



NOVEMBRE 2017

**RISQUE DE LIQUIDITÉ ET  
COMPORTEMENT DES  
INVESTISSEURS : ENJEUX,  
DONNÉES ET MODÉLISATION**

Revue du  
Conseil Scientifique

*Cette publication a été coordonnée par la Division Études, Stratégie et Risques de l'Autorité des marchés financiers, et rédigée par les auteurs qui sont cités. Elle repose sur l'utilisation de sources considérées comme fiables mais dont l'exhaustivité et l'exactitude ne peuvent être garanties. Sauf mention contraire, les données chiffrées ou autres informations ne sont valables qu'à la date de publication de chacune des études et sont sujettes à modification dans le temps. Toute copie, diffusion et reproduction de ces études, en totalité ou en partie, sont soumises à l'accord exprès, préalable et écrit de l'AMF.*

*Avertissement — La « Revue du Conseil scientifique » est une publication qui présente des travaux d'étude et de recherche menés par les membres du Conseil scientifique ou des chercheurs issus du monde académique ou professionnel. Chaque étude reflète les vues personnelles de son ou ses auteurs et n'exprime pas nécessairement la position de l'Autorité des marchés financiers ou de l'institution dans laquelle le ou les auteurs travaillent.*

Serge Darolles  
Université Paris-Dauphine

**Résumé**

La liquidité d'un fonds d'investissement est souvent assimilée à la liquidité des instruments financiers qu'il détient. Or la gestion d'un fonds s'apparente par certains côtés à une activité de transformation de liquidité. Une modification de la structure de la clientèle du fonds a des conséquences sur le décalage potentiel entre la liquidité de l'actif et du passif. Une approche actif/passif de la liquidité est donc essentielle et nécessite une modélisation du comportement des clients. Cet article aborde les différents challenges pour les sociétés de gestion - système d'information, modèles de comportement, modèles de risque et réglementation. Il présente pour un nombre significatif de fonds les résultats d'une étude statistique du comportement des clients et de son impact sur l'exposition d'un fonds au risque de liquidité.

## 1. LES ENJEUX

Le risque de liquidité d'un fonds d'investissement est souvent appréhendé uniquement en analysant les instruments financiers détenus par le fonds. La structure de l'actif du fonds expose en effet directement la performance du fonds à une dégradation de la liquidité des instruments financiers. Cette approche suppose implicitement que la taille du fonds reste constante, et que la dégradation de la liquidité n'engendre aucun mouvement de sortie des clients. Or, en pratique, la taille du fonds change en fonction de la réaction des investisseurs, selon les éventuelles restrictions de liquidité qui leur sont imposées. Ces éléments, liés à la composition du passif du fonds, ont donc une conséquence directe sur le niveau du risque de liquidité, notamment quand le fonds investit dans des actifs illiquides et propose simultanément une liquidité avantageuse à ses clients. La gestion d'un fonds s'apparente alors à une activité de transformation de liquidité, et toute dégradation de la liquidité des titres du portefeuille peut accroître le décalage de liquidité entre l'actif et le passif du fonds. Le risque de liquidité est une conséquence directe de ce décalage, et le quantifier passe par une analyse jointe de l'actif et du passif du fonds. Cela suggère donc une gestion actif/passif de la liquidité qui n'est envisageable que si le gérant intègre dans son analyse des informations sur le comportement de ses clients.

Dans ce contexte, et compte tenu des pratiques courantes de leur département des risques, les sociétés de gestion ont différents challenges à relever. Quatre enjeux principaux nous semblent importants. Le premier concerne leur système d'information. Les données historiques relatives aux instruments financiers détenus par le fonds sont stockées dans ce système. Elles permettent de calculer l'exposition du fonds aux différents types de risques – marché, crédit, liquidité, ..., via des calculs statistiques basés sur les séries historiques de prix de transaction, de fourchettes, de volumes ou de notations. Par contre, les informations relatives aux clients du fonds ne sont généralement pas aussi facilement accessibles, notamment si l'on désire travailler au niveau de granularité le plus fin, i.e. le client. Dans de nombreux cas, seul l'actif sous gestion global est stocké dans le système d'information car il intervient directement dans le calcul des contraintes de levier ou de niveau de diversification. Les équipes des risques n'ont donc pas les outils pour suivre de manière systématique l'évolution de la structure du passif. Or cette évolution peut avoir différentes causes et différentes implications en termes de liquidité. Prenons l'exemple d'un fonds dont la taille globale est stable entre deux dates. Cette stabilité pourrait laisser croire à une stabilité de l'exposition au risque de liquidité du fonds. Or elle peut également être la conséquence d'un rachat initié par des investissements long terme, et d'une souscription par des clients investissant sur de très courtes périodes. On voit bien que, dans ce cas, la qualité du passif se dégrade et l'exposition du fonds au risque de liquidité augmente. La seule manière d'observer cela est de suivre individuellement le comportement des investisseurs, et d'en déduire les évolutions de la structure du passif du fonds. L'enjeu pour les sociétés de gestion est donc d'enrichir leur système d'information en y incluant le suivi de leur passif. Cela passe nécessairement par une meilleure connaissance de leurs clients. Les systèmes de marquage des ordres permettent par exemple d'identifier le client qui a envoyé un ordre d'achat ou de vente donné, puis de stocker les historiques des souscriptions/rachats par client/profil homogène de clients. L'investissement pour la société de gestion est conséquent et les difficultés techniques nombreuses. Mais ces informations permettent également d'alléger le traitement des distributeurs et de leur rémunération, et d'augmenter la valorisation du portefeuille des clients existants. Cela peut être une seconde justification de l'investissement nécessaire.

