

OPINIONS & DÉBATS

N°21 - Juin 2021

Robo-Advising : Moins d'IA et plus de XAI ?
Améliorer les algorithmes en y impliquant de l'humain

Robo-Advising: Less AI and More XAI?
Augmenting algorithms with humans-in-the-loop

Milo Bianchi & Marie Brière



SOMMAIRE

Résumé	7
I. Introduction	8
II. Pourquoi une telle popularité ?	9
2.1 Révolution Fintech	10
2.2 Les problèmes fondamentaux des investisseurs	10
2.3 Les problèmes fondamentaux des conseillers	11
III. Les promesses	12
3.1 Procédures auditable et recommandations personnalisées	12
3.1.1 Le profilage du client	12
3.1.2 L'allocation d'actifs	13
3.1.3 Hétérogénéité des allocations d'actifs proposées	14
3.2 Améliorer la situation des investisseurs	15
3.3 Atteindre les investisseurs mal desservis	16
IV. Questions ouvertes	17
4.1 Pourquoi pas plus d'IA/Big Data?	17
4.2 Jusqu'où devons-nous aller dans la personnalisation?	19
4.2.1 Difficulté de mesurer les caractéristiques d'un individu	19
4.2.2 Sensibilité des algorithmes à l'incertitude des paramètres	20
4.3 Les humains peuvent-ils faire confiance aux robots ?	21
4.3.1 La confiance est clé dans l'adoption des services de robo-advice	21
4.3.2 Aversion pour les algorithmes	21
4.3.3 Qu'est-ce qui engendre la confiance dans un algorithme ?	23
4.3.4 La confiance envers les robots et les décisions financières	24
4.4 Les robots remplacent-ils ou complètent-ils la prise de décision humaine ?	25
V. La prochaine génération de robo-advisors	26
Notes	29
Références	50

Opinions & Débats N° 21 - Juin 2021

Publication de l'Institut Louis Bachelier

Palais Brongniart - 28 place de la Bourse 75002 Paris ♦ Tél. : 01 73 01 93 40 ♦ www.institutlouisbachelier.org

DIRECTEUR DE LA PUBLICATION : Jean-Michel Beacco ♦ COORDINATION ÉDITORIALE : Ryadh Benlahrech

CONTACT : ryadh.benlahrech@institutlouisbachelier.org

CONCEPTION GRAPHIQUE : Vega Conseil 01 48 85 92 01

CONTENTS

Abstract	30
I. Introduction	31
II. Why So Popular?	32
2.1 Fintech Revolution	32
2.2 Fundamental Problems with Investors	33
2.3 Fundamental Problems with Advisors	34
III. Promises	35
3.1 Accountable Procedures and personalized recommendations	35
3.1.1 Client profiling	35
3.1.2 Asset Allocation	35
3.1.3 Heterogeneity in the proposed asset allocations	36
3.2 Improving investor's experience	37
3.3 Reach Under-served Investors	38
IV. Open Questions	39
4.1 Why Not More AI/Big Data?	39
4.2 How Far Should We Go Into Personalization?	40
4.2.1 Difficulty of measuring an individual's characteristics	40
4.2.2 Algorithm sensitivity to parameter uncertainty	41
4.3 Can Humans Trust Robots?	42
4.3.1 Trust is key for robo-advice adoption	42
4.3.2 Algorithm aversion	43
4.3.3 What creates trust in an algorithm?	44
4.3.4 Trust in robots and financial decisions	44
4.4 Do Robots Substitute or Complement Human Decision-Making?	45
V. The Next Generation of Robo-Advisors	46
Notes	49
References	50

Les articles publiés dans la série "Opinions & Débats" offrent aux spécialistes, aux universitaires et aux décideurs économiques un accès aux travaux de recherche les plus récents. Ils abordent les principales questions d'actualité économique et financière et fournissent des recommandations en termes de politiques publiques.

The Opinions and Debates series sheds scientific light on current topics in economics and finance. Bringing together several types of expertise (from mathematicians, statisticians, economists, lawyers, etc.) this publication makes recommendations in the formulation and implementation of public economic policy.



EDITO



Jean-Michel Beacco
*Délégué général
de l'Institut Louis Bachelier*

Depuis quelques années, la révolution des nouvelles technologies, à l'œuvre dans le secteur financier, s'illustre dans de nombreux domaines et bouscule les acteurs traditionnels que sont les banques, les assureurs et les gestionnaires d'actifs. Il faut dire que leur champ d'application s'est considérablement élargi, notamment avec la démocratisation des méthodes d'apprentissage de l'intelligence artificielle, comme le machine et le deep learning ou encore le développement grandissant des fintechs. Cette mutation technologique a bien entendu été permise grâce à l'afflux, la disponibilité et une meilleure compréhension des données provenant de l'utilisation d'Internet et des applications mobiles.

Et parmi les récentes applications et innovations en finance, figurent le robo-advisor, des conseillers financiers virtuels basés sur des algorithmes, qui permettent aux épargnants d'investir dans des placements censés mieux représenter leur profil de risque et leurs préférences individuelles. Si pour l'heure, ce type de produit n'est pas encore très répandu en France, il est probable qu'il se développe davantage dans les années à venir pour plusieurs raisons non exhaustives : l'analyse plus fine des besoins des investisseurs individuels, des alertes plus systématiques générées par l'automatisation, des frais de gestion plus bas permettant un accès plus large aux populations moins aisées. Toutefois, pour que ces projections se réalisent, les algorithmes développés et utilisés doivent nécessairement inclure de l'explicabilité et éviter ainsi de devenir des boîtes noires. Dans ce contexte, l'humain doit jouer un rôle majeur en servant de relais à ces technologies pour qu'elles puissent bénéficier au plus grand nombre.

C'est pour y voir plus clair sur le sujet, qui mêle à la fois les modélisations mathématiques liées aux nouvelles technologies, les préférences individuelles des investisseurs ou les interactions entre les humains et les machines, que Milo Bianchi et Marie Brière, éminents chercheurs en économie et finance, ont contribué à ce nouveau numéro de la collection Opinions & Débats.

Bonne lecture !

For several years now, the new technologies revolution, currently at play in the financial sector, has become apparent in many fields and is shaking up traditional players like banks, insurers and asset managers. It is worth noting that their scope of application has considerably expanded, particularly with the popularization of artificial intelligence learning methods, such as machine and deep learning, and the increasing development of fintechs. The inflow, availability and better understanding of Internet and mobile application data has of course enabled this technological transformation.

Recent applications and innovations in finance include robo-advisors, algorithm-based virtual financial advisors that allow savers to invest in assets meant to better represent their individual risk profile and individual preferences. Although this type of product is not yet widespread in France, it is likely to develop further in the next years, for several reasons not limited to: a finer analysis of the needs of individual investors, more systematic warnings generated by automation, and lower management fees allowing wider access to less wealthy populations. However, for these projections to be achieved, the algorithms developed and used must necessarily include explicability and avoid becoming black boxes. In this regard, humans must play a major role in supporting these technologies so that they can benefit as many people as possible.

In order to gain a clearer understanding of the subject, which combines mathematical modelling of new technologies, individual investor preferences and interactions between humans and machines, Milo Bianchi and Marie Brière, leading researchers in economics and finance, have contributed to this new issue of the Opinions & Debates collection.

Enjoy your reading!

Jean-Michel Beacco

Delegate General of the Institut Louis Bachelier

BIOGRAPHIE



Milo Bianchi

Milo Bianchi est professeur de finance à la Toulouse School of Economics (TSE). Ses travaux de recherche portent sur les fintechs et la finance durable, avec un accent particulier sur les investisseurs individuels. Ses travaux ont été publiés dans les principales revues d'économie et de finance, notamment *Journal of Finance*, *Review of Economic Studies*, *Journal of Economic Theory* et *Management Science*. Il est également membre junior de l'Institut Universitaire de France, directeur de l'initiative FIT-IN (*Financial Inclusion Through INteroperability Initiative*), et membre des centres de finance durable et de finance numérique à TSE. Il a obtenu son doctorat à la Stockholm School of Economics et a occupé des postes de recherche dans diverses institutions, notamment au MIT, à la Paris School of Economics, à l'University College London et à la Shanghai University of Finance and Economics.

Milo Bianchi is Professor of Finance at the Toulouse School of Economics. His current research agenda focuses on fintech and on sustainable finance, with specific focus on individual investors. His work has been published in leading economics and finance journals including *Journal of Finance*, *Review of Economic Studies*, *Journal of Economic Theory*, and *Management Science*. Milo is junior member of the Institut Universitaire de France, director of the FIT-IN Initiative (*Financial Inclusion Through INteroperability Initiative*), and member of the Sustainable Finance and the Digital Finance Centers at TSE. Milo has received his PhD from the Stockholm School of Economics and he has held research positions at various institutions including MIT, Paris School of Economics, University College London, and Shanghai University of Finance and Economics.

Marie Brière est responsable du Centre de recherche aux investisseurs chez Amundi. Elle enseigne à l'Université Paris Dauphine-PSL et est chercheur associé au Centre Emile Bernheim, Solvay Business School rattaché à l'Université Libre de Bruxelles. A l'Institut Louis Bachelier, elle préside le conseil scientifique du Forum International des Risques Financiers (Risk Forum). Ses recherches portent sur l'allocation d'actifs à long terme et la gestion des risques, afin de conseiller les décisions stratégiques des investisseurs institutionnels et la conception de solutions d'investissement pour les investisseurs individuels. Elle est membre de plusieurs conseils scientifiques, comme celui de l'Institut européen des marchés de capitaux du CEPS ou de l'Observatoire de l'épargne européenne, et fait partie du groupe d'experts qui conseille le comité permanent de l'innovation financière de l'ESMA. Elle est titulaire d'un doctorat en économie de l'Université Paris X et diplômée de l'ENSAE. Ses articles scientifiques ont été publiés dans des revues académiques, notamment le *Journal of Banking and Finance*, *Journal of International Money and Finance*, *Journal of Portfolio Management*, *Financial Analyst Journal*. En 2014, elle a reçu le prix Markowitz pour son article co-écrit avec Zvi Bodie on "Sovereign Wealth and Risk Management: A Framework for Optimal Asset Allocation of Sovereign Wealth", publié dans le *Journal of Investment Management*.



Marie Brière

Marie Brière is Head of the Investor Research Center at Amundi. She teaches at Paris Dauphine University and is a senior associate researcher with the Centre Emile Bernheim, Solvay Business School, Université Libre de Bruxelles. At Institut Louis Bachelier, she chairs the Scientific Committee of the International Financial Risk Forum. Marie conducts research on long term asset allocation and risk management, to advise strategic decisions of institutional investors and the design of investment solutions for individual investors. She is a member of several scientific councils, such as that of the European Capital Market Institute of CEPS or of the European Savings Observatory, and is a member of the expert group advising the ESMA Standing Committee on Financial Innovation. She holds a PhD in economics from the University Paris X and graduated from ENSAE. Her scientific articles have been published in academic journals, including the *Journal of Banking and Finance*, *Journal of International Money and Finance*, *Journal of Portfolio Management*, *Financial Analyst Journal*. She received the Markowitz award for her article with Zvi Bodie on "Sovereign Wealth and Risk Management: A Framework for Optimal Asset Allocation of Sovereign Wealth", published in the *Journal of Investment Management*.

Robo-Advising : Moins d'IA et plus de XAI ?

Améliorer les algorithmes en y impliquant de l'humain*

Milo Bianchi

École d'Économie de Toulouse, TSM, et IUF, Université de Toulouse Capitole

Marie Brière

Amundi, Université Paris-Dauphine et Université Libre de Bruxelles, Louis Bachelier Fellow

Résumé

L'intérêt des universitaires et de l'industrie pour les robo-advisors n'a pas cessé de croître ces dernières années. Nous examinons en quoi le conseil financier robotisé pourrait résoudre certains problèmes fondamentaux dans la prise de décision des investisseurs, ainsi que dans le conseil financier traditionnel, en favorisant l'inclusion financière, et en fournissant des recommandations individualisées sur la base de procédures transparentes et auditable. Nous aborderons ensuite certaines des problématiques concernant l'avenir du robo-advice. Premièrement, le rôle que l'intelligence artificielle joue, et devrait jouer, dans le robo-advice. Deuxièmement, jusqu'où doit aller la personnalisation des recommandations ? Par ailleurs, comment la confiance en des conseils financiers automatisés peut-elle être générée et soutenue ? Enfin, en quatrième point, les robots sont-ils perçus comme étant complémentaires ou substituables à la prise de décision humaine ? Notre conclusion présente quelques réflexions sur ce que pourrait être la prochaine génération de robo-advisors. Nous insistons sur l'importance des récentes découvertes en matière d'intelligence artificielle explicable (XAI), y compris sur la manière dont les nouvelles formes d'IA appliquées aux services financiers bénéficieraient de l'importation de connaissances économiques et psychologiques pour concevoir une meilleure interaction humain-robot.

* Milo Bianchi reconnaît le soutien financier du LTI@Unito, du TSE Sustainable Finance Center et de l'ANR (subvention ANR-17-EURE-0010). Nous n'avons aucun intérêt matériel en lien avec les recherches présentées dans cet article. Cet article a été préparé pour le livre "Machine Learning in Financial Markets: A Guide to Contemporary Practice", édité par Agostino Capponi et Charles-Albert Lehalle pour Cambridge University Press.

I. Introduction

Les conseillers robotisés, communément appelés robo-advisors, suscitent un intérêt croissant, aussi bien dans le milieu universitaire que dans le monde de la gestion d'actifs. Dans cet article, nous nous proposons d'examiner certaines des raisons derrière cet intérêt croissant. Nous soulignerons de quelle manière le développement des robo-advisors peut être appréhendé dans le contexte plus général de ladite révolution Fintech. Nous examinerons également certaines des raisons plus spécifiques qui expliquent l'intérêt pour ces conseils financiers automatisés, en nous appuyant sur les problèmes fondamentaux auxquels les investisseurs individuels sont confrontés lorsqu'ils prennent des décisions financières, ainsi que sur les limites souvent constatées dans le conseil financier traditionnel.

Nous verrons ensuite en quoi le robo-advising pourrait éventuellement résoudre ces problèmes fondamentaux et nous présenterons les principaux avantages des robots : (1) promouvoir l'inclusion financière en touchant les investisseurs les moins favorisés, (2) fournir des recommandations sur mesure en fonction de procédures transparentes et auditées et, (3) améliorer la situation globale des investisseurs. Pour chacune de ces promesses, nous reviendrons sur les raisons pour lesquelles un certain espoir peut être placé dans les robots, à partir de ce que la littérature académique a démontré jusqu'à présent.

Dans la troisième partie de cet article, nous aborderons ce que nous croyons être les questions fondamentales qui se posent aujourd'hui concernant l'avenir du robo-advice. Tout d'abord, nous discuterons du rôle que l'intelligence artificielle (IA) joue, et devrait jouer, dans le robo-advice. Nous soulignerons les contraintes, tant réglementaires qu'en matière de théorie du choix de portefeuille qui peuvent limiter le rôle de l'IA dans le robo-advice. Nous insisterons également sur le fait que la quête de simplicité et d'explicabilité des recommandations d'investissement pourraient rendre l'IA indésirable, même si celle-ci est réalisable. Deuxièmement, nous examinerons jusqu'où nous devrions aller en matière de personnalisation des recommandations robotisées. Nous soulignerons le trade-off entre l'objectif visant à rapprocher un portefeuille des besoins d'un individu et les risques liés aux éventuelles erreurs de mesure concernant les caractéristiques individuelles concernées (par exemple, l'aversion au risque). Nous discuterons de la sensibilité des algorithmes face à l'incertitude liée à l'estimation des paramètres. Troisièmement, nous aborderons la manière dont le robo-advice peut nous éclairer sur les questions plus importantes concernant les interactions homme/robot et les mécanismes de confiance en matière de services financiers automatisés. Nous examinerons les arguments relatifs à l'aversion envers les algorithmes, ainsi que les moyens possibles de la réduire, et comment ils peuvent être appliqués dans le cadre des conseils financiers automatisés. Enfin, nous analyserons si les robots doivent être perçus comme des compléments ou des substituts à la prise de décision humaine.

Nous concluons en présentant nos observations sur ce à quoi pourrait ressembler la prochaine génération de robo-advisors. Plutôt que de poursuivre la tendance actuelle à utiliser davantage de données, des modèles plus complexes et des interactions plus automatisées, nous définirons une voie alternative qui s'appuie sur les principes clés du robo-advisor en termes de responsabilité et d'inclusion financière. Le défi essentiel que les robo-advisors auront à surmonter concerne le renforcement de la confiance dans la technologie financière. Nous soulignerons l'importance des découvertes récentes sur la XAI (Explainable Artificial Intelligence) et nous insisterons sur la nécessité pour les nouvelles formes d'IA appliquées aux services financiers d'importer des connaissances issues des sciences sociales, telles que l'économie et la psychologie. Cette revue n'a pas pour objectif d'être exhaustive. Elle doit plutôt être perçue comme étant complémentaire aux revues existantes (telles que D'Acunto et Rossi, 2020).

II. Pourquoi une telle popularité ?

Les robo-advisors utilisent des procédures automatisées, allant d'algorithmes relativement simples qui utilisent des informations partielles sur le client, à des systèmes d'intelligence artificielle construits autour du big data, dans le but de recommander la manière d'allouer des fonds entre différents types d'actifs. Tout d'abord, une technique de profilage du client est utilisée pour évaluer les caractéristiques d'un investisseur (aversion au risque, connaissances financières, horizon d'investissement...) et ses objectifs. Ensuite, un univers d'investissement est défini et, enfin, un portefeuille est proposé en tenant compte des objectifs d'investissement et du niveau de risque souhaité. Comme le documente Beketov, Lehmann et Wittke (2018), dans la plupart des cas, le portefeuille optimal proposé par les robo-advisors, s'appuie sur la théorie moderne du portefeuille, qui remonte à Markowitz (1952). En plus de recommander une allocation initiale de fonds, les algorithmes peuvent être conçus pour surveiller en permanence les portefeuilles et détecter les écarts par rapport au profil ciblé. Lorsque des déviations sont identifiées, le client est alerté et/ou le portefeuille est automatiquement rééquilibré. Le portefeuille peut également être automatiquement rééquilibré pour réduire le risque au fil du temps ou lorsque l'investisseur modifie sa tolérance au risque ou ses objectifs d'investissement. Certains robots aux Etats-Unis proposent également de mettre en œuvre des techniques d'optimisation fiscale : vendre des actifs qui subissent une perte et utiliser le produit de la vente pour acheter des actifs présentant un risque similaire, afin de diminuer les gains en capital et le revenu imposable sans affecter l'exposition du portefeuille au risque. Les robots offrent enfin au client la possibilité de visualiser la répartition de son portefeuille, ainsi que des statistiques intéressantes, telles que le rendement annuel attendu et la volatilité, ainsi que la projection de sa richesse accumulée dans le temps, en utilisant des simulations de Monte Carlo basées sur des hypothèses sur les rendements attendus (souvent considérés identiques aux performances historiques des actifs).

