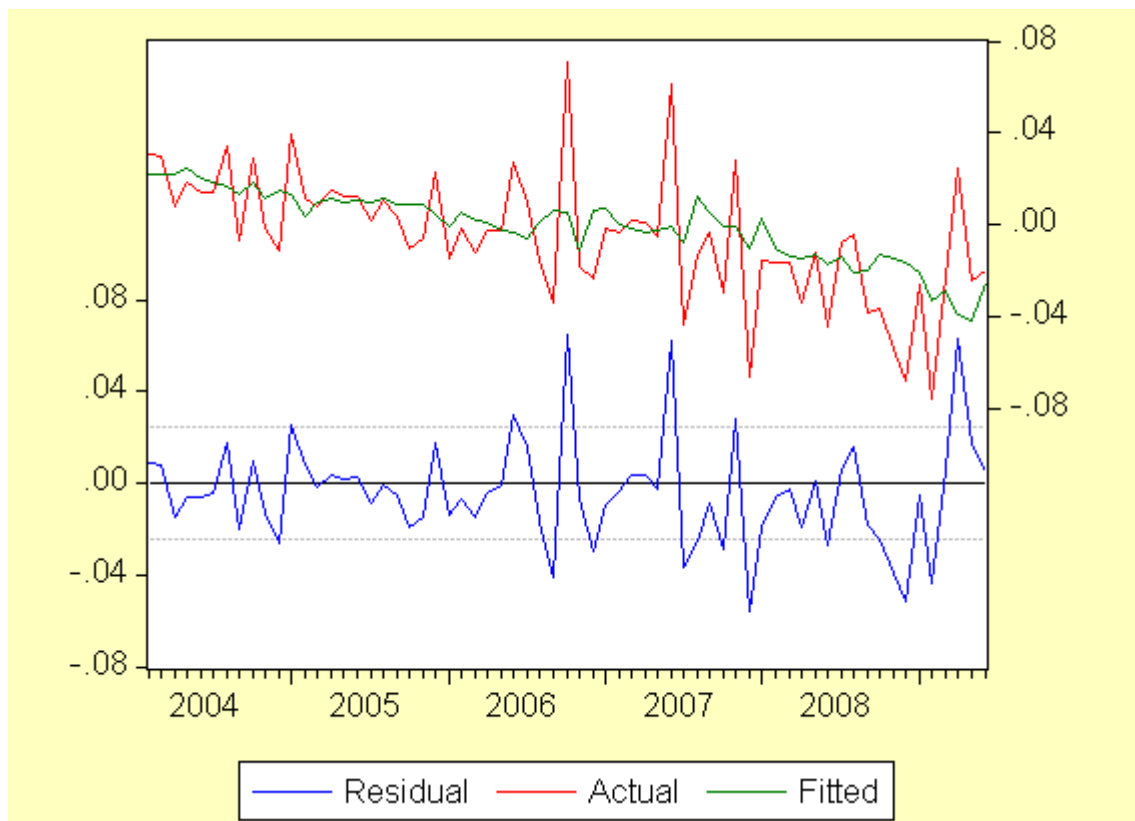
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.954532	0.059575	16.02222	0.0000
MA(1)	-1.054057	0.128417	-8.208102	0.0000
MA(2)	0.263588	0.130646	2.017577	0.0480
R-squared	0.268931	Mean dependent var		-0.004621
Adjusted R-squared	0.245348	S.D. dependent var		0.027851
St.E. of regression	0.024195	Akaike info criterion		-4.560319
Sum squared resid	0.036293	Schwarz criterion		-4.459963
Log likelihood	151.2104	Durbin-Watson stat		1.986412
Estimated AR Roots	.95			
Estimated MA Roots	.65	.41		

Le modèle est donc régi par l'équation suivante :

$$X_t + 0,954532 \times X_{t-1} = \varepsilon_t - 1,054057 \times \varepsilon_{t-1} + 0,263588 \times \varepsilon_{t-2}$$

qui peut également s'écrire sous la forme $\Phi(L)X_t = \Theta(L)\varepsilon_t$ avec $\Phi(L) = I + 0,954532 \times L$ et $\Theta(L) = I - 1,054057 \times L + 0,263588 \times L^2$.

On présente le graphique de la série des pourcentages de souscription/rachat ainsi que sa modélisation par le processus autorégressif. En dessous on peut remarquer le graphe des résidus du modèle. Ces résidus font l'objet de la partie suivante.



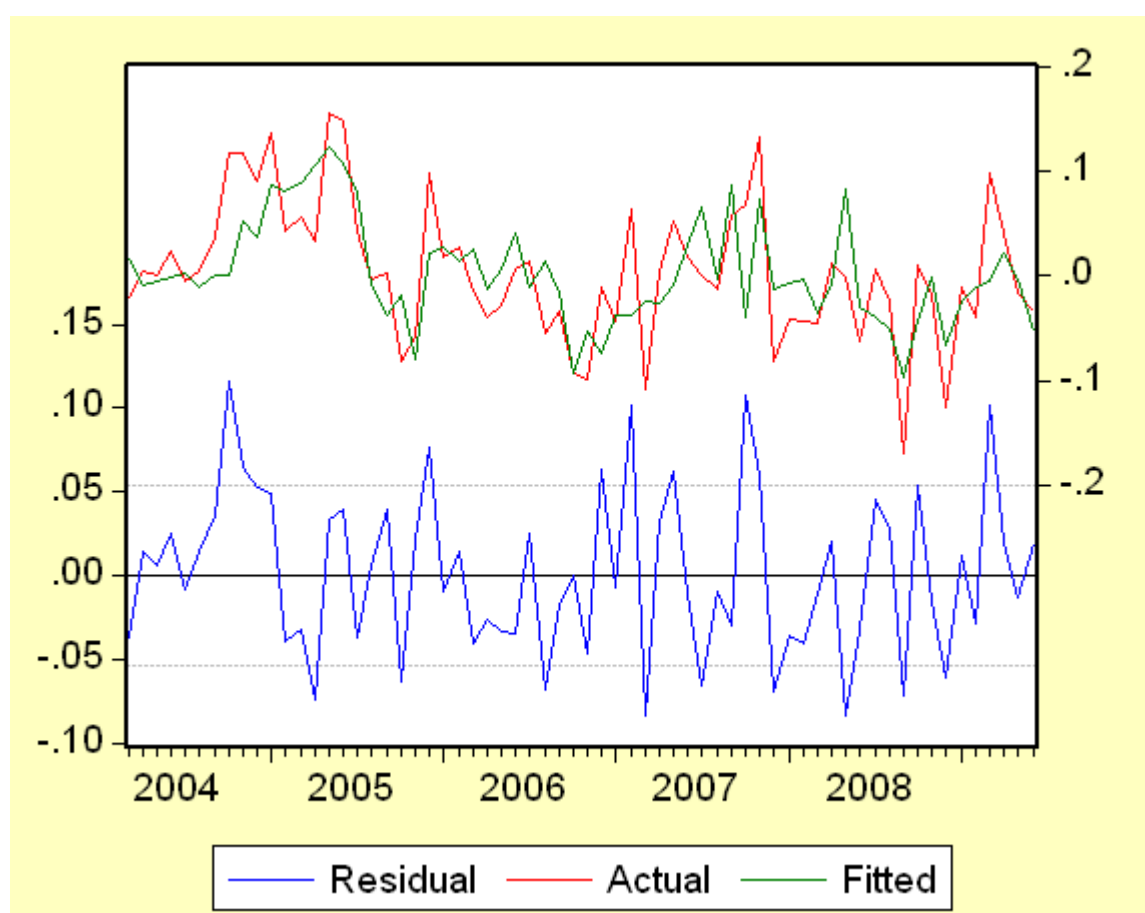
Graphique de la modélisation du pourcentage de souscription/rachat des investisseurs particuliers par un ARMA(1,2)

On remarque que si la modélisation de la tendance est à peu près correcte, les pics quant à eux ne sont que rarement pris en compte par le modèle mis en place.

Concernant la série des pourcentages de souscription/rachat des compagnies financières, on obtient l'estimation des coefficients suivants :

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.414078	0.237848	-1.740934	0.0875
AR(2)	-0.402639	0.239542	-1.680873	0.0987
MA(1)	0.711372	0.180487	3.941397	0.0002
MA(2)	0.573129	0.211741	2.706742	0.0091
MA(3)	0.337531	0.109962	3.069533	0.0034
MA(4)	0.371431	0.086197	4.309089	0.0001
MA(5)	0.546876	0.085048	6.430227	0.0000
MA(6)	0.684684	0.115591	5.923338	0.0000
MA(7)	1.010332	0.146753	6.884581	0.0000
MA(8)	0.634261	0.199742	3.175395	0.0025
MA(9)	0.740788	0.153510	4.825681	0.0000

On en déduit alors l'équation régissant le modèle. On obtient alors la courbe suivante avec les données de base et les données modélisées par le processus ARMA(2,9) :



Données originales et données modélisées de la clientèle GBM

En étudiant ce graphique, il apparaît que la modélisation semble correcte dans l'ensemble et que les résidus du modèle paraissent normalement distribués autour de l'origine. On remarque

une légère hausse de la volatilité des résidus pendant l'année 2007, mais cette hausse reste modérée.

Nous allons maintenant tester l'adéquation des différents modèles utilisés dans le but d'établir des prévisions des souscriptions et des rachats pour les différentes catégories de clients. Ces tests sont basés sur les résidus de ces modèles.

On notera que les graphiques des modélisations des séries de pourcentages de souscription/rachat des autres clients se trouvent en annexe.

III.2 f/ Vérifications à posteriori : Tests d'adéquation de modèles

III.2 f/ i. Test de bruit blanc et de stationnarité

L'hypothèse $(X_t) \sim \text{ARMA}(p, q)$ peut s'écrire $\Phi(L)X_t = \Theta(L)\varepsilon_t$ ou $\varepsilon_t = \Theta(L)^{-1}\Phi(L)X_t$.

Une fois estimés les paramètres p et q et l'ensemble des ϕ_i et θ_j , on obtient des polynômes estimés $\hat{\Phi}(L)$ et $\hat{\Theta}(L)$, qui permettent d'obtenir les résidus estimés : $\hat{\varepsilon}_t = \hat{\Theta}(L)^{-1}\hat{\Phi}(L)X_t$

Pour que les modèles obtenus précédemment soient valides, il convient de vérifier que les résidus estimés suivent bien un bruit blanc.

Pour tester si les résidus sont des bruits blancs, nous utilisons divers tests correspondants à la définition des bruits blancs.

Analyse des fonctions d'autocorrélation

L'autocorrélation des résidus peut être estimée par :

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \quad \text{où} \quad \hat{\gamma}(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (\hat{\varepsilon}_t - \bar{\varepsilon})(\hat{\varepsilon}_{t+h} - \bar{\varepsilon}) \quad \text{et} \quad \bar{\varepsilon} = \frac{1}{n} \sum_{t=1}^n \varepsilon_t$$

L'intervalle de confiance de $\hat{\rho}(h)$ est, dans le cas d'un bruit blanc gaussien $\left[\frac{-t_{\alpha/2}}{\sqrt{T}}; \frac{t_{\alpha/2}}{\sqrt{T}} \right]$ où

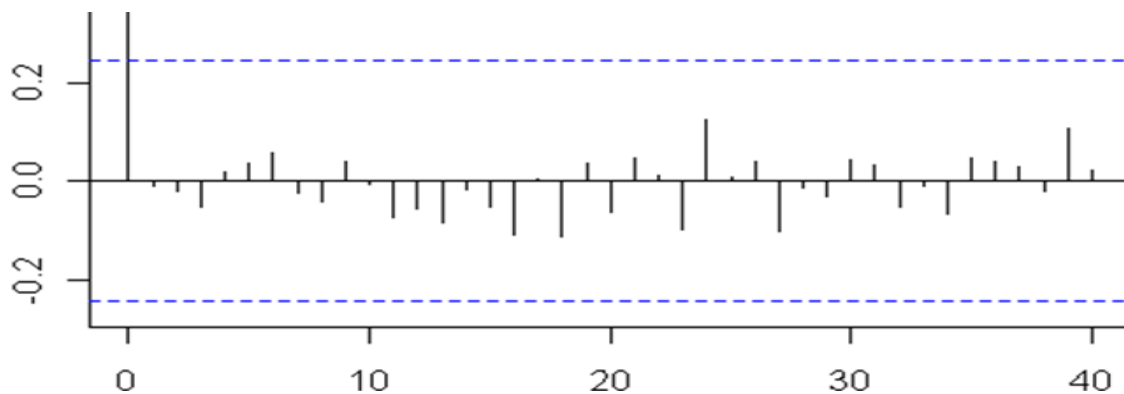
$t_{\alpha/2}$ est le quantile d'ordre $\alpha/2$ de la loi de Student (1.96 pour $\alpha=5\%$). Pour avoir un bruit blanc, il est nécessaire qu'aucune valeur de l'autocorrélogramme ne soit significativement non nulle.