Le second enjeu concerne la compréhension du comportement des clients. La création d'une base de données des souscriptions/rachats donne une vision uniquement historique du comportement des clients. Une bonne maîtrise du risque de liquidité à une date donnée demande une vision prospective permettant d'anticiper les souscriptions/rachats futurs. En effet, sans idée très précise des scénarios de rachats probables, le gérant doit investir une partie de son fonds dans des instruments liquides immédiatement disponibles, cela pour couvrir une partie des éventuels rachats sans avoir à vendre à perte des actifs illiquides<sup>1</sup>. Une meilleure connaissance du comportement des clients permet de mieux gérer l'exposition du fonds au risque de liquidité et d'investir un pourcentage plus important du portefeuille en instruments financiers illiquides. Anticiper le comportement des clients passe par le développement de modèles statistiques plus ou moins sophistiqués. La première approche,

---

<sup>1</sup> cf. Darolles, Roussellet (2017) pour une analyse de la gestion du cash dans un fonds d'investissement.

basique, est de supposer un comportement des clients stable dans le temps. Les données historiques permettent d'estimer la distribution marginale des souscriptions/rachats, puis de calculer la probabilité d'événements extrêmes, par exemple une diminution de 50% des actifs sous gestion. À l'image de ce qui est fait pour la volatilité, il est également possible d'inclure une composante temporelle dans ces modèles simples. L'exploitation des données historiques permet alors de mesurer si, par exemple, d'importantes sorties à une date sont en moyenne suivies par d'importantes sorties aux dates suivantes. Les modèles de séries temporelles permettent de lier statistiquement les observations faites à différentes dates et de capter ces effets temporels. Enfin, une dernière piste consiste à identifier les facteurs économiques liés aux souscriptions/rachats, puis inclure ces facteurs explicatifs exogènes dans le modèle de comportement. Au-delà de la mesure du risque de liquidité, le développement de ces modèles peut également servir à mettre en cohérence le suivi du passif d'un fonds et le calcul d'indicateurs de performance commerciale. Typiquement, tous les clients n'ont pas la même valeur pour une société de gestion, et ceux qui ont tendance à rester investis longtemps sont plus intéressants que les autres. Ainsi, la meilleure compréhension du comportement des clients offre également des outils efficaces de pilotage et de stimulation du développement commercial.

Le troisième enjeu pour les sociétés de gestion concerne le lien entre connaissance du comportement des clients et gestion du risque de liquidité. Dans l'absolu, les modèles de risque devraient tenir compte à la fois des caractéristiques du portefeuille – le risque de liquidité de marché, et de la manière dont ce portefeuille est financé par les clients – le risque de liquidité de financement. Un fonds peut investir dans des actifs illiquides sans souffrir d'une exposition trop grande à la liquidité si son financement s'appuie sur l'apport de clients de long terme. Inversement, un fonds investi en actifs très liquides peut porter un risque de liquidité si l'objectif de placement de ses clients est court. Il est donc essentiel de bien prendre en compte les deux dimensions et de trouver la manière de lier le calcul final du risque à la structure du passif. Prenons l'exemple simple du calcul de la valeur à risque d'un fonds. On peut facilement tenir compte de la liquidité de marché des titres détenus par le fonds et pénaliser les plus illiquides. Mais si l'on s'en tient à cela, on obtient la même valeur à risque pour deux fonds détenant des portefeuilles identiques, mais ayant des structures de clientèles très différentes. Or, un fonds dont les clients investissent à court terme ou sont très sensibles à la performance du fonds devrait être pénalisé, car un choc sur la valeur de marché du fonds implique des sorties aux dates suivantes, et donc des ventes futures d'actifs à des prix potentiellement désavantageux<sup>2</sup>. Cet exemple illustre bien que seule une approche actif/passif permet de maîtriser le risque de liquidité d'un fonds d'investissement. Mesurer le risque de liquidité de l'actif en tenant compte de la structure du passif permet ainsi de mettre en place des mécanismes de gestion de la liquidité afin d'aligner la part des investissements illiquides et la structure du passif du fonds.

Le quatrième concerne la réglementation des fonds ouverts. Il est en effet possible de réduire l'écart de liquidité entre l'actif et le passif d'un fonds en demandant au gérant d'imposer aux clients des restrictions de liquidité. Différents mécanismes permettent d'allonger l'horizon moyen d'investissement d'un client. On peut par exemple lui imposer une durée minimale de détention, ou interdire de racheter à une date donnée plus d'un certain pourcentage de sa position dans le fonds. En réduisant la liquidité offerte, on protège le client d'un écart trop important entre la liquidité de l'actif et du passif, et on permet au gérant de continuer à investir dans des actifs de long terme. Ces restrictions ne doivent être imposées que si elles sont nécessaires, i.e. quand les calculs faits sur un fonds donné montre qu'il y a effectivement un écart important entre la liquidité à l'actif et au passif du fonds. Elles nécessitent donc une meilleure connaissance du passif et de sa structure à un instant donné, ainsi que la dynamique du passif, notamment des effets de contagion que pourraient déclencher des sorties massives suite à des chocs de marché. Dans cette optique, l'accès à l'information sur le passif est bien crucial, et il est important que ces informations soient historisées afin de permettre l'estimation de modèles statistiques du comportement des clients.

