Conseil national
de l'information statistique

Montrouge, le 5 juillet 2021 – N° 92/H030

COMMISSION « SYSTÈME FINANCIER ET FINANCEMENT DE L'ÉCONOMIE »

_____

Réunion du 27 mai 2021

PROJET D'AVIS 2

COMPTE RENDU DE LA RÉUNION 4

RAPPEL DE L'ORDRE DU JOUR 4

LISTE DES PARTICIPANTS 5

LISTE DES DOCUMENTS PRÉPARATOIRES ET COMPLÉMENTAIRES 32

PROJET D'AVIS

COMMISSION « SYSTÈME FINANCIER ET FINANCEMMENT DE L'ÉCONOMIE »
du 27mai 2021

Avis

La commission « Système financier et financement de l'économie » a pris connaissance de quelques avancées et expériences en termes de statistiques économiques et financières, menées dans le contexte de la crise sanitaire, pour en mesurer les conséquences.

Les travaux menés à l'Insee, visant à fournir un diagnostic rapide pour que les politiques publiques puissent s'adapter au mieux à la situation économique et sociale, ont révélé des difficultés à mesurer certaines activités, notamment celles issues des services non marchands, ainsi que la nécessité de décliner un diagnostic par catégorie, aussi bien au niveau des entreprises que des ménages. Le contexte a par ailleurs confirmé l'importance de pouvoir comparer les performances nationales avec celles d'autres pays. Les enquêtes mensuelles de conjoncture de la Banque de France ont été enrichies de questions supplémentaires et les commentaires libres qu'elles comportent ont été exploités en utilisant de nouvelles techniques de traitement de l'information. Le recours à l'exploitation des données massives a été mis en œuvre pour le suivi conjoncturel dans plusieurs domaines, en France et à l'étranger. Un suivi à haute fréquence du marché immobilier au Royaume-Uni a été appréhendé par la Banque de France à partir de données issues du *webscraping* d'informations disponibles sur les sites immobiliers, permettant d'obtenir des prix en niveau et des stocks de patrimoine comparables aux données officielles, pour un certain nombre de pays de l'OCDE. L'OCDE fait des prévisions économiques en temps réel à un rythme hebdomadaire pour 45 pays à partir de données de recherche Google agrégées par *Google Trends*. L'exploitation des données de plateformes comme *Indeed* ou de réseaux sociaux comme *Linked In* par les opérateurs propriétaires fournit une estimation en temps quasi réel des évolutions du marché du travail, au niveau de l'offre ou de la demande. Aux États-Unis, Harvard coordonne des travaux d'exploitation menés dans différents champs affectés par la crise à partir de plusieurs sources privées, qui alimentent une base de données individuelles anonymisées accessible à tous en temps réel. L'exploitation de données d'un panel de comptes bancaires, menée par l'Insee, a permis une observation de la situation financière des ménages en 2020.

Des exploitations nouvelles de sources existantes ont été menées pour suivre l'épargne financière des ménages à partir des données de la Banque de France, ou de manière plus originale à partir d'une mise en commun inhabituelle de données entre l'Insee et la Banque de France pour mesurer l'impact de la crise sur l'activité des entreprises.

La commission salue les initiatives de la statistique publique et encourage la consolidation des travaux en cours et la poursuite des investigations. Concernant l'utilisation des données massives, la commission note que ces données sont granulaires, disponibles rapidement, « homogènes » d'un pays à l'autre, avec un coût marginal proche de zéro et une charge de collecte faible. Elles nécessitent toutefois des travaux qui peuvent être conséquents pour les statisticiens (transformation d'informations de formats parfois hétérogènes en signaux interprétables statistiquement). Dans certains cas, les obtenir impose une tarification, ce qui soulève un problème de fond puisque la contribution à la statistique publique est par principe une contribution citoyenne à la création d'un bien public. En outre, en tant que traces numériques des activités humaines, ces données ne sont pas conçues à des fins scientifiques et ne présentent pas de garantie en termes de représentativité, d'historique, de stabilité de définition, voire de pérennité.

C’est pourquoi la commission retient que l’utilisation de ces données peut être un appui utile aux données traditionnelles, lorsque la priorité est la rapidité. Mais elle recommande que ces données soient systématiquement rapprochées des statistiques publiques et de leurs concepts et que des études de validation soient menées, pour celles qui présentent un potentiel plus pérenne, une vigilance particulière étant accordée à la connaissance du niveau de vie et de la composition des ménages Elle préconise de plus qu’une réflexion soit menée sur les possibilités d’actualisation du cadre juridique de la transmission des bases détenues par des organismes privés vers les services de statistique publique, afin que celui-ci puisse couvrir des travaux tels que ceux réalisés en 2020. La commission encourage par ailleurs le rapprochement entre producteurs, la mise en commun de leurs sources respectives et la poursuite des études croisées permettant d’enrichir les analyses des données individuelles, à l’instar des expériences menées conjointement par l’Insee et la Banque de France. Elle souhaite que ces études et leurs résultats, une fois consolidés, soient rendus publics et facilement accessibles. Enfin, la commission souhaite que les comparaisons internationales des statistiques macro-économiques puissent bénéficier d’une coordination des méthodes et d’une convergence des pratiques, au niveau européen, mais aussi mondial.

Référence : avis généraux de moyen terme n° 4 et n° 7, avis de moyen terme n° 8 de la commission « Système financier et financement de l’économie »

COMPTE RENDU DE LA RÉUNION DE

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

- 27 mai 2021 -

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 générale des statistiques, des études et de l'international (DGSEI), 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 6

I. PRÉSENTATION DU PROGRAMME ANNUEL DE TRAVAIL (PAT) 2021 DE LA STATISTIQUE PUBLIQUE 6

II. LA STATISTIQUE PUBLIQUE EN TEMPS DE CRISE : ENSEIGNEMENTS DE LA CRISE COVID 9

CONCLUSION 31

Liste des participants

ANXIONNAZ	Isabelle	Conseil national de l'information statistique (Cnis)
AUVRAY	Emmanuel	Agence nationale pour la cohésion des territoires (ANCT)
BEN HASSINE	Haithem	France Stratégie
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
BONNET	Odran	Institut National de la statistique et des études économiques (Insee) – Direction des études et des synthèses économiques (Dese)
BOULO	Yann	Ecole des hautes études en sciences sociales (EHESS)
BOURGEOIS	Alexandre	Institut national de la statistique et des études économiques (Insee)
BRICONGNE	Jean-Charles	Banque de France (BdF)
BRUN	Matthieu	Banque publique d'investissement France (Bpifrance)
CARAES	Didier	Assemblée permanente des chambres d'agriculture (APCA)
D'ALESSANDRO	Cristina	Conseil national de l'information statistique (Cnis)
DELATTE	Anne-Laure	Université Paris Dauphine
DEMUYNCK	Julien	Banque de France (BdF)
DUBOIS	Marie-Michèle	Conseil national de l'information statistique (Cnis)
DUQUERROY	Anne	Banque de France (BdF)
GOLFIER	Cécile	Banque de France (BdF)
GUILLAUMAT-TAILLIET	François	Institut National de la statistique et des études économiques (Insee) - Direction des statistiques démographiques et sociales (DSDS)
GUINOUARD	François	Banque de France (BdF)
HAIK	Elisabeth	Ubiqus
HOURIEZ	Guillaume	Institut National de la statistique et des études économiques (Insee) – Direction des études et des synthèses économiques (Dese)
JOLIVET	Patrick	Agence de l'environnement et de la maîtrise de l'énergie (Ademe)
KEMPF	Nicolas	Insee Nouvelle-Aquitaine
KREMP	Elisabeth	Banque de France (BdF)
LALANNE	Adrien	Insee Provence-Alpes-Côte d'azur
LIOGIER	Valérie	Ministère de l'Education nationale, de la jeunesse et des sports - Direction de l'évaluation, de la prospective et de la performance (Depp)
MAUREL	Françoise	Conseil national de l'information statistique (Cnis)
MONTUS	Arnaud	Conseil national de l'information statistique (Cnis)
MOREAU	Sylvain	Institut National de la statistique et des études économiques (Insee) - Direction des statistiques d'entreprises (DSE)
OLIVIA	Tom	Institut National de la statistique et des études économiques (Insee) – Direction des études et des synthèses économiques (Dese)
PERRON-BAILLY	Etienne	Insee PACA
POILON	Gwenaëlle	Mouvement des entreprises de France (Medef)
RAYMOND	Olivier	Comité national routier
SCOTT	Suzanne	Institut National de la statistique et des études économiques (Insee) – Direction des études et des synthèses économiques (Dese)
SERIEYX	Yvon	Union nationale des associations familiales (Unaf)
TAGNANI	Stéphane	Conseil national de l'information statistique (Cnis)
TAZI	Amine	Fédération bancaire française (FBF)
WICKY	Yann	Banque de France (BdF)
WILNER	Lionel	Institut National de la statistique et des études économiques (Insee) – Direction des études et des synthèses économiques (Dese)
WOLOSZKO	Nicolas	Organisation de coopération et de développement économique (OCDE)

