



## COMMISSION « SYSTÈME FINANCIER ET FINANCEMENT DE L'ÉCONOMIE »

---

Réunion du 19 mai 2022

PROJET D'AVIS .....	2
COMPTE RENDU DE LA RÉUNION .....	3
RAPPEL DE L'ORDRE DU JOUR .....	3
LISTE DES PARTICIPANTS .....	4
LISTE DES DOCUMENTS PRÉPARATOIRES ET COMPLÉMENTAIRES.....	28

---

## PROJET D'AVIS

---

### COMMISSION « SYSTÈME FINANCIER ET FINANCEMENT DE L'ÉCONOMIE » du 19 mai 2022

L'avis n° 5 du moyen terme en cours de la commission invite à « mesurer l'importance de la fintech ». Il souligne qu'une « nouvelle industrie financière, la fintech, s'est ainsi récemment développée qui déploie les TIC pour améliorer les activités financières. Elle englobe des dispositifs qui concurrencent les banques traditionnelles, avec par exemple les prêts entre particuliers, mais aussi les systèmes de paiement (paiement par mobile, vérification des transactions financières par exemple) et les crypto-actifs ». Force est de constater qu'il n'existe pas de définition légale de la fintech qui est une notion évolutive, dont les contours et la nécessaire prise en compte font l'objet de discussions tant pour ce qui concerne la rénovation en cours des nomenclatures d'activités et de produits qu'entre comptables nationaux, sans parler du Conseil de stabilité financière. De taille encore modeste, les volumes en jeu sont en forte croissance et sont surveillés de près par les banques centrales dès lors qu'ils sont susceptibles d'avoir un impact sur la stabilité financière et la transmission de la politique monétaire. Le suivi des fintechs est par ailleurs un des axes de la *Data Gaps Initiative III* du G20. La Banque de France a entrepris de se doter d'un outil dynamique de détection et de *monitoring* des fintechs pour répondre aux demandes des utilisateurs de données. D'autres pays mènent des expérimentations dans ce sens, notamment l'Allemagne.

S'agissant des plateformes de financement participatif par le prêt, où ce sont les prêteurs qui choisissent les prêts, la transparence est souhaitable sur les informations relatives aux prêts (Loan book) et autres instruments. La commission note que le niveau de transparence sur les données de crédit immobilier est supérieur aux Etats Unis.

La commission prend acte du fait que la fintech est un domaine non encore stabilisé, en pleine évolution, à l'origine de nouveaux outils en devenir. Elle considère que le travail mené au niveau international sur les nomenclatures constitue un enjeu majeur pour permettre à la statistique publique de bien identifier le développement des fintechs. La commission recommande aux parties prenantes de concevoir ces outils de sorte qu'ils alimentent, en aval, la statistique publique et que l'accès à ces nouvelles sources de données soit facilité pour les chercheurs. La commission souligne l'importance de pouvoir apparier les données qui étaient accessibles *via* l'ODR avec les données accessibles au CASD. Elle demande à être tenue régulièrement informée des avancées dans ce domaine.

Par ailleurs, suivant l'avis général de moyen terme du Cnis n° 2 sur la mesure des inégalités de toutes natures, la commission a réalisé un état des lieux des données permettant de documenter la place des femmes dans les métiers de la finance d'une part et l'accès des femmes au financement de leur activité professionnelle d'autre part. La commission note que les recherches sur le genre et la finance sont en plein développement et que les enjeux de données sont nombreux et méritent d'être approfondis. Elle encourage les appariements de données innovantes (commerciales et administratives) et l'accès aux données les plus désagrégées possible pour poursuivre l'exploration de ces sujets.

*Références : avis de moyen terme n° 5 de la commission « Système financier et financement de l'économie » et avis général de moyen terme n° 2*

---

COMPTES RENDUS DE LA RÉUNION DE  
LA COMMISSION « SYSTÈME FINANCIER ET FINANCEMENT DE L'ÉCONOMIE »

- 19 mai 2022 -

---

Présidente : Anne-Laure DELATTE, chargée de recherche au CNRS, rattachée au Leda, Paris Dauphine

Rapporteurs : Guillaume HOURIEZ, chef du Département des comptes nationaux à l'Insee  
Cécile GOLFIER, chargée de mission à la Direction des statistiques monétaires et financières (DSMF) à la Banque de France

Responsable de la commission : Stéphane TAGNANI (01 87 69 57 03)

*RAPPEL DE L'ORDRE DU JOUR*

INTRODUCTION.....	5
1. PRÉSENTATION DES NOUVEAUTÉS DU PROGRAMME DE TRAVAIL 2022 DES PRODUCTEURS DE LA STATISTIQUE PUBLIQUE .....	5
2. LES « FINTECHS », OU COMMENT APPRÉHENDER LA TRANSFORMATION NUMÉRIQUE DE L'ACTIVITÉ BANCAIRE ET FINANCIÈRE .....	7
3. FINANCE ET GENRE : QUELQUES APPROCHES EXPLORATOIRES .....	14

## Liste des participants

ALEXIS	Caroline	Cour des comptes
ALLAM	Délila	Université Paris 1 Panthéon-Sorbonne
BARDAJI	José	France Assureurs
BAYART	Pierre	Institut national de la statistique et des études économiques (Insee)
BECHETOILLE	Maylis	France Assureurs
BELLY	Guillaume	Banque de France
BERTHOLON	Raphaëlle	Confédération française de l'encadrement – Confédération générale des cadres (CFE-CGC)
BERTIAUX	Juliette	Chambre de commerce et d'industrie de Paris – Ile-de-France
BIARNAIX ROCHE	Véronique	Confédération française de l'encadrement – Confédération générale des cadres (CFE-CGC)
BIDAUD	Fabrice	Banque de France
BUSI	Paul	Confédération française démocratique du travail (CFDT)
BUTHIOT	Enzo	Banque de France – Région Ile-de-France
CAPELLE- BLANCARD	Gunther	Université Paris 1 Panthéon-Sorbonne
D'ALESSANDRO	Cristina	Conseil national de l'information statistique (Cnis)
DELATTE	Anne- Laure	Université Paris Dauphine
DUBOIS	Marie- Michèle	Conseil national de l'information statistique (Cnis)
DUC	Cindy	Institut national de la statistique et des études économiques (Insee) – Direction des statistiques d'entreprise (DSE)
DURAN	Patrice	Ecole normale supérieure
DUSSUD	François- Xavier	Institut national de la statistique et des études économiques (Insee) – Direction des statistiques d'entreprise (DSE)
GOLFIER	Cécile	Banque de France (BdF)
GUILLAUMAT- TAILLIET	François	Conseil national de l'information statistique (Cnis)
HAVRYLCHYK	Olena	Université Paris 1 Panthéon-Sorbonne
HEBERT	Camille	Université de Toronto
HOURIEZ	Guillaume	Institut national de la statistique et des études économiques (Insee) – Direction des études et des synthèses économiques (Dese)
JOUINI	Elyes	Université Paris Dauphine
LACAILLE	Yves	Union nationale des professions libérales (UNAPL)
LANTOINE	Perrine	Groupe BPCE
LAWNICZAK	Anne- Sophie	
LEVESQUE	Céline	Ubiquis
MAUREL	Françoise	Conseil national de l'information statistique (Cnis)
MOREAU	Sylvain	Institut national de la statistique et des études économiques (Insee) – Direction des statistiques d'entreprise (DSE)
POILON	Gwénaelle	Mouvement des entreprises de France (Medef)
TAGNANI	Stéphane	Conseil national de l'information statistique (Cnis)
TASSONE	Laurence	Banque publique d'investissement France (Bpifrance)
THEODOSE	Teddy	Université Paris 13
VALLA	Hugo	Fédération bancaire française (FBF)

La séance est ouverte à 14 heures 30, sous la présidence d'Anne-Laure DELATTE.

## INTRODUCTION

**Anne-Laure DELATTE, Présidente**

Bienvenue à cette nouvelle réunion de la Commission « Système financier et financement de l'économie » dont l'ordre du jour portera sur trois points : les nouveautés du programme annuel de travail pour 2022 présentés par les rapporteurs de la commission à la Banque de France et à l'Insee, à savoir Cécile Golfier et Guillaume Houriez, les « fintechs » et la thématique liant finance et genre. Pour les deux derniers points, nous avons invité des intervenants de la Banque de France et des chercheurs que je remercie pour leur participation.

### 1. PRÉSENTATION DES NOUVEAUTÉS DU PROGRAMME DE TRAVAIL 2022 DES PRODUCTEURS DE LA STATISTIQUE PUBLIQUE

**Cécile GOLFIER, Banque de France**

Je propose de présenter les principaux éléments du programme statistique de la Banque de France et de l'Autorité de contrôle prudentiel et de résolution (ACPR).

Pour ce qui concerne la Banque de France, les nouveautés du programme de travail concernent les points suivants : les moyens de paiement, les statistiques monétaires et financières, la balance des paiements et la position extérieure, la thématique activité-immobilier-inflation, la diffusion statistique et *Data Gaps Initiative* (phase III).

Dans le domaine des moyens de paiement, la directive sur les services de paiement de 2015 (DSP2) avait élargi les attentes en matière de statistiques sur les moyens de paiement. La collecte semestrielle sur les moyens de paiement scripturaux menée par la Banque de France a évolué le 1<sup>er</sup> janvier 2022 pour prendre en compte le règlement (UE) n° 2020/59 relatif aux statistiques de paiement. Cette collecte est enrichie par l'intégration d'éléments sur les paiements instantanés (paiements sans contact, paiement mobile, etc.). Cette évolution vient aussi enrichir les données existantes sur les cartes de paiement : elle introduit le code marchand qui indique le bien ou le service vendu par le bénéficiaire du paiement et ventile les données par réseau de carte (VISA, Mastercard, etc.) et elle introduit des statistiques sur les opérations de paiement frauduleuses en précisant l'origine de la fraude.

Dans le domaine des statistiques monétaires et financières, deux grandes rétroprojections sont au programme. La première concernera les comptes nationaux financiers jusqu'à 1978 tandis que la deuxième portera sur les statistiques monétaires jusqu'en 1945.

Les statistiques sur le secteur de l'assurance et les fonds de pension évolueront avec le nouveau classement des organismes de retraite professionnelle supplémentaire (ORPS) qui étaient classés dans le secteur assurance et qui rejoignent le secteur des fonds de pension. Il est à noter que, en comptabilité nationale, les ORPS demeurent classés dans le secteur des assurances jusqu'au changement de base de 2024. Depuis le 11 mars 2022, la Banque de France diffuse sur son site Internet un référentiel des fonds d'investissement français labellisés par les cinq labels suivants : ISR, Greenfin, Relance, Finansol et CIES.

Dans le domaine de la balance des paiements et de la position extérieure, la Banque de France travaillera sur l'élaboration de statistiques sur les échanges internationaux de services par catégorie d'entreprises. Cette démarche est initiée par Eurostat : il s'agit de la démarche STEC. 2022 sera l'année des travaux préparatoires mais aussi du début de la collecte puisque la première diffusion est prévue pour le millésime 2022 (en 2024). En 2022, une réflexion sera menée sur la bascule de l'unité légale à l'entreprise. Il sera également introduit des statistiques par mode d'offre selon la terminologie du GATS pour une première diffusion prévue pour le millésime 2023 (en 2025). En 2022, la Banque de France et l'Insee collaboreront en vue de l'utilisation par la Banque de France des données Ofats pour produire ces statistiques.