L'objet de cet article est de présenter le bilan d'un projet de recherche lancé en 2015 au sein de l'Université Paris-Dauphine, en collaboration avec des acteurs du monde de la gestion d'actifs et portant sur la modélisation du risque de liquidité au passif des fonds d'investissement. La première étape de la réflexion a porté sur la constitution d'une base de données historiques permettant d'archiver les flux de souscription/rachats passés des

---

<sup>2</sup> Bien évidemment, ce raisonnement ne fonctionne que quand il est coûteux de réduire la taille du portefeuille et n'est donc pas valide dans le cas d'un portefeuille investi en actifs infiniment liquides.

clients de différents types de fonds d'investissement. La participation de plusieurs sociétés de gestion a permis l'accès à ces données au niveau des clients individuels, ce qui est inhabituel dans le monde de la recherche académique. Cela nous a notamment permis de mesurer l'avantage de travailler sur des données désagrégées, au niveau du client plutôt qu'au niveau du fonds. Comprendre les liens entre les décisions des clients selon leur type exploite le travail déjà entamé sur le marquage des ordres d'achat/vente. Cette base de données est décrite dans le Section 2 de l'article. Son exploitation a ensuite permis l'étude du comportement des clients, avec pour objectif d'anticiper leurs décisions d'investissements futures à partir de l'observation de leur comportement présent et passé. Le modèle dynamique du comportement des investisseurs est introduit dans le Section 3 de l'article. L'estimation du modèle pour différents fonds a permis d'identifier des effets de clientèle différents selon le type de fonds. Elle a permis de quantifier les différences observés entre les différentes grandes familles d'investisseurs, et donc d'avoir une prévision du montant des sorties éventuelles en fonction de la composition du passif. Les résultats obtenus permettent ainsi d'optimiser la durée des placements à l'actif ou d'anticiper les actions commerciales. Enfin, dans la Section 4, nous abordons les différentes pistes explorées actuellement.

## 2. CONSTITUTION DE LA BASE DE DONNEES

Les sociétés de gestion collectent et stockent à ce jour peu de données relatives aux comportements de leurs clients. Mais avec la transformation digitale et l'évolution de la réglementation, elles auront rapidement accès à des masses de données de plus en plus importantes sur leurs clients et leur manière de gérer leurs investissements. Si ces données sont très utiles pour le développement commercial des fonds, elles doivent également servir aux équipes de gestion à mieux maîtriser le risque de liquidité. Nous décrivons dans cette section les différentes étapes de la réflexion ayant mené à la constitution de la base de données sur le comportement historique des clients de leurs fonds.

### 2.1 Exploiter l'information au passif des fonds

Pour une société de gestion, il est toujours utile d'identifier avec précision qui est l'investisseur final - celui qui prend la décision de souscription ou de rachat. Une meilleure connaissance du passif permet des actions commerciales plus ciblées et une communication plus efficace en temps de crise. Mais, si ce besoin n'est pas nouveau, les solutions existantes pour arriver à une bonne connaissance du passif ne répondent que partiellement aux attentes. Le modèle actuel de distribution des fonds comporte en effet de nombreux intermédiaires entre les sociétés de gestion et les investisseurs finaux, ce qui rend la qualité du suivi des ordres très incertaine.

Malgré cela, on assiste aujourd'hui à un intérêt marqué pour ces questions. La première raison est réglementaire. L'Autorité des marchés financiers (AMF) a rappelé en février 2017 que « la connaissance et l'analyse du passif est une composante essentielle de l'identification des risques par les sociétés de gestion »<sup>3</sup>. Les futures réglementations européennes (MIFID II, Priips) vont également affecter le circuit de distribution et la relation producteur/distributeur, et ainsi pousser à une meilleure connaissance du passif. La directive MIFID II en particulier pourrait être l'occasion de mettre en place un *reporting* des distributeurs vers les producteurs ventilé par critères de risque des investisseurs.

La seconde raison est stratégique. L'émergence des FinTechs et leur modèle de distribution directe pourraient bouleverser les circuits traditionnels et concurrencer fortement les acteurs en place. Le fait d'avoir un accès direct au client final permet de lui demander des informations auxquelles n'ont pas accès les sociétés de gestion traditionnelles. Ces informations peuvent alors servir à mieux cibler l'offre de placement en fonction des caractéristiques de clients. Elles permettent également, une fois agrégées, d'avoir une meilleure vision de la structure du passif d'un fonds, et de mieux anticiper quelle sera son évolution dans des conditions de marché

---

<sup>3</sup> Utilisation des stress-tests dans le cadre de la gestion des risques : Guide pédagogique pour les sociétés de gestion de portefeuille.

stressé. Le fait de ne pouvoir mener ce type d'analyse représente un désavantage certain pour les gérants de fonds dont le mode de distribution coupe le lien direct entre le gérant et ses clients.

Plusieurs pistes pour un suivi efficace du passif sont évoquées dans le guide<sup>4</sup> de l'AMF paru en février 2017. Elles sont essentiellement qualitatives, via une meilleure connaissance de ce lien entre le gérant et l'investisseur. Or, une analyse statistique de l'historique des souscriptions et rachats sur un fonds, en fonction des caractéristiques des clients, permet de mieux évaluer les risques de rachats. L'approche quantitative peut également servir de préalable à une approche plus qualitative, ciblée et au coup par coup en fonction de l'importance des différents clients et de la taille des fonds. Développer une approche quantitative demande cependant l'accès à des ressources, des données et des savoir-faire qu'une société de gestion seule ne possède pas toujours au sein de ses équipes. D'où l'intérêt de travailler ensemble afin d'étendre le champ de l'étude statistique en termes de fonds, de typologies de gestion et de profondeur d'historique grâce à la mise en commun des données, à un important travail de collecte, d'anonymisation et de normalisation.

## 2.2 L'intérêt du consortium pour la recherche académique

La constitution d'une base historique de souscriptions/rachats, à partir des données fournies par les sociétés de gestion, constitue la première étape du projet collaboratif à la base de cette recherche<sup>5</sup>. Aucune base de données sur les choix d'investissement des clients n'était précédemment accessible aux chercheurs. Ces informations sont hautement stratégiques pour les sociétés de gestion, et on comprend qu'elles ne veuillent pas prendre le risque de les partager. Les informations publiques concernent uniquement l'évolution historique de la taille des fonds d'investissement. Si cette information peut s'avérer très utile pour étudier les liens entre la performance et la taille d'un fonds, toutes les analyses sur la liquidité demandent d'avoir une vision plus précise du passif des fonds, et notamment de sa structure. Or, cette structure évolue dans le temps en fonction des entrées et des sorties des clients. Utiliser des données agrégées ne permet pas d'aller au bout de l'analyse.