Le marché est en pleine croissance. La plupart des praticiens estiment que le marché mondial se situe actuellement autour de 1 000 milliards de dollars, contre 100 milliards de dollars en 2016 (S&P Global Market Intelligence, Backend Benchmarking, Aite Group, voir Buisson, 2019). Les actifs sous gestion dans le segment des robo-advisors dans le monde devraient atteindre entre 1700 milliards de dollars et 4 600 milliards de dollars en 2022 (Statista, BI Intelligence). Le nombre d'utilisateurs devrait s'élever à 436 millions d'ici 2024 (Statista 2020). Cette croissance est stimulée par l'entrée de grands opérateurs historiques dans le secteur des services numériques (par exemple, JPMorgan et Goldman Sachs ont annoncé le lancement d'un service numérique de gestion de patrimoine en 2020) et la migration des actifs gérés par les grandes institutions financières vers leurs opérations de robo-advice, qui s'élève à 8 % de leurs actifs sous gestion et à un quart des actifs des comptes de moins de 1 million de dollars. Simultanément, les clients ont multiplié les demandes d'outils d'investissement numériques, et en particulier de gestion de portefeuille à faible coût et de services adjacents tels que la planification financière. Les États-Unis restent, de loin, le premier marché pour le robo-advising (avec plus de 200 robo-advisors enregistrés), mais le nombre de robo-advisors augmente rapidement en Europe (plus de 70), et aussi en Asie, sous l'impulsion d'une classe moyenne émergente et d'une connectivité technologique élevée (Abraham, Schumkler et Tessada (2019)).¹

Nous renvoyons à Grealish et Kolm (2021) pour plus de détails sur le fonctionnement des robo-advisors et sur les tendances récentes du marché, ainsi que pour les raisons qui pourraient motiver une croissance aussi rapide du marché et un intérêt accru dans les milieux universitaires et les instances politiques.

2.1 Révolution Fintech

Une partie de l'intérêt pour le robo-advising vient de la tendance plus large de l'application de nouvelles technologies et de sources de données innovantes dans le secteur financier, un phénomène lié au développement plus général des fintechs. Ce terme a joué un rôle central dans de nombreux débats universitaires et politiques au cours des dernières années. Les enthousiastes des fintechs parlent d'une révolution qui promet de perturber et de remodeler le secteur des services financiers.²

Buchanan (2019) évoque la croissance mondiale de l'industrie de l'IA et son application au secteur de la finance. Citant un rapport de 2017, elle mentionne que 5,154 startups d'IA ont été créées dans le monde au cours des cinq dernières années, ce qui représente une augmentation de 175 % par rapport aux 12 années précédentes. Cette croissance impressionnante a été stimulée par les progrès de la puissance de calcul, qui ont entraîné une baisse du coût du traitement et du stockage des données, et en même temps par la disponibilité de données de taille et de portée plus grandes. De même, les publications de brevets liés à l'IA aux États-Unis sont passées d'environ 50 en 2013 à approximativement 120 en 2017. En Chine, cette croissance a été encore plus spectaculaire, avec environ 120 brevets en 2013 contre 640 en 2018. L'introduction de l'IA transforme le secteur des services financiers, non seulement en termes de robo-advising, mais aussi en matière de détection des fraudes et de conformité, de chatbots et d'algo-trading (Buchanan (2019)).

Dans les milieux universitaires, l'attention grandissante se manifeste, par exemple, par la croissance exponentielle des études académiques en finance principalement centrées sur l'IA. Bartram, Branke et Motahari (2020) analysent le nombre de mots-clés liés à l'IA dans le titre, le résumé ou les mots-clés listés de tous les documents de recherche postés dans le Financial Economics Network (FEN) entre 1996 et 2018.³ En 1996, aucun document de travail comportant un mot-clé lié à l'IA n'a été posté. En 2018, il y en avait 410, soit 3 % de tous les travaux publiés cette année-là.

Les robo-advisors promettent d'appliquer de nouvelles technologies et procédures pour améliorer la prise de décision financière, comme nous le verrons plus loin, et en tant que tels, ils peuvent être considérés comme une composante de la révolution fintech, au même titre que les monnaies numériques qui promettent de redéfinir le rôle de l'argent traditionnel et les prêts sur plateforme qui offrent la possibilité de redéfinir le rôle de l'accès traditionnel au crédit.

2.2 Les problèmes fondamentaux des investisseurs

D'un point de vue plus spécifique, l'un des principaux intérêts du robo-advising est qu'il est désormais communément admis que de nombreux investisseurs pourraient améliorer considérablement leurs décisions financières. Au cours des dernières décennies, la recherche a démontré que les décisions des investisseurs peuvent s'écarter, parfois de manière fondamentale, des hypothèses standard d'un agent économique totalement rationnel, qui connaît l'ensemble des alternatives possibles avec leur résultat au sens probabiliste, et qui est capable d'associer correctement toutes les informations afin de maximiser son espérance d'utilité à vie.

Si l'on se limite au domaine de l'investissement, qui a été jusqu'à présent l'objet type des robo-advisors, on a constaté que les investisseurs affichent une faible participation au marché actions (Mankiw et Zeldes, 1991), une diversification insuffisante de leur portefeuille (Grinblatt, Keloharju et Linnainmaa, 2011 ; Goetzmann et Kumar, 2008 ; Bianchi et Tallon, 2019), et différents biais comme le biais lié au choix par défaut (Benartzi et Thaler, 2007), l'inertie du portefeuille (Agnew, Balduzzi et Sunden, 2003 ; Biliias, Georgarakos et Haliassos, 2010), les transactions excessives (Odean, 1999), la chasse aux tendances (Greenwood et Nagel, 2009) et une mauvaise compréhension du mécanisme d'abondement (Choi, Laibson et Madrian, 2009). Nombre de ces comportements d'investissement sont associés à une mauvaise compréhension des principes financiers de base (Lusardi, Michaud et Mitchell, 2017 ; Lusardi et Mitchell, 2014 ; Bianchi, 2018).

Plusieurs travaux fournissent une liste exhaustive des biais et des erreurs de trading associées (voir par exemple Guiso et Sodini, 2013, Barber et Odean, 2013, Beshears, Choi, Laibson et Madrian (2018)). Dans le cadre de cet article, deux points méritent d'être soulignés. Premièrement, ces erreurs ne sont pas minimes ; au contraire, leurs répercussions sur le plan du bien-être de l'individu peuvent être considérables (Campbell, 2006, Campbell, Jackson, Madrian et Tufano (2011)). Deuxièmement, elles ne s'annulent pas à l'équilibre ; elles ont plutôt des effets importants sur le fonctionnement des marchés financiers et sur des questions macroéconomiques plus larges telles que les inégalités de richesse (Vissing-Jorgensen, 2004 ; Lusardi, Michaud et Mitchell, 2017 ; Bach, Calvet et Sodini, 2020 ; Fagereng, Guiso, Malacrino et Pistaferri, 2020).

Compte tenu de ces éléments, il est clair que l'amélioration de la prise de décision financière peut être considérée comme un objectif majeur de l'innovation financière, et une partie de l'intérêt pour le robo-advising réside dans sa promesse d'aider les investisseurs dans cette direction.

2.3 Les problèmes fondamentaux des conseillers

Une réaction naturelle aux mauvaises décisions financières des investisseurs est de déléguer la tâche à des experts professionnels, qui ont le temps et les compétences nécessaires pour servir au mieux les intérêts des investisseurs. L'argument repose sur quelques hypothèses essentielles, qui peuvent parfois être difficiles à respecter dans la pratique. Tout d'abord, les conseillers doivent être capables de reconnaître et d'adapter leurs stratégies en fonction des préférences et des besoins de leurs clients. Cela est loin d'être évident et des données récentes suggèrent que les conseillers peuvent eux-mêmes avoir des croyances erronées. Foerster, Linnainmaa, Melzer et Previtero (2017) analysent les décisions de trading et de portefeuille d'environ 10 000 conseillers financiers et 800 000 clients dans quatre institutions financières canadiennes. Ils montrent que les caractéristiques visibles des clients (tolérance au risque, âge, revenu, patrimoine, profession, connaissances financières) n'expliquent conjointement que 12 % de la variation transversale de l'exposition au risque des clients. Ce chiffre est remarquablement bas, surtout si on le compare à celui de l'effet d'être servi par un conseiller donné, qui explique 22 % de la variation de l'exposition au risque des clients. En termes de pouvoir explicatif supplémentaire, l'ajout des effets du conseiller à un modèle dans lequel l'exposition au risque des investisseurs est attribuée à leurs caractéristiques observables améliore le R^2 ajusté de 12 à 30 %. Cette preuve suggère que, dans certains cas, les recommandations financières sont plus proches du modèle

“one size fits all” que de la personnalisation en fonction des préférences et des besoins spécifiques des clients. En outre, Linnainmaa, Melzer et Previtero (2020) révèlent que certains conseillers, lorsqu'ils placent leur propre argent, affichent des biais très similaires à ceux de leurs clients : ils préfèrent la gestion active, “chassent” les rendements et ne sont pas bien diversifiés.

Un deuxième élément clé est que les conseillers doivent être incités à agir dans le meilleur intérêt des clients, plutôt que de poursuivre leurs propres intérêts. Là encore, des preuves récentes suggèrent que ce n'est pas forcément le cas. Mullainathan, Noeth et Schoar (2012) ont mené une étude en formant des auditeurs, en se faisant passer pour des clients de divers conseillers financiers et en leur demandant (au hasard) de représenter différentes stratégies d'investissement et différents biais. Ils démontrent que les conseillers affichent un biais significatif en faveur de la gestion active, ils soutiennent initialement les demandes des clients, mais leurs recommandations finales sont éloignées des préférences déclarées des clients. Finalement, les recommandations des conseillers ne corrigent pas les biais des clients et même aggravent leur situation. De manière analogue, Foà, Gambacorta, Guiso et Mistrulli (2019) décrivent les comportements stratégiques des banques dans leur offre de contrats hypothécaires. Un examen plus approfondi des conseils contradictoires des conseillers est présenté dans Beshears *et al.* (2018).

Un troisième aspect fondamental est que, même en éliminant les problèmes précédents, le conseil financier est coûteux, et une partie importante du coût a une composante fixe (le temps du conseiller). Cela implique que les conseils financiers sont difficilement accessibles aux investisseurs ayant des niveaux de patrimoine plus faibles, alors même qu'ils peuvent être ceux qui en ont le plus besoin.

III. Les promesses

3.1 Procédures auditable et recommandations personnalisées

Les services des robo-advisors offrent généralement une procédure transparente et auditable permettant de répartir le portefeuille d'un individu entre diverses classes d'actifs et différents types de fonds, en fonction de ses caractéristiques individuelles. Deux étapes du processus sont cruciales à cet égard : (1) l'établissement du profil du client, et (2) la répartition des actifs. Si des recommandations sur mesure sont proposées aux clients, il existe une hétérogénéité considérable dans les allocations recommandées, et les algorithmes exacts utilisés par les robo-advisors ne sont généralement pas transparents.

3.1.1 Le profilage du client

Les robo-advisors utilisent généralement un questionnaire en ligne pour évaluer la situation financière, les caractéristiques et les objectifs d'investissement des investisseurs. Ce questionnaire est une exigence réglementaire en vertu des directives de la SEC aux États-Unis

(SEC, 2006 ; SEC, 2019). Une “évaluation du caractère adéquat” des produits proposés est également obligatoire en vertu de la réglementation MiFID (Markets in Financial Instruments Directive) en Europe.⁴

Les caractéristiques individuelles, telles que l'âge, l'état civil, la valeur nette, l'horizon d'investissement et la tolérance au risque sont utilisées pour évaluer la situation de l'investisseur. Il est intéressant de noter qu'une grande variété de questions peut être utilisée pour estimer une caractéristique particulière. Par exemple, si l'on considère la tolérance au risque, la plupart des robo-profileurs utilisent des mesures subjectives de l'aversion au risque basées sur une auto-évaluation. Certains robo-profileurs utilisent des mesures de la capacité en risque (mesure de la capacité à supporter des pertes), estimée à partir des contraintes de pertes du portefeuille, des obligations ou dépenses financières, des informations du bilan, etc. En Europe, selon la directive MiFID II, les conseillers évaluent également “l'expérience et les connaissances” des clients pour comprendre les risques liés au produit ou service d'investissement proposé.⁵ Les robo-advisors posent donc des questions sur les connaissances financières des clients et réduisent la tolérance au risque des individus lorsque les connaissances financières sont faibles.

Les robo-conseillers proposent généralement à leurs clients de choisir un objectif (par exemple, la retraite, l'achat d'une maison, un legs à des membres de la famille, le financement d'études supérieures pour les enfants ou la constitution d'un matelas de sécurité) parmi plusieurs possibilités. Cet objectif peut définir l'horizon d'investissement ou la capacité à prendre des risques, permettant ensuite de définir l'allocation optimale du portefeuille. D'autres robo-advisors permettent à leurs clients de nommer leur objectif avant ou en dehors du processus de profilage du risque, et ne l'intègrent pas nécessairement dans l'allocation du portefeuille. Enfin, quelques robo-advisors permettent à leurs clients de fixer plusieurs objectifs, offrant ainsi à leurs clients la possibilité de ségréguer leur portefeuille dans des “comptes mentaux” (Das, Markowitz, Scheid et Statman, 2010). L'une des contraintes est que les robo-advisors manquent fréquemment d'une vision globale de la situation financière d'un investisseur, car les économies réalisées en dehors de la plateforme robotique sont rarement prises en compte. Certains d'entre eux ont une vision plus large de la situation financière des clients grâce à des partenariats avec des agrégateurs de comptes financiers ou des plateformes numériques d'investissement.⁶

3.1.2 L'allocation d'actifs

Dans un deuxième temps, le robo-advisor propose de structurer un portefeuille en tenant compte des objectifs d'investissement et du niveau de risque souhaité. Beketov *et al.* (2018) ont analysé 219 robo advisors de 28 pays (30 % aux États-Unis, 20 % en Allemagne, 14 % au Royaume-Uni), qui ont été créés entre 1997 et 2017. Comme le montre la figure 1, qui représente la fréquence d'utilisation par les robo advisors des différentes méthodes, on constate une grande variété de techniques de construction de portefeuille. Beketov *et al.* (2018) montrent que la plupart des robo advisors utilisent l'optimisation simple de Markowitz ou une variante de celle-ci, comme Black Litterman (40 %), des grilles d'allocation prédéfinies (27 %) ou des poids constants (14 %). Une minorité de robo advisors utilisent des techniques alternatives de construction de portefeuille, telles que la gestion actif-passif, les techniques de “large scale optimization”, de “risk parity” ou d'assurance de portefeuille.

Figure 1 : Nuage de mots représentant la fréquence des différentes méthodes utilisées par les robo-advisors existants.



Source: Beketov et al. (2018)

Si la plupart des robo-advisors effectuent l'allocation d'actifs en utilisant une analyse moyenne-variance, ou une variante de celle-ci, ils divulguent rarement la manière dont ils choisissent leur univers d'investissement ou leur méthode d'estimation des variances et covariances entre les classes d'actifs. Il est encore plus rare qu'ils divulguent explicitement leurs hypothèses sur les rendements attendus et les risques des classes d'actifs considérées. Parmi les acteurs dominants aux États-Unis, Wealthfront fait exception. Il divulgue, sur son site Web, sa méthode d'optimisation de portefeuille (Black-Litterman), mais aussi ses rendements attendus, ses volatilités et ses

matrices de corrélation, ainsi que la manière dont ils sont estimés.⁷ Betterment est également relativement transparent. Il fournit des justifications et des détails sur le choix de son univers d'investissement, sa méthode d'optimisation de portefeuille (Black-Litterman) et la façon dont il a calculé les rendements attendus et les risques, sans les divulguer explicitement.⁸ Schwab Intelligent Portfolios divulgue également sa méthode d'optimisation de portefeuille, une variante de l'approche de Markowitz (utilisant la Value at Risk conditionnelle au lieu de la variance). Cependant, ils sont moins transparents sur leur méthodologie de simulation Monte-Carlo et leurs hypothèses de rendement attendu.⁹

3.1.3 Hétérogénéité des allocations d'actifs proposées

En théorie, ces procédures rigoureuses et leur caractère systématique devraient permettre de pallier les insuffisances des conseillers humains, en réduisant les biais involontaires et en simplifiant l'interaction avec le client. Le rebalancement, par exemple, est facilité par les plates-formes de robo-advising qui le mettent en œuvre automatiquement ou exigent une simple validation par le client. De même, si les caractéristiques individuelles sont mesurées avec suffisamment de précision, les services de robo-advising devraient permettre de proposer des recommandations d'investissement adaptées à la situation de chaque investisseur.

En pratique, une grande disparité dans les allocations d'actifs proposées a été documentée, pour le même profil d'investisseur. Par exemple, Boreiko et Massarotti (2020) analysent 53 robo-advisors opérant aux États-Unis et en Allemagne en 2019. Ils montrent qu'un profil "modéré" investit en moyenne 56 % en actions, mais l'écart type de l'exposition aux actions proposée est important (23 %). Celle-ci peut varier de 14 % à 100 %, selon le robo-advisor considéré. Les allocations d'actifs agressives ou prudentes présentent des caractéristiques similaires, avec une exposition moyenne aux actions de 73 % et 35 % respectivement, mais varient de 18 % à 100 % pour les allocations agressives, et de 0 à 100 % pour les allocations prudentes.

Cette disparité dans les allocations proposées peut avoir plusieurs sources. Elle peut provenir de différentes méthodologies de construction de portefeuille ou de différentes hypothèses

de risque/rendement attendu. Elle peut également refléter les conflits d'intérêts des robo-advisors. Boreiko et Massarotti (2020) montrent que l'expertise des gestionnaires d'actifs dans une classe d'actifs donnée (représentée par le pourcentage de fonds dans une classe d'actifs donnée dans l'univers total des fonds proposés par le robo-advisor) est le principal facteur explicatif. Des conflits d'intérêts ont également été mis en évidence dans le cas de Schwab Intelligent Portfolio, qui recommandait qu'une partie importante du portefeuille des clients soit investie dans des fonds monétaires. Cette allocation permettait à Schwab Intelligent Portfolios de déléguer la gestion des liquidités à Schwab Bank, qui profitait de la différence de taux d'intérêt entre les taux de prêt et le taux de rendement payé (Lam, 2016 ; Fisch, Laboure et Turner, 2019).

3.2 Améliorer la situation des investisseurs

Comme pour de nombreux services financiers innovants, l'une des promesses clés du robo-advice est d'améliorer la situation globale des investisseurs. Cette assertion est évidemment difficile à vérifier, car elle nécessite une bonne compréhension des préférences des investisseurs, des contraintes et des opportunités extérieures (par exemple, comment ils auraient utilisé leur capital s'ils ne l'avaient pas investi auprès du robo-advisor), ainsi qu'une image exhaustive de leur patrimoine. De plus, même si l'on peut faire des approximations raisonnables sur la manière dont les investisseurs arbitrent entre risque et rendement, les investisseurs peuvent se préoccuper d'autres éléments. Par exemple, certains investisseurs peuvent avoir recours à des conseils financiers uniquement pour acquérir une certaine tranquillité d'esprit. Gennaioli, Shleifer et Vishny (2015) proposent un modèle dans lequel un conseiller financier agit comme un "médecin de l'argent", qui permet aux investisseurs de diminuer efficacement leur réticence à prendre des risques. Rossi et Utkus (2019a) établissent que l'acquisition de la tranquillité d'esprit est l'un des principaux moteurs de la demande de conseils financiers.

La plupart des études académiques ne s'aventurent pas à développer une analyse de bien-être à part entière, elles adoptent une vision plus restreinte et vérifient si avoir accès au robot augmente les rendements des investisseurs, après avoir contrôlé certaines mesures du risque du portefeuille. Une amélioration des rendements ajustés au risque peut provenir de changements statiques dans les choix de portefeuille, par exemple en améliorant la diversification et donc en réduisant le risque pour un niveau donné de rendements attendus. Ou bien elle peut se produire au fil du temps, en permettant aux investisseurs de rebalancer leurs portefeuilles de manière à se rapprocher de leur cible correspondant à leur profil risque-rendement cible.