Dans notre cas, comme l'historique des données étudiées remonte à janvier 2004 et s'arrête en mai 2009, on a $T = 68$.

L'intervalle de confiance à 95% concernant la nullité de la corrélation des résidus est donc :

$$\frac{t_{\alpha/2}}{\sqrt{T}} = \frac{1.96}{\sqrt{68}} = 0.243, \text{ d'où } IC_{95\%} = [-0.243, 0.243].$$

Voici l'autocorrélogramme des résidus pour la modélisation du pourcentage de souscription/rachat de la catégorie de clients *particuliers* (les pointillés désignent l'intervalle de confiance défini ci-dessus):



Autocorrélogramme des résidus de la modélisation des %S/R des particuliers par un ARMA(1,2)

Il semble alors que les résidus ne sont pas corrélés entre eux, car aucun coefficient n'est significativement non nul. On complète cette étude graphique par les tests dits du "portmanteau".

Les autocorrélogrammes des résidus des autres modélisations sont semblables au précédent : les résidus ne semblent à priori pas corrélés entre eux.

Statistique de Box-Pierce, ou test de "portmanteau"

Le test de Box-Pierce permet d'identifier les processus de bruit blanc (i.e. les processus aléatoires de moyenne nulle, de variance constante et non autocorrélés). Il vient confirmer ou désavouer l'analyse graphique établie précédemment. On teste ici $Cov(\varepsilon_t, \varepsilon_{t-h}) = 0$ pour tout h , soit $\rho(h) = 0$ pour tout h . Ce test s'écrit :

$$\begin{cases} H_0 : \rho(1) = \rho(2) = \dots = \rho(h) = 0 \\ H_a : \exists i \in \{1, 2, \dots, h\}, \rho(i) \neq 0 \end{cases}$$

Pour effectuer ce test, on utilise la statistique de Box et Pierce (1970) Q , donnée par :

$$Q_h = T \sum_{k=1}^h \hat{\rho}_k^2$$

où h est le nombre de retards, T est le nombre d'observations et $\hat{\rho}_k$ l'autocorrélation empirique. Asymptotiquement, sous H_0 , Q_h suit un χ^2 à h degrés de liberté. Nous rejetons l'hypothèse de bruit blanc au seuil h si Q_h est supérieure au quantile d'ordre $\alpha/2$ de la loi du χ^2 à h degrés de liberté.

Une statistique ayant de meilleures propriétés pour les échantillons de faibles effectifs sera également utilisée :

$$Q_h = T(T+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{T-k}$$

Ce test est alors appelé test de Ljung-box, il est particulièrement adapté à nos données, car il reste puissant malgré nos échantillons de tailles modérées.

Pour la série des pourcentages de souscription/rachat de la clientèle bancaire, nous obtenons les résultats suivants pour ces deux tests :

Box-Pierce test

```
data: arima(CG, c(2, 0, 9))$resid
X-squared = 0.0015, df = 1, p-value = 0.9692
```

Box-Ljung test

```
data: arima(CG, c(2, 0, 9))$resid
X-squared = 0.0016, df = 1, p-value = 0.9685
```

D'après les p-values de ces tests, on valide l'hypothèse H_0 selon laquelle les fonctions d'autocorrélation sont nulles.

Les résidus de cette modélisation apparaissent donc être des bruits blancs.

III.2 f/ ii. Normalité des résidus

Dans le cadre de prévisions, il convient de vérifier la normalité des résidus pour tester l'adéquation de nos modèles. Il existe de nombreux tests pour tester l'adéquation d'une distribution donnée à la loi normale (Kolmogorov-Smirnov, Jarque-Bera,...).

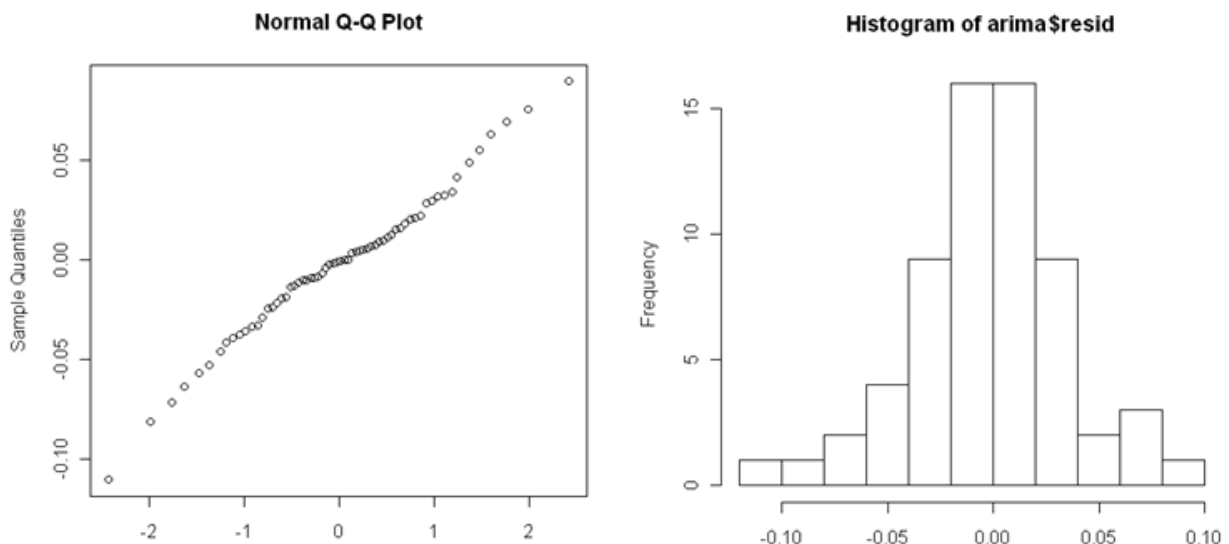
Même si dans la plupart des cas, la normalité peut être détectée de façon graphique (histogramme des fréquences, boîtes à moustaches ou droite de Henry), les tests statistiques amènent un point de vue objectif de la distribution des données.

Nous allons donc, dans un premier temps, utiliser des méthodes graphiques pour déterminer l'éventuel caractère gaussien des résidus des modélisations. Nous vérifierons ensuite les intuitions développées par ces méthodes graphiques par des tests statistiques. On utilisera en particuliers les tests de Jarque-Bera et de Shapiro-Wilk.

Détermination graphique de la normalité

Nous commençons ici par les résultats graphiques concernant la normalité des résidus. En effet, si les résidus d'un modèle ne sont pas gaussiens, cela apparaîtra clairement sur les différents graphiques.

On se propose dans un premier temps d'étudier l'histogramme des résidus de la modélisation des pourcentages de souscription/rachat des clients particuliers par un ARMA(2,1) :



On peut s'assurer graphiquement de la normalité des résidus grâce à ces deux sorties issues du logiciel R.

Le premier graphique, dit Q-Q plot (pour Quantile-Quantile plot) permet de comparer la distribution de deux ensembles de données. On utilise ici la fonction `qqnorm` du logiciel R qui compare la distribution des résidus de notre modèle à celle d'une loi normale.

Aux vues de ces graphes, il ne semble pas illusoire de penser que les résidus issus de la modélisation des pourcentages de souscription/rachat des clients particuliers par un ARMA(2,1) sont gaussiens. En effet, on remarque que le nuage de points du graphique Q-Q plot a une forme de droite diagonale relativement prononcée. De même, l'histogramme de ces résidus semble clairement être celui d'une loi normale.

Au regard de ces différents graphes, on valide a priori l'hypothèse de normalité des résidus pour cette modélisation. En revanche, selon les catégories de clients, tous les résidus ne semblent pas gaussiens.

Dans le but d'apporter un point de vue objectif quant à la détermination de la distribution des résidus de nos modèles, nous allons utiliser des tests statistiques.

Détermination de la normalité des résidus à partir des tests statistiques

Test de Jarque-Bera

Nous commençons par utiliser le test de Jarque-Bera qui est un des plus utilisés pour déterminer le caractère gaussien d'un échantillon donné. Nous proposons dans un premier temps des rappels théoriques concernant ce test.

Le test de normalité de Jarque-Bera est fondé sur les coefficients d'asymétrie et d'aplatissement. Il évalue les écarts simultanés de ces coefficients avec les valeurs de référence de la loi normale.

Prenons les coefficients d'asymétrie et d'aplatissement de Pearson ($\beta_1 = \frac{\mu_3}{\sigma^3}$ et $\beta_2 = \frac{\mu_4}{\sigma^4}$), on en propose les estimateurs suivants :

$$b_1 = \frac{\frac{1}{n} \sum_i (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_i (x_i - \bar{x})^2\right)^{\frac{3}{2}}}$$

$$b_2 = \frac{\frac{1}{n} \sum_i (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_i (x_i - \bar{x})^2\right)^2}$$

La loi conjointe de ces estimateurs est normale bivariée :

$$\sqrt{n} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \sim \mathcal{N} \left[\begin{pmatrix} 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 6 & 0 \\ 0 & 24 \end{pmatrix} \right]$$

La forme quadratique associée permet de produire la statistique de Jarque-Bera T qui s'écrit :

$$T = n \left(\frac{b_1^2}{6} + \frac{(b_2 - 3)^2}{24} \right)$$

Elle est distribuée asymptotiquement selon une loi du χ^2 à 2 degrés de liberté. La statistique T prend des valeurs d'autant plus élevées que l'écart entre la distribution empirique et la loi normale est manifeste. La région critique pour un risque α du test est définie par :

$$R.C. : T > \chi^2_{1-\alpha}(2)$$

Pour un risque $\alpha = 0.05$, le seuil critique est $\chi^2_{0.95}(2) = 5.99$.

Grâce au logiciel R, on obtient les résultats suivants pour les résidus de la modélisation de la série des pourcentages de souscription/rachat de la clientèle bancaire par un ARMA(2,9) :

```
> jarqueberaTest(arima(CG,c(2,0,9))$resid)

Title:
Jarque - Bera Normalality Test

Test Results:
STATISTIC:
X-squared: 0.1693
P VALUE:
Asymptotic p Value: 0.9188
```

On peut alors conclure au vu de la statistique et de sa p-value que l'on ne rejette pas l'hypothèse de normalité des résidus.

Voici les résultats du test de Jarque-Bera pour les différentes modélisations :

Catégorie de clients	Modèle	Stat T	P-value
Particuliers	ARMA(1,2)	1.8507	0.3964
Banques	ARMA(2,9)	0.1693	0.9188
Banque privée	ARIMA(1,1,1)	2.6849	0.2612
Investisseurs institutionnels	ARMA(5,2)	29.3352	4.265e-07

Au vu de ce tableau, on constate que les résidus de tous les modèles apparaissent comme normaux, excepté le modèle ARMA(5,2) retenu pour modéliser le comportement des investisseurs de type institutionnels. Ces résultats sont donc globalement positifs quant à la validité des modélisations.

Dans la pratique, les écarts de puissance s'amenuisent à mesure que les effectifs augmentent. Cependant, la simplicité des calculs, très faciles à appréhender et à mettre en œuvre sur des outils simples tels qu'un tableur, militent en faveur du test de Jarque-Bera.

En vérité, ce test est toujours moins puissant que le test de Shapiro-Wilk, dans le sens où il a une propension plus élevée à conclure à la compatibilité avec la loi normale. On devrait donc toujours préférer ce dernier.

Nous allons donc dans un deuxième temps utiliser le test de Shapiro-Wilk, plus robuste et plus adapté à nos données. En effet, le fait que nos séries temporelles ne soient composées que de 68 données, nécessite d'utiliser des tests relativement plus complexes.

Test de Shapiro-Wilk

Très populaire, le test de Shapiro-Wilk est basé sur la statistique W . En comparaison des autres tests, il est particulièrement puissant pour les petits effectifs ($n < 100$), et donc particulièrement adapté à nos données. La statistique du test s'écrit :

$$W = \frac{\left[\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} a_i (x_{(n-i+1)} - x_{(i)}) \right]^2}{\sum_i (x_i - \bar{x})^2}$$

Où

- _ $x_{(i)}$ correspond à la série des données triées ;
- _ a_i sont des constantes générées à partir de la moyenne et de la matrice de variance covariance des quantiles d'un échantillon de taille n suivant la loi normale. Ces constantes sont fournies dans des tables spécifiques.

La statistique W peut donc être interprétée comme le coefficient de détermination (le carré du coefficient de corrélation) entre la série des quantiles générées à partir de la loi normale et les quantiles empiriques obtenues à partir des données. Plus W est élevé, plus la compatibilité avec la loi normale est crédible. La région critique, rejet de la normalité, s'écrit :

$$R:C: : W < W_{crit}$$

Les valeurs seuils W_{crit} pour différents risques α et effectifs n sont lues dans la table de Shapiro-Wilk (voir en annexe).