La première étape de la constitution de la base s'est concentrée sur la collecte des données sur le comportement des clients d'un fonds. Si des historiques de souscriptions/rachats ont rapidement pu être obtenus auprès des sociétés de gestion partenaires, le challenge a été de qualifier le type de chaque client individuel. En effet, un client institutionnel ne réagit pas comme un client particulier à des chocs sur la valeur de leurs placements. Il a pour cela fallu travailler étroitement avec chacune des sociétés de gestion partenaires. Tout ce travail n'aurait pu être fait hors du cadre du consortium associant chercheurs et professionnels. De plus, modéliser les probabilités d'entrées/sorties nécessite de pouvoir travailler sur un grand volume de données. La collaboration de plusieurs sociétés de gestion prêtes à confier leurs historiques a permis de constituer cette large base de données commune et de la mettre à la disposition d'une équipe de chercheurs académiques.

Une fois la base de données constituée, la seconde étape s'est attachée à définir une nomenclature commune des clients. Nous nous sommes en effet rendu compte qu'aucune des sociétés de gestion n'utilisait la même typologie clients. Nous avons donc dû créer cette typologie unique, ainsi que des règles permettant d'affecter les clients à l'une des classifications de cette nouvelle nomenclature.

## 2.3 Les principales caractéristiques de la base de données

L'accès aux données issues des sociétés de gestion partenaires a permis de sélectionner au sein d'un univers très large de fonds un ensemble de 15 fonds appartenant à différents styles de gestion. L'objectif principal de cette sélection est de travailler sur des fonds dont l'exposition au risque de liquidité à l'actif est variable. Les fonds

<sup>4</sup> Utilisation des stress-tests dans le cadre de la gestion des risques : Guide pédagogique pour les sociétés de gestion de portefeuille.

<sup>5</sup> Le projet baptisé « MGPF » pour « Modélisation de la Gestion du Passif des Fonds » a ainsi été labellisé par le Pôle Finance Innovation en 2013. Il a ensuite obtenu un financement par l'Etat et la Région Ile de France via un « FUI » pour « Fond Unique Interministériel » en juin 2014, permettant au Consortium d'initier concrètement les travaux en novembre 2014. Voir Darolles, Le Fol, Lu, Sun (2017) pour une présentation détaillée de ce projet.

monétaires par exemple sont investis dans des actifs liquides, et dont très peu exposés à une dégradation de la liquidité. Les fonds actions grandes capitalisations sont également peu exposés au risque de liquidité de marché, et cette exposition augmente pour les fonds actions de petites et de moyennes capitalisations. A l'extrême, les fonds obligataires sont eux beaucoup plus sensibles à une dégradation de la liquidité car investis dans des actifs peu liquides. Coté passif, tous les fonds sélectionnés offrent en théorie une liquidité journalière. Mais l'analyse du comportement de la clientèle de chaque fonds permettra de mesurer quel est le niveau de consommation effectif de cette liquidité.

Le total des actifs sous gestion sur l'ensemble des 15 fonds sélectionnés dépasse les 10 milliards d'euros, avec certains fonds de taille importante due à la présence de plusieurs fonds monétaires.

**Tableau 1 : Statistiques descriptives**

Société de gestion	Fonds	Catégorie	Nombre parts	Inception	AUM
<b>Firme 1</b>			<b>16</b>		
	Fonds 1	Euro equity large cap	3	02/10/1998	329 723 439
	Fonds 2	Euro equity mid/ sma	2	06/09/1991	376 326 122
	Fonds 3	Euro Fixed Income	2	03/07/1992	255 141 000
	Fonds 4	Euro Fixed Income	3	24/02/1982	375 685 999
	Fonds 5	Euro Fixed Income	3	05/02/1990	935 044 376
	Fonds 6	Euro Money Market	2	31/12/1985	450 074 000
	Fonds 7	Euro Money Market	1	07/07/1995	1 570 300 000
<b>Firme 2</b>			<b>19</b>		
	Fonds 8	Euro equity large cap	5	20/11/2001	280 424 002
	Fonds 9	Euro equity mid/ sma	5	11/05/1994	333 368 999
	Fonds 10	Euro Fixed Income	5	25/10/2000	354 900 000
	Fonds 11	Euro Money Market	4	08/03/2006	3 415 839 000
<b>Firme 3</b>			<b>8</b>		
	Fonds 12	Euro Money Market	2	04/01/2013	1 319 876 994
	Fonds 13	Euro equity large cap	3	09/01/2001	295 271 161
	Fonds 14	Euro equity mid/ sma	2	14/02/1997	34 287 000
	Fonds 15	Euro Fixed Income	1	30/11/2001	388 074 000
<b>Total</b>			<b>43</b>		<b>10 714 336 092</b>

Par exemple, le plus gros fonds de notre échantillon gère plus de 3.4 milliards, alors que le plus petit appartient à la catégorie Europ Equity Mid/Small Cap avec 34 millions sous gestion. Au niveau des transactions, nous avons sur la période d'étude recueilli et analysé plus de 930 000 mouvements d'achat et de vente (578 000 achats et 357 000 ventes). Le tableau montre que le niveau d'activité sur les fonds est très hétérogène, avec un nombre de transactions par jour très important sur certains fonds et très faible sur d'autres. L'analyse de ce tableau montre également que l'activité journalière est ainsi très importante sur les fonds monétaires alors qu'elle est beaucoup plus faible sur les fonds obligataires. Nous remarquons également sur le Tableau 2 que les niveaux d'activité peuvent également varier d'une société de gestion à l'autre.