Des études universitaires récentes montrent que les services de robo-advising ont tendance à améliorer la diversification et la performance ajustée au risque des investisseurs. Par exemple, D'Acunto, Prabhala et Rossi (2019) analysent un optimiseur de portefeuille ciblant les actions indiennes et constatent que le robo-advising a été bénéfique aux investisseurs a priori sous-diversifiés, en augmentant la diversification de leur portefeuille, en réduisant leur risque et en augmentant leurs rendements moyens a posteriori. Cependant, le robo-advisor n'a pas amélioré la performance des investisseurs déjà diversifiés. Rossi et Utkus (2019b) étudient les effets d'un grand robo-advisor américain sur une population d'investisseurs qui géraient précédemment eux-mêmes leur épargne. Ils constatent que, tous investisseurs confondus, les utilisateurs de robo-advice ont réduit leurs investissements sur le marché

monétaire et augmenté leurs avoirs en obligations. L'introduction du robo-advice a également réduit le risque idiosyncratique du portefeuille, en diminuant les avoirs en actions individuelles et en fonds communs de placement actifs, tout en augmentant l'exposition aux fonds indiciels à faible coût. Le biais "domestique" a été également réduit, grâce à une meilleure diversification internationale. Au final, l'introduction du robot a augmenté la performance globale ajustée au risque des investisseurs. Dans un échantillon différent, Reher et Sun (2019) ont également souligné, dans le cas d'un grand robo-advisor américain, une amélioration de la diversification des portefeuilles des utilisateurs. Bianchi et Brière (2020) étudient l'introduction d'un grand robo-advisor français sur les plans d'épargne salariale. Ils constatent que, par rapport à l'autogestion, l'accès aux services du robo-advisor est associé à une hausse de l'épargne et des rendements ajustés au risque. Les investisseurs prennent plus de risques et rebalancent leurs portefeuilles de manière à maintenir leur allocation plus proche de sa cible. Cette prise de risque accrue est également constatée par Hong, Lu et Pan (2020), qui étudient un robo-advisor chinois et utilisent des données uniques par compte sur la consommation et les investissements des clients dans le cadre du groupe Ant. L'adoption du robot a aidé les ménages à se rapprocher d'une prise de risque optimale, adaptée à la volatilité de leur consommation.

3.3 Atteindre les investisseurs mal desservis

L'une des promesses les plus importantes de la révolution fintech est liée à l'inclusion financière. Comme mentionné précédemment, l'offre de services financiers implique souvent des coûts fixes importants, ce qui peut rendre non rentable le fait de servir des consommateurs plus pauvres. Les nouvelles technologies permettent une diminution spectaculaire des coûts de transaction (Goldfarb et Tucker, 2019). En réduisant ces coûts, les nouvelles technologies peuvent atteindre ceux qui ont été traditionnellement mal desservis (Philippon (2019)).

Les robo-advisors participent à cette promesse. Premièrement, ils exigent généralement un capital initial plus faible pour ouvrir un compte. Par exemple, Bank of America exige 25 000 dollars américains pour ouvrir un compte avec un conseiller financier privé, mais seulement 5 000 dollars pour ouvrir un compte avec son robo-advisor. Certains robo-advisors, tels que Betterment, n'exigent pas d'investissement minimum. Ensuite, ils facturent généralement des frais moins élevés que les conseillers humains. L'automatisation du processus de conseil réduit les coûts fixes liés au conseil. Par exemple, un robo-advisor entièrement automatisé aux États-Unis facture généralement des frais compris entre 0,25 % et 0,50 % des actifs gérés (ces frais sont compris entre 0,25 % et 0,75 % en Europe),¹⁰ alors que les frais des conseillers humains traditionnels sont rarement inférieurs à 0,75 % et peuvent même atteindre 1,5 % (Lopez, Babcic et De La Ossa, 2015 ; Better-Finance, 2020).

Les études universitaires testant l'impact du développement du robo advising sur l'inclusion financière sont rares, mais les premiers résultats semblent soutenir les affirmations précédentes. Hong et al. (2020) montrent que l'adoption d'une plateforme fintech populaire en Chine est associée à une prise de risque accrue, et l'effet est particulièrement important pour les ménages résidant dans des zones à faible couverture de services financiers. Reher et Sokolinski (2020) analysent les effets d'une réduction du seuil minimum d'investissement de 5 000 à 500 dollars par un grand robo-advisor américain. Ils montrent que, grâce à cette réduction, il y a eu une augmentation de 59 % de la part des participants de la

“classe moyenne” (dont la richesse se situe entre 1 000 et 42 000 dollars), mais aucune augmentation de la participation des ménages dont la richesse est inférieure à 1 000 dollars. La majorité des nouveaux participants sont également nouveaux sur le marché boursier et, par rapport aux participants de la classe aisée, ils augmentent leur part risquée de 13 % et leur rendement total de 1,2 % . Bianchi et Brière (2020) montrent également que les participants au robo-advice augmentent leur exposition au risque et leur rendement ajusté au risque. Fait important, l’augmentation de l’exposition au risque est plus importante pour les investisseurs ayant des portefeuilles plus petits et une exposition aux actions plus faible au départ. L’augmentation des performances est également plus importante pour les petits investisseurs et pour ceux ayant des rendements plus faibles au départ. Ces résultats suggèrent que l’accès à un robo advisor peut être particulièrement important pour les investisseurs qui sont moins susceptibles de recevoir des conseils traditionnels et, en cela, il peut être considéré comme un instrument important d’inclusion financière.

IV. Questions ouvertes

4.1 Pourquoi pas plus d’IA/Big Data?

Comme nous l’avons mentionné, la plupart des robo-advisors s’appuient actuellement sur des procédures plutôt simples, tant en ce qui concerne les informations employées pour établir le profil du client, que la manière dont ces informations sont utilisées pour construire un portefeuille optimal. Comme le soulignent Beketov *et al.* (2018), la théorie moderne du portefeuille reste prédominante et les différentes formes d’intelligence artificielle ne sont guère employées. Cela peut sembler surprenant, vu l’intérêt accru pour l’IA et le Big Data mentionné précédemment, d’autant que les robo-advisors sont souvent présentés comme intégrant ces tendances. On peut se demander pourquoi nous ne voyons pas davantage d’IA intégrée au robo-advising.

Une première raison peut être que, bien qu’une telle inclusion soit souhaitable, elle n’est pas réalisable en raison de contraintes technologiques ou de connaissances insuffisantes. En d’autres termes, la théorie financière n’a pas suffisamment progressé pour pouvoir émettre des recommandations sur la manière d’intégrer l’IA aux modèles financiers. Certains chercheurs ne sont pas d’accord. Bartram *et al.* (2020) résumant les lacunes des techniques classiques de construction de portefeuille et soulignent comment les techniques d’IA améliorent la pratique. En particulier, ils affirment que l’IA peut produire de meilleures estimations du risque et du rendement, résoudre des problèmes d’optimisation de portefeuille avec des contraintes complexes et obtenir de meilleures performances hors échantillon par rapport aux approches traditionnelles.

Une deuxième raison pourrait être que le fait d’inclure davantage d’IA violerait les contraintes réglementaires. Selon la réglementation actuelle, en tant que conseiller en investissement enregistré, un robo-advisor a une obligation fiduciaire envers ses clients. Comme l’expliquent Grealish et Kolm (2021), l’obligation fiduciaire aux États-Unis repose sur l’Advisers Act de 1940, adaptée par la SEC en 2017, afin de tenir compte des spécificités du robo-advising.

En particulier, les robo-advisors sont tenus d'obtenir suffisamment d'informations sur le client, d'utiliser des algorithmes correctement testés et contrôlés, et de divulguer pleinement les limites éventuelles des algorithmes.

Les juristes débattent pour savoir dans quelle mesure un robo-advisor peut et doit être soumis à une obligation fiduciaire. Fein (2017) soutient que les robo-advisors ne peuvent pas être pleinement considérés comme fiduciairement responsables puisqu'ils sont programmés pour servir un objectif spécifique du client, par opposition à la prise en compte de son intérêt général. En tant que tels, ils ne peuvent pas respecter la norme de diligence de l'investisseur prudent requise pour les conseillers humains. De même, Strzelczyk (2017) souligne que les robo-advisors ne peuvent pas agir en tant que responsable fiduciaire puisqu'ils ne fournissent pas d'analyse de portefeuille individualisée, mais basent plutôt leurs recommandations sur une connaissance partielle du client. En revanche, Ji (2017) affirme que les robo-advisors peuvent être capables d'exercer le devoir de loyauté envers leurs clients de manière à respecter les normes de l'Advisers Act. Dans le même esprit, Clarke (2020) affirme que l'obligation fiduciaire peut être gérée en fondant les recommandations sur la théorie de la finance et en divulguant pleinement tout conflit d'intérêts éventuel.

Une troisième raison peut être que le fait d'avoir plus d'IA dans le robo-advice n'est tout simplement pas souhaitable. L'incorporation de l'IA ferait de ces robots, au moins en partie, une boîte noire et il serait plus difficile de fournir aux investisseurs des explications claires sur les raisons pour lesquelles certaines recommandations sont données. Patel et Lincoln (2019) identifient trois sources principales de risque associées aux applications de l'IA : (1) une plus grande opacité et complexité ; (2) l'éloignement des humains de la prise de décision ; et (3) un changement des structures d'incitation (avec par exemple plus d'efforts mis dans la collecte de données). Ils examinent les implications de ces sources de risque dans plusieurs domaines, et montrent qu'elles peuvent éroder la confiance dans les services financiers, propager certains biais, voire causer un préjudice à certains groupes de clients, traités de manière injuste. Ils examinent également les implications de ces risques sur le marché, allant de la stabilité financière à la cybersécurité, en passant par les nouveaux défis réglementaires.

La complexité des algorithmes pourrait être particulièrement problématique en période de crise. Un rapport du Financial Stability Board (2017) affirme que l'utilisation croissante de l'IA dans les services financiers peut menacer la stabilité financière. L'une des raisons est que l'IA peut créer de nouvelles formes d'interconnexion entre les marchés et les institutions financières, puisque diverses institutions peuvent employer des sources de données précédemment sans rapport. En outre, l'opacité des méthodes d'apprentissage de l'IA pourrait involontairement devenir une source de risque au niveau macro. La complexité des algorithmes est également particulièrement problématique pour les personnes ayant de faibles connaissances financières. Il a été démontré que les produits financiers complexes sont particulièrement nuisibles pour les investisseurs moins avertis (voir par exemple Bianchi et Jehiel (2020) pour une étude théorique, Ryan, Trumbull et Tufano (2011) et Lerner et Tufano (2011) pour des preuves historiques, et Célérier et Vallée (2017) pour des données plus récentes). Comme pour de nombreuses innovations financières, il existe un risque qu'elles n'atteignent pas ceux qui en auraient le plus besoin, ou qu'elles finissent par être mal utilisées.

Ainsi, certaines promesses clés du robo-advising, notamment sur l'amélioration de l'inclusion financière et de la responsabilité, peuvent être menacées par l'utilisation généralisée de modèles opaques.

4.2 Jusqu'où devons-nous aller dans la personnalisation?

Le potentiel du robo-advice est de combiner la technologie financière et l'intelligence artificielle pour offrir à chaque investisseur des conseils personnalisés en fonction de ses objectifs et de ses préférences. Une difficulté importante réside dans la mesure précise des caractéristiques des investisseurs. Un deuxième problème concerne la sensibilité de l'allocation d'actifs optimale à ces caractéristiques, qui peuvent être sujettes à un grand degré d'incertitude. Cela peut conduire le portefeuille optimal estimé à être sensiblement différent du véritable portefeuille optimal, avec des conséquences dramatiques pour l'investisseur.

4.2.1 Difficulté de mesurer les caractéristiques d'un individu

Lo (2016) appelle au développement d'indices intelligents, qui pourraient être adaptés à la situation et aux caractéristiques des individus. Si nous n'en sommes pas encore là, les robo-advisors pourraient faire un pas dans cette direction, en aidant à définir précisément la situation et les objectifs financiers de chaque investisseur (Gargano et Rossi, 2020). Comme il a été démontré par un grand nombre d'articles universitaires, les choix de portefeuille optimaux dépendent de diverses caractéristiques individuelles telles que le capital humain (Cocco, Gomes et Maenhout, 2005 ; Benzoni, Collin-Dufresne et Goldstein, 2007 ; Bagliano, Fugazza et Nicodano, 2019), l'exposition au marché immobilier (Kraft et Munk, 2011), la préférence temporelle, l'aversion au risque, l'aversion à l'ambiguïté (Dimmock, Kouwenberg, Mitchell et Peijnenburg, 2016 ; Bianchi et Tallon, 2019), etc. Les possibilités d'individualisation sont beaucoup plus importantes que ce qui est actuellement mis en œuvre dans les services de robo-advice.

Cependant, certaines caractéristiques individuelles sont difficiles à mesurer et sujettes à un grand degré d'incertitude. L'aversion au risque est l'une d'entre elles. Différentes méthodes ont été développées par des économistes et des psychologues pour mesurer l'aversion au risque des individus. La plupart d'entre elles sont des mesures expérimentales basées sur des choix hypothétiques. Par exemple, les loteries de Barsky, Juster, Kimball et Shapiro (1997) offrent aux individus le choix entre un emploi avec un salaire sans risque, et un salaire plus élevé mais risqué. D'autres travaux (Holt et Laury, 2002 ; Kapteyn et Teppa, 2011 ; Weber, Weber et Nosić, 2013) mesurent les préférences en fonction d'une série de compromis risque/rendement. Le choix entre un gain certain et une loterie risquée est répété, en augmentant progressivement le gain jusqu'à ce que le sujet choisisse une loterie risquée.

L'une des raisons pour lesquelles il est difficile de mesurer l'aversion au risque pourrait être que les personnes interprètent les résultats comme des gains et les pertes relatives à un point de référence et sont plus sensibles aux pertes qu'aux gains. Kahneman, Knetsch et Thaler (1990) ou Barberis, Huang et Santos (2001) ont apporté des preuves expérimentales de l'aversion aux pertes. L'aversion aux pertes peut également expliquer pourquoi de nombreux investisseurs préfèrent les produits d'assurance de portefeuille offrant des garanties en capital (Calvet, Celerier, Sodini et Vallee, 2020).

En pratique, les robo-advisors évaluent fréquemment la tolérance au risque d'un client sur la base d'une auto-déclaration. On demande aux gens de se noter dans leur capacité à prendre des risques sur une échelle de 1 à 10 (Dohmen, Falk, Huffman, Sunde, Schupp et Wagner, 2005). Ces notations ont l'inconvénient de ne pas être très comparables entre les individus.

Les techniques de notation sont également fréquemment utilisées par les robo-advisors. Ils posent à l'individu un grand nombre de questions, couvrant différents aspects de la vie (consommation, loisirs, santé, loteries financières, travail, retraite et famille). Les scores globaux sont obtenus en additionnant les scores sur différentes dimensions, en ne conservant que les questions qui s'avèrent les plus pertinentes a posteriori pour mesurer l'aversion au risque d'un individu, un critère statistique éliminant les questions les moins pertinentes (Arrondel et Masson, 2013).

En Europe, la mise en œuvre de la réglementation MiFID a donné lieu à plusieurs études académiques évaluant les questionnaires de profilage de risque. La réglementation européenne n'impose pas de solution standardisée, chaque banque reste libre de développer son questionnaire comme elle le souhaite, ce qui explique la grande hétérogénéité des questionnaires distribués en pratique aux clients. Marinelli et Mazzoli (2010) ont envoyé, pour 100 investisseurs potentiels, trois questionnaires différents utilisés par les banques afin de vérifier la cohérence des profils de risque des clients. Seuls 23 % des individus ont été profilés de manière cohérente dans les trois questionnaires, une conséquence probable des différences dans le contenu et les méthodes de notation des questionnaires. D'autres travaux menés dans plusieurs pays européens (De Palma, Picard et Prigent, 2009 ; Marinelli et Mazzoli, 2010 ; Linciano et Soccorso, 2012) sont arrivés à la même conclusion.

4.2.2 Sensibilité des algorithmes à l'incertitude des paramètres

Les allocations optimales sont, en général, très sensibles aux paramètres (rendements attendus, covariance des rendements des actifs) qui sont difficiles à évaluer. Elles dépendent également de manière capitale des caractéristiques de l'investisseur (richesse financière, capital humain, etc.), qui sont souvent connues avec une faible précision. D'un côté, il y a un coût à imposer une allocation d'actifs unique pour tous, car elle va s'avérer sous-optimale dans de nombreux cas. Il y a donc potentiellement des gains substantiels à individualiser (voir Dahlquist, Setty et Vestman, 2018 ; Warren, 2019). Mais d'un autre côté, il existe un risque de surréaction face à des caractéristiques individuelles extrêmes ou variables dans le temps, qui peuvent conduire à des allocations d'actifs "extrêmes", comme l'ont montré les recherches sur l'optimisation de portefeuille tenant compte de l'incertitude des paramètres (voir par exemple Garlappi, Uppal et Wang, 2007). Blake, Cairns et Dowd (2009) affirment qu'une certaine standardisation est nécessaire, comme dans l'industrie aéronautique, pour garantir la sécurité des investisseurs. Le degré de personnalisation nécessaire dépend en grande partie du compromis entre les gains permettant de rapprocher le portefeuille des besoins d'un individu et le risque d'estimer les caractéristiques d'un individu avec erreur.

Le degré de stabilité des caractéristiques d'un individu dans la pratique reste également une question ouverte. Capponi, Olafsson et Zariphopoulou (2019) montrent que si ces profils de risque évoluent dans le temps (en fonction des caractéristiques idiosyncratiques, des rendements du marché ou des conditions économiques), le portefeuille dynamique optimal théorique d'un robo-advisor devrait s'adapter au profil de risque dynamique du client, en ajustant les demandes de couverture inter-temporelles correspondantes. Le robo-advisor est confronté à un trade-off entre la réception des informations du client en temps voulu et l'atténuation des biais comportementaux dans le profil de risque communiqué par le client. Avec une aversion au risque variant dans le temps, le portefeuille optimal proposé par le robo-advisor devrait contrer la tendance du client à réduire son exposition au marché pendant les récessions économiques.

4.3 Les humains peuvent-ils faire confiance aux robots ?

Dans l'interaction entre les humains et les robo-advisors, un élément clé est la confiance, qui détermine la volonté de l'individu d'utiliser le service et de suivre les recommandations du robot. Dans cette section, nous examinons ce qui crée la confiance dans les algorithmes et discutons de l'impact de la confiance sur les décisions financières.

4.3.1 La confiance est clé dans l'adoption des services de robo-advice

Il a été démontré que la confiance est un facteur clé dans les décisions financières (Sapienza, ToldraSimats et Zingales, 2013). Par exemple, les investisseurs confiants sont significativement plus susceptibles d'investir sur le marché boursier (Thakor et Merton, 2018). La confiance est également un facteur clé potentiel dans l'adoption du robo-advice. Comme l'a affirmé Merton (2017), "Ce dont vous avez besoin pour que la technologie fonctionne, c'est de créer la confiance."

La confiance a été étudiée dans diverses disciplines, notamment en sociologie, en psychologie et en économie, pour comprendre comment les humains interagissent entre eux, ou plus récemment avec les machines. La confiance est une "attitude psychologique multidimensionnelle impliquant des croyances et des attentes sur la fiabilité du responsable fiduciaire, dérivées de l'expérience et des interactions avec ce dernier dans des situations d'incertitude et de risque" (Abbass, Scholz et Reid, 2018). On peut également voir la confiance comme une transaction entre deux parties : si A croit que B agira dans le meilleur intérêt de A, et accepte sa vulnérabilité aux actions de B, alors A fait confiance à B (Misztal, 2013). Il est important de noter que la confiance existe pour atténuer l'incertitude et le risque de collaboration en permettant au bénéficiaire d'anticiper que le responsable fiduciaire agira dans son meilleur intérêt.