La séance est ouverte à 14 heures 32, sous la présidence d'Anne-Laure DELATTE.

INTRODUCTION

Anne-Laure DELATTE, Présidente

Merci à tous de votre présence.

Notre séance de ce jour commencera par la présentation du programme annuel de travail, avec les deux rapporteurs de la commission que sont Cécile Golfier (Banque de France) et Guillaume Houriez (Insee). Puis dans une seconde partie, nous évoquerons la statistique publique en temps de crise, avec de nombreux intervenants qui ont bien voulu se mobiliser. Le premier point concernera l'utilisation des données massives. Puis nous ferons un focus sur la façon dont l'épargne des ménages et les comptes des entreprises ont été mesurés pendant la crise, quasiment en temps réel. Les enjeux étant très importants, la mobilisation a été très conséquente pour les suivre.

Comme à l'accoutumée, je remercie très chaleureusement Stéphane Tagnani et Isabelle Anxionnaz d'avoir organisé cette séance de main de maître. J'aurai l'occasion de le préciser en fin de séance, mais il s'agit de la dernière séance d'Isabelle Anxionnaz avec nous.

I. PRÉSENTATION DU PROGRAMME ANNUEL DE TRAVAIL (PAT) 2021 DE LA STATISTIQUE PUBLIQUE

.1 Banque de France et Autorité de contrôle prudentiel et de résolution (APCR)

Cécile GOLFIER, Banque de France

Bonjour à toutes et tous. J'interviendrai pour la Banque de France (BdF) et l'Autorité de contrôle prudentiel et de résolution (ACPR). J'aborderai dans un premier temps les trois domaines qui ont contribué au programme de la statistique à la Banque de France : les statistiques monétaires et financières, les statistiques d'entreprises et la diffusion des statistiques.

.a Statistiques monétaires et financières

Réformes de moyen terme

Les statistiques monétaires et financières mettent en œuvre deux réformes de moyen terme en 2021. Tout d'abord, la Banque prépare la réforme quinquennale des statistiques monétaires, qui intègrera des innovations financières et permettra notamment une meilleure connaissance du financement de l'immobilier d'entreprise.

Par ailleurs, le projet *européen Integrated REporting Framework* (IREF) est en cours ; il concerne à la fois la BdF et l'ACPR. A ce titre, la Banque consulte actuellement la profession bancaire afin de définir le cadre de moyen terme des statistiques monétaires et financières. Le projet IREF permettra une meilleure intégration des déclarations à des fins statistiques, utilisées par la Banque de France, et prudentielles, utilisées par l'ACPR.

Identifier les entreprises ou projets responsables

La Direction des statistiques monétaires et financières va identifier les entreprises ou projets responsables. Pour ce faire, nous définissons tout d'abord les secteurs d'intérêt, la transition énergétique et climatique d'une part, le financement de l'économie sociale et solidaire (ESS) d'autre part.

Le processus s'appuie sur une base de données granulaires sur le crédit collectées auprès des banques (base Anacrédit). Parallèlement, la Banque améliore des référentiels existants pour permettre d'identifier les entreprises ou projets relevant des secteurs qui nous intéressent. En appariant ces éléments à la base Anacredit, nous disposerons de données de référence utiles pour le suivi statistique, l'analyse et la recherche.

Épargne réglementée : évolution de la collecte

La Banque met en œuvre une évolution réglementaire de 2020. Le volet de l'enquête précédemment trimestriel devient mensuel. Cette collecte s'enrichit également d'une ventilation plus fine des emplois de l'épargne réglementée, avec la nouvelle catégorie des projets contribuant à la transition énergétique et à la réduction de l'empreinte climatique.

Observatoire de l'inclusion bancaire : mieux appréhender la population fragile

Il s'agit ici aussi de la mise en œuvre d'une évolution réglementaire de 2020. Dans le cadre de son engagement au service de l'économie et de la société, la Banque entend revoir sa collecte pour mieux identifier les populations en situation de fragilité financière ou bénéficiaires de l'offre spécifique. Un volet trimestriel sera adjoint à la collecte pour dénombrer les personnes concernées et analyser le fonctionnement de leur compte bancaire, notamment les frais bancaires.

En parallèle, une modification de la collecte annuelle est opérée.

Nomenclature des financements selon leur contribution au « verdissement » de l'économie

La Banque de France mettra à disposition de la profession bancaire une nomenclature des financements selon leur contribution au « verdissement » de l'économie. Ceci rejoint le nouveau *reporting* relatif à l'épargne réglementée en 2021. Il s'agit d'un travail transverse, consistant en un suivi rapproché des travaux du groupe d'experts de la taxinomie européenne sur la finance durable. Pour les modalités d'application de cette taxinomie, une coordination des travaux est effectuée avec l'Insee et le Commissariat général au développement durable (CGDD), Service des données et études statistiques (Sdes).

Dans le même temps, la coopération internationale se poursuit. Les réflexions sont mises en commun dans le cadre du *Committee on Monetary, Financial and Balance of Payments Statistics* (CMFB).

.b Les statistiques d'entreprises

Refonte de l'enquête complémentaire sur les échanges internationaux de services des entreprises (ECEIS)

Cette enquête complète par un sondage les collectes exhaustives réalisées auprès des plus grandes entreprises. La refonte comporte trois volets.

En premier lieu, en vue d'alléger la charge déclarative pesant sur les entreprises, la taille des échantillons est réduite d'environ un tiers.

Par ailleurs, la stratification de l'enquête est modifiée, d'une part pour accompagner la réduction de la taille de l'échantillon et d'autre part pour faciliter la production des données *Services Trade by Enterprise Characteristics* (STEC), dont la transmission à Eurostat sera obligatoire dans le cadre du règlement *European Business Statistics* (EBS).

Enfin, la réforme apporte une amélioration de la qualité de la ventilation des contreparties géographiques Désormais, la déclaration inclut toutes les contreparties géographiques. Il n'y aura donc plus d'estimation.

.c Diffusion des statistiques

La modernisation des outils va s'appliquer à trois domaines :

- une amélioration de l'ergonomie du portail statistique *Webstat* ;
- une refonte des publications statistiques *Stat Infos* ;
- un accès des chercheurs aux données individuelles *via* l'*Open Data Room* (ODR) virtuelle, étant rappelé qu'actuellement l'accès se fait à l'ODR « physique », dans les locaux de la BdF.