Voici la sortie R que nous obtenons pour la série des résidus issus de la modélisation des pourcentages de souscription/rachat de la clientèle institutionnelle par un ARMA(2,9) :

```
> shapiro.test(arima(CG,c(2,0,9))$resid)

Shapiro-Wilk normality test

data:  arima(CG, c(2, 0, 9))$resid
W = 0.9728, p-value = 0.143
```


En comparant la p-value de ce test et celle du précédent, on remarque que bien que ce test ne rejette pas non plus l'hypothèse de normalité des résidus (car supérieure au seuil de 5%), ce dernier est tout de même moins catégorique que le précédent.

Voici les résultats du test de Shapiro-Wilk pour les différentes modélisations :

Catégorie de clients	Modèle	Stat W	P-value
Particuliers	ARMA(1,2)	0,9556	0,01652
Banques	ARMA(2,9)	0,9728	0,143
Banque privée	ARIMA(1,1,1)	0,8215	1,29E-07
Investisseurs institutionnels	ARMA(5,2)	0,9297	0,00088

On remarque que ce test rejette la normalité des résidus de certaines modélisations, là où le test de Jarque-Bera ne rejette pas cette hypothèse, ce qui est le cas notamment pour les particuliers. Cependant, on aura tendance à s'appuyer sur les résultats issus de ce dernier, dans la mesure où il est plus puissant que le précédent.

Ce test confirme également le rejet de l'hypothèse de normalité des résidus pour les clients institutionnels et les banques.

L'adéquation des modèles retenus aux données originales semble relativement faible pour ceux dont on ne peut conclure quant à la normalité des résidus. Pour autant, ces modélisations serviront dans tout les cas à établir des prévisions.

Avant de procéder à la phase de prévisions basées sur les modèles trouvés, nous avons souhaité analyser l'éventuel impact de la récente crise financière sur les comportements des clients en terme de rachat. Plus précisément, nous avons souhaité savoir si celle-ci a provoquée une quelconque rupture dans les séries temporelles des pourcentages de souscription/rachat des différentes catégories de clients.

L'impact de la crise financière sur nos modélisations peut se traduire par une instabilité des paramètres des modèles utilisés.

Nous allons donc étudier la stabilité de nos modèles à travers des tests de rupture de tendance.

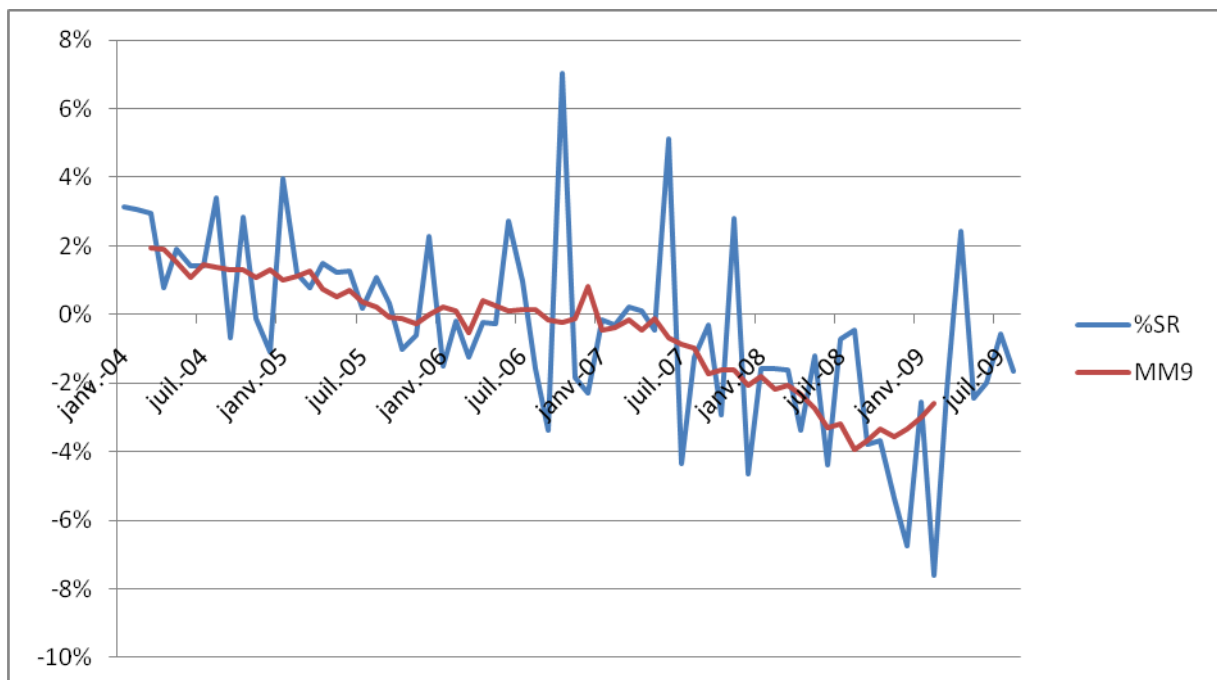
III.2 g/ Etude de la stabilité temporelle des modèles utilisés

L'étude de la stabilité des modèles peut être une manière d'analyser l'impact probable de la récente crise économique-financière sur les comportements des clients d'une société de gestion.

Cependant, il est intéressant de remarquer qu'il est également possible de détecter une éventuelle influence du contexte économique ou financier sur le comportement des clients sans pour autant constater une rupture de tendance dans le modèle retenu.

III.2 g/ i. Détermination graphique de la stabilité

Prenons l'exemple de la série des pourcentages de souscription/rachat des clients particuliers :



Graphes des %SR et de la moyenne mobile d'ordre 9 des clients particuliers

On remarque, en étudiant la moyenne mobile calculée sur 9 mois, la tendance baissière de cette série de pourcentages. Pour autant, cette baisse apparaît comme régulière tout au long de la série, il ne semble pas exister de rupture de tendance pour cette catégorie de clients.

On peut alors penser que la crise financière ne s'est pas répercutée de façon brutale sur le comportement des investisseurs particuliers.

Malgré que l'on constate que le rachat du mois d'août 2007 (début de la crise financière pour certains analystes) de -4,4% est le plus important depuis le début de l'historique,

graphiquement, on ne décèle aucune rupture de tendance, ni à cette date là, ni aux autres dates clés de la crise.

L'hypothèse selon laquelle les comportements des investisseurs ont été influencés par un contexte économique relativement tendu semble légitime, et demande à ce qu'on s'y intéresse de manière plus précise et moins subjective.

Pour cela, nous allons mettre en place un test de rupture de tendance, qui permettra de déterminer la stabilité de nos modèles.

Il est important d'avoir des modèles stables au cours du temps pour réaliser des prévisions. Si un modèle comporte un ou plusieurs « breaks », alors les prévisions relatives à cette série ne devront être basées que sur la dernière partie de cette série.

III.2 g/ ii. Détermination de la stabilité des modèles à l'aide de tests statistiques

Il existe de nombreux tests statistiques portant sur la stabilité de modélisation de séries temporelles par processus autorégressifs : test du CUSUM, du CUSUM carré, de Chow,...

Le test de Chow (1960) est un test de rupture de tendance basé sur les résidus du modèle, grâce à ce dernier, on pourra tester la présence de ruptures (breakpoints) dans nos modélisations à des dates données. Après avoir décrit son fonctionnement, on l'appliquera à nos différentes séries.

Le test de Chow peut s'écrire sous la forme d'un test de plusieurs contraintes. On considère une date susceptible d'avoir induit une modification de l'équation de comportement, on note t_0 cette date et on cherche à tester si les paramètres sont identiques sur l'ensemble de la période $1, \dots, T$ ou si les paramètres sont différents sur les sous-périodes $1, \dots, t_0$ et $t_0+1, \dots, T$, autrement dit on effectue le test suivant :

$$H_0 : B_1 = B_2 = B$$

$$H_a : B_1 \neq B_2$$

avec B_1 le vecteur de paramètre sur la première période et B_2 le vecteur sur la seconde. Sous l'hypothèse nulle il y a stabilité, il s'agit d'un modèle avec k contraintes car chaque paramètre doit être identique sur les deux sous-périodes. Sous l'hypothèse alternative, on a un modèle sans contraintes, il faut dans ce cas estimer le modèle sur les deux sous-périodes pour obtenir $2k$ paramètres estimés.

On calcule la statistique de Fisher :

$$F = \frac{(\tilde{u}'\tilde{u} - (u_1'u_1 + u_2'u_2))/k}{(u_1'u_1 + u_2'u_2)/(T - 2k)},$$

Où $\sum \hat{u}_i^2$ est la somme des carrés des résidus du modèle, $\sum \hat{u}_i^2 / n_i$ est la somme des carrés des résidus des sous-échantillons, T le nombre total d'observations, k le nombre de paramètres de l'équation.

Asymptotiquement, F tend vers une loi de Fisher à k et T-2k paramètres.

Si la statistique calculée est inférieure à la valeur tabulée on accepte l'hypothèse de stabilité. La date t_0 ne constitue alors pas une date de rupture, il n'y a pas de changement structurel.

En pratique, on n'aura une idée de la date t_0 à tester qu'après avoir analysé graphiquement la série temporelle. Dans notre cas, les dates t_0 que nous testerons pour chaque série seront les suivantes :

- Août 2007 (BNP Paribas ferme 3 de ses fonds)
- Septembre 2008 (Faillite de Lehman Brothers)
- Octobre 2008 (Annonce des différents plans de soutien aux établissements bancaires)

On pourra être amené à considérer, pour chaque date et selon les cas, un décalage d'un mois concernant l'impact de l'événement sur le comportement des clients.

Concernant les comportements des investisseurs particuliers, dont les pourcentages de souscription/rachat ont été modélisés par un ARMA(2,1), il apparaît différentes ruptures de tendances aux dates testées.

En effet, les différents tests de Chow effectués présentent cette forme :

Chow Breakpoint Test: 2008M09

F-statistic	6.317638	Probability	0.000840
Log likelihood ratio	18.12779	Probability	0.000414

Dans ce cas, la statistique F est supérieure à la valeur tabulée de la loi de Fisher qui est de 2,76 pour k=3 et T =68.

Partant de ce constat, on va tenter de déceler où commence la rupture pour extraire le plus grand nombre de données possible.

Dans le cas des investisseurs particuliers, on décèle une rupture dès le mois de juin 2007 (ce qui correspond à la faillite de la banque Bear Stearns). Ce mois est en réalité le premier pour lequel la statistique de Chow est supérieure à la valeur critique.

Pour effectuer des prévisions plus précises, nous allons donc, dans ce cas, proposer une nouvelle modélisation de la série, ou plus précisément de la sous-série qui débute en juin 2007.

Le nouveau modèle pour la série des pourcentages de souscription/rachat des particuliers est un ARMA(3,2), dont voici la modélisation :



Modélisation du pourcentage de souscription/rachat des investisseurs particuliers (sur les 24 derniers mois) par un ARMA (3,2)

Les résidus apparaissent comme étant gaussiens, ce qui est confirmé par la p-value du test de Shapiro-Wilk de 0,671. Les prévisions des pourcentages de souscription/rachats des investisseurs particuliers seront donc basées sur ce modèle et non sur le précédent.

On note que dans l'éventualité où la première rupture de tendance se soit produite après le mois de janvier 2008, il aurait été impossible de proposer une nouvelle modélisation pour cette série, car un modèle basé sur un échantillon de moins de 20 données n'a pas de sens statistique.

Concernant les autres modèles, on ne décèle aucune rupture de tendance. On conservera donc ces modèles pour établir des prévisions.

Après avoir effectué cette étude, il semble que les investisseurs particuliers sont plus sensibles que les autres investisseurs à différents facteurs. Une analyse de ces facteurs pouvant influencer la gestion des particuliers fera l'objet d'une dernière partie.

III. 2 g/ Prévisions à l'aide d'un modèle ARMA (p, q)

Cette partie représente le but final de la modélisation des séries de pourcentages de souscription/rachat par les processus autorégressifs. Nous en présentons ici comment obtenir ces prévisions à partir des modèles établis précédemment.

Etant donné une série stationnaire (X_t) observée entre 1 et T, on cherche à effectuer des prévisions à un horizon donné h , et donc à prévoir $X_{T+1}, \dots, X_{T+h}$.

L'équation : $\Phi(L)X_t = \Theta(L)\varepsilon_t$, peut se réécrire $\varepsilon_t = \Theta(L)^{-1}\Phi(L)X_t$. On a

donc $\varepsilon_t = \Psi(L)X_t$, soit $\varepsilon_t = \sum_{k=0}^{+\infty} \alpha_k X_{t-k} \Leftrightarrow X_t = \sum_{k=1}^{+\infty} \alpha_k X_{t-k} + \varepsilon_t$.