Bien que la confiance ait des caractéristiques à la fois cognitives et affectives, dans la recherche sur l'automatisation, les processus cognitifs (plutôt qu'affectifs) semblent jouer un rôle dominant. La confiance dans les robots présente de multiples facettes. Il a été démontré qu'elle dépend de la fiabilité, de la robustesse, de la prévisibilité, de la clarté, de la transparence et de la responsabilité fiduciaire des robots (Sheridan, 1989 ; Sheridan, 2019 ; Muir et Moray, 1996). L'une des principales caractéristiques des robo-advisors est leur recours à des algorithmes plus ou moins compliqués, dans plusieurs étapes du processus de conseil. Comme nous l'avons vu, un algorithme est utilisé pour établir le profil de l'investisseur, puis pour définir l'allocation d'actifs optimale. Le client qui délègue la décision au robot supporte le risque qu'une mauvaise décision du robot entraîne une mauvaise performance de son épargne. La confiance dans ces algorithmes est donc essentielle dans l'adoption des robo-advisors.

4.3.2 Aversion pour les algorithmes

Les résultats d'une enquête (HSBC, 2019) montrent qu'il existe un manque général de confiance dans les algorithmes. Alors que la plupart des gens semblent faire confiance à leur environnement et à la technologie (68 % des personnes interrogées ont déclaré qu'elles feraient confiance à une personne jusqu'à preuve du contraire, 48 % pensent que

la majorité des gens sont dignes de confiance et 76 % se sentent à l'aise pour utiliser les nouvelles technologies), l'intelligence artificielle ne suscite pas encore la confiance. Seuls 8 % des répondants feraient confiance à un robot programmé par des experts pour offrir des conseils en matière de prêts hypothécaires, contre 41 % qui feraient confiance à un courtier en prêts hypothécaires. À titre de comparaison, 9 % seraient susceptibles d'utiliser un horoscope pour guider leurs choix d'investissement ! 14 % des personnes interrogées feraient confiance à un robot programmé par des chirurgiens de renom pour leur faire subir une opération à cœur ouvert, tandis que 9 % feraient confiance à un membre de leur famille pour effectuer une opération assistée par un chirurgien. Seuls 19 % ont déclaré qu'ils feraient confiance à un robo-advisor pour les aider à faire des choix d'investissement. Il existe toutefois de grandes différences entre les pays. Le pourcentage de personnes interrogées qui font confiance aux robo-advisors atteint 44 % et 39 % en Chine et en Inde respectivement, mais il n'est que de 9 % et 6 % en France et en Allemagne.

Certaines études universitaires ont montré que les décideurs sont souvent peu enclins à utiliser des algorithmes, préférant la plupart du temps un jugement humain moins précis. Par exemple, il a été démontré que les prévisionnistes professionnels n'utilisent pas les algorithmes ou leur accordent un intérêt trop limité (Fildes et Goodwin, 2007). Dietvorst, Simmons et Massey (2015) ont donné aux participants d'une expérience incitative le choix d'utiliser exclusivement les prévisions d'un algorithme ou d'utiliser exclusivement leurs propres prévisions au cours d'une tâche de prévision. Ils ont constaté que la plupart des participants ont choisi d'utiliser exclusivement l'algorithme lorsqu'ils n'avaient aucune information sur les performances de l'algorithme. Cependant, lorsque l'expérimentateur leur indiquait que l'algorithme était imparfait, ils étaient beaucoup plus enclins à choisir la prévision humaine. Cet effet a persisté même lorsqu'ils avaient explicitement vu l'algorithme surpasser les prévisions de l'humain. Cette tendance à ignorer de manière irrationnelle les conseils générés et communiqués par des algorithmes informatiques a été appelée "aversion pour les algorithmes". Dans une étude expérimentale ultérieure (Dietvorst, Simmons et Massey, 2018), les participants ont eu la possibilité de modifier l'algorithme. Les participants étaient considérablement plus enclins à choisir l'algorithme imparfait lorsqu'ils pouvaient modifier ses prévisions, même s'ils étaient fortement contraints dans les modifications qu'ils pouvaient effectuer. Cela suggère que l'aversion pour les algorithmes peut être réduite en donnant aux gens un certain contrôle sur les prévisions d'un algorithme imparfait.

Des preuves expérimentales récentes ont mis en évidence une moindre aversion pour les algorithmes. Niszczoła et Kaszas (2020) ont testé si les gens présentaient une aversion pour les algorithmes lorsqu'on leur demandait s'ils préféraient utiliser un conseil humain ou un réseau de neurones artificiel pour prédire l'évolution du prix des actions. Sans aucune information au préalable sur les performances de l'humain par rapport à celles du robot, ils n'ont constaté aucune aversion générale envers les algorithmes. Lorsqu'il a été précisé que les performances du conseiller humain étaient similaires à celles de l'algorithme, 57 % des participants ont montré une préférence pour le conseil humain. Dans une autre expérience, on a demandé aux sujets de choisir entre un conseiller humain et un robo-advisor pour exclure des actions "controversées". Cette tâche nécessitait un jugement subjectif, lié à la moralité, et il est intéressant de noter que dans ce cas, les personnes ont perçu les algorithmes comme étant moins efficaces que les humains.

Germann et Merkle (2019) n'ont également trouvé aucune preuve d'aversion pour les algorithmes. Dans une expérience en laboratoire (basée principalement sur des étudiants en commerce ou en économie), ils ont demandé aux participants de choisir entre un gestionnaire

de fonds humain et un algorithme d'investissement. Le processus de sélection a été répété dix fois, ce qui leur a permis d'étudier la réaction aux performances du conseiller. À frais identiques pour les deux conseillers, 56 % des participants ont décidé de suivre l'algorithme. Lorsque les frais étaient différents, la plupart des participants (80 %) ont choisi le conseiller dont les frais étaient les plus bas. Les choix ont été fortement influencés par les performances passées cumulées. Mais les investisseurs n'ont pas perdu plus rapidement confiance dans l'algorithme après avoir constaté des erreurs de prévision. Une enquête supplémentaire a fourni des explications qualitatives intéressantes aux résultats. Les participants ont expliqué qu'ils croyaient en la capacité de l'algorithme à être plus à même d'apprendre que les humains, tandis que les humains ont, selon eux, un avantage comparatif dans l'utilisation de données qualitatives et le traitement des valeurs aberrantes. Dans l'ensemble, les algorithmes sont considérés comme un complément plutôt que comme un concurrent du conseiller humain.

4.3.3 Qu'est-ce qui engendre la confiance dans un algorithme ?

Jacovi, Marasović, Miller et Goldberg (2020) distinguent deux types de confiance dans les algorithmes : intrinsèque et extrinsèque. La confiance intrinsèque peut être acquise lorsque le processus de décision observable de l'algorithme correspond aux attentes de l'utilisateur. Les explications du processus de décision derrière l'algorithme peuvent aider à créer une confiance intrinsèque.¹¹ En outre, un algorithme peut devenir digne de confiance de par son comportement réel : dans ce cas, la source de confiance (extrinsèque) n'est pas le processus de décision du modèle, mais l'évaluation de son résultat.

La Commission européenne (2019) a récemment énuméré un certain nombre d'exigences pour les algorithmes dignes de confiance. Les exigences liées à la confiance intrinsèque sont les suivantes : (1) la possibilité de surveillance humaine, (2) la confidentialité et la gouvernance des données, (3) la transparence et la capacité à expliquer l'algorithme. La confiance extrinsèque peut être renforcée par (4) la robustesse technologique et la sécurité de l'algorithme, (5) la capacité d'interpréter ses résultats, (6) sa responsabilité et son audibilité. En outre, les considérations d'éthique et d'équité, telles que (7) éviter la discrimination, promouvoir la diversité et l'équité ou (8) encourager le bien-être sociétal et environnemental, sont également considérées comme un élément clé de la confiance.

La confiance dans les algorithmes dépend aussi de manière capitale de la perception de l'expertise et de la fiabilité des humains ou des institutions offrant le service (Prahl et Van Swol, 2017). "La technologie ne crée pas la confiance par elle-même" (Merton, 2017). Les gens font confiance aux humains qui certifient une technologie, pas nécessairement à la technologie elle-même. Dans le cas spécifique du robo-advice, Lourenco, Dellaert et Donkers (2020) étudient la décision des consommateurs d'adopter un service robotisé et montrent que cette décision est clairement influencée par l'orientation à but lucratif ou non de l'entreprise qui propose le service (par exemple, une entreprise privée d'assurance et de gestion des investissements contre un fonds de pension ou une institution parrainée par l'État). La transparence, l'explicabilité et l'interprétabilité peuvent ne pas être suffisantes en soi pour améliorer les décisions et accroître la confiance. Cependant, l'information sur les hypothèses clés et les défauts potentiels des algorithmes lors de la prise de certaines décisions peut être une dimension fondamentale à développer.

4.3.4 La confiance envers les robots et les décisions financières

Tout le monde ne fait pas confiance aux robo-advisors. Dans un échantillon de 34 000 épargnants dans des plans d'épargne salariale français, Bianchi et Brière (2020) constatent que les individus qui sont plus jeunes, de sexe masculin, et plus attentifs à leur plan d'épargne (mesuré par le temps passé sur le site web du plan d'épargne), ont une probabilité plus élevée d'adopter un service de robo-advice. La probabilité d'adopter le service robotisé est également liée négativement à la taille du portefeuille des investisseurs, ce qui suggère que le robo-advice est capable d'atteindre les investisseurs moins riches,¹² un résultat également confirmé par Brenner et Meyll (2020). Les investisseurs ayant des portefeuilles plus petits sont également plus susceptibles d'attribuer une part plus importante de leurs actifs au robot.

Une caractéristique unique du service de robo-advice étudié par Bianchi et Brière (2020) leur permet d'analyser à la fois les utilisateurs du robot et les individus "curieux", qui ont passé les premières étapes liées à la souscription du robot et ont observé sa recommandation sans finalement souscrire au service. Ils montrent que plus la recommandation du robot est éloignée de l'allocation actuelle de l'individu, et plus la probabilité que l'investisseur souscrive au robot est élevée. Ce résultat est très différent de ce que nous observons avec les conseillers humains. Ceux-ci ont tendance à gagner la confiance de leurs clients en se montrant conciliants, c'est-à-dire en leur proposant une allocation proche de ce que les clients ont déjà en portefeuille (Mullainathan *et al.*, 2012). En outre, les investisseurs plus jeunes, les femmes, ceux qui ont une plus grande exposition au risque et des rendements passés plus faibles, sont plus susceptibles d'accepter une augmentation plus importante de leur exposition aux actifs risqués, tels que les actions.

La confiance peut avoir un impact important sur les décisions des investisseurs. Bianchi et Brière (2020) et Hong *et al.* (2020) mettent en évidence une prise de risque accrue après souscription au service robotisé, un résultat cohérent avec une confiance accrue. Par exemple, Bianchi et Brière (2020) constatent une augmentation de 7 % de l'exposition aux actions après l'adoption du robo-advice (par rapport à une exposition moyenne de 16 %). Hong *et al.* (2020) relèvent une augmentation de 14 % (par rapport à une exposition moyenne aux actifs risqués de 37 % sur leur échantillon de 50 000 consommateurs chinois clients de Alibaba). Il est intéressant de noter que Hong *et al.* (2020) montrent, en outre, que ce résultat n'est probablement pas dû à une augmentation de la tolérance au risque de l'individu induite par le soutien du robot. Il semble plutôt refléter un meilleur alignement du portefeuille d'investissement avec la tolérance au risque réelle de l'individu. En particulier, ils démontrent qu'après l'adoption du robot, l'exposition aux actifs risqués est plus conforme à la tolérance au risque de l'individu estimée à partir de la volatilité de la croissance de sa consommation (Merton, 1971), mesurée à partir de la plateforme d'achat en ligne Taobao d'Alibaba. Le robo-advisor semble aider les individus à se rapprocher de l'alignement optimal entre leur prise de risque et leur consommation. Ces résultats doivent toutefois être considérés avec précaution, car les deux études se concentrent sur une période d'investissement relativement courte (en l'absence de tout krach boursier grave) et ne disposent pas d'une vision globale des portefeuilles des individus. Des travaux supplémentaires seraient nécessaires pour déterminer l'impact global à plus long terme.

4.4 Les robots remplacent-ils ou complètent-ils la prise de décision humaine ?

Les systèmes autonomes sont en train de se développer dans de vastes domaines de notre vie quotidienne. Comprendre la manière dont les humains vont interagir avec eux est une question primordiale. En particulier, devons-nous nous attendre à ce que les robots se substituent aux humains ou plutôt à ce qu'ils soient complémentaires ? Dans le cas particulier du conseil financier, les robots sont-ils susceptibles de remplacer les conseillers humains ?

À l'aide d'un échantillon représentatif d'investisseurs américains, Brenner et Meyll (2020) cherchent à savoir si les robo-advisors réduisent la demande des investisseurs pour les conseils financiers humains offerts par les prestataires de services financiers. Ils constatent un important effet de substitution et montrent que cet effet est dû aux investisseurs, qui craignent d'être victimes de fraudes en matière d'investissement ou qui s'inquiètent de conflits d'intérêts potentiels. Dans la pratique cependant, un certain nombre de plateformes, qui étaient entièrement numériques ont décidé de réintroduire des conseillers humains. Par exemple, Scalable Capital, la société européenne de robo-advice en ligne soutenue par BlackRock, ou Nutmeg, ont réintroduit les consultations par téléphone et en face à face après avoir constaté qu'un certain nombre de clients préféraient parler à des conseillers humains plutôt que de répondre uniquement à des questionnaires en ligne.

Une autre question relative à ce sujet est de comprendre de quelle manière les gens vont interagir avec les robots. Délègueront-ils toute la décision au robot ou le surveilleront-ils, pour observer le processus et intervenir si nécessaire ? Dans certaines expériences, les utilisateurs ont accordé trop de confiance aux robots. Robinette, Li, Allen, Howard et Wagner (2016) ont conçu une expérience dans laquelle les participants devaient choisir de suivre, ou non, les instructions d'un robot dans un processus d'évacuation d'urgence. Tous les participants ont suivi le robot pendant la situation d'urgence, même si la moitié des participants ont observé que le même robot avait obtenu de mauvais résultats dans une tâche de guidage de navigation non urgente quelques minutes auparavant. Même lorsque le robot pointait vers une pièce sombre sans issue discernable, la majorité des personnes ont suivi le robot plutôt que de sortir en toute sécurité par le chemin par lequel elles étaient entrées. Andersen, Koslich, Pedersen, Weigelin et Jensen (2017) ont développé ces travaux et montrent que cette confiance excessive peut également affecter les interactions homme/robot en dehors des situations d'urgence.

Dans le contexte des décisions financières, Bianchi et Brière (2020) montrent que l'adoption d'un robot-conseiller entraîne une augmentation significative de l'attention portée aux plans d'épargne, dans les mois qui suivent l'adoption. Les individus sont en général plus attentifs à leur plan d'épargne, en particulier lorsqu'ils reçoivent une rémunération variable et doivent prendre une décision d'investissement. Cela semble indiquer que les individus n'utilisent pas le robot comme un substitut à leur propre attention.

V. La prochaine génération de robo-advisors

Il n'est pas facile de savoir à quelle génération de robo-advisors nous sommes actuellement confrontés. Beketov *et al.* (2018) se concentrent sur les robots de troisième et quatrième génération, qui diffèrent des générations précédentes car ils utilisent davantage d'automatisation et des méthodes plus sophistiquées pour construire et rebalancer les portefeuilles. Une possibilité est que la prochaine génération de robots poursuive la tendance actuelle à utiliser davantage de données et des modèles plus complexes. On peut toutefois imaginer une autre voie. Comme nous l'avons vu précédemment, l'intégration de plus d'IA dans le conseil robotisé (et plus généralement dans les services financiers) se heurte à trois défis majeurs. Premièrement, si les allocations d'actifs hautement personnalisées ont le potentiel de répondre aux besoins d'un individu, elles sont également plus exposées aux erreurs de mesure relatives aux caractéristiques individuelles. Deuxièmement, dans la mesure où l'augmentation de l'IA est associée à une plus grande opacité, le risque est de manquer certaines promesses clés d'une plus grande responsabilité et d'une meilleure inclusion financière. Troisièmement, la confiance est essentielle pour l'adoption de la technologie, et encore plus dans le domaine du conseil financier. Ces défis, à notre avis, exigent des algorithmes qui peuvent être facilement interprétés et évalués. Toreini, Aitken, Coopamootoo, Elliott, Zelaya et van Moorsel (2020) expliquent comment le développement de la confiance dans les technologies (d'apprentissage automatique) exige qu'elles soient justes, explicables, responsables et sûres (FEAS en anglais).

De ce point de vue, les récentes avancées dans ce qu'on appelle le XAI (explainable artificial intelligence), peuvent être particulièrement utiles pour réfléchir à l'avenir des robo-advisors. "L'explicabilité" fait référence à la possibilité d'expliquer une prédiction ou une recommandation donnée, même si elle est basée sur un modèle très compliqué, par exemple en évaluant la sensibilité de la prédiction lors de la modification de l'une des entrées. Elle fait également référence à la possibilité d'expliquer le modèle lui-même. Les explications peuvent aider les humains à effectuer une tâche particulière et, en même temps, à évaluer un modèle donné (voir par exemple Biran et Cotton (2017) pour une étude récente). Comme mentionné par Doshi-Velez et Kim (2017), l'explicabilité peut être considérée à la fois comme une fin en soi, notamment vis-à-vis des questions de confiance et de responsabilité exprimées ci-dessus, mais également comme un outil pour évaluer si d'autres objectifs, tels que l'équité, la confidentialité, la fiabilité, la robustesse, la causalité, l'utilisabilité, sont respectés.

Il existe un grand nombre de publications universitaires examinant si l'intelligence artificielle explicable peut améliorer la prise de décision humaine. La question de savoir dans quelle mesure l'explicabilité est nécessaire pour le fonctionnement réel d'un système automatisé reste ouverte et est souvent débattue dans le contexte des voitures à conduite autonome, par exemple. D'une part, les recherches psychologiques sur la prise de décision suggèrent que, lorsque les décisions impliquent un raisonnement complexe, le fait d'ignorer une partie des informations disponibles et d'utiliser des heuristiques peut aider à prendre des décisions en incertitude de manière plus robuste que de s'appuyer sur des stratégies exigeant une forte intensité de ressources (Gigerenzer et Brighton, 2009). D'autre part, des études expérimentales montrent qu'il est important de fournir au conducteur des informations sur le pourquoi et le comment du comportement d'un véhicule autonome, afin de maintenir une

expérience de conduite sûre (Koo, Kwac, Ju, Steinert, Leifer et Nass, 2015). Ces informations sont particulièrement essentielles dans les situations d'urgence. Les conducteurs qui reçoivent de telles informations ont tendance à moins faire confiance à la voiture et sont plus rapides à prendre le contrôle de la voiture lorsqu'une situation dangereuse se produit (Helldin, Falkman, Riveiro et Davidsson, 2013). Il faut également être particulièrement attentif au risque de surcharge d'informations. Un algorithme est plus facile à interpréter et à utiliser lorsqu'il se concentre sur quelques caractéristiques, il est également plus facile à corriger en cas d'erreur (Poursabzi-Sangdeh, Goldstein, Hofman, Vaughan et Wallach, 2018).

Dans le contexte des robo-advisors, l'explicabilité n'est pas une tâche aisée. L'évaluation de la performance d'une recommandation robotisée n'est pas simple, surtout si l'on utilise l'IA pour évoluer vers des allocations entièrement personnalisées, qui doivent elles-mêmes être évaluées en fonction de benchmarks entièrement personnalisés (comme décrit dans Lo (2016)). Il devient alors plus difficile pour le client de construire des modèles contre-factuels de performance. Et probablement encore plus difficile est d'apprécier le modèle financier sous-jacent qui régit l'algorithme, surtout si l'on souhaite répondre aux besoins d'investisseurs moins expérimentés.