L'APCR

Le Ruba (*reporting* unifié des banques et assimilés) remplace le Surfi (système unifié de *reporting* financier) Son principal objectif est de rationaliser et enrichir les collectes prudentielles (ACPR) et statistiques (BdF). Il s'agit donc d'un travail conjoint entre les deux institutions, dans l'optique de supprimer les doublons et d'alléger les remises des établissements contrôlés dans le cadre du mécanisme de supervision unique (MSU), d'introduire des seuils et périodicités différenciés pour les autres.

.2 Insee

Guillaume HOURIEZ, Insee

Bonjour à tous. Je vais présenter le programme 2021 du service statistique ministériel (SSM) Direction générale des finances publiques (DGFiP) et du Département des comptes nationaux de l'Insee. Les programmes de travail détaillés sont publiés sur le site du Cnis.

Au Département des comptes nationaux, nous avons pour ambition de produire des tableaux entrées-sorties (TES) en émission de gaz à effet de serre, en utilisant le modèle Avionic que nous avons développé. Ce travail sera mené en collaboration avec le Service de la donnée et des études statistiques (SDES). Une étude de faisabilité est attendue fin 2021, avant de lancer à proprement parler les travaux de production.

.a Département des études statistiques et fiscales (DGFiP)

A la DGFiP, la réorganisation du SSM est intervenue. Le Bureau GF-3C est devenu le Département des études statistiques et fiscales. Il est composé d'un pôle de soutien à la gestion fiscale et d'un pôle dédié à la statistique publique, qui rassemble les activités d'études, de production, de diffusion et informatiques. La publication d'études liées aux finances publiques ou à des sujets économiques est attendue dès 2021.

Au titre des nouvelles sources statistiques, les données du fonds de solidarité ont été mises à disposition des chercheurs.

Par ailleurs, deux sources ne sont pas encore disponibles mais le seront prochainement. La première est liée à la nouvelle application Gestion de mes biens immobiliers (GMBI) qui fournira, à partir de 2023, des informations sur les occupants des locaux à usage d'habitation. Ces données seront précieuses puisque des informations seront perdues à la suite de la suppression de la taxe d'habitation.

En outre, le fichier Ecollab sur les plateformes en ligne qui fournissent des informations sur leurs opérations, opèrera une mise à disposition progressive pour la statistique publique.

.b Département des comptes nationaux (Insee)

Notre activité est interne car nous développons des outils informatiques et avons donc une charge de reprogrammation de l'ensemble des processus métiers, à la fois pour les comptes annuels et trimestriels.

De plus, nous commençons à préparer la nouvelle base des comptes nationaux (2020, publication en 2024), en menant un travail de rebasage des agrégats de comptabilité nationale.

Nous nous consacrons également à des travaux concernant l'inventaire RNB (revenu national brut), dans la mesure où nous avons eu une réserve générale sur l'inventaire de la base 2010. Il a donc été nécessaire de rendre un nouvel inventaire à la Commission européenne, ce qui a été fait en mars 2020 et a lancé un cycle de vérifications. Il est vraisemblable qu'en septembre 2021, cette réserve générale sera levée. Néanmoins, nous serons tenus de rendre un nouvel inventaire avant la fin de l'année, ainsi qu'un cycle de vérifications.

Finalement, les travaux envisagés sur la mesure du PIB démarreront plutôt en 2022.

Anne-Laure DELATTE

Avez-vous des questions sur le programme de travail ?

Les données relatives au fonds de solidarité sont disponibles sur Internet, ce que j'ai pu vérifier pendant la présentation. Les détails sont en effet nombreux.

Pour cette deuxième partie, nous disposons de près de deux heures trente pour évoquer la statistique publique pendant la crise de la Covid. Je demanderai à chacun d'éviter au maximum les acronymes et de les définir à chaque fois.

II. LA STATISTIQUE PUBLIQUE EN TEMPS DE CRISE : ENSEIGNEMENTS DE LA CRISE COVID

.1 Les enjeux mis en évidence pendant le premier confinement

.a Une illustration dans le champ des statistiques macroéconomiques

Guillaume HOURIEZ

La crise sanitaire, par son ampleur et sa soudaineté, a mis en évidence de nombreux enjeux pour la production de statistiques macroéconomiques. Le premier est celui de la capacité à fournir un diagnostic rapide pour que les politiques publiques puissent s'adapter au mieux à la situation économique et sociale. De plus, la crise sanitaire a révélé des enjeux de mesure de certaines activités, en particulier celles faisant l'objet de conventions, telles que les services non marchands. Par ailleurs, la crise a mis en évidence la nécessité de comparer les performances nationales avec celles des autres pays. Plus que jamais, elle a démontré le besoin de disposer d'un diagnostic allant au-delà de celui fourni par les grandes catégories d'agents du cadre central des comptes nationaux (entreprises, ménages, administrations publiques), tant les inégalités face à la crise ont été fortes.

Pour la plupart, ces enjeux ne sont pas nouveaux, mais ont été rendus encore plus visibles en raison de la crise, qui a aussi fait apparaître des nouveaux concepts. Par conséquent, un travail d'inventaire doit être mené pour tirer l'ensemble des leçons de la crise sanitaire. Ce travail n'étant pas encore abouti, je ne serai pas exhaustif dans ma présentation.

Capacité à fournir un diagnostic rapide pour que les politiques publiques puissent s'adapter à la situation économique et sociale

A l'annonce du premier confinement, la note de conjoncture de l'Insee - qui donne une prévision de croissance à six mois - était prête à paraître le 26 mars. Très rapidement cependant, Le directeur général de l'Insee a décidé d'y renoncer car elle était devenue non pertinente. A la place, ont donc été publiées des mesures de baisse instantanée de l'activité économique (le PIB et la consommation) à intervalles réguliers, tous les quinze jours environ.

Pour réaliser ces prévisions, les conjoncturistes de l'Insee ont mobilisé de nouvelles données : distribution d'électricité, données de cartes bancaires, de fret ferroviaire, etc. Je renvoie à cet égard à la présentation de Julien Pouget lors de la session précédente. Ces prévisions ont été utiles en tant que telles puisqu'elles ont fourni des informations quasiment en temps réel aux décideurs publics, mais aussi pour les comptables nationaux. En particulier pour la première estimation du premier trimestre, fin avril, les comptes trimestriels ne disposaient pas de données sur le mois de mars pour les services, et les méthodes habituelles d'extrapolation du mois de mars n'étaient plus pertinentes.

Finalement, les premières estimations se sont révélées de bonne qualité. Les exercices de *nowcasting* ont montré leur grande utilité, même s'il reste encore un travail à mener pour améliorer la précision des données. En tout état de cause, le système statistique doit être réactif afin de mobiliser rapidement des données privées. Se pose par exemple la question d'une éventuelle mise en production des données de transactions par carte bancaire pour compléter les indicateurs sous-jacents aux comptes trimestriels.

Plus largement, l'Insee continuera à examiner l'exploitation de nouvelles données et méthodes, notamment à travers un partenariat avec *Paris School of Economics* (PSE) dans le cadre d'une chaire comprenant deux volets : *nowcasting* et mesure du PIB. Cette chaire devrait démarrer à l'automne prochain.

Mieux mesurer certaines activités

La crise sanitaire a mis en évidence les difficultés de mesure de l'activité, par exemple pour les services non marchands, pas seulement du point de vue de la mesure du PIB, mais aussi en tant que telle. Il est donc intéressant de s'interroger pour savoir si les services d'éducation ont-ils été maintenus ou s'ils ont baissé. Si oui, dans quelle proportion ? Il y a évidemment un enjeu important à établir le bon diagnostic. Or aujourd'hui, les comptables nationaux ne disposent pas des informations pour répondre précisément à ces questions, de sorte qu'il est nécessaire de développer les systèmes d'information permettant d'apporter ces réponses.