Aussi ${}_T \tilde{X}_{T+h}$ peut s'écrire de façon itérative : ${}_T \tilde{X}_{T+h} = \sum_{k=1}^{h-1} \alpha_k {}_T \tilde{X}_{T+h-k} + \sum_{k=h}^{+\infty} \alpha_k {}_T \tilde{X}_{T+h-k}$.

Toutefois, un des problèmes est que les (X_t) ne sont pas observés, en pratique, pour $t < 0$. On utilise alors l'écriture suivante :

$$X_{T+h} = \sum_{k=1}^{+\infty} \alpha_k X_{T+h-k} + \varepsilon_{T+h} = \sum_{k=1}^{h-1} \alpha_k X_{T+h-k} + \underbrace{\sum_{k=h}^{+\infty} \alpha_k X_{T+h-k}}_{\text{Reste d'une série ACV}} + \varepsilon_{T+h}$$

où le reste de la série absolument convergente tend (au sens de L^2) vers 0 quand T tend vers l'infini. On peut alors considérer, quand T est suffisamment grand que :

$${}_T \tilde{X}_{T+h} = \sum_{k=1}^{h-1} \alpha_k \cdot {}_T \tilde{X}_{T+h-k} + \sum_{k=h}^{T+h} \alpha_k \cdot X_{T+h-k} + \underbrace{\sum_{k=T+h+1}^{+\infty} \alpha_k \cdot X_{T+h-k}}_{\text{Négligeable (par hyp.)}}$$

Et on approxime ${}_T \tilde{X}_{T+h}$ par ${}_T \hat{\tilde{X}}_{T+h}$: ${}_T \hat{\tilde{X}}_{T+h} = \sum_{k=1}^{h-1} \alpha_k \cdot {}_T \hat{\tilde{X}}_{T+h-k} + \sum_{k=h}^{T+h} \alpha_k \cdot X_{T+h-k}$.

En pratique, grâce au logiciel R, on obtient les prédictions suivantes pour les différentes catégories de clients :

Catégorie de clients	Modèle	Prévision n+1	Prévision n+2
Particuliers	ARMA(1,2)	-0,01433	-0,01333
Particuliers	ARMA(3,2)	0,00572	-0,02289
Banques	ARMA(2,9)	0,06198	0,11701
Banque privée	ARIMA(1,1,1)	-0,01091	-0,01005
Investisseurs institutionnels	ARMA(5,2)	-0,00091	0,01125

On a effectué des prévisions à l'horizon n+2, car au delà, la précision se dégrade fortement. Cette dégradation de la précision des prévisions est causée en partie par la taille relativement faible de nos séries.

On remarque les différences significatives entre les prévisions des deux modèles utilisés pour les particuliers (modèle utilisant l'intégralité de l'historique : ARMA(1,2) et modèle utilisant uniquement les données depuis le début de la crise financière : ARMA(3,2)).

Parmi les cinq prévisions que nous avons effectuées, on remarque que l'on a trois rachats (signe -) et deux souscriptions (signe +). Ces souscriptions concernent les clients particuliers et les entreprises multinationales. Pour ces catégories on relativisera alors le niveau élevé des rachats proposé par les stress tests établis auparavant.

Intervalle de confiance de la prévision :

L'erreur de prévision est la différence entre la prévision de $X(t+h)$, calculée en t et la réalisation de X en $t+h$.

Il est possible de déterminer un intervalle de confiance de la prévision ${}_T \tilde{X}_{T+h}$. En effet, sous l'hypothèse de normalité des résidus, on peut montrer que :

$$\frac{X_{T+h} - {}_T \tilde{X}_{T+h}}{V(X_{T+h} - {}_T \tilde{X}_{T+h})^{1/2}} \xrightarrow[T \rightarrow \infty]{L} N(0,1)$$

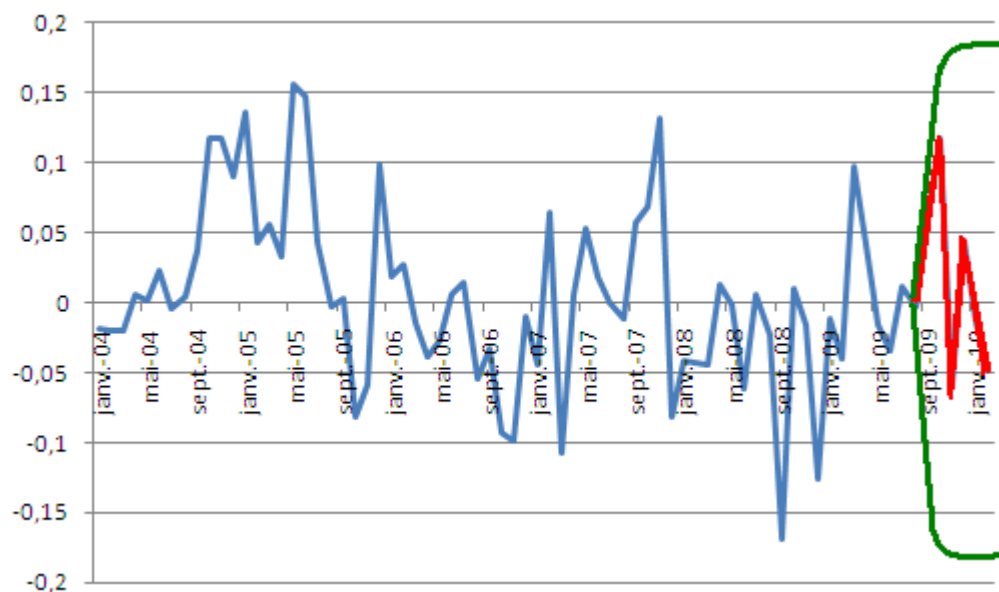
Or on sait que $E\left((X_{T+h} - \tilde{X}_{T+h})^2\right) = E\left[\left(\sum_{k=1}^{h-1} \alpha_k \varepsilon_{T+h-k}\right)^2\right] = \sum_{k=1}^{h-1} \alpha_k^2 \sigma_\varepsilon^2$.

D'où
$$\frac{X_{T+h} - \tilde{X}_{T+h}}{\sigma_\varepsilon^2 \left(\sum_{k=1}^{h-1} \alpha_k^2\right)^{1/2}} \xrightarrow[T \rightarrow \infty]{L} N(0,1)$$

On peut donc construire un intervalle de confiance sous la forme :

$$IC = \left[\tilde{X}_{T+h} \pm t^{\alpha/2} \left(\sum_{k=1}^{h-1} \alpha_k^2\right)^{1/2} \hat{\sigma}_\varepsilon \right]^*$$

Voici ce que l'on obtient pour les prévisions à l'horizon $h=4$ pour la série des pourcentages de souscription/rachat des banques :



Prévisions avec intervalle de confiance des pourcentages de souscription/rachat des banques

On remarque que l'intervalle de confiance (en vert sur le graphique) est relativement large. D'autre part, les prévisions semblent en adéquation avec l'allure générale du graphique. Ce qui confirme la validité du modèle statistique utilisé.

Au terme de cette étude, on obtient donc plusieurs scénarios de rachat pour les différentes catégories de clients. Nous allons désormais, dans une dernière partie, faire la synthèse de tous ces scénarios et confronter les stress tests aux scénarios de type prévisionnels.

III. 3 Synthèse

Voici un tableau qui résume les différents pourcentages de souscription/rachat obtenus par les différents scénarios développés lors de notre étude.

Customer Group	Scénario 1	Scénario 2	Prévisions
Banque privée	-16,50%	-17,00%	-2,88%
Investisseurs institutionnels	-29,90%	-16,50%	0,57%
Banques	-22,50%	-15,70%	6,20%
Particuliers	-	-12,40%	-1,09%

On remarque alors que les scénarios de type prévisionnels (basés sur les modèles statistiques) présentent des pourcentages de souscription/rachat relativement moins stressés que les autres scénarios (stress tests). En effet, les scénarios de type prévisionnels proposent également des souscriptions dans leurs prévisions.

Cependant, il faut rester méfiant à l'égard de certaines prévisions. En effet, on se souvient notamment, que le modèle retenu pour modéliser l'évolution des pourcentages de souscription/rachat de la clientèle bancaire s'adaptait mal aux données (ce qui a été confirmé par les tests sur les bruits blancs issus de ce modèle), il faut donc rester critique quant à la prévision d'une souscription de 6,20% pour le mois à venir.

En croisant ces scénarios avec d'autres scénarios déjà utilisés par le département de la gestion des risques, on peut mettre en lumière les fonds pour lesquels il existe un risque prononcé en terme de juste valeur du remboursement des encours des clients.

Le but final étant de mettre l'accent sur les fonds qui présentent réellement un enjeu quant à la gestion de leur risque de liquidité. Pour que la gestion du risque de rachat soit pertinente, il est important de ne pas mettre en lumière un nombre trop important de fonds, c'est dans cette optique que l'on s'emploie à croiser divers scénarios pour finalement se pencher sur les fonds pour lesquels il paraît exister un risque potentiel en terme de liquidité.

Quatrième partie

Etude de l'influence de la performance des fonds sur le comportement des investisseurs

IV. 1 Contexte

Toujours dans le but de déterminer des scénarios de rachat et ainsi participer à la gestion du risque de liquidité, nous avons souhaité, dans cette partie, étudier l'impact que pouvait avoir la performance (passée) des fonds sur les comportements des clients en matière de souscription et de rachat.

S'il est possible d'établir une relation temporelle entre la performance d'un fonds et les pourcentages de souscription/rachat des différentes catégories de clients de ce fonds, on pourra alors établir des prévisions de souscription ou de rachats à un horizon donné, en mettant en place des modèles statistiques s'appuyant sur l'historique des performances passées.

Dans cette partie, on fait donc l'hypothèse sous-jacente que les décisions d'investissements des clients (collectes et décollectes) sont prises consécutivement à l'analyse des performances des fonds.

Au delà de la performance absolue d'un fonds, il existe d'autres outils à la disposition de l'investisseur pour juger de la qualité d'un fonds, nous allons rappeler quels sont ces principaux instruments de mesure dans le paragraphe suivant.

IV. 2 Outils de mesure de performance d'un OPCVM

IV. 2 a/ La performance nette d'un OPCVM

Il faut différencier la performance réalisée par l'OPCVM depuis sa création et celle réalisée par l'investisseur. Ce dernier a la possibilité d'investir ou de se désinvestir d'un même OPCVM à plusieurs moments. Dans le cadre de l'épargne programmée par exemple, l'investissement se fait en plusieurs fois.

- Financièrement, la bonne façon de procéder pour apprécier la performance de son investissement est de calculer son TRI (Taux de Rendement Interne). Pratiquement, cette approche est difficile à mettre en œuvre. L'alternative passe par le calcul d'un prix de revient historique moyen. Cette méthode, certes plus approximative, est plus facile à réaliser puisqu'elle nécessite « seulement » de connaître les prix et les quantités d'achat des titres et d'en faire la moyenne.
- La performance de gestion d'un OPCVM est mesurée par l'évolution de la valeur liquidative entre deux dates.

L'évolution des performances d'un OPCVM doit être analysée dans le temps, et prendre en compte plusieurs phases de marché. Les gestions doivent s'apprécier, en fonction de leur

horizon d'investissement, et en prenant bien en compte l'alternance des phases de hausse et de baisse.

IV. 2 b/ La comparaison avec le benchmark

Un benchmark (ou indice de référence) est l'indice choisi par un gérant pour lui servir de référence dans la gestion de son fonds.

Il existe plusieurs milliers d'indices de marché à travers le monde.

Ces indices sont représentatifs :

- de l'ensemble du marché
- d'un secteur d'activité (énergie, matières premières, valeurs technologiques...),
- d'une population de valeurs spécifiques (univers des valeurs de moyennes capitalisations, des valeurs de petites capitalisations...).

Ils sont calculés par des bourses (Euronext, Liffe), par des journaux financiers (Financial Times, Dow Jones), par des banques (MSCI, Citigroup) ou par d'autres acteurs (Morningstar,...)

Le benchmark peut être considéré comme une sorte de tuteur pour la gestion de l'OPCVM. Il donne un cap et inscrit la gestion dans un univers d'investissement spécifique.

Ainsi, pour analyser objectivement la performance d'un OPCVM, un des bons points de comparaison est la performance du benchmark. Certains facteurs doivent toutefois être pris en compte dans la comparaison, comme le niveau de risque pris par le gérant par rapport au benchmark.

IV. 2 c/ La comparaison avec d'autres OPCVM

Des bases de données indépendantes permettent en France de comparer les performances des OPCVM entre eux (Europerformance, Lipper, APT, Morningstar...).

- **Les classements**

Chaque fonds est rattaché à une catégorie. Au sein de chaque catégorie, les fonds sont classés sur la base de leurs performances sur des périodes données : un mois, trois mois, un an, etc...