À cet égard, il ne s'agit pas de rechercher une transparence totale de l'algorithme potentiellement compliquée qui sous-tend le processus de conseil personnalisé, en divulguant, par exemple, tous les détails de la méthodologie d'optimisation du portefeuille. Il serait probablement plus efficace de divulguer, par exemple, les scénarios économiques susceptibles de rendre l'algorithme moins performant, et donc a posteriori sous-optimal, et d'informer les clients des limites potentielles de l'algorithme.

Une autre évolution potentiellement intéressante consisterait à renforcer les interactions avec les clients. Par exemple, certains robo-advisors envoient des alertes lorsque le portefeuille d'un client s'écarte de manière significative de l'allocation d'actifs cible (voir par exemple Bianchi et Brière (2020)). Ces alertes peuvent également être considérées comme une occasion d'interagir avec le client. Par exemple, les alertes pourraient être utilisées pour expliquer pourquoi une déviation s'est produite (mouvements du marché, changement des caractéristiques personnelles, etc.) ou pourquoi un rebalancement est recommandé. Un autre développement possible serait de solliciter les perceptions des clients concernant la qualité de la réponse fournie par l'algorithme et intégrer ce retour d'information dans le cadre de l'évaluation du service de robo-advice (Dupont, 2020).

Ces questions ne sont pas nouvelles dans le domaine de l'IA. Biran et Cotton (2017) évoquent les approches antérieures de l'explicabilité des décisions dans les systèmes experts des années 1970 et plus récemment dans les systèmes de recommandation. On peut toutefois affirmer qu'aujourd'hui, les modèles sont probablement plus complexes, plus autonomes et qu'ils couvrent un plus grand nombre de décisions à travers un ensemble plus large d'agents (y compris peut-être des moins sophistiqués), ce qui rend ces questions particulièrement pertinentes dans les débats actuels. En effet, l'amélioration de la transparence est également au cœur du débat politique, comme le récent règlement européen sur la protection des données (RGPD). Comme évoqué dans Goodman et Flaxman (2017), la loi définit un droit à l'explication, par lequel les utilisateurs peuvent s'enquérir de la logique impliquée dans une décision algorithmique les concernant (par exemple, par le biais du profilage), et cela demande des algorithmes aussi explicables qu'efficaces.

Certains éminents spécialistes affirment que la révolution de l'IA n'a pas encore eu lieu. Au lieu d'imiter la pensée ou les interactions humaines sophistiquées, la révolution de l'IA se produira lorsque de nouvelles formes d'intelligence seront envisagées (Jordan (2019a)). Dans ce contexte, l'importation d'idées issues des sciences sociales semble capitale. L'IA a besoin de la psychologie pour saisir la façon dont les humains pensent et se comportent réellement, ou, pour faire référence à Lo (2019), pour inclure des formes de "stupidité artificielle". Les connaissances issues de la philosophie, de la psychologie et des sciences cognitives sont également essentielles pour déterminer comment les explications sont et devraient être communiquées. Miller (2019) examine la grande quantité de travaux dans ces domaines et souligne l'importance de fournir des explications sélectives, basées sur des relations causales et contrefactuelles plutôt que sur des relations statistiques probables. Il insiste sur l'importance de maintenir une dimension sociale, dans laquelle ceux qui reçoivent et ceux qui fournissent l'explication peuvent interagir. L'IA a également besoin de l'économie non seulement pour aider à aborder la causalité et à discuter des contrefactuels, mais aussi pour aider à concevoir de nouvelles formes d'intelligence collective. Ces nouvelles formes peuvent aller au-delà d'une approche purement anthropocentrique et s'appuyer sur une certaine compréhension du fonctionnement des marchés et de leurs éventuelles déficiences (Jordan (2019b)). Nous partageons l'enthousiasme de ces chercheurs lorsqu'ils imaginent des avancées dans ces directions, et nous avons hâte de voir plus de sciences sociales dans la prochaine génération de robo-advisors !

NOTES

- ¹ Les robo-advisors sont déjà présents en Chine, en Inde, au Japon, à Singapour, en Thaïlande et au Vietnam.
- ² Voir par exemple The Economist (2015) : “The Fintech Revolution”, ou The World Economic Forum (2017): “Beyond FinTech: A pragmatic assessment of disruptive potential in financial service.”
- ³ Les mots-clés comprenaient : l’intelligence artificielle, machine learning, l’analyse des clusters, l’algorithme génétique ou l’algorithme évolutionniste, le lasso, le traitement du langage naturel, le réseau neuronal ou le deep learning, la forêt aléatoire ou les arbres décisionnels, et les machines à vecteurs support.
- ⁴ En Europe, la réglementation MiFID a fixé comme objectif une protection accrue et harmonisée des investisseurs individuels, en fonction de leur niveau de connaissances financières. MiFID I (2004/39 / 3C), mise en œuvre en novembre 2007, impose aux sociétés d’investissement d’envoyer à leurs clients un questionnaire pour déterminer leur niveau de connaissances financières, leur patrimoine et leurs objectifs d’investissement. MiFID I a été remplacée en janvier 2018 par MiFID II (2014/65 / UE), qui a exigé un renforcement de la législation dans plusieurs domaines, notamment dans les exigences d’indépendance et de transparence du conseil (sur les coûts, l’offre disponible, etc.).
- ⁵ Voir l’article 25, paragraphe 3, et l’article 56.
- ⁶ Par exemple, Wealthfront a récemment développé une intégration avec des plateformes d’investissement numériques (Venmo, Redfin, Coinbase), de prêt (Lending Club) et de calcul d’impôts (turbotax).
- ⁷ <https://research.wealthfront.com/whitepapers/investment-methodology/>
- ⁸ En ce qui concerne l’univers d’investissement, ils ont exclu les classes d’actifs telles que le capital-investissement, les matières premières et les ressources naturelles, car “les estimations de leur capitalisation boursière ne sont pas fiables et il y a un manque de données pour étayer leurs performances historiques”. Les rendements attendus sont dérivés des pondérations du marché, par le biais d’un exercice classique d’optimisation inverse, qui utilise la matrice de variance-covariance entre toutes les classes d’actifs. Une estimation de cette matrice est effectuée à l’aide de données historiques, combinée à une matrice cible, et en utilisant la méthode de “shrinkage” de Ledoit et Wolf (2004) pour réduire l’erreur d’estimation. Les portefeuilles peuvent également être orientés vers les facteurs “taille” et “valeur” de Fama et French (1992), l’importance de l’orientation étant librement paramétrée par la confiance que Betterment accorde à ces vues. Voir <https://www.betterment.com/resources/bettermentportfolio-strategy/citations>
- ⁹ Ils simulent 10 000 réalisations futures hypothétiques de rendements, en utilisant des hypothèses de distribution à queue épaisse pour les rendements des actifs, permettant également des corrélations dynamiques modélisées avec une approche Copula. Voir <https://intelligent.schwab.com/page/our-approach-to-portfolio-construction>.
- ¹⁰ Nous ne tenons compte ici que des frais de gestion, et non des frais des ETFs ou des fonds sous-jacents.
- ¹¹ Par exemple, un robo-advisor peut divulguer sa méthodologie de profilage des risques, sa méthode d’optimisation et ses hypothèses de risque/rendement, ou révéler les signaux conduisant au rebalancement du portefeuille.
- ¹² À l’inverse, les investisseurs les plus riches sont plus susceptibles d’acquérir des informations sur le robot sans s’abonner au service.

Robo-Advising: Less AI and More XAI?

Augmenting algorithms with humans-in-the-loop*

Milo Bianchi

Toulouse School of Economics, TSM, and IUF, University of Toulouse Capitole

Marie Brière

Amundi, Paris Dauphine University, and Université Libre de Bruxelles

Abstract

We start by considering some of the key reasons behind the academic and industry interest in robo-advisors. We discuss how robo-advice could potentially address some fundamental problems in investors' decision-making, as well as in traditional financial advice by promoting financial inclusion, providing personalised recommendations based on accountable procedures and, ultimately, by aiming to make investors better off. We then move on to some of the ongoing issues regarding the future of robo-advice. Firstly, the role Artificial Intelligence plays, and should play, in robo-advice. Secondly, how far should the personalization of robo-advice recommendations go. Third, how trust in automated financial advice can be generated and maintained. Fourth, whether robots are perceived as complements or substitutes to human decision-making. Our conclusion outlines some thoughts on what the next generation of robo-advisors might look like. We highlight the importance of recent insights in Explainable Artificial Intelligence (AI) and how new forms of AI applied to financial services would benefit from importing insights from economics and psychology to design effective human/robot interaction.

* This chapter was prepared for the book *Machine Learning in Financial Markets: A Guide to Contemporary Practice* edited by Agostino Capponi and Charles-Albert Lehalle for Cambridge University Press. Milo Bianchi acknowledges funding from LTI@Unio, TSE Sustainable Finance Center and ANR (ANR-17-EURE-0010 grant). We have no material interest that relates to the research described in the paper.

The study reflects the personal views of its authors and does not necessarily express the position of the Institute Louis Bachelier and the Louis Bachelier Finance and Sustainable Growth Laboratory of Excellence.

I. Introduction

Automated portfolio managers, commonly known as robo-advisors, are attracting a growing interest both in academia and across the investment industry. In this article, we aim to review some of the reasons behind this growing interest. We emphasise how robo-advice can be seen in the broader context of the so-called Fintech revolution. We also consider some more specific reasons of interest in automated financial advice, building on the fundamental problems that individual investors face in taking financial decisions, and on the limits often observed in traditional financial advice.

We then discuss how robo-advising could potentially address these fundamental problems and highlight robots' main promises. First, promote financial inclusion by reaching underserved investors; second, provide personalized recommendations based on accountable procedures, and, finally, make investors better off. For each of these promises, we revisit the reasons why some hope can be placed on robots and we take a stand on what the academic literature has shown so far.

In the third part of the article, we address what we believe are fundamental ongoing issues in the future of robo-advice. Firstly, we discuss what role Artificial Intelligence (AI) plays, and should play, in robo-advice. We stress the constraints, in terms of the regulatory challenges, as well as the conceptual advances of portfolio theory that may limit how much AI can be placed into robo-advice. We also emphasise how the quest for simplicity and the ease of explanation in recommendations could make AI undesirable even if feasible. Secondly, we discuss how far we should go into the personalization of robo-recommendations, highlighting the trade-off between aiming to bring a portfolio closer to a specific individual's needs and the risks related to possible measurement errors of relevant individual characteristics (say, risk aversion) and the sensitivity of algorithms to parameter uncertainty. Third, we discuss how robo-advice can shed light on the broader issues of human/robot interactions and on the mechanics of trust in automated financial services. We revisit the arguments of algorithm aversion, and the possible ways to reduce it, and how these can be applied in the context of automated financial advice. Finally, we discuss some evidence of whether robots are perceived as complements or substitutes to human decision-making.

We conclude with our thoughts on what the next generation of robo-advisors may look like. Rather than continuing the trend of using more data, more complex models and more automated interactions, we define an alternative path that builds on the key premise of robo-advice in terms of increased accountability and financial inclusion, and on the key challenge of developing trust in financial technology. We highlight the importance of recent insights on XAI (Explainable Artificial Intelligence) and stress how new forms of AI applied to financial services can benefit from importing insights from social sciences, such as economics and psychology. This review does not aim to be comprehensive. Rather, it should be seen as complementary to existing reviews (such as D'Acunto and Rossi, 2020) and to the other chapters in this book.

II. Why So Popular?

Robo-advisors use automated procedures, ranging from relatively simple algorithms that use limited information on the client, to artificial intelligence systems built around big data, with the purpose of recommending how to allocate funds across different types of assets. First, a client profiling technique is used to assess an investor's characteristics (risk aversion, financial knowledge, investment horizon...) and goals. Second, an investment universe is defined and, third, a portfolio is proposed by taking into account the investment goals and desired risk level. As documented in Beketov, Lehmann and Wittke (2018), in most cases, the optimal portfolio builds on modern portfolio theory, dating back to Markowitz (1952). In addition to recommending an initial allocation of funds, algorithms can be designed to continuously monitor portfolios and detect deviations from the targeted profile. Whenever deviations are identified, the client is alerted and/or the portfolio is automatically rebalanced. The portfolio can also be automatically rebalanced to reduce risk as time goes by or when the investor changes their risk tolerance or investment goals. Some robots also propose to implement "tax harvesting" techniques: selling assets that experience a loss and using the proceeds to buy assets with similar risk, to decrease capital gains and taxable income without affecting the portfolio's exposure to risk. Apart from the portfolio allocation, the robot can display statistics of interest to the client, such as the expected annual return and volatility, often by using historic performance and Monte Carlo simulations of the possible future outcome of the portfolio allocation.

The market is growing rapidly. Most practitioners estimate that the global market is currently around \$400-500bn, as compared to \$100bn in 2016 (S&P Global Market Intelligence, Backend Benchmarking, Aite Group - see Buisson (2019)). Worldwide Assets under management in the robo-advice segment worldwide are projected to reach between \$1.7trn and \$4.6trn in 2022 (Statista, BI Intelligence). The number of users is expected to amount to 436m by 2024 (Statista 2020). This growth is driven by the entry of large incumbents in the digital service arena (for example, JPMorgan and Goldman Sachs announced the launch of a digital wealth management services in 2020) and the migration of assets managed by large financial institutions to their robo-advice operations, which amounts to 8% of their AUM and to one-quarter of the assets in accounts with less than \$1m. At the same time, clients have increased their demand for digital investment tools, and particularly for low-cost portfolio management and adjacent services such as financial planning. The United States remains, by far, the leading market for robo-advising (with more than 200 robo-advisors registered), but the number of robo-advisors is growing rapidly in Europe (more than 70), and also in Asia, driven by an emerging middle class and high technological connectivity (Abraham, Schumkler and Tessada (2019)).¹

We refer to Grealish and Kolm (2021) for more details on the functioning of robo-advisors and on recent market trends, as well as highlighting a few reasons which may be motivating such a rapid market growth and increased interest in academic and policy circles.

2.1 Fintech Révolution

Part of the interest in robo-advising comes from the broader trend of applying new technologies and novel sources of data in the financial domain, a phenomenon often dubbed as fintech. The word has played a central role in many academic and policy debates in the past few years. Enthusiasts about fintech talk about a revolution that promises to disrupt and reshape the financial service industry.²

Buchanan (2019) discusses the global growth of the AI industry and its application to the finance industry. Quoting a 2017 report, she mentions that 5,154 AI startups have been established globally during the past five years, representing a 175% increase relative to the previous 12 years. This impressive growth has been driven by the advances in computing power, leading to a decline in the cost of processing and storing data, while at the same time by the availability of data of increased size and scope. Similarly, AI related patent publications (denoted by the AI keyword) in the US have grown from around 50 in 2013 to around 120 in 2017. In China, such growth has been even more dramatic, with around 120 patents in 2013 rising to 640 patents in 2018. Buchanan (2019) also discusses the broad range of ways in which AI is changing the financial services industry, not only in terms of robo-advising but also for fraud detection and compliance, chatbots, and algo-trading.

In academic circles, the increased attention can be seen for example from the exponential growth in finance academic studies primarily centered around AI. Bartram, Branke and Motahari (2020) analyze the number of AI-related keywords in the title, abstract, or listed keywords of all working papers posted in the Financial Economics Network (FEN) between 1996 and 2018.³ In 1996, no working paper with any AI-related keyword was uploaded, in 2018 the number of posted papers including such keywords were 410, accounting for 3% of all papers posted that year.

Robo-advisors promise to apply new technologies and procedures to improve financial decision making, as we discuss below, and as such they can be seen as a piece of the broader fintech revolution, just like digital currencies that promise to redefine the role of traditional money and platform lending which promises to redefine the role of traditional access to credit.

2.2 Fundamental Problems with Investors

From a more specific perspective, one key interest in robo-advising is that it is now commonly understood that many investors could substantially improve their financial decisions. In the past decades, the literature has documented various ways in which investors' decisions may deviate, sometimes in a fundamental way, from the standard premises of a fully rational economic agent, who knows the entire set of possible alternatives, the associated outcomes in a probabilistic sense, and can correctly match all the information in order to maximize life-time utility.

Restricting to the investment domain, which has so far been the typical focus of robo-advisors, investors have been found to display low participation (Mankiw and Zeldes, 1991), underdiversification (Grinblatt, Keloharju and Linnainmaa, 2011; Goetzmann and Kumar, 2008; Bianchi and Tallon, 2019), default bias (Benartzi and Thaler, 2007), portfolio inertia (Agnew, Balduzzi and Sunden, 2003; Biliias, Georgarakos and Haliassos, 2010), excessive trading (Odean, 1999), trend chasing (Greenwood and Nagel, 2009), and a poor understanding of matching mechanism (Choi, Laibson and Madrian, 2009). Many of those investment behavior are associated with a poor understanding of basic financial principles (Lusardi, Michaud and Mitchell, 2017; Lusardi and Mitchell, 2014; Bianchi, 2018).

Several surveys provide a comprehensive list of biases and associated trading mistakes (see e.g. Guiso and Sodini, 2013, Barber and Odean, 2013, Beshears, Choi, Laibson and Madrian (2018)). For the purpose of this article, two points are worth stressing. First, these mistakes are not small ; on the contrary, their welfare implications can be substantial

(Campbell, 2006, Campbell, Jackson, Madrian and Tufano (2011)). Second, they do not cancel out in equilibrium; rather, they have important effects on the functioning of financial markets and on broader macroeconomic issues such as wealth inequality (Vissing-Jorgensen, 2004 ; Lusardi, Michaud and Mitchell, 2017; Bach, Calvet and Sodini, 2020 ; Fagereng, Guiso, Malacrino and Pistaferri, 2020).

Motivated by this evidence, it is clear that improving financial decision making can be seen as a major goal of financial innovation, and part of the interest in robo advising lies in its promise to help investors in this dimension.

2.3 Fundamental Problems with Advisors

A natural response to investors' poor financial decision-making is to delegate the task to professional experts, who have the time and skill to serve investors' best interest. The argument relies on a few important assumptions, which may sometimes be difficult to meet in practice. First, advisors are required to be able to recognize and adapt their strategies to match their clients' preferences and needs. This is far from obvious and recent evidence suggests that advisors may themselves have misguided beliefs. Foerster, Linnainmaa, Melzer and Previtero (2017) analyse the trading and portfolio decisions of around 10,000 financial advisors and 800,000 clients in four Canadian financial institutions. They show that clients' observable characteristics (risk tolerance, age, income, wealth, occupation, financial knowledge) jointly explain only 12% of the cross-sectional variation in clients' risk exposure. This is remarkably low, especially compared to the effect of being served by a given advisor, which explains 22% of the variation in clients' risk exposure. In terms of incremental explanatory power, adding advisor effects to a model in which investors' risk exposure is explained by their observable characteristics improves the adjusted R2 from 12% to 30%. This evidence suggests that, in some cases, financial recommendations are closer to "one size fits all" than being fully personalised to clients' specific preferences and needs. Furthermore, Linnainmaa, Melzer and Previtero (2020) show that some advisors, when trading with their own money, display very similar trading biases to their clients: they prefer active management, chase returns and they are not well diversified.

A second key aspect is that advisors need to have the incentives to act in clients' best interests, rather than pursuing their own goals. Again, recent evidence suggests this need not be the case. Mullainathan, Noeth and Schoar (2012) conducted a study by training auditors, posing as customers of various financial advisors, and (randomly) asking them to represent different investment strategies and biases. They show that advisors display a significant bias towards active management, they initially support clients' requests but their final recommendations are orthogonal to clients' stated preferences. At the end, advisors failed to correct clients' biases and even made clients worse off. Similarly, Foà, Gambacorta, Guiso and Mistrulli (2019) document banks' strategic behavior in their offer of mortgage contracts. A more extensive review of advisors' conflicted advice is provided in Beshears *et al.* (2018).