D'autres questions de mesure se posent, en particulier concernant le pouvoir d'achat. Il s'agit notamment de savoir ce que représente la mesure du pouvoir d'achat, alors même que durant le confinement certains produits sont devenus essentiels et d'autres non essentiels. On peut par conséquent se demander si les prix sont encore représentatifs des préférences des ménages.

Bien entendu, d'aucuns considèrent ce problème comme transitoire et lié au confinement. Néanmoins, il semble que le problème de fond demeure, ainsi que les questions théoriques.

La crise a aussi accéléré le développement de l'économie numérique. Les débats sur le *mismeasurement*, c'est-à-dire la difficulté à mesurer l'activité d'économies de plus en plus dématérialisées et globalisées (sous-estimation des gains de productivité), sont loin d'être clos. Il existe ici un enjeu de compréhension et de mesure, se traduisant à la fois sur le plan conceptuel et sur la mise en œuvre d'outils. En la matière, les organisations internationales auront un rôle important à tenir, en particulier les travaux de l'OCDE sur l'économie numérique.

Besoin d'un diagnostic allant au-delà de celui fourni à partir des grandes catégories d'agents du cadre central des comptes nationaux

Les comptes d'agents publiés par les comptes trimestriels ont mis en évidence la baisse de revenu liée à la crise, et le fait que les administrations publiques aient absorbé l'essentiel du choc. De ce fait, le pouvoir d'achat a bien résisté et le besoin de financement des entreprises a été limité grâce au dispositif d'aides. De ce point de vue, les comptes nationaux ont bien rempli leur rôle.

Néanmoins, la crise a touché de façon très différenciée à la fois les entreprises et les ménages. Sur ce point, les statistiques macroéconomiques ont, en quelque sorte, été défaillantes. On peut penser toutefois que les études *ad hoc* sont plus adaptées pour y répondre, mettant en œuvre des outils d'analyse plus souples que ceux des comptes nationaux et permettant de faire parler des ensembles particuliers de données.

Finalement, il est possible d'estimer que la réactivité de l'Insee, de la Banque de France et des chercheurs a été satisfaisante. Il existe cependant un enjeu dans notre capacité à relier les analyses microéconomiques et l'histoire macroéconomique. La réponse appartiendra par conséquent aux comptables nationaux, mais elle semble compliquée à la fois du point de vue de la disponibilité des données et des processus de production. Certains travaux ont été engagés depuis longtemps, par exemple sur les comptes des ménages par catégories, que l'Insee a développés au début des années 2010. De même, les travaux plus récents des banques centrales, en particulier de la BdF, sur les comptes de patrimoine par catégories de ménages, devront être poursuivis. Les progrès ne pourront être raisonnablement attendus qu'à moyen ou long terme.

Nécessité de pouvoir comparer les performances nationales avec celles des autres pays

Dans cette crise, tous les pays ont mis en place des mesures sanitaires et des politiques publiques très fortes. Il apparaissait donc indispensable de comparer la performance de chacun des pays, afin de mieux s'adapter aux situations.

La crise a exhibé des écarts temporaires de méthode, et aussi sans doute d'accès à certaines données. Eurostat a produit un grand nombre de guides méthodologiques à destination des comptes trimestriels afin de faire en sorte que les résultats soient les plus comparables possibles. Toutefois, il n'a pas été possible d'aller jusqu'au bout. Il existe par conséquent un enjeu de coordination des méthodes, en faisant en sorte que les échanges entre les pays soient plus axés sur la convergence des pratiques. Les organisations internationales sont les premières concernées. Eurostat se trouve en première ligne au sein de l'Union européenne, mais aussi l'Organisation de coopération et de développement économiques (OCDE) et le

Fonds monétaire international (FMI). Nous rencontrerons bientôt l'OCDE, d'ailleurs, sur le sujet de la mesure du non-marchand pendant la crise sanitaire.

Vos questions sont les bienvenues.

Nicolas WOLOSZKO, OCDE

Qui sera en charge de la chaire PSE-Insee ?

Anne-Laure DELATTE

Je serais également intéressée de connaître le programme de travail de cette chaire, si les détails en sont déjà connus.

Sur les difficultés à harmoniser le non-marchand au niveau européen, où en est-on ? Les problèmes avaient en particulier porté sur la mesure du PIB et l'éducation, les mesures du PIB allemand et français ayant divergé. Des corrections ont-elles été apportées ?

Yvon SERIEYX, Union nationale des associations familiales (Unaf)

Nous avons ressenti ce manque d'information sur les impacts différenciés sur les ménages. Nous n'étions pas en mesure de connaître la quantité de personnes impactées, car nous n'avions pas beaucoup d'informations différenciées selon les tailles des ménages. Par conséquent, au-delà de très gros travaux de réflexion pour disposer d'indicateurs en temps réel, existe-t-il des petites modélisations simples de façon immédiate, pour garder l'œil sur l'aspect « ménages » ? Deux ans plus tard, nous avons ces chiffres, mais dans l'immédiat nous devrions pouvoir connaître le nombre d'individus impactés. Sont également passées sous silence les difficultés des familles avec enfants, notamment lorsqu'un des membres perd son emploi. Nous avons une vision très partielle des impacts de la crise sur les ménages.

Guillaume HOURIEZ

Je ne pourrai pas vous en dire davantage sur la chaire, car nous travaillons actuellement sur la convention. Cette semaine, nous avons travaillé sur la constitution du programme de travail, que je pourrai donc vous présenter à la prochaine session de la Commission.

Sur le sujet du non-marchand (éducation, santé notamment), les questions méthodologiques sont nombreuses, mais également les questions de données. Sur la santé, nous sommes très équipés pour mesurer l'activité des hôpitaux grâce au *reporting* détaillé des maladies et des soins. A l'inverse sur l'éducation, il n'est pas simple de trouver des informations concernant le suivi réel des cours et le volume d'heures disposées. C'est pourquoi au départ, nous avons mobilisé des données d'heures travaillées des enseignants, avant d'utiliser des enquêtes du SSM de l'Éducation nationale sur le nombre de « décrocheurs ». Pour les élèves n'ayant pas décroché, se posait aussi la question du nombre d'heures de cours suivies réellement, dans la mesure où certains professeurs étaient absents ou donnaient moins de cours. Les estimations en la matière restant fragiles, nous souhaiterions pouvoir les améliorer.

Nous pourrions effectivement travailler sur une modélisation d'agents hétérogènes. Au lieu d'avoir un ménage représentatif de l'ensemble, nous travaillerions sur des catégories de ménages. Il aurait cependant été nécessaire de développer ce modèle avant que la crise n'ait lieu. Cela étant, le sujet fait partie des réflexions menées en interne à l'Insee.

.b Exemples d'utilisation des données massives

Les innovations apportées à l'enquête mensuelle de conjoncture (EMC)

Julien DEMUYNCK, Banque de France

Merci de me donner m'opportunité de présenter quelques éléments sur l'enquête de conjoncture en temps de Covid. Il y a un avant et un après.

Modalités des enquêtes

La Banque de France réalise des enquêtes de conjoncture depuis très longtemps, les premières remontant dans nos bases à 1976. L'enquête se caractérise aussi par une certaine stabilité du questionnaire. Il est important de pouvoir comparer au cours du temps les différents signaux et, sur une longue période pour capter plusieurs cycles.

Les enquêtes de conjonctures présentées aujourd'hui couvrent une grande partie de l'économie, et notamment l'industrie manufacturière, les services marchands et le bâtiment. Sont donc exclus la sphère publique, les services non marchands et la sphère financière.