**Tableau 2 : Niveau d'activité**

Fonds	Période	Jours	Nombre total			Nombre journalier		
			Souscription	Rachats	Transactions	Souscription	Rachats	Transactions
			522 015	283 034	805 049			
Fonds 1	2013-2014	497	174 903	22 134	197 037	351,91	44,53	396,44
Fonds 2	2013-2014	497	144 992	20 880	165 872	297,73	42,01	339,74
Fonds 3	2013-2014	497	18 942	6 436	25 378	38,11	12,95	51,06
Fonds 4	2013-2014	497	3 709	5 983	9 692	7,46	12,04	19,50
Fonds 5	2013-2014	497	5 671	7 323	12 994	11,41	14,73	26,14
Fonds 6	2013-2014	497	36 779	54 044	90 823	74,00	108,74	182,74
Fonds 7	2013-2014	497	137 019	166 234	303 253	275,69	334,47	610,17
			52 773	68 955	121 728			
Fonds 8	2010-2014	1252	7 005	6 354	13 359	5,60	10,67	16,27
Fonds 9	2010-2014	1252	5 663	4 037	9 700	4,52	7,75	12,27
Fonds 10	2010-2014	1252	1 399	6 952	8 351	1,12	6,67	7,79
Fonds 11	2010-2014	1252	38 706	51 612	90 318	30,92	72,14	103,05
			45 768	62 601	108 369			
Fonds 12	2013-2014	493	210	312	522	0,43	0,63	1,06
Fonds 13	2010-2014	1249	1 468	1 400	2 868	1,18	1,12	2,30
Fonds 14	2010-2014	1115	1 877	3 023	4 900	1,68	2,71	4,39
Fonds 15	2010-2014	1233	564	564	1 128	0,46	0,46	0,91
			620 556	414 590	1 035 146			

La seconde étape a consisté à regrouper les clients en groupes homogènes. Pour chacun des fonds, nous avons accès à plusieurs informations : l'ensemble des mouvements achats/ventes effectués par les investisseurs du fonds, un identifiant du client<sup>6</sup>, le nombre de parts concernées, le cours correspondant et la date de l'opération. Chaque société de gestion classe l'ensemble de ses clients en différents groupes, mais cette classification varie d'une société de gestion à l'autre, et il nous a fallu créer une nouvelle classification commune à l'ensemble des sociétés et l'appliquer à l'ensemble des investisseurs traitant sur les fonds sélectionnés. Ce travail nous a amené à définir 16 types d'investisseurs, allant des institutionnels au grand public en passant par les banques privées ou les conseillers en gestion de patrimoine indépendants. Cette classification nous a permis d'observer de manière agrégée ce comportement d'un grand nombre d'investisseurs et donc d'en tirer des conclusions sur leur comportement moyen. Par exemple, observe-t-on un horizon d'investissement proche pour les clients institutionnels et grands publics des fonds de la catégorie actions petites et larges capitalisations ? Il est évident que la seule manière de répondre à cette question est de travailler sur les données désagrégées. L'observation des flux d'investissement au niveau des fonds ne permet tout simplement pas de calculer des durées moyennes de détention. Un second avantage de l'approche désagrégée est de pouvoir calculer séparément le montant des souscriptions et des rachats. Un fonds peut voir son actif sous gestion rester stable, alors qu'une proportion significative des clients institutionnels est remplacée par des clients grand public. Cette modification de la structure du passif du fonds n'est observable qu'à partir des flux individuels et impossible à détecter en exploitant les flux agrégés. Enfin, le traitement statistique des données désagrégées permet également de suivre l'historique d'un client donné au sein d'une même société de gestion<sup>7</sup>. On est ainsi capable de suivre les éventuels arbitrages entre classes d'actifs ou types de fonds d'un client ou d'un groupe de clients donné. Par exemple, une dégradation de la liquidité de marché incite-t-elle certains investisseurs à diminuer leur exposition à ce risque et à investir dans des fonds monétaires ? Là encore, seule une exploitation des données désagrégées permet de répondre à ce type de question.

L'accès aux données désagrégées permet également d'aborder les questions concernant les phénomènes de contagion. Prenons l'exemple de 2 types de clients – les investisseurs institutionnels et grand public, investis dans un fonds donné. On peut imaginer que les clients les mieux informés vont réagir rapidement à une dégradation de l'environnement financier et donc diminuer leur risque plus rapidement, notamment en diminuant leurs investissements dans les actifs les plus risqués. Cette réaction préfigure-t-elle des sorties futures de la part des clients grand public ? Et quelle anticipation peut-on faire du montant de ces sorties futures ? Ces questions peuvent également être posées au niveau de l'univers d'investissement dans son ensemble ou uniquement pour un fonds donné. Au final, l'objectif est d'exploiter les séries temporelles d'achats/ventes des différents

<sup>6</sup> Pour des raisons de confidentialité, un identifiant a été affecté à chaque client au sein de la société de gestion.

<sup>7</sup> Les identifiants utilisés dans l'anonymisation des clients sont définis au sein de chaque société de gestion. Il est donc impossible de suivre les éventuels transferts d'un client entre deux sociétés de gestion différentes.

investisseurs en fonction de leur type. On peut aussi évaluer les risques liés au caractère ouvert des fonds d'investissement, et ceux relatifs à la présence des mêmes investisseurs dans différents fonds. Cela peut créer des canaux de contagion du côté du passif des fonds, que seule une approche désagrégée peut mesurer.

Cependant, et nous le voyons déjà pour l'univers très restreint de 15 fonds retenus pour cette première étude, la taille des échantillons à traiter est très vite énorme – près d'un million de transactions dans notre cas. Et la généralisation de l'étude à l'ensemble des fonds nous amène à traiter plusieurs milliards de transactions. L'exploitation efficace de ces données nécessite le développement de modèles – simples au moins au départ pour les rendre accessibles au plus grand nombre. Ces modèles sont présentés dans la prochaine section.