A third key aspect is that, even removing the previous concerns, financial advising is costly, and a significant part of the cost has a fixed component (say, the advisor's time). This implies that financial advice may not be accessible to investors with lower levels of wealth, who may in fact be those who need it most.

III. Promises

3.1 Accountable Procedures and personalized recommendations

Robo-advisors' services offer accountable procedure to allocate an individual's portfolio across various asset classes and different types of funds, depending on their individual characteristics. Two stages of the process are crucial to this : (1) client profiling, and (2) asset allocation. While personalised recommendations are offered to clients, there is considerable heterogeneity in the recommended allocations, and the exact algorithms used by robo-advisors are typically not transparent.

3.1.1 Client profiling

Robo-advisors typically use an online questionnaire to assess investors' financial situation, characteristics and investment goals. This questionnaire is a regulatory requirement under SEC guidelines in the US (SEC, 2006; SEC, 2019). A "suitability assessment" is also mandatory under MiFID (Markets in Financial Instruments Directive) regulation in Europe.⁴

Individual characteristics, such as age, marital status, net worth, investment horizon and risk tolerance are used to assess the investor's situation. Interestingly, a large variety of questions can be used to estimate one particular characteristic. For example, if you consider risk tolerance, most of the robo-profilers use subjective measures of risk aversion based on a self-assessment. Some robo-profilers use risk capacity metrics (measuring the ability to bear losses), estimated from portfolio loss constraints, financial obligations or expenses, balance sheet information, etc. In Europe, under MiFID II, advisors also assess the clients' "experience and knowledge" to understand the risks involved in the investment product or service offered.⁵ Robo-advisors thus ask questions about the clients' financial literacy and reduce the individuals' risk tolerance when financial literacy is low.

Robo-advisors typically propose their clients pick a goal (for example, retirement, buying a house, a bequest to family members, a college/education fund or a safety net) among several possibilities during the risk profiling questionnaire. This goal can define the investment horizon or the risk capacity in the optimal portfolio allocation. Other robo-advisors allow their clients to name their goal before or outside of the risk profiling process, and do not necessarily incorporate it into the portfolio allocation. Finally, a few robo-advisors allow their clients to set multiple goals, thus offering their clients the ability to explicitly put their portfolio in a mental account (Das, Markowitz, Scheid and Statman, 2010). One limitation is that robo-advisors frequently lack a total view of an investor's overall financial situation, as savings outside of the robo platform are rarely taken into account. Some of them have a broader view of the clients' financial situation through partnerships with financial account aggregators or digital platforms of investment.⁶

3.1.2 Asset Allocation

In a second step, the robo-advisor proposes to structure a portfolio by taking into account investment goals and the desired risk level. Beketov *et al.* (2018) analysed 219 robo advisors

from 28 countries (30% in the US, 20% in Germany, 14% in the UK), that were founded between 1997 and 2017. As shown in Figure 1, a word cloud representing the occurrence of different methods used by robo advisors, a large variety of portfolio construction techniques are used. Beketov *et al.* (2018) show that most robo advisors use simple Markowitz optimisation or a variant of it, such as Black Litterman (40%), sample portfolios applying a pre-defined grid (27%) or constant portfolio weights (14%). A minority of robo advisors use alternative portfolio construction techniques such as liability driven investment, full-scale optimisation, risk parity and constant proportion portfolio insurance.

Figure 1: Word cloud representing the occurrence of different methods used by existing robo-advisors



Source: Beketov *et al.* (2018)

If most robo-advisors perform asset allocation by using a mean–variance analysis, or a variant of it, they rarely disclose information on how they choose their asset class investment universe or how they estimate variances and correlations between asset classes. They even more rarely disclose their expected return and risk parameters explicitly. Among the dominant players in the US, Wealthfront is probably one of the few exceptions. It discloses on its website its portfolio optimisation method (Black-Litterman), but also its expected returns, volatilities and correlation matrices and the way these are estimated.⁷ Betterment

is also relatively transparent. It provides justification and detail on the choice of its investment universe, its portfolio optimisation method (Black-Litterman) and the way it calculated expected returns and risk, without disclosing them explicitly.⁸ Schwab Intelligent Portfolios also discloses its portfolio optimisation method, a variant of the Markowitz approach (using Conditional Value at Risk instead of variance). However, it is less transparent on its Monte-Carlo simulation methodology and expected return hypotheses.⁹

3.1.3 Heterogeneity in the proposed asset allocations

In theory, these rigorous procedures and their systematic nature should make it possible to overcome the shortcomings of human advisors, by reducing unintentional biases and simplifying the interaction with the client. Rebalancing for example, is made easier through robo-advising platforms that implement this automatically or require a simple validation by the client. Also, if individual characteristics are measured with sufficient precision, robo-advising services should make it possible to offer investment recommendations that are personalised to each investor’s situation.

In practice, a large disparity in the proposed asset allocations has been documented, for the same investor’s profile. For example, Boreiko and Massarotti (2020) analyses 53 robo-advisors operating in the US and Germany in 2019. They show that a “moderate” profile invests an average 56% in equities, but the standard deviation of the proposed equity exposure is large (23%). Equity exposure can go from 14% to 100%, depending on the robo-advisor. Aggressive or conservative asset allocations have similar features, with an average equity

exposure of 73% and 35% respectively, but range between 18% and 100% for aggressive allocations, and from 0 to 100% for conservative allocations.

This disparity in the proposed allocations can have several sources. It could come from different portfolio construction methodologies or different expected risk/return hypotheses. It may also reflect robo-advisors' conflicts of interest. Boreiko and Massarotti (2020) show that the asset managers' expertise in a given asset class (proxied as the percentage of funds in a given asset class across the total universe of funds proposed by the robo-advisor) is the main driver. Conflicts of interest were also demonstrated in the case of Schwab Intelligent Portfolio, which recommended that a significant portion of the clients' portfolio be invested in money market funds. Lam (2016) argued that this unusually large asset allocation to cash allowed Schwab Intelligent Portfolios to delegate cash management to Schwab Bank, allowing the firm to profit from the interest rate difference between lending rates and the paid rate of return (Fisch, Laboure and Turner, 2019).

3.2 Improving investor's experience

As for many innovative financial services, a key promise of robo-advice is to make investors better off. This claim is obviously difficult to test, as it requires having a good understanding of investors' preferences, constraints and outside opportunities (for example, how they would otherwise have used the capital invested with the robo-advisor), as well as a complete picture of investors' assets. Moreover, even if one can have reasonable approximations of how investors trade-off risk and returns, investors may care about other elements. For example, some investors may just use financial advice to acquire peace of mind. Gennaioli, Shleifer and Vishny (2015) propose a model in which a financial advisor acts as a "money doctor" and allows investors to effectively decrease their reluctance to take risk. Rossi and Utkus (2019a) document that acquiring peace of mind is one of the key drivers behind the demand for financial advice.

Most academic studies do not venture into developing a fully-fledged welfare analysis, they take a more limited view and check whether having access to the robot increases investors returns, after having controlled for some measures of portfolio risk. An improvement in risk-adjusted returns can come from static changes in portfolio choices, for example by improving diversification and therefore reducing risk for a given level of expected returns. Or they may occur over time, by allowing investors to rebalance their portfolios in a way that stays closer to their target risk-return profile.

Recent academic studies document that robo-advising services tend to improve investors' diversification and risk-adjusted performance. For example, D'Acunto, Prabhala and Rossi (2019) study a Portfolio Optimizer targeting Indian equities, and find that robo-advice was beneficial to ex-ante under diversified investors, by increasing their portfolio diversification, reducing their risk and increasing their ex-post mean returns. However, the robo-advisor did not improve the performance of already-diversified investors. Rossi and Utkus (2019b) study the effects of a large U.S. robo-advisor on a population of previously self-directed investors. They find that, across all investors, robo-advice users reduced their money market investment and increased their bond holdings. The introduction of robo-advice also reduced idiosyncratic risk by lowering the holdings of individual stocks and active mutual funds and raising exposure to low-cost indexed mutual funds. It also reduced portfolios' home bias by

significantly increasing international equity and fixed income diversification. The introduction of the robot increased individuals' overall risk-adjusted performance. In a different sample, Reher and Sun (2019) also pointed to a diversification improvement of robo-advice users generated by a large US robo-advisor. Bianchi and Brière (2021) study the introduction of a large French robo-advisor on employee savings's plans. They find that relative to self-management, accessing the robo-advice services is associated with an increase in individuals' investment and risk-adjusted returns. Investors bear more risk, and rebalance their portfolios in a way to that keeps their allocation closer to the target. This increased risk taking is also found by Hong, Lu and Pan (2020), studying a Chinese robo-advisor, and using unique account-level data on consumption and investments from Ant Group. Robo-adoption helped households move towards optimal risk-taking, reducing their consumption volatility.

3.3 Reach Under-served Investors

One the most important promises of the fintech revolution is linked to financial inclusion. As mentioned previously, offering financial services often involves substantial fixed costs, which can make it unprofitable to serve poorer consumers. New technologies allow a dramatic decrease in transaction costs (Goldfarb and Tucker (2019) identify various ways through which this could happen). By reducing these costs, new technologies can reach those who have been traditionally under-served (Philippon (2019)).

Robo-advisors can be seen as part of this promise. First, they typically require lower initial capital to open an account. For example, Bank of America requires US\$25,000 to open an account with a private financial advisor, but only US\$5,000 to open an account with its robo-advisor. Some robo-advisors, such as Betterment, do not require a minimum investment at all. Second, they typically charge lower fees than human advisors. The automation of the advice process reduces the advice-related fixed costs. For example, a fully automated robo-advisor in the US typically charges a fee between 0.25% and 0.50% of assets managed (this is between 0.25% and 0.75% in Europe),¹⁰ whereas the fees for traditional human advisors rarely fall below 0.75% and can even reach 1.5% (Lopez, Babccic and De La Ossa, 2015; Better Better Finance, 2020).

Academic studies on robo advising and financial inclusion are scarce, but the initial results seem to support the above claims. Hong *et al.* (2020) show that the adoption of a popular fintech platform in China is associated to increased risk taking, and the effect is particularly large for households residing in areas with low financial service coverage. Reher and Sokolinski (2020) analyse the effects of the reduction of the account minimum from \$5,000 to \$500 by a major U.S. robo-advisor. They show that, thanks to this reduction, there was a 59% increase in the share of "middle class" participants (with wealth between \$1,000 and \$42,000), but no increase in participation by households with wealth below \$1,000. The majority of new middle-class robo-advice participants are also new to the stock market and, relative to upper class participants, they increase their risky share by 13 pps and their total return by 1.2 pps. Bianchi and Brière (2021) also show that robo-advice participants increase their risk exposure and their risk adjusted returns. Importantly, the increase in risk exposure is larger for investors with smaller portfolios and lower equity exposure at the baseline, and the increase in returns is larger for smaller investors and for investors with lower returns at the baseline. These results suggest that having access to a robo advisor may be particularly important for investors who are less likely to receive traditional advice and, as such, it can be seen as an important instrument towards financial inclusion.

IV. Open Questions

4.1 Why Not More AI/Big Data?

As mentioned, most robo-advisors currently build on rather simple procedures both in terms of the information employed to profile the client and how this information is used to construct an optimal portfolio. As emphasised in Beketov *et al.* (2018), modern portfolio theory remains dominant and forms of artificial intelligence are hardly employed. This may seem surprising given the increased interest in AI and Big Data mentioned previously, and since that robo-advisors are often presented as incorporating these trends. One may wonder why we fail to see more AI built into robo-advising.

A first reason may be that, while such inclusion would be desirable, it is not feasible due to technological or knowledge constraints. That is, finance theory has not advanced enough to be able to give recommendations on how to incorporate AI into finance models.

Some scholars would not agree. Bartram *et al.* (2020) summarise the shortcomings of classical portfolio construction techniques and highlight how AI techniques improve the practice. In particular, they argue that AI can produce better risk-return estimates, solve portfolio optimisation problems with complex constraints, and yield better out-of-sample performance compared with traditional approaches.

A second reason may be that including more AI would violate regulatory constraints. According to the current discipline, as a registered investment advisor, a robo-advisor has a fiduciary duty to its clients. As discussed by Grealish and Kolm (2021), the fiduciary duty in the U.S. builds on the 1940 Advisers Act and it has been adapted by the SEC in 2017 so as to accommodate the specifics of robo-advising. In particular, robo-advisors are required to elicit enough information on the client, use properly tested and controlled algorithms, and fully disclose the algorithms' possible limitations.

Legal scholars debate how much a robo-advisor can and should be subject to fiduciary duty. Fein (2017) argues that robo-advisors cannot be fully considered as fiduciaries since they are programmed to serve a specific goal of the client, as opposed to considering their broader interest. As such, they cannot meet the standard of care of the prudent investor required for human advisers. Similarly, Strzelczyk (2017) stresses that robo-advisors cannot act as a fiduciary since they do not provide individualised portfolio analysis but rather base their recommendations on a partial knowledge of the client. On the other hand, Ji (2017) argues that robo-advisors can be capable of exercising the duty of loyalty to their clients so as to meet the Advisers Act's standards. In a similar vein, Clarke (2020) argues that the fiduciary duty can be managed by basing recommendations on finance theory and by fully disclosing any possible conflict of interest.

A third reason may be that having more AI in robo-advice is simply not desirable. Incorporating AI would at least partly make these robots a black-box and would make it harder to provide investors with clear explanations of why certain recommendations are given. Patel and Lincoln (2019) identify three key sources of risk associated with AI applications: first, opacity and complexity; second, the distancing of humans from decision-making; and third, changing incentive structures (for example in data collection efforts). They consider the

implications of these sources of risk in several domains, ranging from damaging trust in financial services, propagating biases, harming certain group of customers possibly in an unfair way. They also consider market level risks ranging from financial stability, cybersecurity and new regulatory challenges.

Algorithm complexity could be particularly problematic in bad times. Financial Stability Board (2017) argues that the growing use of AI in financial services can threaten financial stability. One reason is that AI can create new forms of interconnectedness between financial markets and institutions, since various institutions may employ previously unrelated data sources, for example. Moreover, the opacity of AI learning methods could become a source of macro-level risk due to their possibly unintended consequences.

Algorithm complexity is also particularly problematic for those with lower financial capabilities. Complex financial products have been shown to be particularly harmful for less sophisticated investors (see e.g. Bianchi and Jehiel (2020) for a theoretical investigation, Ryan, Trumbull and Tufano (2011) and Lerner and Tufano (2011) for historic evidence, and Célérier and Vallée (2017) for more recent evidence). As for many (financial) innovations, the risk is that they do not reach those who would need it the most, or that they end up being misused.

In this way, some key promises of robo-advising, notably on improved financial inclusion and accountability, can be threatened by the widespread use of opaque models.

4.2 How Far Should We Go Into Personalization?

The potential of robo-advice is to combine financial technology and artificial intelligence and offer to each investor personalised advice based on their objectives and preferences. One important difficulty lies in the precise measurement of investors' characteristics. A second issue relates to the sensitivity of the optimal asset allocation to these characteristics, which can be subject to a large degree uncertainty. This can lead the estimated optimal portfolio to be substantially different from the true optimal one, with dramatic consequences for the investor.

4.2.1 Difficulty of measuring an individual's characteristics

Lo (2016) calls for the development of smart indices, that could be personalised to individuals' circumstances and characteristics. If we are not there yet, robo-advisors could make a step in that direction, by helping to precisely define an investor's financial situation and goals (Gargano and Rossi, 2020). As it has been demonstrated by a large number of academic papers, optimal portfolio choices rely on various individual characteristics such as human capital (Cocco, Gomes and Maenhout, 2005; Benzoni, Collin-Dufresne and Goldstein, 2007; Bagliano, Fugazza and Nicodano, 2019), housing market exposure (Kraft and Munk, 2011), time preference, risk aversion, ambiguity aversion (Dimmock, Kouwenberg, Mitchell and Peijnenburg, 2016; Bianchi and Tallon, 2019), etc. Individualisation possibilities are much wider than what is currently implemented in robo-advice services.

However, some individual characteristics are difficult to measure and subject to a large degree of uncertainty. Risk aversion is one of them. Different methods have been developed

by economists and psychologists to measure individuals' risk aversion. Most of them are experimental measurements based on hypothetical choices. For example, the lotteries of Barsky, Juster, Kimball and Shapiro (1997) offer individuals the choice between employment with a risk-free salary, and a higher but risky salary. Other work (Holt and Laury, 2002; Kapteyn and Teppa, 2011; Weber, Weber and Nosić, 2013) measure preferences based on a series of risk/return tradeoffs. The choice between a certain gain and a risky lottery is repeated, gradually increasing the prize until the subject picks one lottery.

One reason it is difficult to measure risk aversion might be that people interpret outcomes as gains and losses relative to a reference point and are more sensitive to losses than to gains. Kahneman, Knetsch and Thaler (1990) or Barberis, Huang and Santos (2001) report experimental evidence of loss aversion. Loss aversion can also explain why many investors prefer portfolio insurance products offering capital guarantees (Calvet, C  l  rier, Sodini and Vall  e, 2020).

In practice, robo-advisors frequently assess a clients' risk tolerance based on a self-declaration. People are asked to rate themselves in their ability to take risks on a scale of 1 to 10 (Dohmen, Falk, Huffman, Sunde, Schupp and Wagner, 2005). These ratings have the disadvantage of not being very comparable across individuals. Scoring techniques are also frequently used by robo-advisors. They ask the individual a large number of questions, covering different aspects of life (consumption, leisure, health, financial lotteries, work, retirement and family). Global scores are obtained by adding the scores across various dimensions, keeping only those questions which prove to be the most relevant ex-post to measure an individual's risk aversion, a statistical criterion which eliminates the least relevant questions (Arrondel and Masson, 2013).

In Europe, the implementation of MiFID regulation led to several academic studies assessing risk profiling questionnaires. European regulation does not impose a standardised solution, each investment company remains free to develop its questionnaire as it wishes, which explains the great heterogeneity of the questionnaires distributed in practice to clients. Marinelli and Mazzoli (2010) sent three different questionnaires used by banks to 100 potential investors to verify the consistency of the clients' risk profiles. Only 23% of individuals were profiled in a consistent way across the three questionnaires, a likely consequence of the differences in the contents and scoring methods of the questionnaires. Other work carried out in several European countries (De Palma, Picard and Prigent, 2009; Marinelli and Mazzoli, 2010; Linciano and Soccorso, 2012) arrived to the same conclusion.

4.2.2 Algorithm sensitivity to parameter uncertainty

Optimal allocations are usually very sensitive to parameters (expected returns, covariance of assets' returns) which are hard to estimate. They also depend crucially on investor's characteristics (financial wealth, human capital, etc.) often known with poor accuracy. On one hand, there is a cost for suboptimal asset allocation (one size does not fit all) and substantial gains to individualise (see Dahlquist, Setty and Vestman, 2018; Warren, 2019). On the other hand, there is a risk of overreaction to extreme/time-varying individual characteristics, potentially leading to "extreme" asset allocations, as it has been shown by the literature on optimisation with parameter uncertainty (see for example Garlappi, Uppal and Wang, 2007). Blake, Cairns and Dowd (2009) claim that some standardisation is needed, like in the aircraft industry, to

guarantee investors' security. How much customisation is needed depends largely on the trade-off between the gains to bring the portfolio closer to an individual's needs and the risk of estimating an individual's characteristics with a large degree of error.

How stable an individual characteristics are in practice also remains an open question. Capponi, Olafsson and Zariphopoulou (2019) show that if these risk profiles are changing through time (depending on idiosyncratic characteristics, market returns or economic conditions), the theoretical optimal dynamic portfolio of a robo-advisor should adapt to the client's dynamic risk profile, by adjusting the corresponding inter-temporal hedging demands. The robo-advisor faces a trade-off between receiving client information in a timely manner and mitigating behavioural biases in the risk profile communicated by the client. They show that with time-varying risk aversion, the optimal portfolio proposed by the robo-advisor should counter the client's tendency to reduce market exposure during economic contractions.