Les définitions de catégories peuvent varier d'une base de données à l'autre. Les classements peuvent faire apparaître des résultats différents. Ils doivent donc être utilisés à la lumière de cette précision.

- **Les notations**

Ces sociétés d'information financières ont mis en place des systèmes de notation qui intègrent de nombreux paramètres leur permettant de donner un avis sur la qualité d'un fonds et de sa gestion.

Dans un premier temps, nous allons nous intéresser à la performance absolue des fonds comme mesure de performance.

Nous nous intéresserons par la suite à l'évolution conjointe de la performance des fonds et de celle des benchmarks associés.

IV. 3 Présentations des données

IV.3 a/ Fonds retenus pour l'étude

Sur tous les fonds existants en septembre 2009, seul un tiers a été étudié dans cette partie. En effet, pour les autres fonds, l'historique disponible n'était pas suffisamment profond pour pouvoir étudier l'éventuelle influence de la performance sur les comportements des investisseurs de ces fonds. Certains fonds n'ont également pas été retenus à cause de leur faible nombre de clients.

On note que l'on dispose d'un historique remontant à début 2005 pour les performances, contrairement à celui des pourcentages de souscription/rachat qui remonte à début 2004, On étudiera donc des séries composées d'une cinquantaine de données (selon la date de création des fonds).

On ne dispose pas de l'historique des valeurs liquidatives pour tous les historiques des fonds, en revanche, on dispose de toutes les performances calculées à un mois. On pourra alors calculer différentes performances glissantes à partir de la performance un mois.

IV.3 b/ Calcul des performances glissantes

La performance absolue sur une période d'un titre financier ou d'un indice est calculée en pourcentage et correspond à la différence entre la valeur liquidative à la fin de la période considérée et la valeur à l'achat en début de période.

On parle de performance à propos d'un fonds ou d'un gestionnaire de fonds pour mesurer la pertinence de sa politique d'investissement.

Pour les fonds retenus, on dispose de la performance mensuelle sur un historique de 5 ans. Certaines valeurs liquidatives étant manquantes, nous utiliserons la performance mensuelle calculée à l'instant i , pour retrouver les performances glissantes sur différentes périodes h selon la formule :

$$Perf(h)_i = \left(1 + \frac{\prod_{k=1}^i (1 + Perf(1)_k) - \prod_{k=1}^{i-h} (1 + Perf(1)_k)}{\prod_{k=1}^{i-h} (1 + Perf(1)_k)} \right)^{1/h} - 1$$

ou encore, en simplifiant :

$$Perf(h)_i = \left(\prod_{k=i-h+1}^i (1 + Perf(1)_k) \right)^{1/h} - 1$$

Par exemple, on pourra calculer la performance glissante sur 3 mois du fonds suivant, pour le mois de mai :

Mois	Janvier	Février	Mars	Avril	Mai
Perf	1%	1,5%	2%	-1%	1%

On a donc le choix de calculer cette performance avec la première ou avec la deuxième formule.

Avec la première, on obtient :

$$\begin{aligned}
 Perf(3M) &= \left(1 + \frac{(1+1\%)*(1+1,5%)*(1+2%)*(1-1%)*(1+1\%)-(1+1\%)}{(1+1\%)} \right)^{1/3} - 1 \\
 &= \left(1 + \frac{1,0455-1,01}{1,01} \right)^{1/3} - 1 \\
 &= 1,16\%.
 \end{aligned}$$

Avec la deuxième formule, on obtient :

$$\begin{aligned}
 Perf(3M) &= ((1 + 1,5\%) * (1 + 2\%) * (1 - 1\%) * (1 + 1\%))^{1/3} - 1 \\
 &= 1,16\%.
 \end{aligned}$$

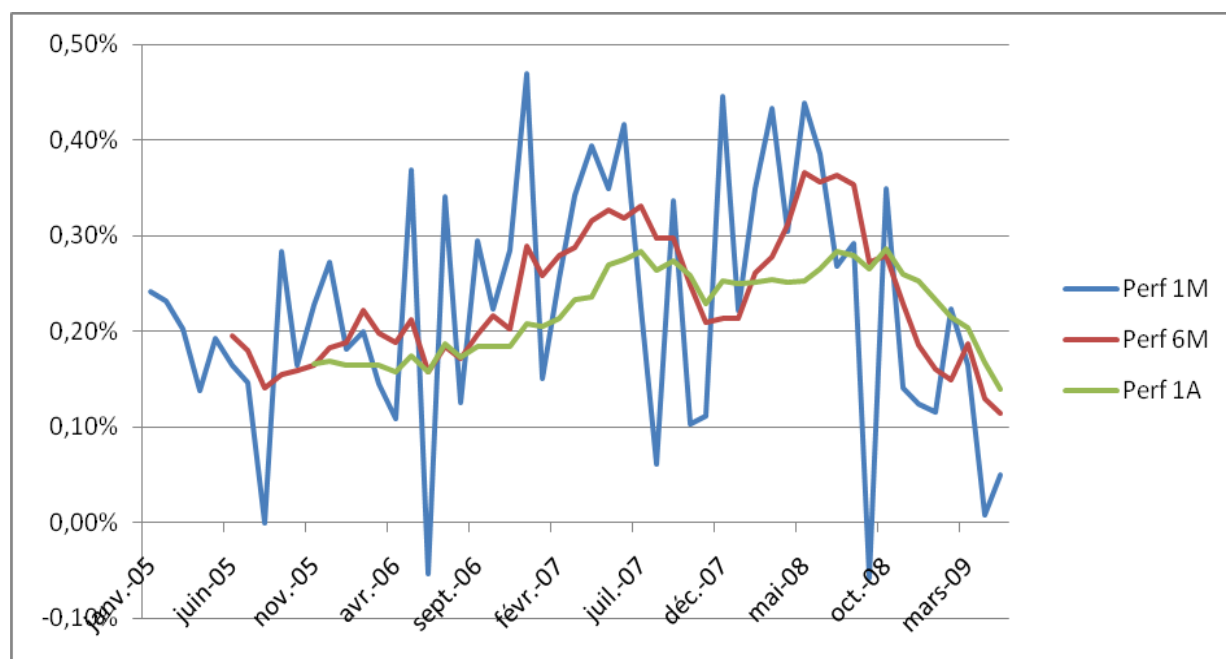
Nous avons alors calculé les performances glissantes sur des périodes de : 2 mois, 3 mois, 6 mois, 9 mois, 1 an, 18 mois et 2 ans.

A partir des performances glissantes, nous avons déterminé, pour chaque fonds et pour chaque catégorie de client, les coefficients de corrélation entre le pourcentage de souscription/rachat et les différentes performances glissantes.

Nous avons effectué les mêmes calculs pour les performances des Benchmarks associés aux fonds, on dispose alors de données homogènes.

On note que l'on ne dispose pas d'échantillons de même taille selon les performances considérées. En effet, pour la variable « performance 3 mois », on dispose de 3 données de moins que pour la variable « performance 1 mois » dues au décalage nécessaire au calcul.

Voici la représentation graphique des performances glissantes un mois, six mois et un an d'un fonds de type alternatif :



Historique des performances glissantes d'un fonds action

On observe le même phénomène d'atténuation de l'amplitude que sur toutes les moyennes mobiles, à savoir que plus la période de calcul n est grande, plus l'amplitude des variations de la série est faible.

Dans ce cas, un investisseur pour qui le coefficient de souscription/rachat sera significativement corrélé avec la performance un mois, sera un investisseur pour qui l'horizon de gestion sera relativement court, contrairement à un investisseur qui se basera sur des performances glissantes calculées sur des périodes plus longues pour prendre ses décisions d'investissement.

Malgré tout, on ne peut nier que le type de fonds (actions, obligations, garantis,...) joue un rôle prépondérant dans la corrélation entre ces deux variables.

IV.4 Etude de la corrélation entre les performances glissantes des fonds et le pourcentage de souscription/rachat

Nous avons calculé, pour chaque catégorie de client de chaque fonds, les coefficients de corrélation entre le pourcentage de souscription/rachat et les différentes performances glissantes du fonds en question.

Pour chaque Customer group de chaque fonds, nous avons produit trois résultats différents : avec des décalages entre les deux séries de un, deux et trois mois, pour tenir compte du temps nécessaire aux décisions des investisseurs, qui sont susceptibles de différer selon la catégorie de client.

Pour chaque catégorie de fonds, on calcule le coefficient de corrélation entre le pourcentage de souscription/rachat (mensuel) et les différentes performances glissantes du fonds. On retient in fine le maximum (en valeur absolue) de ce coefficient.

	Coef. Corrélation décalé d'un mois	Coef. Corrélation décalé de 2 mois	Coef. Corrélation décalé de 3 mois
Perf. 1M	9,81%	9,61%	4,88%
Perf. 2M	34,78%	31,56%	27,65%
Perf. 3M	28,19%	29,43%	18,17%
Perf. 6M	11,35%	10,87%	11,35%
Perf. 9M	2,80%	16,78%	7,42%
Perf. 1A	8,92%	14,42%	18,56%
Perf. 18M	9,01%	5,83%	8,41%
Perf. 2A	29,14%	15,55%	22,54%

Coefficients de corrélation entre le pourcentage de souscription/rachat des clients particuliers et différentes performances glissantes d'un fonds action.

Dans cet exemple, on retiendra que le pourcentage de souscription/rachat de ce fonds est le plus corrélé avec la performance glissante calculée sur 2 mois avec un décalage d'un mois entre les deux variables.

Cette corrélation positive avec un décalage d'un mois signifie que les investisseurs ont tendance à racheter leurs parts de ce fonds lorsqu'ils jugent que la performance du mois précédent est mauvaise et inversement, qu'ils ont tendance à souscrire lorsque celle-ci est positive. Il faut cependant considérer ces résultats avec précaution, car cette corrélation, bien que maximale reste relativement faible.

Comme cet exemple l'illustre, le coefficient de corrélation retenu, bien que maximal pour une performance glissante donnée peut ne pas être suffisamment élevé pour que l'on considère qu'il existe une corrélation entre cette performance et le pourcentage de souscription/rachat. Usuellement, on considère que deux variables sont corrélées lorsque leur coefficient de corrélation est supérieur à 70%.

Au final, seul 7% des fonds retenus présentent une corrélation significative entre la variable de performance et celle de pourcentage de souscription/rachat, ce qui est relativement faible pour pouvoir établir des statistiques fiables.

Nous avons alors décidé de présenter les résultats de tous les fonds et pas uniquement ceux pour lesquels le coefficient de corrélation est statistiquement significatif.

Pour la catégorie des particuliers, on a déterminé avec quelle performance glissante le pourcentage de souscription/rachat était le plus corrélé (tous décalages temporels confondus). On a effectué cette étude pour chaque fonds et on a considéré un regroupement par classe d'actif pour tenter de dégager des tendances.

Corrélation	Actions	Alternatifs	Garantis	Obligations	Total
Perf 1M	22%	6%	20%	22%	19%
Perf 2M	17%	6%	5%	0%	9%
Perf 3M	21%	6%	13%	6%	14%
Perf 6M	5%	17%	13%	6%	9%
Perf 9M	3%	22%	15%	11%	11%
Perf 1A	3%	11%	5%	17%	7%
Perf 18M	17%	11%	11%	17%	14%
Perf 2A	10%	22%	18%	22%	16%

Répartition des fonds en fonction de la corrélation entre le pourcentage de souscription/rachat et de la performance glissante.

On remarque alors que les pourcentages de souscription/rachat des fonds actions ont tendance à être plus corrélés avec les performances glissantes que l'on peut qualifier de court terme (1, 2 et 3 mois pour 60% des fonds actions) comparativement aux autres catégories de fonds.

On peut en déduire que l'horizon d'investissement des clients des fonds de type actions est relativement court, notamment par rapport aux fonds de type obligataire, pour lesquels le pourcentage de souscription/rachat est « corrélé » avec les performances glissantes à long terme (1 an, 18 mois et 2 ans pour 56% des fonds obligataires)

Cette étude est à considérer avec précaution au regard de la significativité relative des corrélations. Cependant, elle n'en est pas moins intéressante, dans la mesure où, en considérant tous les fonds possibles, il est possible de dégager certaines tendances, ce qui est difficile à réaliser en ne considérant que les fonds pour lesquels le coefficient de corrélation est statistiquement significatif.

Pour les autres catégories d'investisseurs, on remarque peu ou prou les mêmes tendances. Concernant les investisseurs institutionnels notamment, on note que les coefficients de corrélation entre les pourcentages de souscription/rachat et les performances glissantes sont maximaux pour les performances à long terme pour la très grande majorité des fonds.

Au terme de cette étude sur la corrélation entre la performance des fonds et le comportement des investisseurs en terme de souscription et de rachat, il apparaît que la performance ne semble pas être une variable pouvant à elle seule expliquer les investissements des clients. Cependant, nous allons tenter d'analyser un autre type de relation entre ces deux variables, à savoir la relation de causalité.