### 3. MODELISATION DU COMPORTEMENT DES INVESTISSEURS

L'objectif de cette section est de présenter les différentes étapes de la modélisation retenue. Notre volonté a été dans un premier temps de proposer des modèles simples, puis de rajouter pas à pas des niveaux de complexité en fonction des faits stylisés observés. Le but est de développer une méthode standard, en gardant ensuite la possibilité d'utiliser des modèles internes plus sophistiqués. L'analogie peut être faite avec le calcul de la valeur à risque d'un portefeuille. On peut voir les différentes étapes de la modélisation comme le calcul d'une valeur à risque gaussienne, puis d'une valeur à risque ARCH ... Une telle approche permet d'évaluer à chaque étape l'apport des nouveaux paramètres inclus dans le modèle.

Notre but se limite dans cet article à modéliser le nombre de souscriptions et de rachats observés pour un fonds donné, pour l'ensemble des clients. Il est évidemment possible d'appliquer le modèle aux données issues d'un ensemble de fonds – les fonds actions grandes capitalisations par exemple, ou de l'univers des fonds tout entier. A titre d'illustration, nous ne discutons que les résultats obtenus sur des fonds bien identifiés de notre échantillon, en fonction de la stratégie d'investissement suivie et du degré d'exposition au risque de liquidité de marché. Il est également possible d'appliquer cette approche à différents types d'investisseurs, et d'étudier ainsi des effets de causalité dans le comportement de ces investisseurs. Nous ne présentons ici que les premiers résultats obtenus sur l'ensemble des investisseurs sans distinction de types.

Appliquer cette approche à des historiques de souscriptions/rachats observés sur différents types de fonds et pour tous types de clients nous a permis d'obtenir des résultats significatifs. On a pu par exemple constater un comportement très différent des clients en fonction de la liquidité des fonds dans lesquels ils investissent, même quand les fonds affichent des conditions de liquidité identiques.

#### 3.1. Le modèle de Poisson

Le modèle statistique de comptage le plus simple est le modèle de Poisson. Il suppose que le nombre de souscriptions ou de rachats observés chaque jour, noté  $N_t$ , est la réalisation d'une variable aléatoire indépendante et identiquement distribuée selon une loi de Poisson de paramètre  $\lambda$ . L'un des faits stylisés observés sur les données de transactions va cependant à l'encontre des propriétés théoriques de la loi de Poisson – en particulier, la variance est égale à la moyenne<sup>8</sup>. En effet, on observe empiriquement que la variance du nombre de souscriptions ou de rachats est bien supérieure à sa moyenne. Utiliser une simple loi de poisson aurait donc l'inconvénient de mal calibrer la variabilité des séries d'intérêt, et de sous-estimer par exemple la probabilité des scénarios de stress de liquidité. Il est donc essentiel de tenir compte de cela en travaillant avec des modèles permettant de reproduire les caractéristiques empiriques des séries. Pour cela, nous utilisons un modèle de Poisson avec sur-dispersion, i.e. les observations sont supposées tirées dans une loi de Poisson dont le paramètre est égal à  $\lambda F_t$ , où  $F_t$  désigne la valeur prise à la date  $t$  par un facteur inobservable tiré selon une loi binomiale négative d'espérance 1 et de paramètre<sup>9</sup>  $\gamma$ . Le facteur latent  $F_t$  permet de créer de la sur-dispersion et

<sup>8</sup> Voir par exemple Cox (1983).

<sup>9</sup> Voir Johnson et al. (1992).

donc de mieux calibrer les moments de la distribution empirique. L'estimation des deux paramètres  $\lambda$  et  $\gamma$  pour les différents fonds de l'échantillon montre bien que la sur-dispersion est effective, avec des paramètres  $\gamma$  statistiquement différents de zéro. Les valeurs estimées de ce paramètre, et donc les niveaux de sur-dispersion, sont supérieures pour les rachats que pour les souscriptions. Il y a donc beaucoup plus de variabilité temporelle sur ces derniers.

Ce modèle a l'avantage d'être très simple à estimer, mais il a aussi l'inconvénient d'être purement statique, i.e. la loi de l'observation à une date  $t$  ne dépend pas des observations aux dates précédentes. Ceci a une implication directe gênante. La meilleure prévision du nombre de souscriptions ou de rachats à la date suivante est égale à la moyenne des souscriptions ou des rachats, constante dans le temps. Or une étude des séries temporelles correspondantes montre qu'il existe des phénomènes de concentration de souscriptions ou de rachats sur des périodes données. Comme c'est le cas pour la volatilité, une période de forts rachats semble augmenter la probabilité d'observer des rachats importants à suivre. Il y a donc des phénomènes de persistance dans les souscriptions et les rachats qu'il est impossible de capter à l'aide d'un modèle statique. La solution est d'inclure une composante dynamique dans le modèle.

### 3.2. Le modèle autorégressif univarié

La composante dynamique du modèle est incluse en adoptant la philosophie des modèles ARCH<sup>10</sup>, où la volatilité à une date donnée dépend du carré des rendements observés aux dates précédentes. Dans notre cadre, nous proposons de remplacer le paramètre constant de la loi de Poisson  $\lambda$  par un paramètre  $\lambda_t$  qui varie dans le temps comme fonction de l'observation précédente de la série temporelle. La spécification retenue est la suivante :