D'autres travaux concerneront les investissements directs étrangers (IDE). Un chantier portera sur l'identification des *Special Purpose Entities* (SPEs). Jusqu'à présent, les SPEs n'étaient pas identifiées en

5

France. Un deuxième chantier portera sur la ventilation des IDE par finalité. Ce travail vise notamment à mettre en évidence les investissements *Greenfield* (qui désignent la création de capacités de production).

D'autres travaux concerneront la mesure de l'activité, l'immobilier et la mesure de l'inflation. Durant la pandémie, la Banque de France avait effectué une mesure de l'activité en s'appuyant sur le lien entre activité et pollution au dioxyde d'azote. Les travaux menés en 2022 viseront à évaluer la pertinence de cette méthode en dehors de périodes de forts retournements mais et à étendre cette méthode aux pays émergents et en développement.

Dans le domaine de l'immobilier, les travaux statistiques porteront sur le suivi des prix et loyers par *webscraping* sur les sites immobiliers, sur le suivi des zones bâties et non bâties au sol à partir d'images satellite et sur le suivi de l'activité des cimenteries à partir des relevés infrarouges des fours sur images satellite. L'objectif de cette démarche est de présenter des statistiques complémentaires en 2022 pour enrichir les statistiques officielles. Sur ce domaine, des travaux viseront également à mieux couvrir l'activité des pays émergents et en développement.

Les travaux viseront enfin à enrichir la mesure de l'inflation grâce à une analyse textuelle menée sur les réseaux sociaux et la presse.

La Banque de France s'est engagée dans la phase 3 de la *Data Gaps Initiative* du G20 en étroite collaboration avec l'Insee. Cette initiative vise quatre grands thèmes émergents : l'accès aux données privées et administratives et le partage des données ; le changement climatique ; l'information distributionnelle sur les ménagers ; les fintechs et l'inclusion financière. Les deux premières phases avaient été lancées en 2009 et 2015 et portaient sur des problématiques financières, suite à la crise financière de 2008 – la seconde phase, en 2015, a notamment introduit la problématique des interconnexions financières entre pays. La phase 3 couvre donc un registre différent.

Concernant la diffusion statistique, il est à noter que l'*Open Data Room* (ODR) de la Banque de France a fusionné avec le Centre d'accès sécurisé aux données (CASD). Cette fusion a permis d'améliorer l'appariement des données individuelles.

Le nouveau site Internet de la Banque de France sera accessible en novembre 2022.

Du côté de l'ACPR, l'autorité déploiera en 2022 la taxinomie du nouveau système de *reporting* des banques et assimilés [Ruba qui a succédé au Système unifié de *reporting* financier (Surfi)]. Par ailleurs, le LEI (*Legal Entity Identifier*) remplace désormais le code interbancaire (CIB).

Dans le domaine immobilier, l'ACPR revoit son enquête portant sur le financement des professionnels de l'immobilier et ajuste le *reporting* Credithab relatif au crédit à l'habitat. L'ACPR réouvrera également son rapport statistique et mettra à disposition ses séries statistiques sur un portail Internet.

## **Guillaume HOURIEZ, Insee**

Mon intervention couvre le programme de travail du département des études et statistiques fiscales de la Direction générale des finances publiques (DGFIP) et du département des comptes nationaux de l'Insee.

Deux nouveautés sont à signaler en 2022 pour ce qui concerne la DGFIP : d'une part, l'exploitation des données relatives au prélèvement à la source qui fournissent des informations à fréquence mensuelle sur les revenus imposables des foyers fiscaux et qui constituent une source prometteuse d'études et de production statistique ; d'autre part, l'appariement des déclarations fiscales des professionnels et des particuliers. Ce deuxième chantier sera également prometteur, par exemple pour travailler sur les patrimoines professionnels.

Du côté du département des comptes nationaux, l'année 2022 sera essentiellement consacrée au changement de base. Celle-ci sera publiée pour la première fois en mai 2024.

Les résultats détaillés du premier trimestre 2022 seront publiés sur une nouvelle application. Une partie des comptes trimestriels pourront ensuite être proposés en *open source*. Les comptes annuels ont également changé d'outils.

## 2. LES « FINTECHS », OU COMMENT APPRÉHENDER LA TRANSFORMATION NUMÉRIQUE DE L'ACTIVITÉ BANCAIRE ET FINANCIÈRE

Anne-Laure DELATTE

Nous allons commencer par deux présentations de la Banque de France avant d'écouter Olena Havrylchyk faire un point sur les travaux de la recherche sur les fintechs et sur les éléments qui peuvent manquer encore à la recherche.

### 2.1. Distinction des « fintechs » dans les nomenclatures et dans les statistiques macro-économiques

Fabrice BIDAUD, Banque de France

Il n'existe pas de définition normée des fintechs qui fasse consensus ou qui ait valeur légale. La définition des fintechs selon le Conseil de stabilité financière (*Financial Stability Board* ou FSB) est la suivante : « *L'innovation technologique dans les services financiers qui pourrait déboucher sur de nouveaux modèles d'entreprise, applications, processus ou produits, avec un effet matériel associé sur les institutions et marchés financiers et sur la fourniture de services financiers* ». Cette définition complexe tente d'embrasser largement. Les termes « *innovation* », « *technologique* » et « *nouveaux* » laissent entrevoir qu'il ne sera pas vraiment question d'un périmètre ou d'une classification stable dans le temps mais que ce qui était innovant ou nouveau à une époque donnée ne le sera peut-être plus à l'avenir.

En dépit des difficultés à poser un périmètre stable, il est possible de distinguer trois grandes familles de fintechs.

La première catégorie englobe les start-ups ou nouvelles entreprises qui opèrent principalement dans le secteur des services financiers. Ces entreprises produisent des innovations dans divers segments : services de paiement, services de compensation et de règlement, crédit (néobanques), assurance (assurtech), gestion d'actifs (conseil en investissement, *robo-advisers*) et également activités liées aux cryptoactifs.

La deuxième catégorie regroupe les grandes entreprises, souvent technologiques, aussi dénommées *big tech*, qui offrent des services financiers dans le cadre d'activités beaucoup plus vastes.

La troisième catégorie couvre les institutions financières traditionnelles qui modifient leurs modèles d'affaires pour ne pas se laisser distancer par la concurrence et qui peuvent ainsi proposer des services à contenu innovant. Ces institutions financières peuvent proposer ces nouveautés en faisant appel à des équipes internes ou en partenariat.

Sont traditionnellement exclus de ce périmètre les « Tech facilitators », c'est-à-dire les entités en charge de construire l'infrastructure informatique pour les fintechs ou les fournisseurs de services (serveurs, cloud, outils d'intelligence artificielle...).

Le sujet de l'inscription des fintechs dans les nomenclatures est un sujet d'actualité puisque des discussions internationales sont ouvertes sur l'intégration de la notion de fintechs dans les manuels de statistiques macroéconomiques (comme le SNA pour la comptabilité nationale ou le BPM pour la balance des paiements) ou dans les nomenclatures relatives à l'activité économique (comme la CITI). Les versions actuelles de ces classifications datent de plus de dix ans et ne couvrent donc pas ces entreprises.

L'enjeu principal de ces discussions est de déterminer si les fintechs opèrent sur des activités nouvelles nécessitant de faire évoluer la nomenclature pour les prendre en compte, en proposant de nouvelles définitions ou une nouvelle ventilation, ou si les fintechs proposent les mêmes services que les sociétés financières traditionnelles et devraient alors pouvoir s'insérer dans les nomenclatures existantes. L'intégration des fintechs dépendra aussi de la maille de la nomenclature : si cette maille est lâche, il sera plus aisé d'y intégrer ces nouvelles entités tandis qu'une description plus fine conduira à poser davantage de questions. Les enjeux dépendront également de la taille des fintechs et de leur dynamique de croissance. En effet, si la taille de ces structures est modeste, créer une nomenclature dédiée pourrait ne pas être une décision pertinente. Cependant, les manuels doivent aussi être en mesure de prendre en considération des entités dont la croissance est importante et qui pourraient se développer fortement.

Auquel cas, les prendre en compte dans les nomenclatures serait pertinent sans avoir à attendre une prochaine révision des manuels.

Les manuels internationaux de comptabilité nationale (SNA sous l'égide de l'ONU et BPM sous l'égide du FMI) sont en cours de révision et ces travaux devraient aboutir en 2025. Des *task teams* travaillent actuellement sur différentes problématiques et soumettent leurs travaux à consultation publique. Une *guidance note* a déjà paru sur les fintechs et semble obtenir consensus. Dans cette note, il est précisé que les fintechs ne constituent pas un secteur de comptabilité nationale à distinguer, considérant qu'elles opèrent dans des activités qui rejoignent des sous-secteurs existants. Il en ressort que les sous-secteurs financiers (banques, sociétés d'assurance, auxiliaires financiers...) restent à une maille suffisamment haute pour pouvoir englober les fintech sans difficulté. La question qui reste à trancher demeure de déterminer s'il est pertinent de faire figurer des lignes de décomposition « dont fintechs » par rapport à l'existant.

Les discussions sur la révision des manuels sont animées concernant la prise en considération des cryptoactifs. Ces discussions ne portent pas tant sur le classement des entités, puisqu'une plate-forme d'échange de cryptoactifs pourrait être classée parmi les auxiliaires financiers, mais plutôt sur le classement des cryptoactifs en tant qu'instruments (financiers, non financiers, hybrides...).

Concernant la nomenclature des activités économiques, des discussions sont en cours pour les nomenclatures CITI au niveau international et NACE au niveau européen. En 2021, ces discussions ont porté sur le niveau groupe (codes à 3 chiffres). En 2022, ces travaux porteront sur les classes (codes à 4 chiffres). Au niveau groupe, les fintechs ne bénéficieraient pas d'une catégorisation nouvelle. Comme pour les statistiques macroéconomiques, un compromis reste à trouver entre la taille encore relativement modeste de certains types d'activités des fintechs (qui ne pousse pas à la création d'une nouvelle catégorie) et la dynamique de croissance de ce secteur ainsi que l'importance du sujet pour les banques centrales (notamment au titre de leur mandat de stabilité financière puisque certaines fintechs ne sont pas régulées, ce qui soulève des enjeux d'identification et de suivi, et pour l'analyse de la transmission de la politique monétaire). Cependant, le niveau de détail de la nomenclature CITI est bien plus fin que celui des statistiques macro-économiques et cela nécessite de se poser bien plus de questions. Pour approfondir le sujet de l'enjeu des fintechs pour les banques centrales, je vous renvoie vers le rapport *ad hoc* de l'*Irving Fisher Committee* qui pousse à opérer une distinction dans la nomenclature pour les fintechs. Par ailleurs, pour rappel, le suivi des fintechs est l'un des quatre axes de *la Data Gaps Initiative III* du G20.