IV.5 Etude de la relation de causalité entre performance et souscription/rachat

Dans cette partie, on étudie les liens de causalité qui peuvent exister entre les variables de performances et la variable de pourcentage de souscription/rachat de chaque catégorie de client de chaque fonds. On prêter une attention particulière aux fonds pour lesquels on a pu établir l'existence d'une corrélation entre ces deux variables.

Si avec la notion de corrélation on peut déterminer s'il existe un lien entre deux variables, la causalité précise ce lien en examinant quelle variable influence l'autre.

Il est important de comprendre que ces deux concepts sont différents. Par exemple, La taille moyenne des japonais a augmenté de 15 cm depuis la fin de la 2ème guerre mondiale alors que la distance entre le Japon et les Etats-Unis augmente de 2 ou 3 cm par an à cause de la dérive des continents. Il y a corrélation, mais il n'y a bien évidemment pas causalité.

Dans le cadre de notre étude, nous allons nous intéresser à la relation de causalité entre la performance d'un fonds et les pourcentages de souscription/rachat des différentes catégories d'investisseurs de ce fonds.

Si on démontre que les performances « causent » le comportement des investisseurs (en matière de souscriptions et de rachats des parts qu'ils détiennent), alors il sera intéressant d'étudier en détail les performances des fonds pour être en mesure de proposer un modèle statistique pour modéliser le comportement des investisseurs en utilisant la performance comme variable explicative.

IV.5 a/ Définition de la causalité au sens de Granger

Avant toute chose, il est nécessaire de définir convenablement ce que l'on appelle causalité. Afin de fournir une définition, Granger (1969) a développé un nouveau concept de lien de causalité qui est désormais connu sous l'appellation « lien de causalité dans le sens de Granger ». Expliqué d'une manière synthétique, il s'agit simplement de déterminer si une variable x « cause selon Granger » une variable y en observant tout d'abord dans quelle mesure les valeurs passées de y arrivent à expliquer la valeur actuelle de y et de voir par la suite l'amélioration de l'estimation grâce à la prise en compte de valeurs retardées de la variable x . Y peut être considérée comme « causée selon Granger » si la variable x est déterminante dans l'estimation de y ou encore, d'une manière équivalente, si les coefficients des valeurs retardées de la variable x sont significativement différents de zéro.

Cette notion de causalité est basée sur deux principes fondamentaux :

- La cause précède l'effet (principe d'antériorité) ou lui est contemporaine
- La série causale contient de l'information sur l'effet qui n'est contenue dans aucune autre série

Il est important de signaler qu'une double causalité n'est jamais à exclure lors de ce test. Il faut tout de même préciser que la théorie de Granger a subi un certain nombre de critiques de la part du monde académique. Il ne serait, par exemple, pas possible de déterminer la direction du lien de causalité dégagé d'un test (Jacobs, Leamer et Ward, 1979). Ce problème serait principalement dû au fait que l'hypothèse nulle testée constitue une condition nécessaire mais non suffisante, et que ces tests sont extrêmement sensibles aux erreurs de spécification.

Cependant, dans notre étude, il nous paraît valable d'utiliser ces tests de causalité en respectant naturellement les limites du modèle lors de la phase d'interprétation des résultats.

On peut donc écrire la régression suivante représentant le point de départ du test :

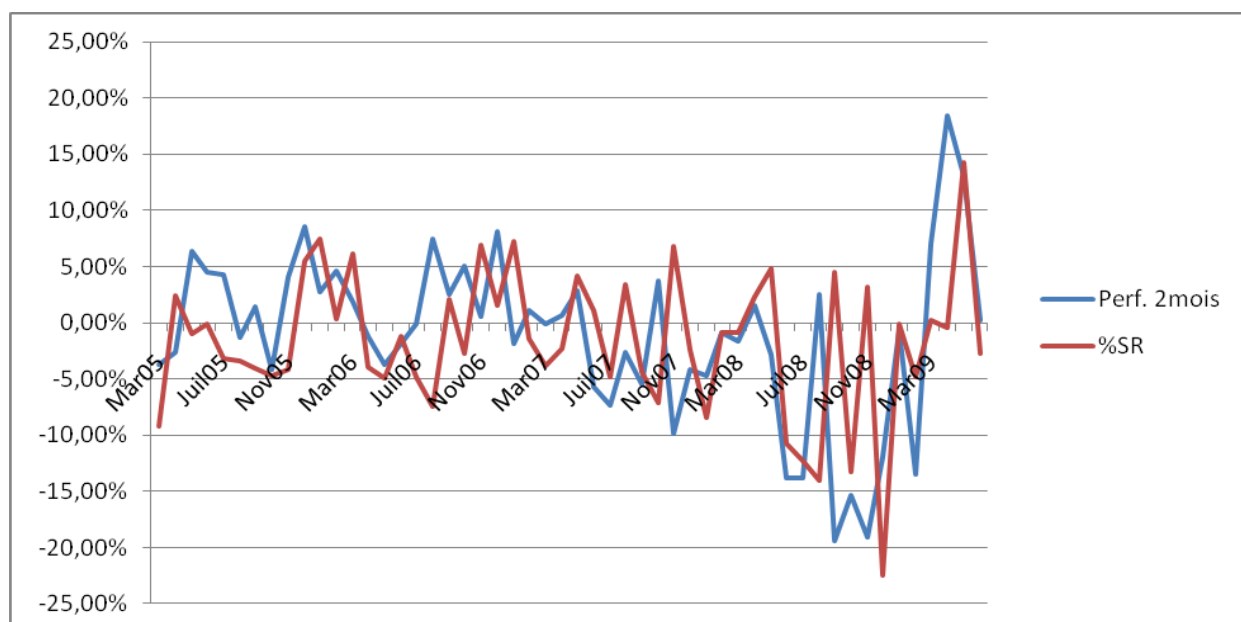
$$Y_t = \sum_{i=1}^n \mu_i Y_{t-i} + \sum_{j=1}^n \lambda_j X_{t-j} + \varepsilon_t \quad (1)$$

Si les coefficients λ_j sont significativement différents de zéro, alors on peut dire que la prise en compte des valeurs retardées de la variable X dans l'équation (1) donne une meilleure estimation des valeurs futures de Y. Comme il a été précisé plus haut, il va de soit que l'on peut refaire la même régression en inversant les variables, cependant dans notre étude, il semblerait étrange que les souscriptions ou les rachats des clients « causent » la performance d'un fonds. Nous nous limiterons donc à étudier la causalité de la performance sur les souscriptions ou les rachats (avec un, deux ou trois mois de décalage), bien qu'il soit intéressant d'analyser, dans un cadre théorique, dans quelle mesure un investisseur peut influencer la performance d'un actif (cette étude ferait l'objet d'un autre mémoire).

En testant la causalité de la performance sur les souscriptions et les rachats des investisseurs, on teste l'hypothèse sous jacente que les investisseurs se basent sur les performances passées d'un fonds pour souscrire ou racheter leurs parts. On suppose par exemple qu'un fonds affichant de mauvaises performances aura tendance à voir ses investisseurs racheter leurs parts et vice-versa. Cette tendance étant propre à chaque type de fonds (car à données égales, une performance pourra être considérée comme mauvaise pour un fonds et bonne pour un autre) et à chaque catégorie de client (car on a vu que les clients ont des comportements différents en terme d'investissement, ils sont donc plus ou moins exigeants quant à la performance).

Nous allons étudier l'exemple d'un fonds de type action pour lequel les investisseurs particuliers sont majoritaires. Nous avons donc utilisé le logiciel Eviews pour tester la causalité de la performance glissante 2 mois (coefficient de corrélation maximal de 65%) sur le pourcentage de souscription/rachat des investisseurs particuliers de ce fonds avec un décalage de deux mois entre ces deux variables.

Nous présentons le graphe sur lequel les deux variables sont représentées :



Graphique de l'évolution conjointe de la performance glissante 2 mois d'un fonds action et du pourcentage de souscription/rachat des investisseurs particuliers de ce fonds

Grâce à ce graphe, on remarque que l'évolution du pourcentage de souscription/rachat semble suivre celle de la performance glissante 2 mois. Cette tendance est plus marquée à certaines périodes de l'historique qu'à d'autres.

On remarque surtout, après comparaison avec d'autres fonds, que l'amplitude des deux variables est relativement proche. En effet, dans de nombreux cas, l'amplitude de la performance est faible comparée à celle du pourcentage de souscription/rachat. On remarque ce phénomène pour les fonds garantis ou obligataires pour lesquels la performance semble relativement constante comparativement au pourcentage de souscription/rachat.

Pour le fonds étudié plus haut, on obtient les résultats de causalité suivants :

Pairwise Granger Causality Tests
Date: 01/03/10 Time: 18:51
Sample: 2005M02 2009M06
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Probability
PERF does not Granger Cause SR	50	0.33069	0.30317
SR does not Granger Cause PERF		2.37160	0.00360

Test de causalité au sens de Granger entre les variables performance glissante 2 mois et la variable de souscription/rachat des investisseurs particuliers

On peut conclure, en observant cette sortie issue du logiciel, qu'on ne peut pas rejeter l'hypothèse selon laquelle la performance de ce fonds ne cause pas (au sens de Granger) le pourcentage de souscription/rachat des investisseurs particuliers de ce fonds. En revanche, on

rejette l'hypothèse selon laquelle le pourcentage de souscription/rachat cause la performance du fonds, ce qui semble logique.

IV.5 b/ Conclusion sur la causalité

Que peut-on alors dire pratiquement sur ces liens de causalité entre nos variables ? Il est vrai qu'ils existent et que de ce point de vue les tests sont concluants. Ces tests ne sont cependant pas assez affinés pour en tirer des conclusions lourdes de conséquences. Il est donc tout à fait judicieux de les réexaminer pour les confirmer ou les affirmer. Ceci étant fait, les liens valables pour les fonds étudiés peuvent être considérés plus sérieusement en tant qu'outils de prévisions mais, à notre avis il serait faux de prendre des décisions uniquement sur la base de ces outils. Ces liens peuvent donc constituer une aide à la décision mais, ils ne devraient en aucun cas représenter le facteur déterminant de la prise de décision.

De plus, le fait que les liens ne soient pas immuables dans le temps corrobore exactement la prudence qui doit être de mise. En effet, les liens peuvent évoluer et il serait donc risqué de prétendre avoir trouvé l'outil de prévision général et applicable à toutes les situations. Il faut replacer tous les liens de causalité dans leur contexte et être bien conscient que les causalités en vigueur à un moment donné peuvent rapidement devenir obsolètes. Ces liens sont la proie des tendances qui frappent les marchés financiers, ce qui rend l'application directe de ces derniers en tant qu'outils de prévision encore plus difficile.

En ce sens, il ne paraît pas pertinent de proposer une modélisation statistique du pourcentage de souscription/rachat basée sur la performance.

IV.5 c/ Etude de la relation entre les comportements d'investissement et la différence de performance du fonds et du benchmark

Au terme de notre étude, la relation de causalité entre la performance d'un fonds et le pourcentage de souscription/rachat des différentes catégories d'investisseurs. Les quelques fonds pour lesquels cette relation de causalité est validée par le test de Granger sont en majorité des fonds de type action.

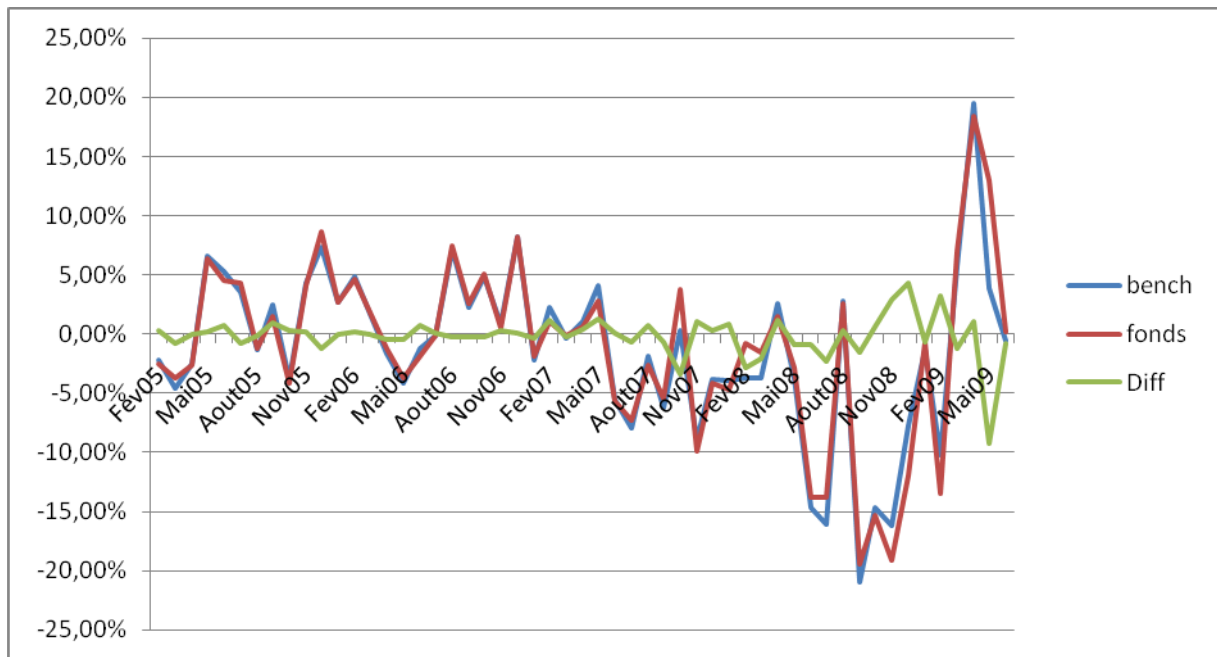
Dans une seconde phase d'analyse, il semblait intéressant de mesurer la performance d'un fonds non pas en utilisant le seul indice de performance de ce dernier, mais en analysant conjointement la performance du fonds et celle du benchmark associé à celui-ci.