$$\lambda_t = \lambda_0 + \rho N_{t-1},$$

où le paramètre  $\lambda_0$  est une intensité constante et le paramètre additionnel  $\rho$  capte la persistance temporelle dans les souscriptions ou les rachats. Si  $\rho > 0$ , on observe bien qu'une augmentation du nombre de transactions à la date passée va avoir un impact positif sur l'intensité  $\lambda_t$ , et donc augmenter le nombre moyen de transactions à la date suivante. Ce canal crée à la fois de la persistance et des clusters dans la série temporelle des transactions, à l'image de ce que l'on obtient pour la volatilité à l'aide des modèles ARCH. L'inclusion dans la spécification ci-dessus de retards supplémentaires se fait sans difficulté. Nous nous limitons dans cette étude à l'inclusion d'un seul retard afin de garder le modèle le plus simple et le plus parcimonieux possible. Dans le cas où  $\rho = 0$ , on se retrouve dans le cas simple précédemment décrit – le passé n'a pas d'influence sur le nombre de transactions à la date courante. Il est donc très simple d'évaluer s'il y a ou non de la persistance dans les séries observées. Il suffit d'estimer le paramètre  $\rho$  est de tester s'il est statistiquement différent de zéro. L'estimation de cette spécification sur les séries de souscriptions et de rachats de notre échantillon montre que pour la grande majorité des fonds, les paramètres de persistance  $\rho$  sont significativement différents de zéro, et que les niveaux de persistance les plus importants sont observés pour les rachats. Le risque d'observer des concentrations d'ordres sur des périodes courtes est donc plus important pour les rachats que pour les souscriptions.

D'un point de vue pratique, l'intérêt des modèles dynamiques est de fournir des prévisions de souscriptions ou de rachat non constantes. Dès que le paramètre  $\rho$  est significatif, la prévision des rachats futurs dépend des rachats courants, et celle des souscriptions futures des souscriptions courantes. Cette prévision peut naturellement être utilisée par le gérant d'un fonds qui a la possibilité d'anticiper quel sera le montant des remboursements à la date suivante. Il lui est donc possible de commencer à ajuster la taille de son portefeuille afin de pouvoir faire face sans difficulté aux ordres de remboursements de ses clients.

<sup>10</sup> Voir Engle (1982) pour une introduction aux modèles ARCH.

### 3.3. Le modèle autorégressif avec effets croisés

Dans les approches précédentes, les séries des souscriptions et de rachats sont modélisées séparément. Il est intéressant de voir s'il n'existe pas des effets croisés entre ces deux séries. Pour cela, nous pouvons étendre l'approche proposée en incluant de nouveaux paramètres. Nous notons par la suite  $N_t^{in}$  le nombre de souscriptions dans un fonds à la date  $t$ , et  $N_t^{out}$  le nombre de rachats pour le même fonds toujours à la date  $t$ . Nous supposons alors que l'intensité des lois de Poisson décrivant les souscriptions à la date  $t$ ,  $\lambda_t^{in}$  et celle décrivant les rachats  $\lambda_t^{out}$  satisfont les deux équations suivantes :

$$\lambda_t^{in} = \lambda_0^{in} + \rho^{in-in} N_{t-1}^{in} + \rho^{in-out} N_{t-1}^{out},$$

$$\lambda_t^{out} = \lambda_0^{out} + \rho^{out-in} N_{t-1}^{in} + \rho^{out-out} N_{t-1}^{out},$$

où les deux paramètres  $\rho^{in-out}$  et  $\rho^{out-in}$  permettent de capturer les dépendances entre les souscriptions (resp. les rachats) à la date  $t$  et les rachats (resp. les souscriptions) à la date précédente. Tous les autres paramètres du modèle gardent leurs interprétations précédentes.

Il est maintenant possible de discuter de l'interprétation financière des quatre paramètres  $\rho$  inclus dans la version la plus générale du modèle. Le premier paramètre  $\rho^{in-in}$  peut facilement s'interpréter en termes de réputation. Des souscriptions passées augmentent en moyenne le nombre de souscriptions courantes. Le paramètre  $\rho^{out-out}$  capture lui les effets de panique. Les investisseurs, voyant un nombre de sorties importantes se réaliser interprètent cela comme un signal négatif et ont tendance à également sortir du fonds. Les effets croisés peuvent également être interprétés. Le paramètre  $\rho^{in-out}$  mesure la capacité du gérant à stabiliser la taille de son fonds, par exemple en déclenchant des actions commerciales pour compenser par un nombre plus important de souscriptions les sorties passées. Enfin, le dernier paramètre  $\rho^{out-in}$  capture lui le comportement des investisseurs qui sortent du fonds suite à l'entrée massive d'autres investisseurs. On peut voir cela comme un comportement *contrarian* de certains investisseurs qui quittent le fonds quand de trop bonnes performances attirent un nombre trop important de nouveaux investisseurs. Ils anticipent des problèmes de capacité et une détérioration de la performance du fait de l'accroissement de la taille du fonds.

En termes de gestion du risque de liquidité, la présence d'effets croisés positifs est plutôt bénéfique puisqu'elle a tendance à stabiliser l'actif sous gestion du fonds. Le cas le plus critique correspond à une valeur  $\rho^{in-out} < 0$ , qui signifie moins de souscriptions en moyenne quand les rachats augmentent. Grâce à cette modélisation, on peut séparer en deux les effets négatifs des sorties passées. Elles peuvent en effet augmenter les sorties futures ou diminuer les entrées futures. L'analyse empirique permettra de dire lequel des deux effets est le plus important.

**Tableau 3 : Estimateurs par type de fonds**

Catégorie	$\lambda^{in}$	$\lambda^{out}$	In-In	In-Out	Out-In	Out-Out	$g^{in}$	$g^{out}$
Fonds Monétaire	8.1***	10.1***	0.32***	0.30***	0.19***	0.46***	35.5***	44.9***
Fonds Actions Large Capitalisation	4.3***	4.8*	0.22***	0.00	0.78***	0.18***	35.9***	68.8***
Fonds Actions Petite/Moyenne Capitalisation	0.5***	1.1***	0.52***	0.06**	0.16***	0.19***	2.5***	4.4***
Fonds Obligataire	0.4***	0.4***	0.10*	0.00	0.00	0.02	0.5***	0.7***

Le Tableau 3 présente les estimateurs<sup>11</sup> de ce dernier modèle pour 4 fonds appartenant à des catégories différentes, et donc avec une exposition variée au risque de liquidité de marché. L'objectif de l'exercice est de comparer les différences dans le comportement de la clientèle de ces fonds en fonction de leur liquidité, alors qu'ils offrent tous une liquidité journalière.