La Banque de France remercie l'Insee qui porte les préoccupations françaises au niveau international. Ces pourparlers n'empêchent pas de prendre des décisions à la maille nationale, mais un consensus plus large sur la terminologie permettra de comparer plus facilement les données. Les projets de collecte des données sont également plus aisés avec un encadrement par des règlements ou normes internationales.

Les banques centrales ont proposé de possibles nouvelles ventilations sur la nomenclature CITI. A titre d'exemples, les *Pay Tech* (paiements mobiles, portefeuilles numériques, paiements de pair à pair...) pourraient relever d'une sous-ventilation du secteur 66.1. Cependant, cette proposition peut encore évoluer. Les discussions avaient également conduit à proposer que les activités liées aux cryptoactifs (émission, activités de validation et d'extraction, services de paiement, etc.) ainsi que le financement participatif (par don ou par recherche de profit, *via* des prêts ou des actions, etc.) fassent l'objet d'une catégorie distincte. D'autres catégories ont également fait l'objet de questionnements : « Robo-financial », « Banktech », « Insurtech », etc.

Cet état des lieux des discussions en cours montrent qu'un consensus reste à trouver entre les partisans d'une vision prudente, pour ne pas céder aux effets de mode, et ceux qui portent une vision proactive.

**Anne-Laure DELATTE**

Merci beaucoup. Je vous propose que nos échanges aient lieu en fin de présentation.

## **2.2. Identification des « fintechs » par l'intelligence artificielle**

**Guillaume BELLY, Banque de France**

L'objectif de la présentation porte sur les moyens de détecter les fintechs dans la mesure où les fintechs ne sont pas clairement identifiées dans les classifications de comptabilité nationale. Par exemple, il n'existe pas



de catégorie « fintechs » dans les classifications ISIC ou NACE et ces entreprises sont plutôt logées dans des classes liées à la programmation informatique ou au développement de logiciel ce qui ne correspond pas à leur véritable activité. De plus, ce secteur est marqué par une très forte innovation, ce qui rend d'autant plus compliqué sa définition. Mieux les détecter permet de répondre aux demandes des utilisateurs en vue de disposer de jeux de données de référence sur un domaine innovant du secteur financier. Pour cela, l'objectif est de disposer d'un outil permettant de mettre à jour la liste des fintechs, de les identifier et de les suivre.

Dans cet objectif, il est proposé de mettre en place un algorithme d'intelligence artificielle entraîné à reconnaître les fintechs. La démarche n'est pas de s'appuyer sur des sous-classes de fintechs mais de partir d'une liste de fintechs déterminée à dire d'expert et d'une importante base de données des entreprises françaises pour obtenir une base de données des fintechs en France qui puisse se rafraîchir de manière fréquente. L'algorithme semi-supervisé d'intelligence artificielle pourra ainsi apprendre la typologie d'une fintechs et être en mesure de détecter ces entreprises. Les tests de performance seront effectués sur un jeu de fintechs et d'entreprises non-fintechs qui n'aura jamais été soumis à l'algorithme.

Une fois que la liste des fintechs connues a été fournie par les experts, des variables devront être attribuées pour les caractériser. Pour cela, nous avons eu recours à un prestataire externe dans le domaine de l'intelligence économique qui dispose d'une base semi-structurée de type *Big Data* qui regroupe un peu plus de 10 millions d'entreprises. Nous en avons retenu un sous-ensemble de 10 000 entreprises non-fintechs que nous avons associé aux 350 fintechs que nous avons sélectionnées dans un premier temps. Puis, nous avons utilisé les 1 400 variables disponibles dans la base. Ces variables sont issues de bases de données légales utilisées par *webscrapping* ou issues d'autres bases de données (comme les articles de presse). Le polymorphisme de ces variables est très riche puisque la base propose des variables binaires, comme des dates de création, des informations sur les propriétaires des entreprises, des informations sur les métiers occupés dans ces entreprises, la description de l'activité, des articles de presse (titre, résumé, etc.), des données numériques, la longitude et la latitude des établissements et des codes (NAF, etc.). Pour chaque entreprise, ces variables seront exploitées comme vecteur de caractéristique, vecteur qui sera ensuite appris par l'algorithme.

A partir de ces 1 400 variables, trois grands jeux ont été définis : un jeu initial de 52 variables, un jeu intermédiaire de 81 variables et un jeu complet de 147 variables. Entre le jeu intermédiaire et le jeu complet, nous avons ajouté des données financières.

Sur cette base, deux algorithmes ont été préparés : un algorithme standard et un algorithme modifié.

Dans le graphique qui figure au dossier, chaque entreprise est représentée par un point. Les points rouges représentent les entreprises non-fintechs correctement détectées. Les points verts sont les fintechs correctement détectées. Les points bleus sont les faux positifs (fintechs non détectées) et les points noirs les faux négatifs (non-fintechs détectées comme fintechs). Avec l'algorithme standard, c'est le jeu intermédiaire qui affiche les meilleures performances. L'apport des données financières entre le jeu intermédiaire et le jeu complet vient finalement déstabiliser l'algorithme, car les données financières sont très peu renseignées. Par ailleurs, la précision reste faible en raison du déséquilibre au sein du jeu de données qui se caractérise par peu de fintechs et beaucoup d'entreprises non-fintechs.

Un deuxième algorithme a donc été mis en place avec stratégie de génération synthétique qui correspond plus ou moins à un *sur-sampling* conduisant à ajouter de nouvelles fintechs qui possèdent des caractéristiques équivalentes celles de la base originelle sans pour autant être exactement les mêmes. Cette stratégie permet de mieux entraîner l'algorithme mais aussi de mieux calculer les indicateurs de performance. Cet algorithme affiche des indicateurs très satisfaisants (supérieurs à 99 %).

De ces travaux, il ressort que les variables qui permettent de prédire quelles sont les fintechs sont issues des articles de presse (thème, description, source), des variables administratives (code d'activité, taille), des variables liées au type d'emploi et de compétences dans l'entreprise (par exemple la présence de *data scientists*, etc.), des noms des personnes ou des fonds au Conseil d'administration et des éléments de la marque déposée.

Une des grandes difficultés est de prendre en compte le volume, la variété et la qualité des données. Avec un historique de trois ans pour les 10 000 entreprises de la base, il faut en effet manipuler des données de plusieurs milliers de giga-octets. L'obtention de toutes les données sur les entreprises est une étape

préliminaire, extrêmement importante, qui requiert beaucoup de temps et de moyens, raison pour laquelle il a été fait appel à un agrégateur privé. Pour aller plus loin dans la démarche, par exemple au niveau européen, il faudrait disposer d'infrastructures très robustes compte tenu du volume de données à traiter.

### **Enzo BUTHIOT, Banque de France**

Sous réserve de l'obtention d'un budget, le projet de la Banque de France serait de lancer un outil en production permettant de déployer l'algorithme sur une base plus large mais aussi de tester l'aspect dynamique de l'outil, c'est-à-dire sa capacité à détecter de nouvelles fintechs, à être mis à jour facilement et à découvrir de nouvelles variables, même dans des domaines plus lointains que ceux initialement identifiés. En fonction des besoins, il pourrait aussi être créé une interface utilisateur qui remonterait les informations les plus utiles pour le *monitoring* des fintechs. Cette interface pourrait prendre une forme simple comme un fichier Excel précisant les variables les plus importantes. Éventuellement, il pourrait être aussi intéressant d'associer de nouvelles données (par exemple des données financières, des données disponibles à la Banque de France ou dans d'autres instituts publics) permettant ainsi de ne pas avoir recours à un prestataire externe, et donc de réduire les coûts. Cette démarche pourrait aussi être dupliquée dans d'autres organisations.

Une extension de ce projet pourrait être envisagée *via* une mise à l'échelle de l'algorithme au niveau européen par l'intermédiaire d'un réseau de banques centrales. Ce réseau permettrait de spécifier la structure de la base de données et des sources, ce qui garantirait une pérennité des sources de données et des méthodes. A ce sujet, il convient de noter que la Bundesbank, avec l'aide d'une start-up berlinoise, a conduit une expérimentation avec une approche par graphe, démarche qui a obtenu de très bons résultats sur les fintechs allemandes. Cette méthode pourrait être complémentaire à celle de la Banque de France, puisque la Bundesbank utilise une méthode de *webscraping* tandis que la Banque de France utilise une méthode *flat*.

Un réseau européen ou international permettrait également la mise en place de registres nationaux interconnectés sur les fintechs. Ces registres permettraient de comparer les données, comparaison qui serait d'autant plus utile que les *business models* des fintechs sont souvent internationaux.

Ce mode de création de statistiques publiques pourrait ensuite donner lieu à d'autres applications

### **2.3. Les données de la recherche pour comprendre les « fintechs »**

#### **Olena HAVRYLCHYK, Université Paris 1 Panthéon-Sorbonne**

Mon exposé portera sur les données utilisées par les chercheurs pour leurs travaux sur les fintechs. Ma présentation sera centrée sur les fintechs qui proposent des prêts aux particuliers ou aux petites et moyennes entreprises (PME) mais englobera également les *big techs* car la plupart des prêts ne sont pas accordés par les fintechs mais par des *big techs*.

Il existe aujourd'hui plusieurs plates-formes participatives par le prêt. Les chercheurs qui étudient ce nouveau modèle d'intermédiation financière exploitent les données de deux plates-formes aux États-Unis (*Lending Club et Prosper*) et une plate-forme en Chine (*Renrendai*).

Le concept du financement participatif suppose une grande transparence puisque ce sont les prêteurs eux-mêmes qui choisissent les projets auxquels ils souhaitent participer. La plate-forme doit donc fournir un maximum d'informations. Aux États-Unis, les plates-formes publient historiquement des *loan books* qui fournissent des informations très détaillées sur les emprunteurs. Les chercheurs qui étudient ce champ peuvent s'inscrire sur la plate-forme en tant que prêteur potentiel et télécharger le *loan book* qui décrit l'emprunteur (photo, texte de l'annonce, adresse, revenus, Fico score, utilisation du prêt, etc.). Des informations supplémentaires sont à la disposition des seuls chercheurs mais celles-ci sont plus difficiles à obtenir. Ces informations supplémentaires portent sur le lien entre les prêteurs et les emprunteurs et précisent si le prêteur est un particulier ou institutionnel (banque, *hedge fund*, fonds de pension, etc.).

Ce domaine a suscité de nombreux projets de recherche qui montrent que les investisseurs institutionnels ont un rendement plus important que les investisseurs individuels, qu'il existe une discrimination à l'endroit des emprunteurs femmes et des emprunteurs issus des minorités ethniques, qu'il existe une discrimination en faveur des personnes belles. De nombreuses recherches ont également porté sur les liens sociaux et ont

démontré que les rendements étaient plus élevés lorsque les prêteurs finançaient leurs amis. D'autres travaux de recherche ont porté sur le texte qui accompagne le projet de prêt et ont démontré que les textes simples rédigés par les emprunteurs sont associés à un taux de défaut plus faible. Il ressort aussi que les plates-formes sont utilisées comme un outil d'arbitrage réglementaire. Ceci ne vaut pas que pour les plates-formes de financement participatif, mais également pour d'autres fintechs.