4.3 Can Humans Trust Robots?

In the interaction between humans and robo-advisors, a key ingredient is trust, determining the individual's willingness to use the service and to follow the robo recommendations. We review what creates trust in algorithms and discuss the impact of trust on financial decisions.

4.3.1 Trust is key for robo-advice adoption

Trust has been shown to be a key driver of financial decisions (see Sapienza, Toldra-Simats and Zingales (2013) for a review). For example, trustful investors are significantly more likely to invest in the stock market (Thakor and Merton, 2018). Trust is also a potential key driver of robo-advice adoption. As stated by Merton (2017), "What you need to make technology work is to create trust."

Trust has been studied across a variety of disciplines, including sociology, psychology and economics, to understand how humans interact with other humans, or more recently with machines. Trust is a "multidimensional psychological attitude involving beliefs and expectations about the trustee's trustworthiness, derived from experience and interactions with the trustee in situations involving uncertainty and risk" (Abbass, Scholz and Reid, 2018). One can also see trust as a transaction between two parties: if A believes that B will act in A's best interest, and accepts vulnerability to B's actions, then A trusts B (Misztal, 2013). Importantly, trust exists to mitigate uncertainty and the risk of collaboration by enabling the trustor to anticipate that the trustee will act in the trustor's best interests.

While trust has both cognitive and affective features, in the automation literature, cognitive (rather than affective) processes seem to play a dominant role. Trust in robots is multifaceted. It has been shown to depend on robot reliability, robustness, predictability, understandability, transparency, and fiduciary responsibility (Sheridan, 1989 ; Sheridan, 2019; Muir and Moray, 1996). One key feature of robo-advisors is their reliance on more or less complicated algorithms, in several steps of the advisory process. An algorithm is used to profile the investor, and then to define the optimal asset allocation. A client delegating the decision to the robot bears the risk that a wrong decision by the robot will lead to poor performance of their savings. Trust in these algorithms is thus key for robo-advisor adoption.

4.3.2 Algorithm aversion

Survey evidence (HSBC, 2019) shows that there is a general lack of trust in algorithms. While most people seem to trust their general environment and technology (68% of the survey respondents said they will trust a person until prove otherwise, 48% believe the majority of people are trustworthy and 76% that they feel comfortable using new technology), artificial intelligence is not yet trusted. Only 8% of respondents would trust a robot programmed by experts to offer mortgage advice, compared to 41% trusting a mortgage broker. As a comparison, 9% would be likely to use a horoscope to guide investment choices! 14% would trust a robot programmed by leading surgeons to conduct open heart surgery on them, while 9% would trust a family member to do an operation supported by a surgeon. Only 19% said they would trust a robo-advisor to help them make investment choices. There are large differences across countries however. The percentage of respondents who trust robo-advisors rises to 44% and 39% in China and India respectively, but it is only 9% and 6% in France and Germany.

Some academic studies have shown that decision makers are often averse to using algorithms, most of the time preferring less accurate human judgment. For example, professional forecasters have been shown not to use algorithms or give them insufficient weight (Fildes and Goodwin, 2007). Dietvorst, Simmons and Massey (2015) gave participants the choice of either exclusively using an algorithm's forecasts or exclusively using their own forecasts during an incentivised forecasting task. They found that most participants chose to use the algorithm exclusively only when they had no information about the algorithm's performance. However, when the experimenter told them it was imperfect, they were much more likely to choose the human forecast. This effect persisted even when they had explicitly seen the algorithm outperform the human's forecasts. This tendency to irrationally discount advice that is generated and communicated by computer algorithms has been called "algorithm aversion". In a later experimental study (Dietvorst, Simmons and Massey, 2018), participants were given the possibility to modify the algorithm. Participants were considerably more likely to choose the imperfect algorithm when they could modify its forecasts, even if they were severely restricted in the modifications they could make. This suggests that algorithm aversion can be reduced by giving people some control over an imperfect algorithm's forecast.

Recent experimental evidence shows less algorithm aversion. Niszczoła and Kaszás (2020) tested if people exhibited algorithm aversion when asked to decide whether they would use human advice or an artificial neural network to predict stock price evolution. Without any prior information on the human versus robot's performance, they found no general aversion towards algorithms. When it was made explicit that the performances of the human advisor was similar to that of the algorithm, 57% of the participants showed a preference for the human advice. In another experiment, subjects were asked to choose a human or a robo-advisor to exclude stocks that were controversial. Interestingly, people perceived algorithms as being less effective than humans when the tasks required a subjective judgment to be made, such as morality.

Germann and Merkle (2019) also found no evidence of algorithm aversion. In a laboratory experiment (mostly based on business or economics' students), they asked participants to choose between a human fund manager and an investment algorithm. The selection process was repeated ten times, which allowed them to study the reaction to the advisor's performance. With equal fees for both advisors, 56% of participants decided to follow the algorithm. When fees differed, most participants (80%) chose the advisor with the lower fees.

Choices were strongly influenced by the cumulative past performance. But investors did not lose confidence in the algorithm more quickly after seeing forecasting errors. An additional survey provided interesting qualitative explanations to the results. Participants believed in the ability of the algorithm to be better able to learn than humans. They viewed humans as having a comparative advantage in using qualitative data and dealing with outliers. All in all, the algorithms are viewed as a complement rather than a competitor to a human advisor.

4.3.3 What creates trust in an algorithm?

Jacovi, Marasović, Miller and Goldberg (2020) distinguish two sources of trust in algorithm: intrinsic and extrinsic. Intrinsic trust can be gained when the observable decision process of the algorithm matches the user priors. Explanations of the decision process behind the algorithm can help create intrinsic trust.¹¹ Additionally, an algorithm can become trustworthy through its actual behaviour: in this case, the source of trust is not the decision process of the model, but the evaluation of its output.

The European Commission (2019) recently listed a number of requirements for trustworthy algorithms. Related to intrinsic trust are the requirements of (1) the user's agency and human oversight, (2) privacy and data governance, (3) transparency and the ability to explain the algorithm. Extrinsic trust can be increased by (4) the technical robustness and safety of the algorithm, (5) the ability to interpret its output, (6) its accountability and auditability. In addition, ethical and fairness considerations, such as (7) avoiding discrimination, promoting diversity and fairness or (8) encouraging societal and environmental well-being are also considered to be a key component of trust.

Trust in algorithms also crucially depends on the perception of the expertise and reliability of the humans or institutions offering the service (Prahl and Van Swol, 2017). "Technology doesn't create trust on its own" (Merton, 2017). People trust humans certifying a technology, not necessarily the technology itself. In the specific case of robo-advice, Lourenço, Dellaert and Donkers (2020) study the decision of consumers to adopt the service and show that this decision is clearly influenced by the for-profit versus not-for-profit orientation of the firm offering the service (for example private insurance and investment management firm versus pension fund or government-sponsored institution). Transparency, explainability and interpretability may not be sufficient by themselves for enhancing decisions and increasing trust. However, informing key hypotheses and potential shortcomings of algorithms when making certain decisions, may be a fundamental dimension to be worked on.

4.3.4 Trust in robots and financial decisions

Not everyone trusts robo-advisors. In a sample of 34,000 savers in French employee savings' plans, Bianchi and Brière (2021) document that individuals who are young, male, and more attentive to their saving plans (measured by the time spent on the savings plan website), have a higher probability of adopting a robo-advice service. The probability of taking up the robo-advice option is also negatively related to the size of the investors' portfolio, which suggests that the robo-advisor is able to reach less wealthy investors,¹² a result also confirmed by Brenner and Meyll (2020). Investors with smaller portfolios are also more likely to assign a larger fraction of their assets to the robot.

A unique feature of the robo-advice service studied by Bianchi and Brière (2021) allows them to analyse both "robo-takers" and the "robo-curious," i.e., individuals who observe the robot's recommendation without actually subscribing to it. Interestingly, the further away is the robot's recommendation relative to the current allocation, the larger the probability that the investor will subscribe to the robot. This finding can be contrasted with the observation that human advisors tend to gain trust from their clients by being accommodating with clients (Mullainathan *et al.*, 2012). Moreover, investors who are younger, female, those who have larger risk exposure and lower past returns, as well as less attentive investors are more likely to accept a larger increase in their exposure to risky assets, such as equities.

Trust can have a large impact on investor's decisions. Bianchi and Brière (2021) and Hong *et al.* (2020) show evidence of increased risk taking, a result consistent with increased trust. For example, Bianchi and Brière (2021) document a 7% increase in equity exposure after robo-advice adoption (relative to an average 16% exposure). Hong *et al.* (2020) document a 14% increase (relative to an average risky exposure of 37% on their sample of 50,000 Chinese consumer clients of Alibaba). Interestingly, Hong *et al.* (2020) additionally show that this result is likely not to be driven by an increase in the individual's risk tolerance driven by robot support. Rather, it seems to reflect a better alignment of the investment portfolio with the actual risk tolerance of the individual. In particular, they show that after robo-advice adoption, exposure to risky assets is more in line with the individual's risk tolerance estimated from their consumption growth volatility (Merton, 1971), measured from Alibaba's Taobao online shopping platform. The robo-advisor seems to help individuals to move closer to their optimal alignment of risk-taking and consumption. These results should, however, be used with caution, as both studies concentrate on a relatively short period of investment (absent any serious market crash) and lack a global view on the individuals' overall portfolios. More work would need to be done to document a long term impact.

4.4 Do Robots Substitute or Complement Human Decision-Making?

Autonomous systems are being developed across large areas of our everyday life. Understanding how humans will interact with them is a key issue. In particular, should we expect that robots will become substitutes to humans or rather be complementary? In the special case of financial advice, are they likely to replace human advisors?

Using a representative sample of US investor, Brenner and Meyll (2020) investigate whether robo-advisors, reduce investors' demand for human financial advice offered by financial service providers. They document a large substitution effect and show that this effect is driven by investors who fear to be victimised by investment fraud or worry about potential conflicts of interest. In practice however, a number of platforms that were entirely digital decided to reintroduce human advisors. For example, Scalable Capital, the European online robo-advice company backed by BlackRock, or Nutmeg, reintroduced over-the-phone and face-to-face consultations after finding that a number of clients preferred talking to human advisors rather than solely answering online questionnaires.

Another related question is understanding how people will interact with robots. Will they delegate the entire decision to the robot or will they keep an eye on it, to monitor the process and intervene if necessary? In certain experiments, users put too much faith in robots. Robinette, Li, Allen, Howard and Wagner (2016) designed an experiment where participants

were asked choose to follow, or not to follow, a robot's instructions in an emergency. All participants followed the robot during the emergency, even though half of the participants observed the same robot perform poorly in a non-emergency navigation guidance task just a few minutes before. Even when the robot pointed to a dark room with no discernible exit the majority of people did not choose to safely exit the way they had entered. Andersen, Kösllich, Pedersen, Weigelin and Jensen (2017) expand on this work and show that such over-trust can also affect human/robot interactions that are not set in an emergency situation.

In the context of financial decisions, Bianchi and Brière (2021) document that robo-advisor adoption leads to significantly increase attention on savings plans, during the months following the adoption. Individuals are in general more attentive to their saving plan, particularly when they receive variable remuneration and need to make an investment decision. This seems to indicate that people do not use the robot as a substitute for their own attention.

V. The Next Generation of Robo-Advisors

It is not clear which generation of robo-advisors we are currently facing. Beketov *et al.* (2018) focus on robots of third and fourth generation, which differ from earlier generations as they use more automation and more sophisticated methods to construct and rebalance portfolios. One possibility is that the next generation of robots would continue the trend of using more data and more complex models. One may, however, imagine an alternative path. As discussed previously, incorporating more AI into robo-advice (and more generally into financial services) faces three key challenges. Firstly, while highly personalised asset allocations have the great potential of accommodating an individual's needs, they are also more exposed to measurement errors of relevant individual characteristics and to parameter uncertainty. Secondly, to the extent that increased AI is associated to increased opacity, the risk is to miss some key promises of increased accountability and financial inclusion. Third, trust is key for technology adoption, even more so in the domain of financial advice. These challenges, in our view, call for algorithms that can be easily interpreted and evaluated. Toreini, Aitken, Coopamootoo, Elliott, Zelaya and van Moorsel (2020) discusses how developing trust in (machine learning) technologies requires them to be fair, explainable, accountable, and safe (FEAS).

From this perspective, recent advances in so-called XAI (explainable artificial intelligence), can be particularly useful when thinking about the future of robo-advisors. "Explainability" refers to the possibility of explaining a given prediction or recommendation, even if based on a very complicated model, for example by evaluating the sensitivity of the prediction when changing one of the inputs. It also refers to how much a given model can itself be explained. Explanations can help humans perform a given task and, at the same time, evaluate a given model (see e.g. Biran and Cotton (2017) for a recent survey). As discussed in Doshi-Velez and Kim (2017), explainability can be considered a desiderata both in itself, in relation to the issues of trust and accountability expressed above, and also as a tool to assess whether other desiderata, such as fairness, privacy, reliability, robustness, causality, usability, are met.

There is a large amount of academic literature examining whether explainable artificial intelligence can improve human decision-making. How much explainability is needed for the actual functioning of an automated system remains an open question and it is often debated in the context of self-driving cars, for example. On the one hand, psychological research regarding decision-making suggests that when decisions involve complex reasoning, ignoring part of the available information and using heuristics can help to deal more robustly with uncertainty than relying on resource-intensive processing strategies (Gigerenzer and Brighton, 2009). On the other hand, experimental studies show that providing the driver with information on why and how an autonomous vehicle acts, is important to maintain a safe driving experience (Koo, Kwac, Ju, Steinert, Leifer and Nass, 2015). This information is particularly key in emergency situations. Drivers receiving such information tend to trust the car less and are faster to take control of the car when a dangerous situation occurs (Helldin, Falkman, Riveiro and Davidsson, 2013). One should also be particularly attentive to the risk of information overload. An algorithm is easier to interpret and to use when it focuses on a few features, it is also easier to correct in case of mistakes (Poursabzi-Sangdeh, Goldstein, Hofman, Vaughan and Wallach, 2018).

In the context of robo-advisors, explainability is not an easy task. Evaluating the performance of a robo-recommendation is not straightforward, especially if one uses AI to move towards fully personalised allocations to be evaluated against fully personalised benchmarks (as described in Lo (2016)). Even more difficult for the client is to build counterfactuals of performance. And probably even more difficult is to appreciate the underlying finance model which governs the algorithm, especially if one wishes to serve less experienced investors.

In that respect, the quest is not for full transparency of the potentially complicated algorithm underlying the robo-advice process, disclosing for example all the details of the portfolio optimisation methodology or the covariance matrix estimates. It would probably be more effective to disclose, for example, which economic scenarios may cause the algorithm to perform less accurately, possibly proving ex-post sub-optimal, and informing clients about the potential limitations of the algorithm.

Another potentially interesting development would be to strengthen the interactions with clients. For example, some robo-advisors send alerts when a clients' portfolio deviates significantly from the target asset allocation (see e.g. Bianchi and Brière (2021)). These alerts could also be seen as an opportunity to interact with the client. For example, alerts could be used to explain why a deviation occurred (market movements, change in personal characteristics, etc.) or why rebalancing is recommended. Another example is that one could elicit customers' perceptions regarding the quality of the response provided by the algorithm and integrate this feedback as part of the evaluation of the robo-advice service (Dupont, 2020).

These issues are not new in AI. Biran and Cotton (2017) discuss earlier approaches of explainability of decisions in expert systems in the 1970s and more recently in recommender systems. One may argue, however, that today models are probably more complex, more autonomous and they span a larger set of decisions across a larger set of agents (including possibly less sophisticated ones), which make these issues particularly relevant in current debates. Indeed, improving transparency is also central to the policy domain, such as the recent EU regulation on data protection (GDPR). As discussed in Goodman and Flaxman (2017), the law defines a right to explanation, whereby users can inquire about the logic involved in an algorithmic decision affecting them (say, through profiling), and this calls for algorithms which are as explainable as they are efficient.

Some prominent scholars argue that the AI revolution has not happened yet. Instead of mimicking human interactions or more sophisticated human thinking, the AI revolution will happen when new forms of intelligence are considered (Jordan (2019a)). In this context, importing insights from social sciences seems crucial. AI needs psychology to capture how humans actually think and behave, or, to reference Lo (2019), to include forms of "artificial stupidity". Insights from philosophy, psychology and cognitive sciences are also key to informing how explanations are and should be communicated. Miller (2019) reviews the large amount of literature in these fields and emphasises the importance of providing selective explanations, based on causal relations and counterfactuals rather than likely statistical relations, and of allowing a social dimension in which explainers and "explainees" may interact. AI also needs economics not only to help address causality and discuss counterfactuals, but also to help design new forms of collective intelligence. These new forms may go beyond a purely anthropocentric approach, and build on some understanding of how markets function and how they may fail (Jordan (2019b)). We share the enthusiasm of these scholars when imagining advances in these directions, we look forward to seeing more social sciences in the next generation of robo-advisors!

NOTES

¹ Robo-advisors are already present in China, India, Japan, Singapore, Thailand and Vietnam.

² See e.g. The Economist (2015) on The Fintech Revolution or The World Economic Forum (2017) on Beyond Fintech: A pragmatic assessment of disruptive potential in financial service.

³ Keywords included artificial intelligence, machine learning, cluster analysis, genetic algorithm or evolutionary algorithm, lasso, natural language processing, neural network or deep learning, random forest or decision tree, and support vector machine.

⁴ In Europe, the MiFID regulation has set the objective of increased and harmonized individual investors' protection, according to their level of financial knowledge. MiFID I (2004/39 / 3C), implemented in November 2007, requires investment companies to send their clients a questionnaire to determine their level of financial knowledge, their assets and their investment objectives. MiFID I has been replaced in January 2018 by MiFID II (2014/65 / EU), which has demanded a strengthening of legislation in several areas, in particular in the requirements of advice independence and transparency (on costs, available offering, etc.).

⁵ See Article 25(3) and 56.

⁶ For example, Wealthfront recently featured direct integrations with digital investment platforms (Venmo, Redfin, Coinbase), lending (Lending Club) and tax calculation (turbotax).

⁷ <https://research.wealthfront.com/whitepapers/investment-methodology/>

⁸ On the investment universe, they excluded asset classes such as private equity, commodities and natural resources, since "estimates of their market capitalization is unreliable and there is a lack of data to support their historical performance". Expected returns are derived from market weights, through a classical reverse optimisation exercise that uses the variance covariance matrix between all asset classes. An estimation of this covariance matrix is made using historical data, combined with a target matrix, and using the Ledoit and Wolf (2004) shrinkage method to reduce estimation error. Portfolios can also be tilted towards Fama and French (1992) value and size factors, the size of the tilt being freely parametrised by the confidence that Betterment has in these views. See <https://www.betterment.com/resources/betterment-portfolio-strategy/citations>

⁹ They simulate 10,000 hypothetical future realisations of returns, using fat-tailed assumptions for the distribution of asset returns, also allowing for changing correlations modeled with a Copula approach. See <https://intelligent.schwab.com/page/our-approach-to-portfolio-construction>.

¹⁰ We consider here management fees only, not underlying ETFs or funds' fees.

¹¹ For example, a robo-advisor may disclose its risk profiling methodology, its optimization method and risk/return hypotheses, or reveal the signals leading to portfolio rebalancing.

¹² Conversely, wealthier investors are more likely to acquire information about the robot without subscribing to the service.