Nous avons alors simplement défini un indice basé la différence entre ces deux performances. Cet indice étant positif lorsque le fonds surperforme le marché et négative lorsque le fonds sousperforme le marché.

En définissant cet indice, on fait l'hypothèse sous jacente que l'investisseur compare la performance du fonds avec celle de son benchmark avant de souscrire ou de racheter les parts qu'il possède dans ce fonds.

Cependant, il est rapidement apparu que cet indice est quasiment constant tout au long de l'historique pour la très grande majorité des fonds.

Voici la représentation graphique de l'historique de la performance d'un fonds action, celle du Benchmark associé à ce fonds et de la différence :



Graphique des performances d'un fonds action, son Benchmark et de la différence.

On remarque que la performance de ce fonds est toujours très proche de celle du Benchmark. Le gérant a donc réellement utilisé le Benchmark comme outils de guidage de sa gestion. En effet, le coefficient de corrélation des deux variables est de 97% ce qui est très élevé.

L'indice calculé (en vert sur le graphique) apparaît comme relativement constant alors que l'on a choisi un fonds dont l'amplitude de la performance est assez élevée comparativement à celle des fonds obligataires ou garantis.

Il en résulte que cet indice ne peut être utilisé dans le cadre de la modélisation des comportements des investisseurs après l'analyse de la performance.

Conclusion générale

La modélisation du comportement des investisseurs est une problématique encore peu étudiée mais dont le développement est nécessaire en vue d'une amélioration du suivi du risque de liquidité et de la mise en place des nouvelles normes comptables et prudentielles.

Au terme de cette étude, nous avons été à même de proposer des scénarios de rachat impactant de façon plus ou moins importante la liquidité des différents fonds. Les scénarios de type prévisionnels basés sur des modélisations statistiques par processus autorégressifs permettent de tempérer les résultats souvent très stressés obtenus par les scénarios catastrophes. Ces scénarios catastrophes étant intéressants dans le cadre du suivi du risque de liquidité dans la mesure où ils permettent d'étudier l'impact d'événements plus ou moins susceptible de se produire sur les comportements des investisseurs et en particuliers sur leurs comportement en terme de souscription, mais surtout (ce qui nous préoccupe le plus), de rachat.

En étudiant le comportement des différentes catégories d'investisseurs (particuliers, institutionnels, banques,...) nous avons mis en avant l'importance de la diversification de la clientèle d'un fonds. On opère, en procédant ainsi, à une mutualisation des risques et on atténue en particulier le risque de rachat.

Les scénarios développés, basés sur les modèles statistiques, montrent néanmoins certaines limites, notamment dues à la faible profondeur de l'historique des données disponible. Il aurait été possible de palier ce défaut si nous avions disposé de données établies sur une base journalière, la précision des modèles aurait ainsi pu être améliorée.

C'est en partie pour ces raisons que l'on ne peut obtenir des prévisions statistiquement viables avec ces modélisations, au delà d'un horizon de deux à trois mois.

L'étude sur le lien entre les performances des fonds et le comportement des investisseurs en terme de souscription/rachat, si elle s'était avérée pertinente au départ, elle a rapidement montré ses limites. En effet, au terme de cette étude, on peut conclure que les investisseurs ne se basent pas uniquement sur la performance des fonds pour prendre leurs décisions en matière d'investissement, même si on peut penser que cette variable joue un rôle très important.

Cependant, nous avons vu que les rachats et les souscriptions de catégories de clients de certains fonds sont statistiquement « causés » par la performance de ces fonds. Même s'il paraît illusoire de fonder un modèle de prédiction du comportement des investisseurs sur la seule variable de performance des fonds, il conviendra d'analyser le lien existant entre la performance du fonds et la variable de souscription/rachat pour les clients des fonds jugés exposés au risque de liquidité

Au terme de ce mémoire, on remarque qu'il aurait été particulièrement intéressant d'étudier le comportement des investisseurs de manière plus précise. Une étude qualitative préalable, menée auprès des commerciaux en charge de la gestion des différents clients ou directement auprès d'un échantillon représentatif de clients aurait permis de cerner certaines de leurs habitudes en terme d'investissement et aurait ainsi permis de guider notre élaboration de scénarios.

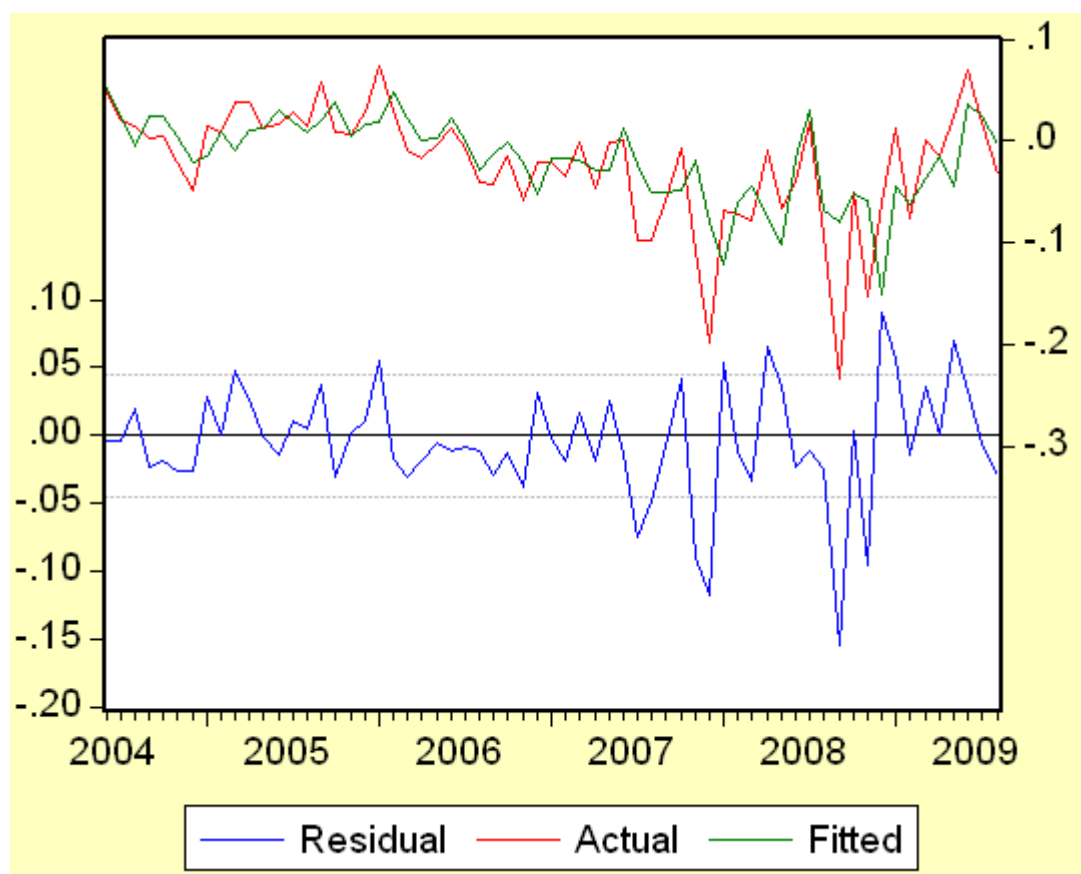
Il aurait été également pertinent de se pencher sur les transferts de liquidité d'une catégorie de fonds à une autre, pour chacune des classes de clients. Cette étude aurait été d'autant plus intéressante en période de crise financière, notamment pour anticiper les attentes des investisseurs.

Annexes

Annexe N°1

On présente dans cette annexe les résultats des différentes modélisations effectuées pour les séries de pourcentages de souscription/rachat qui n'ont pas été insérées dans la partie principale de ce mémoire par soucis de clarté.

Voici le graphique de la modélisation du pourcentage de souscription/rachat de la clientèle institutionnelle par un ARMA(5,2) :

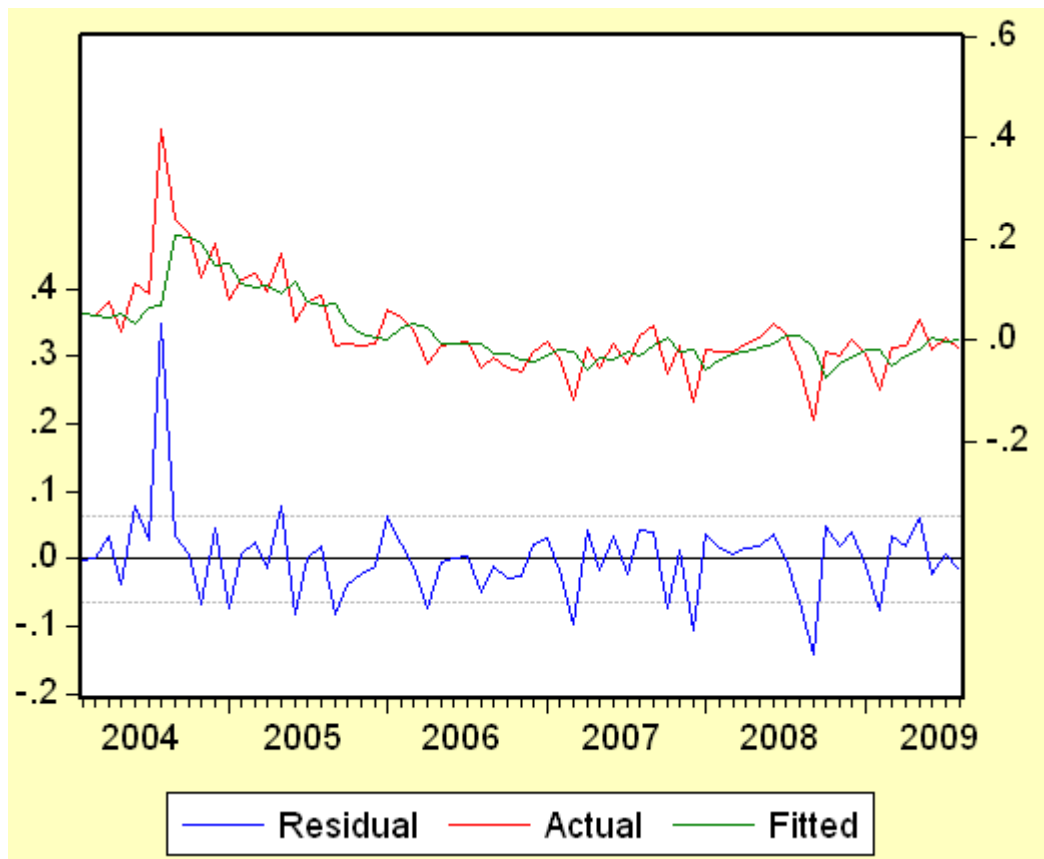


On remarque que le modèle « fit » relativement bien aux données sur l'ensemble de l'historique. Cependant on remarque une légère désynchronisation avec les données réelles entre fin 2007 et mi-2008 (période où on pouvait constater les premières tensions liées à la crise financière). Le rachat de 21% de novembre 2007, surement trop prononcé, n'a pas été suffisamment pris en compte par le modèle et a ainsi entraîné une erreur de près de 12%. En revanche, l'adéquation du modèle aux données redevient intéressante dès le début de l'année 2009.

Même si les différents tests de Chow appliqués à la modélisation rejette l'hypothèse d'une (ou plusieurs) rupture de tendance, on note une fois de plus les effets de la crise financière sur les

pourcentages de souscription/rachat et la difficulté d'obtenir des modélisations qui prennent en compte avec précision ces périodes d'augmentation de la volatilité.