La richesse des informations disponibles a été exploitée pour alimenter de nombreux travaux de recherche, cependant il est noté une évolution vers une moindre transparence depuis que la part des investisseurs institutionnels a augmenté. *Lending Club* n'accepte plus les investisseurs individuels depuis 2020. Les plates-formes de financement participatif ne méritent donc plus leur appellation. Du fait de ce virage, il n'est plus nécessaire de publier des *loan books*. Au Royaume-Uni, la plate-forme *Funding Circle* a aussi changé de *business model* et ne permet plus aux investisseurs individuels de choisir les prêts.

En France, les plates-formes de financement participatif ne peuvent concerner que les PME. Les plates-formes de prêts en direction des particuliers ne sont pas autorisées. Uniquement *October* et *PretUp* publient les *loan books* sur leurs sites. D'autres plates-formes françaises peuvent envoyer des informations sur demande, mais le processus d'obtention des informations est plus difficile et les données sont incomplètes. Une des plates-formes que j'ai sollicitées m'a aussi répondu que l'envoi des données représentait un coût pour elle, ce qui peut sembler surprenant.

Au-delà des informations sollicitées directement auprès des plates-formes, il est aussi possible de récupérer ces données *via* la Banque de France. Les informations fournies à la banque centrale sont le numéro de Siren, le montant du prêt, le taux d'intérêt et la date. En revanche, il n'est communiqué aucune donnée sur le défaut de paiement ou le remboursement, par exemple. Il était jusqu'alors possible d'accéder à ces données *via Open Data Room* et maintenant *via* le CASD (Centre d'accès sécurisé aux données). Cependant, les données du CASD, si elles peuvent être appariées à celles de la Banque de France, ne peuvent pas être appariées à d'autres données, ce qui représente une gêne pour les chercheurs. Pour améliorer la transparence, une piste serait que la Banque de France collecte mieux les données et les partage avec les chercheurs, d'autant que ces données ne sont pas confidentielles. En effet, les données sur le défaut de paiement, par exemple, sont communiquées aux prêteurs. Une autre piste serait d'obliger les plates-formes de financement participatif à publier les *loan books* sur leurs sites.

Après ce premier tour d'horizon sur les plates-formes de prêt participatif, je souhaiterais évoquer devant vous le *shadow fintech banking*. Il existe une grande transparence sur les données de tous les crédits immobiliers aux États-Unis avec une information individuelle sur la demande de crédit, le crédit accordé, le montant, les taux d'intérêt, la maturité du crédit, etc. Les informations disponibles portent aussi sur le profil du prêteur (nom). Cette information permet de déterminer si le prêteur est une banque, une *shadow bank* ou une fintech. Les travaux menés par les chercheurs dans ce domaine ont mis en évidence l'utilisation de ces données pour des arbitrages réglementaires et ont montré qu'il existait une discrimination des minorités par les fintechs comme par les non-fintechs.

Je terminerai mon exposé par un zoom sur les *big techs* qui utilisent, comme d'autres entreprises, les données alternatives que sont le *digital footprint* et les données de paiement. Par exemple, le *digital footprint* peut permettre de prédire le taux de défaut d'un particulier. Ainsi, il a été démontré que l'utilisation d'une adresse mail Yahoo *versus* une adresse T-online donnait lieu à un taux de défaut plus important. Le taux de défaut peut même être affiné en utilisant deux variables.

Les fintechs ont promis plus de transparence dans l'intermédiation financière, mais cette transparence est en train de disparaître à la suite d'un changement de *business model* et à l'entrée des GAFAs qui disposent d'informations confidentielles auxquelles les chercheurs ont rarement accès. Les *big techs* ont donc augmenté l'opacité des relations financières avec des conséquences sur la concurrence, la régulation et la recherche.

## Échanges

**Anne-Laure DELATTE**

Merci pour ces trois présentations.

## **Cécile GOLFIER**

Dans sa présentation, Olena Havrylchuk a mentionné la difficulté à utiliser les données qui transitent par la Banque de France et l'ODR. Pourriez-vous nous préciser pour quels « jeux de données » vous y avez été confrontée, afin que nous puissions l'analyser ?

## **Olena HAVRYLCHYK**

La Banque de France anonymise toutes les données, ce que je comprends parfaitement, mais la logique du CASD est différente. Le Centre d'accès sécurisé aux données n'anonymise pas les données accessibles aux chercheurs, c'est-à-dire qu'une confiance est accordée aux chercheurs dans ce cadre. Je ne propose donc pas de modifier la politique de la Banque de France mais préconise que les plates-formes soient dans l'obligation de publier leurs données non confidentielles pour les appairer avec les données du CASD qui sont disponibles, notamment les liasses fiscales. Cette possibilité permettrait de mener des recherches sur les entreprises qui s'endettent auprès des plates-formes de prêt participatif. Pour cela, nous n'aurions pas besoin des données de la Banque de France.

## **Cécile GOLFIER**

Il est exact que l'adresse des entreprises n'est pas une donnée confidentielle. Cependant, si cette adresse figure dans un jeu de données confidentiel, elle le deviendra également, car son affichage en clair est incompatible avec la confidentialité des données.

Pour ce qui concerne votre demande, c'est au CASD de procéder à l'appariement et de vous restituer les données. Il n'existe pas une volonté de la Banque de France de masquer des données qui sont publiques par ailleurs.

## **Olena HAVRYLCHYK**

Bien entendu, ce n'est pas une volonté. Cependant, les données de la Banque de France ne sont pas en l'état utilisables pour les chercheurs alors qu'elles existent et qu'elles ne sont pas confidentielles. C'est fort dommage.

## **Anne-Laure DELATTE**

Il me semble que cette demande va au-delà du sujet particulier des fintechs car le rapprochement CASD et ODR pose un enjeu de pseudonymisation pour l'ensemble des données.

## **Olena HAVRYLCHYK**

En effet, la difficulté couvre toutes les données de la Banque de France qui ne peuvent pas être appariées à d'autres données accessibles au niveau du CASD et qui ne sont pas des données de la Banque de France. À titre d'exemple, je pourrai évoquer ici les lignes de crédit des entreprises.

## **Anne-Laure DELATTE**

Nous serons confrontés à la même question sur la thématique de la finance et du genre. Il me semble donc important de la soulever, car le rapprochement CASD-ODR n'est pas un moyen de développer la recherche tant que ce point de blocage perdure. C'est un point qu'il nous faudrait discuter avec la Banque de France car la pseudonymisation ne permet pas d'appairer les données.

## **Cécile GOLFIER**

Le rapprochement CASD-ODR visait précisément à permettre l'appariement de données produites par la Banque de France – accessibles uniquement *via* l'ODR – avec les données mises à disposition des chercheurs *via* le CASD.

## **Anne-Laure DELATTE**

Je pense que le problème concerne toutes les données pseudonymisées, c'est-à-dire toute la base ODR.

## **Olena HAVRYLCHYK**

Les chercheurs n'ont pas besoin des données nominatives et la pseudonymisation ne nous pose pas problème.

## **Cécile GOLFIER**

Je crois savoir que le CASD procède à l'appariement mais je dois encore le vérifier auprès des sachants.

## **Anne-Laure DELATTE**

Si c'est le cas, le problème est résolu. Olena, d'où vient l'information que l'appariement est impossible ?

## **Olena HAVRYLCHYK**

Lorsque je demande l'appariement de données du CASD avec d'autres données, il m'est répondu que c'est impossible pour le moment.

## **Cécile GOLFIER**

Si c'est pour le moment, je suis dans une certaine mesure rassurée car je pense que c'est une question de temps. Lors de la présentation du projet par Jérémie Montornès, celui-ci avait précisé que la montée en puissance serait progressive. Cependant, l'objet même de cette initiative est bien de favoriser les appariements. Je suis donc surprise par vos retours. Toutefois, je comprends maintenant que c'est une question de temps, mais le point reste à creuser.

## **Françoise MAUREL, Cnis**

Dans la mesure où nous n'avons pas de spécialistes de la question dans l'assemblée, je préconise que nous nous montrions prudents dans la rédaction de notre avis. Les représentants du Cnis peuvent toutefois porter ce message auprès des autorités compétentes. La Banque de France peut également relayer ces messages vers des experts. En revanche, le Cnis n'est pas impliqué dans le rapprochement CASD-ODR et ne peut pas se montrer trop précis, au moins à ce stade. De plus, le Cnis ne rentre pas en général dans les aspects techniques, mais se contente d'exprimer un besoin. Or le besoin est que les données ne soient pas tronquées ou trop pseudonymisées afin que les travaux de recherche puissent être menés. Par ailleurs, le Cnis couvre le champ de la statistique publique en général et ne se focalise pas sur les données dont les chercheurs ont besoin. Nos travaux doivent donc porter sur la transparence des données en général pour tous les publics et non porter uniquement sur les données confidentielles dont l'usage est limité à certains publics.

## **Anne-Laure DELATTE**

Je souhaiterais poser une question sur la présentation de la Banque de France. Les algorithmes que vous nous avez présentés complètent-ils une collecte de données existante ? Ce projet vise-t-il à remplacer les formulaires que les fintechs auraient à renseigner ? Pourquoi avons-nous besoin de ces algorithmes ? Par ailleurs, pourquoi établissez-vous des algorithmes pour les fintechs alors que ces algorithmes ne visent pas d'autres entreprises ? En posant ces questions, je souhaite m'assurer que nous ne lançons pas un travail très coûteux de collecte de données pour alléger la charge des fintechs alors que l'obligation de renseignement porterait sur les entreprises non-fintechs.

## **Guillaume BELLY**

Comme nous avons du mal à identifier les fintechs puisqu'elles ne sont pas classifiées, il convient de les détecter de manière différente pour suivre le secteur et c'est tout l'objet de l'algorithme sur lequel nous travaillons. Ce travail doit permettre de les identifier sans avoir besoin d'une classification comme un code NAF, par exemple.

Nos travaux ont utilisé une base externe d'un prestataire car nous n'avons pas de données sur les fintechs. Cependant, cette base concerne toutes les entreprises, c'est-à-dire que nous disposerons des mêmes variables pour toutes les entreprises. Cette base peut être utilisée par un directeur marketing pour élaborer un plan de développement économique et identifier des prospects. Toutes les entreprises françaises sont

identifiées dans cette base, pas d'un point de vue Banque de France mais en fonction de l'activité des dites entreprises. Pour nous, il était intéressant d'utiliser ces données non collectées par la Banque de France pour vérifier si nous pouvions classer les fintechs et les suivre car la Banque de France ne dispose pas de données suffisantes pour monitorer les fintechs.

Avec l'algorithme sur lequel nous travaillons, nous visons en fait à classer les fintechs.