Pour l'industrie manufacturière, les questions posées ont trait essentiellement à la production passée, les prévisions, l'évolution du carnet de commandes, les taux de capacité d'utilisation en pourcentage, les niveaux des stocks et carnets de commande, l'évolution des prix des matières premières et des produits finis, les effectifs et la situation de trésorerie. Ces grands blocs sont suivis au mois le mois, sachant que les questions sont posées à des entreprises de l'industrie ou des services. Pour la plupart des items, nous recueillons le sentiment du chef d'entreprise sur l'évolution du mois passé et celui pour le mois à venir.

Au total pour chacun des blocs, une quinzaine de questions sont posées. Les échantillons concernent environ 8 500 entreprises, dont 3 500 dans le secteur de l'industrie, 3 500 dans les services et 1 500 dans le secteur du bâtiment. Les proportions sont relativement stables au cours du temps, de même que les échantillons.

Méthodologie

Les enquêtes visent à saisir les évolutions mensuelles en réalisant une comparaison entre le mois concerné et le mois précédent, ainsi que le positionnement par rapport à la normale. Une seule question quantitative est posée concernant l'industrie, tandis que le reste des questions portent sur l'évolution au cours du mois par rapport à une situation normale. C'est donc une recherche d'opinion.

Toutes les enquêtes de conjoncture sont réalisées par le réseau de la BdF, qui compte 117 implantations à ce jour. Les testeurs sont au nombre de 2 000 environ, sachant que les questionnaires sont en général réalisés par téléphone ou par mail, sur une durée brève. L'intermédiation nous permet d'avoir un taux de réponse élevé en général, avec peu de pertes en contexte de crise car les liens entre le chef d'entreprise et l'agent de la banque sont solides. Le chef d'entreprise est donc attaché à remonter ses problématiques.

Les échelles de notation portent classiquement sur sept positions, avec pour réponses possibles :

- en forte progression (notation + 2) ;
- en progression (+ 1) ;
- en légère progression (0.5) ;
- identique (0) ;
- en légère diminution (-0.5) ;
- en baisse (–1) ;
- en forte baisse (-2).

La même échelle s'applique en termes de niveau (comparaison avec les niveaux normaux), sauf exception.

Une zone libre de commentaire permet au testeur de synthétiser les informations pertinentes pour expliquer les réponses aux questions. Par ailleurs, le test de conjoncture est réalisé selon un entretien semi-directif, sans lecture *in extenso* d'un questionnaire. Cet entretien vise à cerner la situation de l'entreprise et à répondre aux questions de l'enquête de conjoncture.

Les non-réponses sont très peu nombreuses. Toutefois, la non-réponse d'une entreprise a un impact sur le calcul puisqu'il faut procéder à un recalcul des pondérations. La non-réponse peut donner lieu à une estimation à dire d'expert, de manière individuelle. En pratique, l'impact est faible car le taux d'estimation

des non-réponses est limité à 5 % des réponses. Les soldes d'opinions sont des moyennes pondérées des effectifs des entreprises ou établissements ainsi que par la valeur ajoutée.

Nouvelles exploitations

Le focus portait exclusivement sur la partie « commentaires », puisqu'il s'agit de commentaires libres. Des questions récurrentes portent sur le niveau d'activité depuis avril 2020. A cet égard, il est demandé aux chefs d'entreprise leur sentiment par rapport à une situation normale pré-crise. Nous formalisons ces questions dans une zone de commentaires, et exploitons le résultat *via* une analyse textuelle.

Il est également possible de poser des questions plus ponctuelles sur l'actualité. En début de crise, les questions portaient sur le télétravail ou le recours au prêt garanti par l'État (PGE). A l'automne dernier avant le deuxième confinement, nous avons sondé les chefs d'entreprise sur le retour à la normale. Plus récemment, les questions ont porté sur la mesure des surcoûts induits par la crise, à la fois les coûts directs (achat de matériel) et indirects (réorganisation des process).

Parallèlement à ces questions formalisées, certaines questions relèvent davantage du *text-mining*. Il s'agit de lire le questionnaire et de relever les occurrences de certains mots, ce qui nous a notamment permis de produire, dans l'un des points de conjoncture, une mesure de l'incertitude. Nous avions relevé des mots-clés par exemple sur les difficultés à se projeter, sur l'incertitude face à l'avenir. Une batterie d'expressions permettait de mesurer si le chef d'entreprise se trouvait en situation d'incertitude, et de construire un indicateur au cours du temps. De même sur les difficultés d'approvisionnement, nous avons une publication en cours.

Extraits de publications dans les différents Points de conjoncture

A la question *« Comment positionnez-vous sur l'activité de votre entreprise par rapport à la normale ?* », l'exploitation des données amène à produire un agrégat moyen du niveau d'activité selon les secteurs. Nous examinons comment évolue ce ratio, ce qui permet de se positionner par rapport au premier confinement et d'évaluer le retour à la normale. De même pendant le premier confinement, nous avions publié un indicateur sur le retour à la normale, attestant de l'optimisme des chefs d'entreprise : plus de la moitié d'entre eux anticipaient en effet un retour à la normale dès 2020.

Indicateur d'incertitude

Il ne s'agit pas d'une question formalisée sur l'incertitude mais d'un dénombrement de mots-clés. Les comparaisons au cours du temps sont particulièrement intéressantes. L'effet « Gilets jaunes » est relevé sur le graphique à la fin 2018, tandis que la crise Covid a une ampleur quatre fois plus importante. La perception de l'incertitude progresse ou reflue selon les confinements.

Difficultés d'approvisionnement

Sur cet aspect, nous avons juste procédé à une analyse textuelle des entreprises qui mentionnent l'approvisionnement dans leurs commentaires. L'automobile est concernée au premier chef, ainsi que quelques secteurs de l'industrie et le bâtiment. Les commentaires sont donc d'intérêt pour l'exploitation économique qui en est faite.

Anne-Laure DELATTE

Des questions sont-elles incluses sur l'état financier des entreprises, par exemple sur les difficultés à faire face des remboursements ? Est-il possible d'avoir des informations de nature financière ?

Julien DEMUYNCK

Il existe un volet trimestriel de questions sur l'accès des entreprises au marché du crédit bancaire, concernant le même échantillon d'entreprises que celui évoqué. Il leur est demandé si elles ont sollicité un crédit, et le cas échéant, dans quelles proportions elles l'ont obtenu. C'est donc un volet spécifique « financement bancaire ». Dans l'EMC présenté ce jour, figure une réplique, portant sur la situation de trésorerie. Nous avons à cet égard constaté qu'au moment où les PGE se sont fortement développés, la situation de trésorerie des entreprises s'est très significativement améliorée, de manière historique.

Raphaëlle BERTHOLON, CFE-CGC

Parvient-on à dégager une vision sur le passage à l'économie numérique ? Nous voyons bien que les entreprises purement numériques ont eu des progressions importantes, tandis que d'autres ont dû, en raison du confinement, par exemple passer au *click and collect*. Existe-t-il des éléments permettant d'appréhender la part de ces entreprises ?

Julien DEMUYNCK

C'est une très bonne question. Nous ne mesurons pas ce point directement dans l'enquête. Un grand nombre de secteurs, notamment industriels, n'ont pas été fortement touchés pendant les confinements et se sont adaptés. Une enquête parallèle dénommée « Commerce de détail », sonde plusieurs milliers de points de vente qui ont subi de plein fouet la crise. Ces boutiques ont dû faire face en développant le *click and collect* et la vente à distance

Je vous invite à consulter ces publications ayant donné lieu à une synthèse annuelle, qui mettent bien en évidence le fait que le commerce de détail ait été fortement impacté par la crise. Du fait du changement des comportements d'achat des ménages, le développement de la vente à distance a plus ou moins impacté certains secteurs.