La modélisation du pourcentage de souscription/rachat des clients de la banque privée par un ARIMA(1,1,1) présente la forme ci-dessous :



On remarque que la modélisation en tendance semble correcte, mais que les « sauts » (qui sont assez faibles) ne sont que rarement pris en compte par le modèle. L'exemple le plus flagrant étant le premier pic du mois d'août 2004 (+41,7%) pour lequel le modèle semble mal adapté puisqu'il présente un bruit blanc de près de 30% à cette même période.

Bien que ce modèle ne soit pas entièrement satisfaisant (il ne valide pas certains critères d'adéquation, comme le test de normalité des résidus de Shapiro-Wilk), nous l'avons tout de même utilisé pour établir des prévisions à court terme.

Annexe N° 2 : tables statistiques utilisées

Table de Shapiro-Wilk

NB: Si W est plus grand ou égal au $W_{1-\alpha,n}$ de la table alors on considère la distribution comme normale

	Risque 5%	Risque 1%
n	$W_{0,95}$	$W_{0,99}$
5	0,7 62	0,686
6	0,988	0,713
7	0,803	0,730
8	0,818	0,749
9	0,829	0,764
10	0,842	0,781
11	0,850	0,792
12	0,859	0,805
13	0,866	0,814
14	0,874	0,825
15	0,881	0,835
16	0,887	0,844
17	0,892	0,851
18	0,897	0,858
19	0,901	0,863
20	0,905	0,868
21	0,908	0,873
22	0,911	0,878
23	0,914	0,881
24	0,916	0,884
25	0,918	0,888
26	0,920	0,891
27	0,923	0,894
28	0,924	0,896
29	0,926	0,898
30	0,927	0,900
31	0,929	0,902
32	0,930	0,904
33	0,931	0,906
34	0,933	0,908
35	0,934	0,910
36	0,935	0,912
37	0,936	0,914
38	0,938	0,916
39	0,939	0,917
40	0,940	0,919

41	0,941	0,920
42	0,942	0,922
43	0,943	0,923
44	0,944	0,924
45	0,945	0,926
46	0,945	0,927
47	0,946	0,928
48	0,947	0,929
49	0,947	0,929
50	0,947	0,930

Table de Fisher

On utilise cette table dans le test de Jarque-Bera.

$\alpha = 0,95$

		v ₁																		
		1	2	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	•	
v ₂	1	161	200	216	225	230	234	237	239	241	242	246	248	250	252	253	254	254	254	
	2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5	
	3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70	8,66	8,62	8,58	8,55	8,54	8,53	8,53	
	4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86	5,80	5,75	5,70	5,66	5,65	5,64	5,63	
	5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62	4,56	4,50	4,44	4,41	4,39	4,37	4,37	
	6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94	3,87	3,81	3,75	3,71	3,69	3,68	3,67	
	7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,51	3,44	3,38	3,32	3,27	3,25	3,24	3,23	
	8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22	3,15	3,08	3,02	2,97	2,95	2,94	2,93	
	9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01	2,94	2,86	2,80	2,76	2,73	2,72	2,71	
	10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85	2,77	2,70	2,64	2,59	2,56	2,55	2,54	
	11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,72	2,65	2,57	2,51	2,46	2,43	2,42	2,40	
	12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,62	2,54	2,47	2,40	2,35	2,32	2,31	2,30	
	13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,53	2,46	2,38	2,31	2,26	2,23	2,22	2,21	
	14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,46	2,39	2,31	2,24	2,19	2,16	2,14	2,13	
	15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40	2,33	2,25	2,18	2,12	2,10	2,08	2,07	
	16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,35	2,28	2,19	2,12	2,07	2,04	2,02	2,01	
	17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,31	2,23	2,15	2,08	2,02	1,99	1,97	1,96	
	18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,27	2,19	2,11	2,04	1,98	1,95	1,93	1,92	
	19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,23	2,16	2,07	2,00	1,94	1,91	1,89	1,88	
	20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,20	2,12	2,04	1,97	1,91	1,88	1,86	1,84	
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,15	2,07	1,98	1,91	1,85	1,82	1,80	1,78		
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,11	2,03	1,94	1,86	1,80	1,77	1,75	1,73		
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,07	1,99	1,90	1,82	1,76	1,73	1,71	1,69		
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,04	1,96	1,87	1,79	1,73	1,69	1,67	1,65		
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,01	1,93	1,84	1,76	1,70	1,66	1,64	1,62		
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	1,92	1,84	1,74	1,66	1,59	1,55	1,53	1,51		
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,87	1,78	1,69	1,60	1,52	1,48	1,46	1,44		
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,84	1,75	1,65	1,56	1,48	1,44	1,41	1,39		
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	1,79	1,70	1,60	1,51	1,43	1,38	1,35	1,32		
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,77	1,68	1,57	1,48	1,39	1,34	1,31	1,28		
200	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93	1,88	1,72	1,62	1,52	1,41	1,32	1,26	1,22	1,19		
500	3,86	3,01	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,69	1,59	1,48	1,38	1,28	1,21	1,16	1,11		
•	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,67	1,57	1,46	1,35	1,24	1,17	1,11	1,00		

Annexe N°3

Dans cette annexe, nous présentons le code utilisé pour déterminer les ordres p et q des modèles ARMA qui ont été utilisés dans le but d'établir des prévisions des pourcentages de souscription/rachat des différentes catégories de clients.

Ce code a été utilisé sous le logiciel statistique R.

```
donnees=read.table(file="Data.txt")
Sino=as.matrix(donnees)
CG=Sino[,1] #On modifie ici le nombre de colonne pour chaque Customer Group

plot(CG,type="l") #On commence par afficher le graphique des données (évolution des
pourcentages de soucription/rachat au fil des mois.

p <- 10; q <-10;d<-1; # p=10, q=10 et d=1 #sont les ordres maximums que l'on testera
mat_coef <- (matrix(nrow = p * q, ncol = 7, dimnames = list(c(), c("logLik", "AIC",
"HQ ", "p", "q", "d", "nb_iter"))))
serie <- CG; t = 1;
system.time(
  for(i in 1:p) for(j in 1:q) for(k in 0:d){
    res <- arima(x = serie, order = c(i, k, j))
    L<-logLik(res)
    HQ<-log(L)+2*(i+j)*log(log(n))/n
    mat_coef[t,1] <- res$loglik
    mat_coef[t,2] <- res$aic
    mat_coef[t,3] <- HQ
    mat_coef[t,4] <- i
    mat_coef[t,5] <- j
    mat_coef[t,6] <- k
    mat_coef[t,7] <- t
    t = t + 1 } )
#Dans cette matrice, on calcule pour chaque ordre p, d et q la log-vraisemblance, le critère
AIC et le critère HQ

N<- na.omit(mat_coef[,2])
M<-matrix(0,nrow=length(N),ncol=1)
i=1
while(i<length(N)+1){
  M[i]<-ifelse(N[i]==min(N),i,0)
  i=i+1}
S<-sum(M)
```

#Cette boucle a pour but de récupérer le modèle qui maximise le critère AIC ou le critère HQ (au choix) à partir de la matrice calculée précédemment.

na.omit(mat_coef[S,3]); na.omit(mat_coef[S,4]);na.omit(mat_coef[S,5]); #On donne ici les coefficients du modèle.

Box.test(arima(CG,c(na.omit(mat_coef[S,3]),na.omit(mat_coef[S,5]),na.omit(mat_coef[S,4]))\$residual)

shapiro.test(arima(CG,c(na.omit(mat_coef[S,3]),na.omit(mat_coef[S,5]),na.omit(mat_coef[S,4]))\$residual)

#Ici on effectue des tests d'adéquation du modèle

predict(arima(CG,c(na.omit(mat_coef[S,3]),na.omit(mat_coef[S,5]),na.omit(mat_coef[S,4]))),2) #On obtient finalement les prévisions relatives au modèle trouvé, le chiffre 2 correspondant à l'horizon de prévision

Annexe N°4

On présente dans cette annexe les recommandations de l'AMF relatives à la gestion du risque de liquidité mises en ligne le 9 octobre 2009 sur site de l'AMF (<http://www.amf-france.org>)

Le risque de liquidité correspond au risque qu'un émetteur ne soit pas en mesure de faire face à ses besoins monétaires grâce à ses ressources financières. Les ressources financières comprennent les ressources générées par les activités et celles mobilisables auprès de tiers.

Le risque de liquidité peut être évalué au moyen des deux éléments suivants :

- un échéancier des paiements contractuels au titre des passifs financiers présents au bilan à la date de clôture (IFRS 7 § 39) ;
- les décaissements auxquels l'émetteur devra faire face dans le cadre de son activité.

L'AMF recommande que les émetteurs expliquent en détails la manière dont sont gérés le risque de liquidité et l'exposition à ce risque au sein de leur groupe (IFRS 7 § 33 et 34). Une présentation possible peut être de préciser par exemple :

- la politique de gestion de la liquidité (centralisation, etc.) ;
- la diversification des sources de financement.

Les émetteurs fournissent une information détaillée sur l'existence et la nature des clauses de défaut et d'exigibilité anticipée, quand la probabilité d'occurrence des faits générateurs et leurs impacts possibles sur la situation financière du groupe, notamment en termes de trésorerie, sont significatifs. Les paragraphes 18 et 19 de la norme IFRS 7 demandent une information en cas de défaut des covenants.

Si les passifs financiers ne sont assortis d'aucune clause particulière (dont notamment des covenants) susceptible d'en modifier significativement les termes, l'émetteur le spécifie explicitement.

Si les émetteurs font l'objet de notations financières par les agences de notation, les notes à court et long terme sont précisées par agence de notation et l'évolution de la notation sur l'exercice écoulé est commentée.

Bibliographie

Articles

- [1] Etude de séries chronologiques linéaires à temps discret, F. Borgard & D. Guégan. *Revue de statistique appliquée*, tome 44, N°4, 1996.
http://archive.numdam.org/ARCHIVE/RSA/RSA_1996__44_4/RSA_1996__44_4_59_0/RSA_1996__44_4_59_0.pdf
- [2] L'indice CAC 40 emporté par une spirale baissière, Laure Burrus. *JDF Hebdo*, 25/10/2008.
<http://www.jdf.com/tendance-valeurs/2008/10/25/04010-20081025ARTHBD00084-l-indice-cac-emporte-par-une-spirale-baissiere.php>
- [3] La descente aux enfers continue pour les établissements bancaires, Annelot Huijgen. *JDF Hebdo*, 24/01/2009.
<http://www.jdf.com/enquete/2009/01/24/04004-20090124ARTHBD00078-la-descente-aux-enfers-continue-pour-les-etablissements-bancaires.php>

Mémoires

- [4] Stéphanie Fauvel et Maryse Le Pévédic, Analyse des rachats d'un portefeuille vie individuelle : Approche théorique et Application pratique. *Mémoire Ensaë*, 2007.
- [5] Floriane Vicaine, Actif Passif d'un fonds en euros. Modélisation stochastique et étude d'indicateur de risque. *Mémoire Isup*, 2003.

Ouvrages

- [6] Séries Temporelles Et Modèles Dynamiques, Christian Gouriéroux & Alain Monfort. *Economica*, 1995.
- [7] Les tests statistiques à l'aide de R, Gaël Millot. *De Boeck*, 2009.

- [8] La gestion des risques financiers, Thierry Roncalli. *Economica*, 2004.
- [9] Analyse des séries temporelles en économie, Bourbonnais Régis, Terraza Michel. *Presses Universitaires de France*, 1998.
- [10] Time series analysis: Forecasting and control, George Box et Gwilym Jenkins. *San Francisco: Holden-Day*, 1970.
- [11] Theory of Games and Economic Behavior, John von Neumann et Oskar Morgenstern. *Princeton University Press*, 1947.
- [12] G. Demange et J.C. Rochet, Méthodes mathématiques pour la finance. *Economica*, 1992.
- [13] S. Lardic et V. Mignon, Econométrie des séries temporelles macroéconomiques et financières. *Economica*, 2002.

Supports de cours

- [14] Cours de séries temporelles : Théorie et applications, Arthur Charpentier. *Polycopié de cours ENSAE*, 2008.
- [15] Cours de séries temporelles : Théorie et applications, Jean Pierre Indjehagopian. *Polycopié de cours ISUP*, 2009.
- [16] Tests de normalité : Techniques empiriques et tests statistiques, Ricco Rakotomalala. Université Lyon Lumière II, 2008.
http://eric.univ-lyon2.fr/~ricco/cours/cours/Test_Normalite.pdf
- [17] Cours de Modèles Macroéconomiques, Jean-Christian Lambelet. Université de Lausanne, 2001.

Guide d'utilisation

- [18] Eviews 5 User's guide, Quantitative Micro Software, LLC, 2004.

Sites Internet

www.bloomberg.com : Site d'informations financières

www.boursorma.fr : Site d'informations financières

www.insee.fr : Institut National des Statistiques et des Etudes Economiques