### **Enzo BUTHIOT**

Nous récupérons des données sur internet car la Banque de France ne dispose pas d'un *reporting* réglementaire pas plus qu'il n'en existe au sein de l'Euro-système. Les seules mesures qui existent s'appuient sur des questionnaires diffusés par certaines banques centrales mais rien n'oblige les fintechs à procéder à une déclaration hormis lorsqu'elles demandent un agrément à l'ACPR. Cependant, il existe aussi beaucoup de fintechs sur lesquelles nous n'avons aucune information. L'algorithme vise donc à exploiter les données disponibles sur Internet pour détecter les nouvelles fintechs. A terme, nous souhaiterions qu'il existe des *reportings* réglementaires afin de suivre l'activité des fintechs et correctement classer leur activité.

### **Anne-Laure DELATTE**

Je suis surprise d'entendre que toutes les fintechs ne disposent pas d'une accréditation pour exercer. Est-ce à dire qu'une fintech peut exercer sans autorisation ?

A quel stade est la discussion sur la réglementation de ce nouveau secteur ? Existe-t-il d'autres secteurs que celui des fintechs dont les données ne sont pas encore collectées de manière spécifique ? Devons-nous comprendre de votre réponse que plusieurs autres secteurs émergents ne sont pas suivis de manière adéquate ?

### **Enzo BUTHIOT**

Les codes NAF ne permettent pas d'identifier les fintechs. C'est pour cela que des discussions sont en cours au niveau européen et international afin de mieux prendre en compte les activités des fintechs. Nous collectons bien entendu des informations sur les fintechs dès lors qu'elles ne peuvent pas opérer sans agrément, si elles interviennent dans le domaine du paiement ou du crédit par exemple. Cependant, d'autres fintechs n'ont pas besoin d'agrément, par exemple les comparateurs d'assurances. Aujourd'hui, nous pouvons disposer de données sur leurs métriques, par exemple sur leur chiffre d'affaires, mais nous ne les identifions pas en tant que fintech.

### **Fabrice BIDAUD**

Certaines fintechs sont régulées car leur activité l'exige, mais d'autres fintechs ne le sont pas. Pour ces entreprises non régulées, comme les comparateurs d'assurance par exemple, la Banque de France ne dispose d'aucune collecte statistique. L'objectif est également de pouvoir distinguer les fintechs des non-fintechs parmi le secteur régulé (banque, assurance), notamment pour suivre la dynamique de croissance du secteur ou pour comparer les performances.

### **Anne-Laure DELATTE**

Je vous remercie pour ces éclaircissements.

## **3. FINANCE ET GENRE : QUELQUES APPROCHES EXPLORATOIRES**

### **3.1. Finance et genre en finance d'entreprise : revue de littérature et comparaison internationale des bases de données**

#### **Camille HEBERT, Université de Toronto**

Mon exposé sera centré sur la revue de la littérature sur la thématique de la finance d'entreprise et du genre. Il faut retenir que cette littérature est assez nouvelle. Il y a trois ou quatre ans, il n'y avait aucune session dédiée à la finance et au genre dans les conférences internationales alors que toutes les grandes

conférences consacrent désormais au moins une session à cette thématique. C'est donc un thème émergent.

Je commencerai par dresser un état des lieux de la littérature sur cette thématique et je terminerai par un point sur les thèmes qui doivent encore être approfondis par la recherche. Au cours de ma présentation, je mettrai en avant les thèmes sur lesquels les données françaises pourraient être utiles.

La littérature existante sur le travail montre que les femmes sont moins rémunérées que les hommes. Pour un euro gagné par un homme, une femme percevra en moyenne 80 centimes. On sait aussi que les femmes représentent seulement 5 % des dirigeants des grandes entreprises américaines cotées. Les données en France sont semblables. En outre, uniquement 19 % des membres des conseils d'administration sont des femmes.

Des travaux plus récents montrent que les femmes fondent 30 % des nouvelles entreprises. Ce taux est stable en France comme aux États-Unis ou au Canada. Toutefois, les entreprises fondées par les femmes ne représentent que 10 % des entreprises qui lèvent des fonds externes. Cet écart serait expliqué, du côté de la demande, par des motivations différentes des fondateurs des entreprises, et du côté de l'offre, par une discrimination exercée par les investisseurs. Cette discrimination est fondée sur des croyances qui peuvent être biaisées, par exemple le fait de croire que les hommes ont des compétences dans le domaine informatique tandis que les femmes sont compétentes dans le secteur de la santé et du soin. Ces croyances poussent les investisseurs à prendre des décisions fatalement biaisées.

D'autres publications portent sur des études empiriques. Certaines visent à documenter les disparités homme/femme ; d'autres tendent à proposer une analyse causale. Pour réaliser cette analyse causale, les chercheurs ont utilisé plusieurs éléments. Certains se sont appuyés sur les quotas de femmes dans les *boards*. Ces politiques ont commencé en Norvège en 2003, suivie par l'Espagne, l'Islande, l'Italie, la Finlande et la France en 2014. En France, il a été exigé que 20 % de femmes siègent dans les conseils d'administration des entreprises publiques et des grandes entreprises privées françaises, taux porté à 40 % à partir de 2017. Le Parlement européen a voté une proposition de loi visant à étendre cette mesure à tous les pays européens. Les analyses causales peuvent aussi s'appuyer sur les lois qui imposent la transparence dans la rémunération des salariés obligeant ainsi les entreprises à publier les moyennes de salaire par genre.

Je propose ici de mettre l'accent sur trois publications utilisant des données administratives en Norvège, au Danemark et en France. La publication portant sur le périmètre de la France utilise des données de l'Insee : c'est un travail dont je suis coauteure et pour lequel j'avais demandé une autorisation au Cnis. Les trois publications que je vais relater ici sont diffusées dans les journaux de référence en économie et finance.

L'étude portant sur les données norvégiennes s'appuie sur les quotas de genre au sein des *boards* et questionne le bénéfice apporté par ces quotas au-delà de la présence des femmes dans ces comités. Cette étude s'appuie sur les données issues des DADS norvégiennes entre 1986 et 2010 sachant que le quota a été introduit en Norvège en 2003. Il convient tout d'abord de souligner que les DADS norvégiennes sont beaucoup plus riches que les DADS françaises : ces déclarations fournissent des informations sur le niveau d'éducation, sur les liens familiaux, sur les revenus, sur les impôts, etc. En France, pour obtenir ces mêmes informations, il convient de demander des habilitations assez difficiles à obtenir.

Il ressort de ces travaux que les femmes qui rejoignent les *boards* grâce aux quotas sont plus qualifiées et bénéficient de meilleurs salaires que les femmes qui siégeaient dans les comités avant cette loi. Il est donc constaté une réduction de l'écart lié au genre. Cette conclusion était déjà démontrée par d'autres publications. La nouveauté de cette étude norvégienne est que ces quotas n'entraînent pas de modification dans l'économie, c'est-à-dire qu'il n'est pas noté un phénomène de ruissellement et que cette politique est sans effet sur la réduction de l'écart de rémunération et de promotion entre les femmes et les hommes, sans effet sur le taux d'inscription des femmes dans les programmes de formation amenant à des fonctions dirigeantes et sans effet sur l'évolution des femmes dans d'autres secteurs de l'économie. Cette publication conclut donc à une absence d'externalités positives des quotas dans la composition des *boards*. C'est la première publication qui lie la base de données Boardex avec des données administratives.

La deuxième publication a utilisé les données administratives danoises pour investiguer sur les effets des lois qui imposent une transparence sur les salaires homme/femme. L'étude a porté sur la période 2004-2008 sur la base des DADS danoises. Depuis 2006, au Danemark, les entreprises doivent publier des statistiques

sur les écarts de salaire entre les hommes et les femmes. Cette étude indique que cette mesure a conduit à une réduction de l'écart de salaire selon le genre. Cependant, cette situation tient à une diminution de la rémunération des hommes et non à une augmentation du salaire des femmes. Par ailleurs, cette étude indique que la loi sur la transparence des rémunérations a des conséquences sur la performance des entreprises.

La troisième publication a pour terrain d'étude les différences constatées en France entre les entrepreneurs hommes et femmes et les stéréotypes. Il convient de rappeler ici que les femmes entrepreneurs ont des difficultés à lever des financements externes en *private equity* ou en *capital risk* puisque ce taux est d'environ 15 % pour les femmes contre 30-35 % en moyenne pour toute la population. Face à ce constat, deux explications peuvent être avancées. La première explication est de dire que les femmes lèvent moins de capitaux car leurs projets sont de moins bonne qualité. La deuxième explication avancée par la littérature est que les femmes sont discriminées. Cette publication avait pour objectif de vérifier ces hypothèses. Cependant, pour opérer ces vérifications, il faut obtenir des données sur les femmes ayant réussi à lever des capitaux et sur celles qui auraient pu en lever. Pour mener à bien ce travail, les données de l'Insee ont été très utiles. L'article visait également à identifier les sources potentielles de discrimination ou les stéréotypes conduisant à ce que les femmes lèvent moins de capitaux que leurs homologues masculins.

Pour mener à bien ce travail, nous avons utilisé les données du Sine (Système d'information des nouvelles entreprises) et les liasses fiscales Ficus-Fare. Grâce à ces données, nous observons que les stéréotypes de genre dans le financement en *capital risk* est une explication probable de l'écart constaté entre les hommes et les femmes. Pour rappel, Sine est une enquête représentative des nouvelles entreprises créées en France. Tous les quatre ans, 25 % de la nouvelle cohorte d'entrepreneurs sont destinataires de l'enquête qui est envoyée par l'administration fiscale. Dans ce document, les entrepreneurs doivent renseigner des informations sur leur genre, sur leur niveau d'études, sur leurs expériences professionnelles passées, sur les entreprises éventuelles créées avant celle-ci, sur leur structure familiale, sur l'entreprise elle-même, etc. Pour compléter ces données, nous avons également utilisé des bases de données externes qui portent sur les accords de financement (*VentureXpert*, *Crunchbase* et *Pitchbook*) et précisent les montants investis, fournissent des informations sur les termes du contrat, etc. Comme ces bases ne précisent pas le numéro Siren des entreprises, j'avais créé d'abord un *web crawler* sur le Bodacc et sur [www.societe.com](http://www.societe.com) avant d'utiliser l'API mis en place plus récemment par l'Insee.

Ces travaux aboutissent à la conclusion que les femmes entrepreneurs sont en moyenne moins susceptibles de lever du financement externe que les hommes, ce qui était déjà avancé dans d'autres articles. La nouveauté est que nous pouvons dire que les femmes dans les secteurs dits masculins sont moins susceptibles de lever des fonds que leurs homologues masculins, mais que les femmes dans les secteurs dits féminins (comme le secteur de la santé) lèvent plus de fonds que les hommes. Il existe donc une hétérogénéité dans l'écart homme/femme. Cependant, les secteurs dits masculins représentent 85 % de l'économie française contre une part de 15 % pour les secteurs dits féminins. Par ailleurs, il est relevé que le groupe minoritaire (c'est-à-dire les femmes dans les secteurs dits masculins et les hommes dans les secteurs dits féminins) surperforme l'autre groupe. Cette situation peut tenir au fait que les attentes exprimées sont plus fortes à l'endroit de ce groupe minoritaire, ce qui est à relier aux stéréotypes de genre.