Apports du *web-scraping* pour un suivi à haute fréquence du marché immobilier : le cas du Royaume-Uni (RU)

Jean-Charles BRICONGNE, Banque de France

Le *web-scraping* est le téléchargement massif de données sur Internet. Cette présentation se focalise sur le thème de l'immobilier au Royaume-Uni. Originellement, nous souhaitions couvrir le Brexit, puis nous en sommes arrivés à couvrir le Brexit et le confinement en 2020, et enfin la pandémie.

L'approche retenue pour le Royaume-Uni a été de pratiquer le *web-scraping* tous les jours sur les principaux sites immobiliers, afin d'avoir une évolution en dynamique. Les données collectées de cette façon permettent un suivi granulaire (tout en veillant à la représentativité pour traiter de l'ensemble des régions) et plus régulier des marchés immobiliers. L'objet n'est pas de remplacer les statistiques officielles mais de bénéficier d'un angle complémentaire, avec le point de vue des vendeurs, voire leur psychologie pendant le confinement.

Dans le cas du Royaume-Uni, le marché a été gelé pendant le premier confinement, avec un effondrement temporaire des offres postées mais une relative stabilité des prix (absolus), sauf à Londres. Au Royaume-Uni, peu d'offres indiquent les surfaces pour privilégier l'indication du nombre de pièces. Les prix sont donc communiqués en valeur absolue par logement, et non au mètre carré.

Les données de *scraping*, en appliquant cette technique à d'autres pays, sont cohérentes avec les sources officielles et permettent d'obtenir des prix en niveau (ce qui est intéressant pour les banquiers centraux et les régulateurs), qui apportent de l'information par rapport aux indices. Il est ainsi possible de prévoir de façon plus précise les retournements et d'estimer les stocks de richesse immobilière. Plus spécifiquement au Royaume-Uni, de nombreuses informations publiques sont disponibles en temps réel sur les sites immobiliers : 92 % des agences immobilières publient des annonces sur Internet, de sorte que cette source est représentative.

L'utilisation de données alternatives pour suivre le marché immobilier n'est pas nouvelle. Par exemple, les *Google Trends* ou d'autres études issues du *web-scraping* (par exemple Hanson et Santas, 2014 ou Bricongne et al., 2019) sont disponibles. Le *web-scraping* a le mérite d'offrir le point de vue des vendeurs, alors que les transactions finales résultent de l'interaction entre vendeurs et acheteurs.

En l'espèce, la couverture a concerné uniquement le marché britannique dans cette période particulière de Covid-19 et de Brexit, sachant que l'approche peut être étendue à d'autres zones géographiques. La couverture des sites utilisés a été variable selon les régions, mais s'est trouvée assez élevée dans tous les cas pour assurer une bonne représentativité au niveau infra-national.

La plus grande partie des données étaient issues des sites majeurs d'annonces immobilières au RU : *Rightmove*, *Zoopla* et *OnTheMarket*, ainsi que d'autres sites plus localisés pour l'Irlande du Nord et l'Écosse. Grâce à ces sites Internet, plus de 1,5 million d'offres immobilières sont téléchargées en moyenne chaque jour. Nettoyé des doublons, le nombre d'offres avoisine encore 500 000, avec des URL uniques par jour. Environ les deux tiers de ces dernières sont des offres de vente. Les données pour *Zoopla* sont téléchargées (et nettoyées) depuis début mars 2020 alors que les données d'autres sites Internet le sont depuis juillet 2020. Pour la location, un grand nombre de biens sont également disponibles. Dans le domaine commercial, il existe aussi un fort potentiel.

Le graphique communiqué est fort intéressant. Il laisse apparaître qu'avant le confinement sur le site *Zoopla*, près de 45 000 nouvelles offres hebdomadaires étaient postées. Pendant le confinement, le nombre de ces offres a été divisé par plus de 8, pour retomber à 5 000 pendant un mois environ avant de revenir progressivement à la normale. Pendant le deuxième confinement, on observe une baisse significative mais moins spectaculaire que la première. En définitive, le premier confinement a donné lieu à une quasi-disparition des offres postées.

L'autre graphique relatif aux nouvelles offres immobilières, laisse apparaître que la proportion des biens encore présents un mois après leur première publication sur Internet était encore de 90 % pendant le premier confinement (alors qu'en période normale, la proportion est plutôt des deux tiers).

S'agissant des prix, une relative stabilité – voire une légère hausse dans le sud-est de l'Angleterre – est relevée. En revanche sur les biens en prix absolus à Londres, l'effritement du prix moyen des biens est constaté (d'un million de livres sterling à 900 000 livres sterling en septembre). La raison en est soit une vente accrue des petites surfaces soit une baisse du prix au mètre carré, ce qui nécessiterait de mener une étude complémentaire.

Ces analyses des prix postés permettent de prendre connaissance de la psychologie des vendeurs. De façon générale, les statistiques issues du *web-scraping* sont complémentaires des statistiques officielles et permettent un suivi en temps réel et granulaire. De façon plus structurelle, elles permettent aussi un calcul de prix (ou de loyers) en niveaux. Les deux sources sont utiles notamment pour les banques centrales afin d'identifier les risques de retournement et la stabilité financière.

Le *web-scraping* permet un calcul de prix en niveau

Il y a quelques années à la Commission européenne, les offres postées à un instant T dans un certain nombre de pays de l'OCDE ont été chargées, en utilisant les principaux sites immobiliers identifiés. Un calcul en deux temps a été effectué. Par région, un calcul du prix moyen a été obtenu en considérant le prix au mètre carré, avec une moyenne pondérée par la surface du logement. Puis une fois le prix moyen par mètre carré et région obtenu, une nouvelle pondération a été effectuée par le stock de mètres carrés fournis par le recensement.

Des ordres de grandeur cohérents

Les données de *scraping*, comparées aux données des comptes nationaux de stocks de mètres carrés pour une vingtaine de pays, laissent apparaître un écart maximum de 12 % et un écart médian de 8 %. De ce fait, à partir de données *de scraping* gratuites et librement accessibles, il est très intéressant de savoir que l'ordre de grandeur est correct, à 10 % près. Pour les pays ne publiant aucune donnée immobilière, tels que la Grèce, la Roumanie ou la Bulgarie par exemple, il est donc d'autant plus intéressant d'avoir un chiffrage. Enfin par rapport aux enquêtes publiées dont les périmètres ne couvrent pas toujours tout le parc de logements, les données de *scraping* sont beaucoup plus proches des données de comptes nationaux.

Quelques applications possibles

Le *web-scraping* pourrait permettre d'estimer la richesse immobilière (nationale ou régionale), surtout dans les pays qui ne publient pas de données. Il pourrait également aider à prévoir les retournements grâce aux ratios de prix sur revenu. Au Royaume-Uni, nous avions par exemple montré que les marges de négociation entre le prix posté et le prix de transaction étaient plus faibles à Londres et plus fortes dans les régions en province. Il pourrait donc être intéressant d'observer l'évolution au cours du temps. Si la marge de négociation augmentait à Londres, cette donnée indiquerait sans doute que les offreurs auraient moins de pouvoir de marché, et donc que le marché se retournerait.

Prolongement

La Banque de France n'est pas seule à travailler sur ces données, puisque l'OCDE, la Banque mondiale et le FMI s'y attèlent également. Par conséquent, une mise à disposition de statistiques alternatives sur les prix et les loyers (issues du *web-scraping* ou éventuellement d'autres sources comme les données satellites) devrait intervenir sur un site unique hébergé par le réseau *Real Estate Finance and Economics* (ReFinE) (projet en cours).