REFERENCES

- Abbass, H. A., Scholz, J. and Reid, D. J. (2018), *Foundations of trusted autonomy*, Springer Nature.
- Abraham, F., Schmutz, S. L. and Tessada, J. (2019), 'Robo-advisors: Investing through machines', *World Bank Research and Policy Briefs* (134881).
- Agnew, J., Balduzzi, P. and Sunden, A. (2003), 'Portfolio choice and trading in a large 401 (k) plan', *American Economic Review* 93(1), 193–215.
- Andersen, K. E., Koslich, S., Pedersen, B. K. M. K., Weigelin, B. C. and Jensen, L. C. (2017), Do we blindly trust self-driving cars, in 'Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-robot Interaction', pp. 67–68.
- Arrondel, L. and Masson, A. (2013), Measuring savers' preferences how and why?, Technical report, HAL Working Paper.
- Bach, L., Calvet, L. E. and Sodini, P. (2020), 'Rich pickings? risk, return, and skill in household wealth', *American Economic Review* 110(9), 2703–47.
- Bagliano, F. C., Fugazza, C. and Nicodano, G. (2019), 'Life-cycle portfolios, unemployment and human capital loss', *Journal of Macroeconomics* 60, 325–340.
- Barber, B. M. and Odean, T. (2013), The behavior of individual investors, in 'Handbook of the Economics of Finance', Vol. 2, Elsevier, pp. 1533–1570.
- Barberis, N., Huang, M. and Santos, T. (2001), 'Prospect theory and asset prices', *The Quarterly Journal of Economics* 116(1), 1–53.
- Barsky, R. B., Juster, F. T., Kimball, M. S. and Shapiro, M. D. (1997), 'Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study', *The Quarterly Journal of Economics* 112(2), 537–579.
- Bartram, S. M., Branke, J. and Motahari, M. (2020), *Artificial Intelligence in Asset Management*, number 14525, CFA Institute Research Foundation.
- Beketov, M., Lehmann, K. and Wittke, M. (2018), 'Robo advisors: quantitative methods inside the robots', *Journal of Asset Management* 19(6), 363–370.
- Benartzi, S. and Thaler, R. (2007), 'Heuristics and biases in retirement savings behavior', *Journal of Economic Perspectives* 21(3), 81–104.
- Benzoni, L., Collin-Dufresne, P. and Goldstein, R. S. (2007), 'Portfolio choice over the life-cycle when the stock and labor markets are cointegrated', *Journal of Finance* 62(5), 2123–2167.
- Beshears, J., Choi, J. J., Laibson, D. and Madrian, B. C. (2018), Behavioral household finance, in 'Handbook of Behavioral Economics: Applications and Foundations 1', Vol. 1, Elsevier, pp. 177–276.
- Better Finance (2020), Robo-advice 5.0: Can consumers trust robots?, Technical report.
- Bianchi, M. (2018), 'Financial literacy and portfolio dynamics', *Journal of Finance* 73(2), 831–859.
- Bianchi, M. and Brière, M. (2021), 'Robo-advising for small investors', *Working Paper available at SSRN* (3751620).
- Bianchi, M. and Jehiel, P. (2020), 'Bundlers' dilemmas in financial markets with sampling investors', *Theoretical Economics* 15(2), 545–582.

- Bianchi, M. and Tallon, J.-M. (2019), 'Ambiguity preferences and portfolio choices: Evidence from the field', *Management Science* 65(4), 1486–1501.
- Bilias, Y., Georgarakos, D. and Haliassos, M. (2010), 'Portfolio inertia and stock market fluctuations', *Journal of Money, Credit and Banking* 42(4), 715–742.
- Biran, O. and Cotton, C. (2017), Explanation and justification in machine learning: A survey, in 'IJCAI-17 workshop on explainable AI (XAI)', Vol. 8(1), pp. 8–13.
- Blake, D., Cairns, A. and Dowd, K. (2009), 'Designing a defined-contribution plan: What to learn from aircraft designers', *Financial Analysts Journal* 65(1), 37–42.
- Boreiko, D. and Massarotti, F. (2020), 'How risk profiles of investors affect robo-advised portfolios how risk profiles of investors affect robo-advised portfolios', *Frontiers in Artificial Intelligence* 3, 60.
- Brenner, L. and Meyll, T. (2020), 'Robo-advisors: A substitute for human financial advice?', *Journal of Behavioral and Experimental Finance* 25, 100275.
- Buchanan, B. (2019), 'Artificial intelligence in finance', *Alan Turing Institute Working Paper* available at <http://doi.org/10.5281/zenodo.2612537>.
- Buisson, P. (2019), 'Pure robo-advisors have become viable competitors in the US', *Amundi Digibook*.
- Calvet, L. E., Célérier, C., Sodini, P. and Vallée, B. (2020), 'Can security design foster household risk-taking?', *Working Paper* available at SSRN (3474645).
- Campbell, J. Y. (2006), 'Household finance', *Journal of Finance* 61(4), 1553–1604.
- Campbell, J. Y., Jackson, H. E., Madrian, B. C. and Tufano, P. (2011), 'Consumer financial protection', *Journal of Economic Perspectives* 25(1), 91–114.
- Capponi, A., Olafsson, S. and Zariphopoulou, T. (2019), 'Personalized robo-advising: Enhancing investment through client interaction', *arXiv preprint arXiv:1911.01391* .
- Célérier, C. and Vallée, B. (2017), 'Catering to investors through security design: Headline rate and complexity', *The Quarterly Journal of Economics* 132(3), 1469–1508.
- Choi, J. J., Laibson, D. and Madrian, B. C. (2009), 'Mental accounting in portfolio choice: Evidence from a flypaper effect', *American Economic Review* 99(5), 2085–95.
- Clarke, D. (2020), 'Robo-advisors-market impact and fiduciary duty of care to retail investors', *Working Paper* available at SSRN (3539122).
- Cocco, J. F., Gomes, F. J. and Maenhout, P. J. (2005), 'Consumption and portfolio choice over the life cycle', *Review of Financial Studies* 18(2), 491–533.
- D'Acunto, F., Prabhala, N. and Rossi, A. G. (2019), 'The promises and pitfalls of robo-advising', *Review of Financial Studies* 32(5), 1983–2020.
- D'Acunto, F. and Rossi, A. G. (2020), 'Robo-advising', *CESifo Working Paper* 8225 .
- Dahlquist, M., Setty, O. and Vestman, R. (2018), 'On the asset allocation of a default pension fund', *Journal of Finance* 73(4), 1893–1936.
- Das, S., Markowitz, H., Scheid, J. and Statman, M. (2010), 'Portfolio optimization with mental accounts', *Journal of Financial and Quantitative Analysis* 45(2), 311–334.
- De Palma, A., Picard, N. and Prigent, J.-L. (2009), 'Prise en compte de l'attitude face au risque dans le cadre de la directive mifid', *HAL Working Paper*.

- Dietvorst, B. J., Simmons, J. P. and Massey, C. (2015), 'Algorithm aversion: People erroneously avoid algorithms after seeing them err.', *Journal of Experimental Psychology: General* 144(1), 114.
- Dietvorst, B. J., Simmons, J. P. and Massey, C. (2018), 'Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them', *Management Science* 64(3), 1155–1170.
- Dimmock, S. G., Kouwenberg, R., Mitchell, O. S. and Peijnenburg, K. (2016), 'Ambiguity aversion and household portfolio choice puzzles: Empirical evidence', *Journal of Financial Economics* 119(3), 559–577.
- Dohmen, T. J., Falk, A., Huffman, D., Sunde, U., Schupp, J. and Wagner, G. G. (2005), 'Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey', *Working Paper available at SSRN* (807408).
- Doshi-Velez, F. and Kim, B. (2017), 'Towards a rigorous science of interpretable machine learning', *arXiv preprint arXiv:1702.08608*.
- Dupont, L. (2020), "Gouvernance des algorithmes d'intelligence artificielle dans le secteur financier : Analyse des réponses à la consultation", *Pôle finance innovation de l'ACPR* (December).
- European Commission (2019), 'Ethics guidelines for trustworthy ai', *Available at <https://ec.europa.eu/futurium/en/ai-alliance-consultation>*.
- Fagereng, A., Guiso, L., Malacrino, D. and Pistaferri, L. (2020), 'Heterogeneity and persistence in returns to wealth', *Econometrica* 88(1), 115–170.
- Fein, M. L. (2017), 'Are robo-advisors fiduciaries?', *Working Paper available at SSRN* (3028268).
- Fildes, R. and Goodwin, P. (2007), 'Against your better judgment? how organizations can improve their use of management judgment in forecasting', *Interfaces* 37(6), 570–576.
- Financial Stability Board (2017), 'Artificial intelligence and machine learning in financial services: Market developments and financial stability implications'.
- Fisch, J. E., Laboure, M. and Turner, J. A. (2019), 'The emergence of the robo-advisor', *The Disruptive Impact of FinTech on Retirement Systems* 13.
- Foà, G., Gambacorta, L., Guiso, L. and Mistrulli, P. E. (2019), 'The supply side of household finance', *Review of Financial Studies* 32(10), 3762–3798.
- Foerster, S., Linnainmaa, J. T., Melzer, B. T. and Previtero, A. (2017), 'Retail financial advice: does one size fit all?', *Journal of Finance* 72(4), 1441–1482.
- Gargano, A. and Rossi, A. G. (2020), 'There's an app for that: Goal-setting and saving in the fintech era', *Working Paper available at SSRN* (3579275).
- Garlappi, L., Uppal, R. and Wang, T. (2007), 'Portfolio selection with parameter and model uncertainty: A multi-prior approach', *Review of Financial Studies* 20(1), 41– 81.
- Gennaioli, N., Shleifer, A. and Vishny, R. (2015), 'Money doctors', *Journal of Finance* 70(1), 91–114.
- Germann, M. and Merkle, C. (2019), 'Algorithm aversion in financial investing', *Working Paper available at SSRN* (3364850).
- Gigerenzer, G. and Brighton, H. (2009), 'Homo heuristics: Why biased minds make better inferences', *Topics in cognitive science* 1(1), 107–143.
- Goetzmann, W. N. and Kumar, A. (2008), 'Equity portfolio diversification', *Review of Finance* 12(3), 433–463.

- Goldfarb, A. and Tucker, C. (2019), 'Digital economics', *Journal of Economic Literature* 57(1), 3–43.
- Goodman, B. and Flaxman, S. (2017), 'European union regulations on algorithmic decision-making and a "right to explanation"', *AI magazine* 38(3), 50–57.
- Grealish, A. and Kolm, P. N. (2021), Robo-advisory: From investing principles and algorithms to future developments, in 'Machine Learning in Financial Markets: A Guide to Contemporary Practice', Cambridge University Press.
- Greenwood, R. and Nagel, S. (2009), 'Inexperienced investors and bubbles', *Journal of Financial Economics* 93(2), 239 – 258.
- Grinblatt, M., Keloharju, M. and Linnainmaa, J. (2011), 'Iq and stock market participation', *Journal of Finance* 66(6), 2121–2164.
- Guiso, L. and Sodini, P. (2013), Household finance: An emerging field, in H. R. Stoll, G. M. Constantinides, M. Harris and R. Stulz, eds, 'Handbook of the Economics of Finance', Vol. 2, Elsevier, pp. 1397–1532.
- Helldin, T., Falkman, G., Riveiro, M. and Davidsson, S. (2013), Presenting system uncertainty in automotive uis for supporting trust calibration in autonomous driving, in 'Proceedings of the 5th international conference on automotive user interfaces and interactive vehicular applications', pp. 210–217.
- Holt, C. A. and Laury, S. K. (2002), 'Risk aversion and incentive effects', *American Economic Review* 92(5), 1644–1655.
- Hong, C. Y., Lu, X. and Pan, J. (2020), 'Fintech adoption and household risk-taking', *NBER Working Paper* (28063).
- HSBC (2019), Trust in technology, Technical report.
- Jacovi, A., Marasović, A., Miller, T. and Goldberg, Y. (2020), 'Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in ai', *arXiv preprint arXiv:2010.07487*.
- Ji, M. (2017), 'Are robots good fiduciaries: Regulating robo-advisors under the investment advisers act of 1940', *Colum. L. Rev.* 117, 1543.
- Jordan, M. I. (2019a), 'Artificial intelligence -the revolution hasn't happened yet', *Harvard Data Science Review* 1(1).
- Jordan, M. I. (2019b), 'Dr. ai or: How i learned to stop worrying and love economics', *Harvard Data Science Review* 1(1).
- Kahneman, D., Knetsch, J. L. and Thaler, R. H. (1990), 'Experimental tests of the endowment effect and the coase theorem', *Journal of Political Economy* 98(6), 1325– 1348.
- Kapteyn, A. and Teppa, F. (2011), 'Subjective measures of risk aversion, fixed costs, and portfolio choice', *Journal of Economic Psychology* 32(4), 564–580.
- Koo, J., Kwac, J., Ju, W., Steinert, M., Leifer, L. and Nass, C. (2015), 'Why did my car just do that? explaining semi-autonomous driving actions to improve driver understanding, trust, and performance', *International Journal on Interactive Design and Manufacturing (IJDeM)* 9(4), 269–275.
- Kraft, H. and Munk, C. (2011), 'Optimal housing, consumption, and investment decisions over the life cycle', *Management Science* 57(6), 1025–1041.
- Lerner, J. and Tufano, P. (2011), 'The consequences of financial innovation: a counterfactual research agenda', *Annual Review of Financial Economics* 3(1), 41–85.
- Linciano, N. and Soccorso, P. (2012), 'Assessing investors' risk tolerance through a questionnaire', *Working Paper available at SSRN* (2207958).

- Linnainmaa, J. T., Melzer, B. T. and Previtero, A. (2020), 'The misguided beliefs of financial advisors', *Journal of Finance* .
- Lo, A. W. (2016), 'What is an index?', *Journal of Portfolio Management* 42(2), 21–36.
- Lo, A. W. (2019), 'Why artificial intelligence may not be as useful or as challenging as artificial stupidity', *Harvard Data Science Review* 1(1).
- Lopez, J. C., Babcic, S. and De La Ossa, A. (2015), 'Advice goes virtual: how new digital investment services are changing the wealth management landscape', *Journal of Financial Perspectives* 3(3).
- Lourenço, C. J., Dellaert, B. G. and Donkers, B. (2020), 'Whose algorithm says so: The relationships between type of firm, perceptions of trust and expertise, and the acceptance of financial robo-advice', *Journal of Interactive Marketing* 49, 107–124.
- Lusardi, A., Michaud, P.-C. and Mitchell, O. S. (2017), 'Optimal financial knowledge and wealth inequality', *Journal of Political Economy* 125(2), 431–477.
- Lusardi, A. and Mitchell, O. S. (2014), 'The economic importance of financial literacy: Theory and evidence', *Journal of Economic Literature* 52(1), 5–44.
- Mankiw, N. G. and Zeldes, S. P. (1991), 'The consumption of stockholders and nonstock-holders', *Journal of Financial Economics* 29(1), 97–112.
- Marinelli, N. and Mazzoli, C. (2010), 'Profiling investors with the mi-fid: current practice and future prospects', *Research Paper available at <https://www.ascosim.it/public/19Ric.pdf>*.
- Merton, R. (1971), 'Optimal portfolio and consumption rules in a continuous-time model', *Journal of Economic Theory* 3(4), 373–413.
- Merton, R. C. (2017), 'The future of robo-advisors', *Video available at <https://www.cnn.com/2017/11/05/mit-expert-robert-merton-on-the-future-of-robo-advisors.html>*
- Miller, T. (2019), 'Explanation in artificial intelligence: Insights from the social sciences', *Artificial Intelligence* 267, 1–38.
- Misztal, B. (2013), *Trust in modern societies: The search for the bases of social order*, John Wiley & Sons.
- Muir, B. M. and Moray, N. (1996), 'Trust in automation. part ii. experimental studies of trust and human intervention in a process control simulation', *Ergonomics* 39(3), 429–460.
- Mullainathan, S., Noeth, M. and Schoar, A. (2012), 'The market for financial advice: An audit study', *NBER Working Paper* (17929).
- Niszczoła, P. and Kaszas, D. (2020), 'Robo-investment aversion', *Plos one* 15(9), e0239277.
- Odean, T. (1999), 'Do investors trade too much?', *American Economic Review* 89(5), 1279–1298.
- Patel, K. and Lincoln, M. (2019), 'It's not magic: Weighing the risks of AI in financial services', *Centre for the Study of Financial Innovation Working Paper*.
- Philippon, T. (2019), 'On fintech and financial inclusion', *NBER Working Paper* (26330).
- Poursabzi-Sangdeh, F., Goldstein, D. G., Hofman, J. M., Vaughan, J. W. and Wallach, H. (2018), 'Manipulating and measuring model interpretability', *arXiv preprint arXiv:1802.07810* .
- Prahl, A. and Van Swol, L. (2017), 'Understanding algorithm aversion: When is advice from automation discounted?', *Journal of Forecasting* 36(6), 691–702.

Reher, M. and Sokolinski, S. (2020), 'Does fintech democratize investing?', *Working Paper available at SSRN* (3515707).

Reher, M. and Sun, C. (2019), 'Automated financial management: Diversification and account size flexibility', *Journal of Investment Management* 17(2), 1–13.

Robinette, P., Li, W., Allen, R., Howard, A. M. and Wagner, A. R. (2016), Overtrust of robots in emergency evacuation scenarios, in '2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)', IEEE, pp. 101–108.

Rossi, A. and Utkus, S. (2019a), 'The needs and wants in financial advise: Humans versus robo-advising', *George Washington University Working Paper* .

Rossi, A. and Utkus, S. (2019b), 'Who benefits from robo-advising', *Working Paper available at SSRN* (3552671).

Ryan, A., Trumbull, G. and Tufano, P. (2011), 'A brief postwar history of us consumer finance', *Business History Review* 85(03), 461–498.

Sapienza, P., Toldra-Simats, A. and Zingales, L. (2013), 'Understanding trust', *The Economic Journal* 123(573), 1313–1332.

SEC (2006), 'Questions advisers should ask while establishing or reviewing their compliance programs', Available at https://www.sec.gov/info/cco/advise_compliance_questions.htm.

SEC (2019), 'Commission interpretation regarding standard of conduct for investment advisors', Available at https://www.sec.gov/info/cco/adviser_compliance_questions.htm .

Sheridan, T. (1989), Trustworthiness of command and control systems, in 'Analysis, Design and Evaluation of Man–Machine Systems 1988', Elsevier, pp. 427–431.

Sheridan, T. B. (2019), 'Individual differences in attributes of trust in automation: measurement and application to system design', *Frontiers in Psychology* 10, 1117.

Strzelczyk, B. E. (2017), 'Rise of the machines: the legal implications for investor protection with the rise of robo-advisors', *DePaul Business and Commercial Law Journal* 16, 54.

Thakor, R. T. and Merton, R. C. (2018), 'Trust in lending', *NBER Working Paper* (24778).

Toreini, E., Aitken, M., Coopamootoo, K., Elliott, K., Zelaya, C. G. and van Moorsel, A. (2020), The relationship between trust in ai and trustworthy machine learning technologies, in 'Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency', pp. 272–283.

Vissing-Jorgensen, A. (2004), Perspectives on behavioral finance: Does "irrationality" disappear with wealth? evidence from expectations and actions, in M. Gertler and K. Rogoff, eds, 'NBER Macroeconomics Annual 2003, Volume 18', MIT Press, pp. 139–208.

Warren, G. J. (2019), 'Choosing and using utility functions in forming portfolios', *Financial Analysts Journal* 75(3), 39–69.

Weber, M., Weber, E. U. and Nosić, A. (2013), 'Who takes risks when and why: Determinants of changes in investor risk taking', *Review of Finance* 17(3), 847–883.

Institut Louis Bachelier

Palais Brongniart

28, place de la Bourse

75002 Paris

Tél. : +33 (0)1 73 01 93 40

Fax : +33 (0)1 73 01 93 28

contact@institutlouisbachelier.org



LABEX

Louis Bachelier