Pour cette recherche, j'ai été confrontée à plusieurs difficultés, notamment à celle consistant à identifier le genre dans les données administratives françaises. Le genre figure certes dans la base Sine mais il n'est pas mentionné dans la base Siren, sauf pour une entreprise personnelle. Cependant, il reste possible d'identifier le genre du fondateur de l'entreprise ou de ses dirigeants à partir des DADS avec la PCS23 pour le chef d'entreprise, la PCS37 pour les managers et la PCS38 pour les ingénieurs et cadres.

Avec Paul Beaumont, je mène également une autre étude portant sur les conséquences des quotas dans les conseils en France sur les performances des entreprises privées non cotées. Pour mener ce travail, nous utilisons la base Boardex et Ficus-Fare.

Une autre question qui serait à investiguer porterait sur les contraintes qui pèsent sur l'accès au crédit des entreprises fondées par les femmes. Pour mener cette étude, il faudrait pouvoir lier les données de la base de données SCR de la Banque de France sur le crédit distribué par les banques avec les DADS. Cependant, ce lien est aujourd'hui impossible. Cette recherche serait pourtant très intéressante et permettrait d'objectiver l'accès au crédit entre femmes et hommes entrepreneurs. La mise à disposition de ces données et leur appariement sont nécessaires pour répondre aux questions de société.



### 3.2. Finance de marché et genre

#### Elyès JOUINI, Université Paris Dauphine

Avant de passer à ma présentation, je souhaiterais rebondir brièvement sur les propos tenus précédemment par Camille Hébert concernant la surperformance du genre minoritaire dans certains champs. Cette situation renvoie aux travaux de Thomas Breda et de Son-Thierry Ly qui ont démontré que les femmes passant le concours de l'Ecole normale supérieure étaient surnotées à l'oral dans les disciplines considérées comme masculines par rapport à leurs résultats à l'écrit tandis que les hommes étaient surnotés dans les disciplines présentées comme féminines.

Les thématiques qui nous occupent sont des sujets de société qui posent des questions d'équité. En effet, il n'y a pas de raison que les femmes aient moins accès aux études ou au crédit, par exemple, en fonction de leur genre. Ces thématiques rejoignent également des questions d'efficacité économique, car se priver *a priori* d'une partie de la population pour une activité revient à perdre potentiellement en efficacité. C'est donc sous ces deux angles que ces sujets doivent être regardés.

Mon exposé portera sur la finance de marché et rejoindra, sur de nombreux points, les constats posés précédemment sur la finance d'entreprise. Aujourd'hui, 25 % des diplômés en finance ou en sciences sont des femmes et les femmes ne représentent que 21 % des enseignants en finance. Pourtant, on sait que les femmes *traders* présentent des rendements significativement plus élevés que ceux des hommes. En outre, les FCP et *hedge funds* gérés par des femmes sont plus performants que ceux gérés par les hommes, une fois ajustés du risque. Dans les entreprises, les femmes ont tendance à être de meilleurs gestionnaires financiers, en particulier elles font preuve de moins de sur-confiance dans leurs décisions d'investissement que les hommes. On sait également qu'une plus grande diversité sous toutes ses formes et à tous les niveaux de l'entreprise a un impact positif, notamment dans les structures de gouvernance. Pourtant, malgré de meilleures performances et des décisions moins guidées par la sur-confiance, les femmes sont moins nombreuses que les hommes tant dans l'industrie financière que dans les études en finance. La question reste de savoir si ces différences sont comportementales, si ce sont des différences de goût ou si cette situation est à relier à une culture organisationnelle de l'industrie financière, auquel cas il faudrait que l'industrie financière s'adapte à une société plus égalitaire dans laquelle hommes et femmes doivent avoir les mêmes chances et les mêmes perspectives.

Dans la littérature portant sur les femmes et l'investissement, une publication a porté sur les conséquences d'une plus grande aversion des femmes au risque et sur leur moindre confiance en elle, laissant supposer que les femmes pourraient ne pas saisir certaines opportunités. Pourtant, les travaux de Barber et Odean montrent que leur moindre confiance en elles conduit à une plus grande performance. Finalement, les femmes réaliseront en moyenne moins d'opérations de marché et réajusteront moins souvent leur portefeuille, ce qui reviendra à payer moins de coûts de transaction.

En 2010, les travaux de Wann et Lobo ont montré que les hommes passaient plus de temps à gérer leur portefeuille et passaient davantage de transactions. Les hommes travaillent sur une palette plus large d'actifs avec davantage d'effets de levier. En moyenne, leur performance est équivalente à celle des femmes mais avec un niveau de risque plus élevé.

D'après Meyers et Levy en 1989, « *les hommes ne s'engagent pas dans le traitement complet de toutes les informations disponibles. [...] De plus, le mode de traitement sélectif et heuristique des hommes est souvent dicté par le recours à des indices très disponibles et particulièrement saillants* ». Ce constat conduit à dire que les hommes se focaliseraient sur quelques points saillants tandis que les femmes traiteraient un ensemble plus large d'informations pouvant les conduire à percevoir un niveau de risque plus élevé. Cette vision conduirait à dire que les femmes ne sont pas forcément plus adverses au risque mais qu'elles percevraient davantage les risques du fait d'une vision plus large de la problématique. C'est aussi parce que les femmes ont une conscience accrue de l'ensemble des dimensions d'un problème et qu'elles ont accès à un ensemble plus complet de signaux informatifs, y compris ceux opposés à leur décision, qu'elles apparaîtraient comme moins confiantes. Cependant, la plus grande aversion au risque réelle ou apparente comme la sous-confiance réelle ou apparente seraient au service d'une plus grande efficacité.

Selon les travaux de Fang et Huang, l'appartenance à des réseaux de décideurs a un impact sur la qualité des analyses financières et sur les carrières, mais cet impact n'est pas le même pour les hommes et pour les femmes. En effet, faire partie d'un réseau d'anciens élèves, par exemple, a un effet deux à trois fois plus

important pour les hommes en termes de qualité des analyses financières et en termes de carrière. Pour autant, il ne peut pas être affirmé que les femmes mobilisent moins leur réseau que les hommes. Appartenir à un réseau impacte également la probabilité d'être élu « analyste star ». Appartenir à un bon réseau agit comme un substitut à la performance pour les hommes tandis que cette appartenance agit comme un complément à la performance pour les femmes.

Enfin, selon l'étude de Niessen et Ruenzi, les fonds gérés par les femmes lèvent moins d'argent que ceux gérés par les hommes alors même qu'il n'existe pas de différences de performance. Les investisseurs soumis à des tests mesurant les stéréotypes (comme l'*Implicit Association Test*) montrent que plus le biais de genre est important chez l'investisseur et moins il investit dans un fonds géré par une femme. Les stéréotypes seraient donc la cause d'un moindre investissement dans les fonds gérés par les femmes.

Les bases de données sur la finance de marché sont assez pauvres. Pour les études sur les analyses, les gestionnaires et les traders, c'est sur la base du prénom collecté qu'il est possible de déduire le genre. Les bases ISR/ESG recensent le genre du président directeur général ou le pourcentage de femmes dans les conseils d'administration voire le pourcentage de femmes managers, mais les critères restent très macro. De plus, ces bases se limitent souvent au top management, ce qui introduit un biais car ces informations ne permettent pas de dire si l'évolution de la situation homme/femme dans le top management agit par ruissellement dans les autres strates de l'entreprise.

En conclusion, la diversité de genre dans les conseils d'administration est plus faible en finance que dans les autres secteurs. Il existe également d'importants écarts de genre en sciences technologies ingénierie et mathématique (STIM). La finance n'est pas classée parmi les STIM, mais c'est un secteur intensif en mathématiques, et la finance peut donc être classée dans le champ des sciences sur le plan des compétences mobilisées. Or l'intensité mathématique moyenne des métiers féminins est inférieure à celle des métiers masculins et plus particulièrement en finance. Cela signifie que, même lorsqu'il y a autant d'hommes que de femmes dans les secteurs financiers, les femmes seront davantage positionnées sur les fonctions de communication et des ressources humaines, c'est-à-dire qu'elles seront présentes dans des métiers moins intensifs en sciences, tandis que les hommes occuperont des métiers dans le cœur de métier de la finance. Grâce aux lois promulguées pour promouvoir la parité dans les *boards*, la situation évolue dans les entreprises. Cependant, dans les Comex, les femmes qui y siègent sont souvent issues de la direction de la communication ou de la direction de la RSE, par exemple. Par ailleurs, les écarts sont moins importants dans les pays où les écarts en mathématiques sont moins importants. Toutes industries confondues, les femmes administrateurs sont moins susceptibles de détenir un diplôme scientifique de haut niveau, mais elles sont plus susceptibles d'en détenir un dans les banques. Cette relation est d'autant plus forte dans les pays où les écarts en mathématiques sont plus importants. Manifestement, ces questions sont liées également à la présence de stéréotypes sur le lien entre femmes et disciplines scientifiques. La saillance de ce stéréotype dans les pays peut donc guider le choix des femmes à poursuivre des études scientifiques.

Plusieurs pays ont fait le choix d'instaurer des quotas pour commencer à corriger les inégalités, mais les quotas ne sont pas suffisants. Les écarts constatés tiennent aussi à la persistance de stéréotypes. La correction des écarts passera aussi par la mise en place de politiques garantissant des résultats éducatifs plus égaux. En effet, pour que les entreprises du secteur financier puissent recruter davantage de femmes, il convient aussi que les écoles et universités forment un plus grand nombre de candidates dans ces champs.

## Échanges

### **Gunther CAPELLE-BLANCARD, Université Paris 1 Panthéon-Sorbonne**

L'index sur l'égalité professionnelle est devenu obligatoire en France depuis quelques années. Est-il possible que le Cnis intervienne pour que les données qui sont utilisées pour construire cet index nous soient accessibles plutôt que les seules informations agrégées ?

Dans sa présentation, Camille Hébert a souligné qu'il était difficile d'obtenir des données Banque de France alors que l'*Open Data Room* permet bien de les fournir. Pouvez-vous revenir sur ce point ?

## **Camille HEBERT**

Les données sont disponibles, mais la difficulté est de lier les données banque de France et celles de l'Insee.

## **Gunther CAPELLE-BLANCARD**

Je suis d'accord, mais je pense que cela peut progresser et que la Banque de France et l'Insee y sont favorables même si le processus est peut-être long et coûteux.

## **Anne-Laure DELATTE**

C'est un point que nous allons pointer dans l'avis. Nous avons obtenu le rapprochement de l'ODR et du CASD, mais il reste à traiter le sujet de l'appariement des données.