Conclusions, perspectives, enjeux

Les données issues du *web-scraping* permettent d'approfondir la comparaison avec les statistiques officielles, en gardant à l'esprit les biais possibles. Elles aboutissent en fait à comparer encore davantage les données postées et transactées. Il pourrait être intéressant d'analyser l'évolution entre le dernier prix posté et le prix transacté, avec l'intuition que l'écart serait très faible. Une analyse est en cours pour étendre l'observation aux loyers, au prix des terrains et à l'immobilier commercial.

Par ailleurs, la granularité peut être utilisée pour étudier les effets de certaines politiques réglementaires ou certains chocs touchant des régions ou des villes. Le travail est mené avec la Banque mondiale pour mener une extension de l'approche à d'autres pays.

Enfin, des questions juridiques et de sécurisation des sources liées au *web-craping* se posent, dans la mesure où certains sites ne souhaitent pas être « *scrapés* ». En général, les exceptions existent au bénéfice des chercheurs, dont le travail est bien entendu différent de celui des *scrapeurs* utilisant les données dans un but commercial, pour faire concurrence au site. Il demeure qu'une zone grise doit être levée. De plus, si les sources doivent être utilisées de manière récurrente, il faudra s'assurer qu'elles n'ont pas été modifiées ou qu'elles n'ont pas disparu avec le temps. Par conséquent pour publier des statistiques officielles ou semi-officielles, se posera le sujet de la coopération avec la source d'origine.

Yvon SERIEYX

Sur la question des surfaces et du nombre de pièces, avez-vous aussi une bonne visibilité ?

Jean-Charles BRICONGNE

J'ai choisi précisément le cas du Royaume-Uni car seules 10 à 20 % des offres publient les surfaces. Nous avons préféré ne pas utiliser ces offres, car nous avons constaté qu'elles étaient biaisées. En réalité, les offres publiant les surfaces sont le plus souvent situées à Londres, très certainement à l'intention des expatriés – en particulier des Français - qui recherchent au premier chef tant l'indication du prix que de la surface. Par conséquent, le fait de conserver uniquement les offres publiant les surfaces aurait conduit à surpondérer Londres et sa région.

Cependant, nous avons tenté de creuser ce sujet du nombre de pièces et de la surface, car même en province il existe un certain nombre de biens pour lesquels les surfaces sont indiquées. Nous avons commencé, mais pour l'heure le travail n'est pas entièrement conclusif. Nous avons observé le lien entre le nombre de pièces et les surfaces, en prenant en compte le fait qu'à Londres la taille moyenne des pièces doit être inférieure à la taille en province, avec une valeur constante non comptée comme une pièce (salle de bains, cuisine, parties communes etc…). Il est vrai que le Royaume-Uni est un cas particulier dans la mesure où contrairement à d'autres pays, dont la France, les surfaces ne sont pas systématiquement indiquées. Il faut donc utiliser les prix absolus ou les prix par pièce, dont l'indication sera quelque peu différente entre Londres ou d'autres régions en raison de la différence de taille des pièces.

Anne-Laure DELATTE

Nous aurons un temps d'échanges après le point suivant. Pour toutes les questions relatives au *web-scraping* et aux données massives, nous utiliserons donc ce temps.

Nicolas WOLOSZKO, OCDE

Je propose un tour d'horizon et une sélection parfaitement arbitraire d'un certain nombre de projets qui ont attiré notre attention à l'OCDE depuis le début de la crise, utilisant le *Big Data* pour effectuer de la prévision économique. Je vais donc commencer par le mien, pour évoquer ensuite ceux de personnes avec lesquelles nous avons travaillé.

The *OECD Weekly Tracker*

L'*OECD Weekly Tracker* est un outil créé depuis décembre 2020, fournissant une estimation en temps réel de l'activité économique hebdomadaire, avec deux jours de délai de publication. A la fin de la semaine, il est donc possible de connaître l'activité de la semaine passée. Cet outil utilise les données *Google Trends* et un algorithme de *machine learning* pour extraire du signal environ 250 catégories de mots-clés de recherche effectuées par les internautes sur Google.

Sur le graphique est communiqué l'exemple de la France. En rouge sont représentées les séries officielles de PIB et en noir, l'une des versions du *Weekly Tracker* fournissant une estimation du PIB par rapport à ce qu'il aurait été en l'absence de crise. Nous utilisons des projections de l'OCDE de novembre 2019, pour constater par exemple que l'activité de la semaine dernière était de 4 % inférieure à sa trajectoire prévue.

Google Trends fournit un indice des volumes de recherche (*Search Volume Indices*) à partir de mots-clés ou catégories de mots-clés. Il est intéressant de constater qu'un grand nombre de ces catégories sont directement liées à l'activité économique (par exemple les données de consommation sur les produits et boissons alimentaires, destinations touristiques, marques automobiles). Les marchés de l'emploi et de l'immobilier sont également concernés. L'accès public à *Google Trends* est limité. Il existe un site Internet consultable, sur lequel chacune des séries peut être téléchargée manuellement, mais ce travail est fastidieux et limité en volume. L'OCDE a obtenu un arrangement avec Google en juin 2019 grâce à son chef économiste Hal Varian, qui nous a conféré un accès privilégié à l'*Application Programmer Interface* (API).

Taux de recrutement *LinkedIn* (*Hiring Rate*)

Ce taux correspond au nombre d'utilisateurs de *LinkedIn* déclarant avoir obtenu un nouvel emploi dans le cours d'un mois calendaire, divisé par le nombre total d'utilisateurs de *LinkedIn* par pays.

En termes de gouvernance, *LinkedIn* a constitué une équipe dénommée la *Economic Graph Team,* composée d'environ cinquante chercheurs économistes et ayant vocation à produire de la recherche à partir de la manne de données dont dispose *LinkedIn*, pour soutenir les décisionnaires politiques. *LinkedIn* possède en effet des données en temps réel sur le marché du travail, en particulier sur les offres d'emploi et sur les salariés d'un grand nombre de pays. Le *Hiring Rate* est utilisé par la plateforme dans le cadre du *Workforce Report* consultable sur le site *LinkedIn*, et également par la cinquième édition du *Bulletin économique de la Banque centrale européenne* (BCE).

Pour ma part, j'ai contacté *LinkedIn* en urgence en avril dernier afin de recueillir leurs données, pour suivre le marché du travail pendant la crise. La réponse a été favorable puisque *LinkedIn* a communiqué les données demandées. Les études de validation ont été menées pour déterminer dans quelle mesure les données d'un fournisseur privé étaient concordantes avec des données publiques, en l'occurrence les *Labor Force Studies* d'Eurostat (ELFS).

Les résultats sont quelque peu décevants. Les données sont pertinentes pour capturer les chocs importants, ainsi que dans les quelques pays où les parts de marché de *LinkedIn* dépassent 90 %. En revanche dans les autres pays, les résultats ne sont pas satisfaisants. Même si l'échéant est gigantesque, de l'ordre de 50% de la force de travail, cet échantillon est biaisé car *LinkedIn* réunit plutôt les cadres. L'évolution n'est donc pas représentative, sauf en cas de grand choc macroéconomique affectant toutes les catégories.

Par ailleurs, la collaboration avec *LinkedIn* a été globalement difficile car leurs données ne sont pas publiées librement, à la différence de *Google Trends*. Ces données nous ont été fournies de manière discrétionnaire et non en temps réel, de sorte que nous n'avons pas pu les utiliser réellement.

Indeed

Indeed collecte des informations sur les offres d'emplois. Le site s'est doté d'un laboratoire de recherches (*Hiring Lab*) dont la figure de proue est Tara Sinclair, tandis qu'un autre chercheur dirige les recherches pour la France. Les pays couverts sont la majorité des pays développés. Tous les chercheurs sont salariés du laboratoire. Les rapports ne sont pas publics et se retrouvent dans *le World Economics Outlook* (WEO) du FMI.