<sup>11</sup> Les paramètres statistiquement significatifs à 10% (resp. 5%, 1%) sont indiqués dans le tableau par \* (resp. \*\* ; \*\*\*)

On peut commencer l'analyse du tableau par les composantes marginales capturées via les paramètres  $\beta$  du tableau. On constate que plus on augmente l'exposition au risque de liquidité du fonds, plus la valeur des paramètres diminue. On observe donc bien en moyenne beaucoup moins de mouvements sur les fonds dont l'actif est illiquide – les fonds obligataires par exemple. Si l'on se concentre maintenant sur le risque lié aux effets de panique, il faut faire un focus sur la colonne out-out reprenant les estimateurs des  $\rho^{out-out}$  pour les quatre fonds. On observe que ce paramètre est significatif pour les trois premiers fonds, les moins exposés à la liquidité de marché, alors qu'il devient non significatif pour le dernier. Il semble donc que les clients prennent en compte dans leur comportement d'investisseur la difficulté que peut rencontrer le gérant dans la gestion de la liquidité offerte. Un client d'un fonds obligataire va moins réagir aux sorties des autres investisseurs que le client d'un fonds liquide. On peut interpréter la valeur très importante obtenue sur les fonds monétaires comme des effets de saisonnalité qui expliquent que les clients sortent des fonds sur des périodes similaires. La première conclusion de cette étude est donc que les clients des fonds intègrent bien la dimension liquidité dans leur politique d'investissement et que la liquidité journalière offerte par le fonds n'est pas utilisée de la même manière selon le type de fonds considéré.

Si l'on intéresse maintenant aux effets croisés entre les souscriptions et les rachats, on observe que ces effets sont significativement différents de zéro et positifs pour les 3 fonds les plus liquides. Ils jouent donc un rôle stabilisateur et en particulier compensent les sorties passées par un nombre plus important de nouvelles entrées. Ce mécanisme stabilisateur n'existe par contre pas pour le fonds le plus exposé au risque de liquidité. Enfin, pour finir, nous pouvons commenter les valeurs de paramètre de sur-dispersion  $\beta$ . Les valeurs les plus importantes sont observées pour les fonds liquides avec un maximum atteint pour le fonds actions grandes capitalisations. De plus, les sorties sont systématiquement plus dispersées que les souscriptions.

#### 4. CONCLUSION

Travailler sur le comportement des clients des fonds d'investissement au niveau le plus fin possible nous a permis d'obtenir un certain nombre de résultats nouveaux. Dans un premier temps, nous avons pu mesurer quelle était la difficulté de constituer une base de données unique à partir de l'information obtenue auprès de différentes sociétés de gestion. Les particularités des systèmes de distribution font qu'il est très difficile de caractériser le client final, et donc constituer des groupes homogènes de clients afin d'en étudier le comportement. Nous avons ensuite développé des modèles relativement simples et facilement estimables du comportement des investisseurs. Ces modèles n'exploitent clairement qu'une infime partie de l'information continue dans les données, mais ils permettent cependant de tirer un certain nombre de conclusions sur le comportement des investisseurs en fonction de la liquidité du fonds dans lequel ils placent leur argent. Ces modèles de comportement fournissent un bloc élémentaire essentiel à l'analyse de la transformation de liquidité effectuée par les fonds d'investissement. Elle permet en effet d'évaluer le décalage de liquidité entre le niveau de liquidité du portefeuille et la liquidité effectivement demandée par les clients du fonds.

Nous ne sommes évidemment qu'au début des analyses statistiques du comportement des investisseurs. On pourrait sûrement améliorer les capacités de prédiction des modèles en regardant par exemple les corrélations entre les comportements de différents types d'investisseurs, institutionnels et grand public par exemple. Le fait de travailler sur des échantillons de données beaucoup plus larges permettrait également d'améliorer la qualité des résultats, et d'utiliser des modèles beaucoup plus complexes. Nous pensons cependant qu'avant d'aller plus loin dans la modélisation, il est important de participer à l'éducation des acteurs du monde de la gestion d'actifs et de commencer à bien expliquer ce qu'apporte une meilleure connaissance statistique du comportement des clients d'un fonds. Pour cela, des ateliers ont permis aux équipes des risques des sociétés de gestion partenaires du projet d'échanger sur l'intérêt des travaux présentés dans cet article. Le prototype d'un outil de gestion de risque a également été réalisé. Basé sur une plateforme de big data, l'outil utilise la puissance de stockage et de calcul pour apporter aux gérants, aux commerciaux et aux équipes des risques une connaissance en temps réel des probabilités de rachats au passif et leur permettre de simuler différents scénarios de stress de liquidité en intégrant des chocs sur la structure du passif.

## Références

Autorité des Marchés Financiers (2017), Utilisation des stress-tests dans le cadre de la gestion des risques : Guide pédagogique pour les sociétés de gestion de portefeuille, Février 2017, Publications de l'AMF.

Cox, D. (1983), Some Remarks on Overdispersion, *Biometrika*, 70, 269-274.

Darolles, S., Le Fol, G., Lu, Y., Sun, T. (2017), Self-Exciting Model for Mutual Fund Flows: Investor Behaviour and Liability Risk, Université Paris-Dauphine Working Paper.

Darolles, S., Roussellet, G. (2017), Managing Hedge Fund Liquidity Risks, Université Paris-Dauphine Working Paper.

Engle, R. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation, *Econometrica*, 50, 987-1008.

Johnson, N., Kotz, S., Kemp, A. (1992), *Univariate Discrete Distribution*, Second Edition, Wiley, New York.