### **Avis de la commission**

#### **François GUILLAUMAT-TAILLIET, Cnis**

Je propose de donner lecture de la dernière version de l'avis après plusieurs échanges internes. Il est le suivant :

*« L'avis n° 5 du moyen terme en cours de la commission invite à « mesurer l'importance de la fintech ». Il souligne qu'une « nouvelle industrie financière, la fintech, s'est ainsi récemment développée qui déploie les TIC pour améliorer les activités financières. Elle englobe des dispositifs qui concurrencent les banques traditionnelles, avec par exemple les prêts entre particuliers, mais aussi les systèmes de paiement (paiement par mobile, vérification des transactions financières par exemple) et les cryptoactifs ». Force est de constater qu'il n'existe pas de définition légale de la fintech qui est une notion évolutive, dont les contours et la nécessaire prise en compte font l'objet de discussions tant pour ce qui concerne la rénovation en cours des nomenclatures d'activités et de produits qu'entre comptables nationaux, sans parler du Conseil de stabilité financière. De taille encore modeste, les volumes en jeu sont en forte croissance et sont surveillés de près par les banques centrales dès lors qu'ils sont susceptibles d'avoir un impact sur la transmission de la politique monétaire. L'apport des fintechs en termes d'inclusion bancaire est par ailleurs un des axes de la Data Gaps Initiative III du G20. La Banque de France a entrepris de se doter d'un outil dynamique de détection et de monitoring des fintechs pour répondre aux demandes des utilisateurs de données : une base de données d'entreprises contenant les informations nécessaires et préfigurant un futur registre national, à partir duquel on pourra créer des statistiques publiques. D'autres pays mènent des expérimentations dans ce sens, notamment l'Allemagne.*

*S'agissant des plateformes de financement participatif par le prêt, où ce sont les prêteurs qui choisissent les prêts, la transparence est requise. La Commission note que le niveau de transparence sur les données de crédit immobilier est supérieur aux États-Unis.*

*La commission prend acte du fait que la fintech est un domaine non encore stabilisé, en pleine évolution, à l'origine de nouveaux outils en devenir. Elle considère que le travail mené au niveau international sur les nomenclatures constitue un enjeu majeur pour permettre à la statistique publique de bien identifier le développement des fintechs. La commission recommande aux parties prenantes de concevoir ces outils de sorte qu'ils alimentent, en aval, la statistique publique et que l'accès à ces nouvelles sources de données soit facilité pour les chercheurs. La commission alerte sur la nécessité que les données fintechs accessibles via l'ODR puissent être fusionnées avec les données CASD. Elle demande à être tenue régulièrement informée des avancées dans ce domaine.*

*Par ailleurs, suivant l'avis de moyen terme du Cnis n° 2 sur la mesure des inégalités de toutes natures, la commission a réalisé un état des lieux des données permettant de documenter la place des femmes dans les métiers de la finance, d'une part, et l'accès des femmes au financement de leur activité professionnelle, d'autre part. La commission note que les recherches sur le genre et la finance sont en plein développement et que les enjeux de données sont nombreux et méritent d'être approfondis. Elle encourage l'utilisation de données innovantes (données commerciales notamment pour poursuivre l'exploration de ces sujets). »*

## **Gunther CAPELLE-BLANCARD**

Sur la thématique de la finance et du genre, il me semble que le projet d'avis pourrait aller plus loin et demander *a minima* de revoir l'index d'égalité professionnelle pour qu'il soit véritablement utile grâce à l'accès aux données brutes.

## **Anne-Laure DELATTE**

Le projet d'avis sur la finance et le genre pourrait aussi conduire à pointer qu'il existe un enjeu d'appariement des données existantes.

## **François GUILLAUMAT-TAILLIET, Cnis**

Je souhaite rebondir sur la remarque de Gunther Capelle-Blancard concernant l'index. Pouvez-vous être plus précis sur son objet et sur sa source ?

## **Gunther CAPELLE-BLANCARD**

L'index est produit par les entreprises et est placé sous la responsabilité du ministère du Travail. C'est une nouvelle obligation qui pèse sur les entreprises. Ma demande est que les règles du jeu changent pour rendre cette publication plus utile.

## **François GUILLAUMAT-TAILLIET**

Nous avons abordé ce sujet dans le cadre du groupe de travail sur les discriminations, mais avec circonspection pour ce qui concerne l'intérêt de ce type de mesure pour la statistique publique. Je serai donc prudent, d'autant plus que cet index n'a pas été présenté au sein de notre instance. Je pense qu'il faudrait en savoir plus sur ses conditions de fabrication et sur sa finalité avant de statuer.

## **Camille HEBERT**

Mon intervention portera quant à moi sur l'appariement des bases de données. Selon moi, la base de données essentielle qui serait à appairer à d'autres est la DADS car c'est à ce niveau qu'il est possible d'identifier le genre. C'est la base clé pour le genre.

## **Anne-Laure DELATTE**

*Quid* de Sine ?

## **Camille HEBERT**

Cette base est utile sur les questions portant sur l'entrepreneuriat, mais les DADS couvrent l'univers des entreprises sans être circonscrite à certaines d'entre elles. Si une recommandation doit être formulée, j'insisterai donc sur l'appariement avec les DADS.

## **Anne-Laure DELATTE**

C'est la conclusion à laquelle nous sommes arrivés lors de nos discussions. Si les représentants du Cnis ne jugent pas la proposition trop technique, il serait donc utile d'évoquer l'appariement et de pointer l'appariement avec les DADS sur les questions de genre.

## **Camille HEBERT**

Les DADS contiennent des données sensibles comme les salaires. L'Insee pourrait donc peut-être créer une base de données allégée qui contiendrait seulement les PCS, le genre et le Siren.

## **François GUILLAUMAT-TAILLIET**

Les données de salaires ne sont pas sensibles contrairement aux données de santé par exemple.

**Camille HEBERT**

Mes deux recommandations pour la rédaction de l'avis seraient de faire mention de l'appariement avec les DADS et de faciliter l'importation de données commerciales dans le CASD. Cette dernière opération est déjà possible, mais est chronophage car il faut alors demander une autorisation. L'importation de bases de données commerciales devrait être rendue automatique.

**François GUILLAUMAT-TAILLIET**

Je préconise de ne pas entrer dans le détail dans le texte de l'avis mais de nous contenter de dire que nous encourageons à utiliser des données innovantes, commerciales notamment.

**Anne-Laure DELATTE**

L'avis doit préciser que la commission encourage l'utilisation des données innovantes mais aussi l'appariement de données commerciales et administratives. Se contenter d'évoquer des « *données innovantes* » est selon moi trop vague. C'est aussi ici qu'il serait possible d'évoquer les données des DADS.

**François GUILLAUMAT-TAILLIET**

Pouvons-nous dire « *la commission encourage l'utilisation de données innovantes et la possibilité de les appairer avec les DADS* » ?

**François MAUREL**

L'appariement est une façon d'innover. Nous pourrions donc écrire : « *La commission encourage l'utilisation de données innovantes (notamment l'appariement de données commerciales et administratives)* ».

**Anne-Laure DELATTE**

Il ne faudrait pas que le contenu de la parenthèse soit perçu comme accessoire alors que c'est la problématique elle-même.

**Françoise MAUREL**

Nous pourrions alors écrire : « *La commission encourage les appariements* ».

**François GUILLAUMAT-TAILLIET**

Je vous propose la rédaction suivante : « *Elle encourage les appariements de données innovantes (commerciales et administratives) pour poursuivre l'exploration de ces sujets* ».

**Anne-Laure DELATTE**

Pourrions-nous ajouter ici une mention sur les DADS ?

**François GUILLAUMAT-TAILLIET**

Il n'est peut-être pas utile de pointer une donnée en particulier. De plus, la DADS devient la DSN.

**Anne-Laure DELATTE**

Nous pourrions également ajouter que le souhait est de « *favoriser l'utilisation et la publication des données individuelles non agrégées* » sans pour autant faire référence à l'index d'égalité professionnelle.

**François GUILLAUMAT-TAILLIET**

Il est possible que les appariements permettent d'avoir accès à l'équivalent de l'index sur l'égalité homme/femme.

### **Gunther CAPELLE-BLANCARD**

A l'heure du *big data*, il est étonnant de n'avoir accès qu'à des indices synthétiques et agrégés. Or les entreprises savent très bien les utiliser pour masquer leurs véritables politiques. C'est pour cela qu'il faut avoir accès à des données brutes et non à des données corrigées et pondérées par les entreprises elles-mêmes. Je pense que le Cnis remplit sa mission en formulant cette demande.

### **Camille HEBERT**

La publication de statistiques agrégées sur les salaires des hommes et des femmes répond à une proposition de loi du Parlement européen. Cependant, il y a forcément d'autres éléments qui entrent dans la composition de cet indice et auxquels il serait intéressant d'avoir accès.

### **François GUILLAUMAT-TAILLIET**

Je serais plus à l'aise si nous avions eu une présentation de cet indice, car nous sommes ici en train de parler d'un index dont nous ne connaissons pas précisément la teneur.

### **Françoise MAUREL**

Il est toutefois possible de demander les données les plus désagrégées possible.

### **François GUILLAUMAT-TAILLIET**

Bien sûr, nous pourrions donc écrire : « *Elle encourage les appariements de données innovantes (commerciales et administratives) et la transmission de données les plus désagrégées possible pour poursuivre l'exploration de ces sujets* ».

### **Gunther CAPELLE-BLANCARD**

On pourrait mettre entre parenthèses une référence à l'indice synthétique d'égalité professionnelle. Comme les avis du Cnis font l'objet d'un suivi, nous pourrions ainsi faire référence à cette première demande.

### **François GUILLAUMAT-TAILLIET**

Je suis d'accord avec vous sur le fond, mais je maintiens que cet index ne nous a pas été présenté.

### **Guillaume HOURIEZ**

Le terme « transmission » me dérange dès lors que nous ne précisons pas à qui ces données doivent être transmises.

### **Anne-Laure DELATTE**

Nous pourrions plutôt écrire : « *l'accès aux données les plus désagrégées possible pour poursuivre l'exploration de ces sujets* ».

### **Olena HAVRYLCHYK**

Concernant la première partie de l'avis, au deuxième paragraphe, il est indiqué que la transparence est requise. Cette demande me semble trop floue alors que la demande est de publier le *loan book*.

### **François GUILLAUMAT-TAILLIET**

Je pense que beaucoup de personnes ne savent pas ce que sont les *loan books*. Il faudrait donc préciser cette terminologie.

## **Olena HAVRYLCHYK**

La cellule fintechs de l'ACPR connaît bien ce terme.

## **Anne-Laure DELATTE**

Nous pouvons écrire : « *les opérations relatives aux prêts et aux opérations (par exemple loan books)* ».

## **Fabrice BIDAUD**

Quel texte réglementaire ou législatif impose aux acteurs privés de fournir ces informations ?

## **Olena HAVRYLCHYK**

Aucune réglementation ne prévoit la publication du *loan book*, mais je pense que l'ACPR peut le conseiller et je le conseille en tant que chercheuse.

## **Anne-Laure DELATTE**

C'est notre rôle que de formuler cette recommandation. Qu'en pensent les représentants du Cnis ?

## **Françoise MAUREL**

Le pouvoir du Cnis est assez limité en la matière, mais nous pouvons écrire que la publication de ces données est souhaitée. Si aucune norme n'impose la publication, il faut aussi se montrer réaliste même si le Cnis en émet le vœu.

## **François GUILLAUMAT-TAILLIET**

Le paragraphe serait donc remanié ainsi : « *S'agissant des plates-formes de financement participatif par le prêt, où ce sont les prêteurs qui choisissent les prêts, la transparence est souhaitable sur les informations relatives aux prêts (loan book et autres instruments). La commission note que le niveau de transparence sur les données de crédit immobilier est supérieur aux États-Unis.* »

## **Cécile GOLFIER**

Je souhaite réagir aux propos de Camille Hébert et dire que les données de la centrale de risques hors entrepreneurs individuels sont bien dans le CASD. De mon point de vue, il n'y a donc pas de sujet.

Par ailleurs, je demande le retrait de la phrase suivante : « *La commission alerte sur le nécessité que les données « fintechs » accessibles via l'ODR puissent être fusionnées avec les données CASD* ». En effet, la Banque de France n'a pas pour mission de rediffuser les données fintechs d'autres institutions. De plus, les données spécifiques aux fintechs n'existent pas aujourd'hui puisqu'un travail est en cours pour justement les détecter. Selon moi, il n'y a pas matière à alerter sur un quelconque problème avec les données des fintechs. Par exemple, ce n'est pas à la Banque de France de diffuser l'adresse d'une entreprise. La question de la diffusion de données propres aux fintechs se posera lorsque la Banque de France aura créé des données spécifiques à ce secteur.

## **Olena HAVRYLCHYK**

Il y a peut-être une confusion s'agissant des données des fintechs visées car la demande est de pouvoir disposer du *loan book*. C'est une information collectée par la Banque de France et envoyée par les plates-formes. La difficulté est que les chercheurs peuvent récupérer cette donnée *via* l'ODR mais sans pouvoir l'apparier avec les données du CASD. Pour la question du genre comme pour la question des fintechs, le point bloquant est le même et porte sur l'impossibilité d'apparier les données.

### **Françoise MAUREL**

Nous pouvons dire de manière générale que notre souhait est que les données fintechs (éventuellement *via* l'ODR) puissent fusionner avec les données CASD.

### **Anne-Laure DELATTE**

Le point général soulevé et que nous pourrions inscrire de nouveau à l'ordre du jour est celui de l'appariement de toutes les données administratives sur le CASD. Cécile Golfier a souligné plus tôt que c'était peut-être une question de temps. Nous ne pouvons donc pas lancer une alerte. En revanche, nous pouvons dire qu'il est souhaitable de favoriser l'appariement.

### **François GUILLAUMAT-TAILLIET**

Je propose donc la rédaction suivante : « *La commission souligne l'importance de l'appariement* ».

### **Anne-Laure DELATTE**

Nous pouvons compléter la phrase en ajoutant : « *en particulier dans le cadre du rapprochement ODR/CASD* ».

### **Cécile GOLFIER**

Je n'y suis pas favorable et je pense préférable de rester général, car nous ne savons pas de quelles données nous parlons.

### **Anne-Laure DELATTE**

Si, ce que nous visons est le *loan book*. C'est donc très précis.

### **Olena HAVRYLCHYK**

Le problème existe bel et bien aujourd'hui puisque l'appariement des données m'a été refusé. Peut-être que la difficulté sera levée à l'avenir mais elle existe aujourd'hui. La Banque de France ayant fermé l'ODR, je dispose maintenant de six mois pour travailler sur les données du CASD avant de payer cette utilisation. Cependant, comme ces données ne peuvent pas être appariées, ce travail devra être mené dans un second temps.

### **Cécile GOLFIER**

Ce commentaire relève d'un autre sujet.

### **Anne-Laure DELATTE**

Je suis d'accord. Cependant, je crois que nous pouvons faire un constat et le formuler de manière diplomatique.

### **François GUILLAUMAT-TAILLIET**

Nous pourrions écrire simplement : « *La commission souligne l'importance de pouvoir appairer les données administratives* ».

### **Gunther CAPELLE-BLANCARD**

Nous savons que les données existent mais l'obstacle est uniquement technique, ce qui crée de la frustration. Il faudrait donc insister sur ce point et citer le CASD et la Banque de France.



**Cécile GOLFIER**

Toutefois, ce n'est pas documenté et nous ne savons pas de quelle base il est question.

**Gunther CAPELLE-BLANCARD**

Il s'agit des données disponibles sur l'ODR.

**Françoise MAUREL**

Nous pourrions indiquer que notre souhait est que « *les données accessibles via l'ODR puissent être fusionnées* » sans faire mention des données administratives ou des fintechs. Ainsi, la demande resterait générale.

**Fabrice BIDAUD**

Je rejoins Cécile Golfier, le sujet n'a pas été suffisamment documenté.

**Olena HAVRYLCHYK**

Vous m'avez invité pour apporter un témoignage sur la manière dont les chercheurs utilisent les données. Je vous affirme que j'ai demandé un appariement qui m'a été refusé. Vous ne pouvez donc pas réfuter mon témoignage.

**Anne-Laure DELATTE**

C'est vrai que l'on invite des chercheurs qui prennent du temps gratuitement et que l'on n'écoute pas.

**Fabrice BIDAUD**

Ce n'est pas ce point que je remets en cause, mais je veux juste dire que nous n'avons pas documenté si le problème d'appariement entre le CASD et les nouvelles données ODR se limite à l'ouverture post-fusion ou si le problème est en cours de résolution voire déjà résolu.

**Olena HAVRYLCHYK**

S'il était résolu, j'en serais informée.

**Fabrice BIDAUD**

Dans tous les cas, nous n'avons aucune information sur la persistance de cette difficulté. Mettre l'accent sur un possible problème d'appariement entre les données Banque de France et d'autres bases me semble prématuré.

**Anne-Laure DELATTE**

Nous ne disons pas qu'il existe un problème, mais nous pointons l'importance de l'appariement.

**François GUILLAUMAT-TAILLIET**

En effet, il ne me semble pas choquant de souligner l'importance de l'appariement. Nous pourrions le rédiger ainsi : « *La commission souligne l'importance de pouvoir apparier les données accessibles via l'ODR avec les données se trouvant au CASD* ».

**Cécile GOLFIER**

Aujourd'hui, les données CASD incluent les données ex-ODR.

**Olena HAVRYLCHYK**

Certes, l'ODR n'existe plus mais les données restent segmentées.

**Anne-Laure DELATTE**

Cependant, si nous retirons les mentions « ODR » et « CASD », l'ensemble sera compréhensible.

**Olena HAVRYLCHYK**

Cette formulation pourrait laisser entendre que nous ne sommes pas informés que l'ODR n'existe plus.

**François GUILLAUMAT-TAILLIET**

Nous pourrions dire alors : « *La commission souligne l'importance de pouvoir apparier les données qui étaient accessibles via l'ODR avec les données accessibles au CASD* ».

**Anne-Laure DELATTE**

Je suis désolée mais je vais devoir quitter la séance. Est-ce problématique si la séance se poursuit sans ma présence ?

**François GUILLAUMAT-TAILLIET**

Non, nous renverrons le projet d'avis pour validation.

**Anne-Laure DELATTE**

Très bien, merci.

*Anne-Laure DELATTE quitte la réunion à 17h30.*

**Fabrice BIDAUD**

J'ai une remarque sur le premier paragraphe. Dans la phrase « (...) sont surveillés de près par les banques centrales dès lors qu'ils sont susceptibles d'avoir un impact sur la politique monétaire (...) », j'aurais ajouté « *sur la stabilité financière et la transmission de la politique monétaire* ».

Au paragraphe suivant, j'aurais également retiré la mention sur l'inclusion financière.

**Cécile GOLFIER**

Des modifications sont intervenues suite à la demande d'un pays et la thématique est bien celle des fintechs et de l'inclusion financière (recommandation 12). Cependant, c'est un point que nous pouvons revoir : dans l'avis, nous pouvons retirer la mention relative à l'inclusion bancaire.

**Fabrice BIDAUD**

Je suis également gêné par la dernière phrase qui évoque la préfiguration d'un futur registre national puisque seule une expérimentation a été menée à ce stade.

**Cécile GOLFIER**

Je rejoins cette remarque.

**François GUILLAUMAT-TAILLIET**

Pourrions-nous nous contenter d'écrire ceci : « *La Banque de France a entrepris de se doter d'un outil dynamique de détection et de monitoring des fintechs pour répondre aux demandes des utilisateurs de*

*données : une base de données d'entreprises contenant les informations nécessaires à partir desquelles on pourra créer des statistiques publiques ».*

**Cécile GOLFIER**

Je pense que nous en sommes encore loin. La deuxième partie de la phrase devrait être reformulée pour apparaître comme une préconisation du Cnis. Rédigé ainsi, nous le comprenons comme une ambition affichée de la Banque de France.

**Fabrice BIDAUD**

Nous pourrions en effet conserver la première partie de la phrase, puis partager l'identification de la manière la plus large possible lorsque le temps sera venu.

**Cécile GOLFIER**

Je pense que cette perspective est de moyen terme.

**Françoise MAUREL**

Il faut peut-être dire néanmoins que l'objectif est d'alimenter les statistiques publiques à terme, ce qui n'engage personne sur un horizon rapproché.

**Cécile GOLFIER**

Il ne faudrait pas laisser à penser que c'est inscrit dans le programme de travail de la Banque de France.

**François GUILLAUMAT-TAILLIET**

C'est aussi une demande que nous reprenons plus loin en disant : *« Elle recommande aux parties prenantes de concevoir ces outils de sorte qu'ils alimentent, en aval, la statistique publique et que l'accès à ces nouvelles sources de données soit facilité pour les chercheurs »*. Nous pourrions donc supprimer la première mention qui fait polémique.

Je vous remercie pour vos contributions. Je renverrai le projet d'avis aux rapporteurs et à la présidente.

Je vous souhaite une très bonne fin de journée.

*La séance est levée à 17 heures 40.*

Tous les documents sont disponibles sur le site [cnis.fr](http://cnis.fr) à la [page de la commission](#).

---

## DOCUMENTS PRÉPARATOIRES

---

- [Commission Système financier et financement de l'économie - Bilan 2021 des producteurs de la statistique publique \(pdf - 254 Ko\)](#)
- [Commission Système financier et financement de l'économie - Programme de travail 2022 des producteurs de la statistique publique \(pdf - 244 Ko\)](#)

---

## DOCUMENTS COMPLÉMENTAIRES

---

- [Programme de travail 2022 de la statistique publique - Banque de France et ACPR - Diaporama \(pdf - 1 Mo\)](#)
- [Les "fintechs" dans les nomenclatures et les statistiques macro-économiques - Banque de France - Diaporama \(pdf - 719 Ko\)](#)
- [Identification des "fintechs" par l'intelligence artificielle - Banque de France - Diaporama \(pdf - 818 Ko\)](#)
- [Les données de la recherche pour comprendre les "fintechs" - Université Paris 1 - Diaporama \(pdf - 833 Ko\)](#)
- [Finance et genre en finance d'entreprise - Université de Toronto - Diaporama \(pdf - 284 Ko\)](#)
- [Finance de marché et genre - Université Paris Dauphine - Diaporama \(pdf - 97 Ko\)](#)