La principale différence avec *LinkedIn* tient à la forme d'autorité académique conférée à *Indeed* par la présence de Tara Sinclair et de chercheurs académiques.

Havard Opportunity Insights

Il s'agit de l'un des exemples les plus aboutis d'emploi des données issues du *Big Data* pour le *nowcasting*.

Harvard Opportunity Insights a créé une base de données publique, permettant de suivre en temps réel l'activité économique des Etats-Unis. Les études sont opérées par quartier (*county*) en temps réel, grâce à un grand nombre de partenariats noués avec des firmes privées fournissant des données anonymisées. Les données quotidiennes concernent notamment les dépenses de consommation, les revenus des entreprises, le taux d'emploi.

Cette approche est différente de celle du *webscraping* puisqu'elle tire profit uniquement des partenariats conclus. J'en donnerai quelques exemples. Pour les dépenses de consommation, *Affinity Solutions* est le fait, pour les entreprises, de vendre les données sur les dépenses de cartes de crédit à des services financiers. Pour l'emploi, *Earnin* est le fait, pour une entreprise qui verse à ses salariés leur salaire en début de mois plutôt qu'à la fin, de recueillir les données d'une très large base de salariés.

L'étude de l'évolution des dépenses de consommation laisse apparaître que la consommation a le plus baissé dans les quartiers riches, qui sont les plus nombreux en temps normal au restaurant et au théâtre. Par conséquent, l'emploi a baissé massivement dans les catégories les plus défavorisées (serveurs par exemple), travaillant dans les quartiers riches. Cette donnée a été fournie très tôt aux politiques pendant la crise, ce qui démontre la force de frappe d'Harvard.

Étude sur les données de transactions des cartes bancaires en France (Camille Landais)

Cette étude a notamment révélé la grande hétérogénéité des secteurs.

Development Data Partnership

Cette initiative a été lancée par le FMI, la Banque mondiale et la Banque de coopération interaméricaine. L'objectif est de créer un cadre de coopération juridique et technique permettant le partage de données par des grandes entreprises de technologie avec des acteurs publics.

En synthèse

Le *Big Data* propose des aperçus en temps réel, granulaires et non coûteux. Les difficultés sont néanmoins liées au fait que les données ne sont pas collectées dans le but de produire de la science. Les études de validation sont donc indispensables, même si nombre de publications ne les mettent pas en œuvre. Dans le cadre d'une crise telle que l'actuelle, Tara Sinclair considère que la valeur ajoutée des données réside dans leur temporalité (temps réel) et leur granularité.

Une question importante se pose relativement à la gouvernance de l'accès public des données privées, sachant que *Google* est la seule entreprise ayant construit des outils dans le but de partager les données, sans atteinte à la vie privée. D'autres entreprises telles que *LinkedIn* et *Indeed* ont créé des cellules de recherche qui exploitent ces données, mais n'ont pas vocation à les partager avec les chercheurs et les entités publiques.

Enfin, l'approche par partenariats entre le public et le privé est intéressante. Il reste à concevoir une gouvernance basée sur la notion de bien public et une régulation pour faciliter l'accès à ces données.

Anne-Laure DELATTE

Merci, c'était passionnant. Nous sommes au cœur de notre sujet.

L'exploitation de données bancaires pour observer la situation financière des ménages en 2020

Odran BONNET, Insee

Notre présentation a trait aux données bancaires que nous avons pu utiliser, et qui proviennent du Crédit Mutuel. Ces données nous ont permis de faire une publication dans la *Note de conjoncture* de mars dernier.

Introduction et motivations

La genèse est celle du contexte exceptionnel de l'an dernier, avec une crise sanitaire sans précédent et une crise économique subséquente. A l'Insee, nous avons publié un grand nombre de renseignements agrégés sur l'évolution de la consommation ou du PIB. Pour autant, il nous a manqué de pouvoir désagréger ces évolutions au niveau des ménages. Nous avions une idée générale de la façon dont la consommation avait baissé avant d'augmenter à nouveau à la fin du premier confinement, mais n'étions pas en mesure de déterminer le type de ménages touchés (les plus pauvres, les plus riches, les catégories socioprofessionnelles, les tranches d'âges).

Par conséquent, il existait un besoin de renseigner la situation économique des agents selon leurs caractéristiques socio-démographiques avant que les données fiscales ne soient disponibles. L'Insee est en mesure d'y procéder à plus ou moins long terme, en utilisant généralement des données fiscales. Or le défaut de ces données est d'être disponibles uniquement un an plus tard.

Nous avons eu l'opportunité d'utiliser les données du Crédit Mutuel pour dresser un panorama de la situation financière des ménages en 2020, et de la façon dont ils ont subi la crise. L'échantillon a été pris à partir de janvier 2019, puis mis à jour chaque mois. Les données étaient constituées des soldes mensuels au niveau du « groupe famille » (et non « individu »), des comptes courants détenus, des livrets d'épargne, des comptes titres et des assurances-vie. Nous avons ainsi pu suivre l'évolution du patrimoine financier détenu par un groupe ménage au sein du CIC-Crédit Mutuel.

Nous disposions également des données de transactions, avec le détail de chacune d'entre elles. Le détail des transactions par carte était disponible dans notre base de données, en la liant au groupe ménage qui les avait réalisées. Nous disposions en outre d'informations sur le type d'achats (codes MCC) ainsi que sur les prélèvements et les virements. De plus, nous étions en possession de variables qualitatives telles que la catégorie socio-professionnelle des membres du ménage, l'âge et le code d'habitation.

Forts de ces données, nous avons publié dans la note de conjoncture de mars des informations sur la situation financière des ménages et leur consommation. Il en ressort en particulier :

- une chute de la consommation deux fois plus intense au premier qu'au second confinement ;
- une chute plus prononcée parmi les personnes aisées, qui consommaient davantage auparavant : la consommation est devenue sous contrainte en se concentrant sur les biens nécessaires, sans pouvoir se diriger vers les produits habituels de type loisirs ou restaurants ;
- une forte hausse de l'épargne pour tous les déciles de ménages, mais surtout pour les plus aisés en volume : la consommation a chuté mais les revenus ont été relativement stables, malgré une baisse de revenus mesurée sur l'ensemble des virements entrants sur les comptes ;

Ces données peuvent être utilisées à d'autres fins que des études de conjoncture, par exemple pour évaluer en temps réel la situation des personnes qui connaissent une modification de leur situation personnelle (chômage, veuvage divorce) et l'impact de cette nouvelle situation sur leur richesse.

De plus, le calcul de la propension marginale à consommer pourrait être effectué, en mesurant la part de ce revenu additionnel destinée à la consommation. Cette étude aurait aussi son importance pour les politiques publiques et la mesure de l'effet des plans de relance par la consommation. Enfin, des études pourraient porter sur le patrimoine.

Contenu des données

L'échantillon a porté sur 300 000 groupes-familles depuis janvier 2019, ce groupe étant entendu comme le client du Groupe Crédit Mutuel et éventuellement son conjoint et ses enfants mineurs. Cependant, ces données présentent le défaut de ne pas avoir pour but la réalisation d'études économiques, d'où des problèmes de représentativité. En effet, nous disposions des données du Crédit Mutuel, dont les clients sont spécifiques à cette banque et non représentatifs de la totalité de la population française. Toutefois, cette absence de représentativité doit être nuancée, dans la mesure où l'évolution des soldes des comptes de l'échantillon a été comparé à l'échantillon agrégé des soldes de la Banque de France, après un calage sur marge. Finalement, les évolutions sont apparues très similaires.

En définitive pour étudier des populations très particulières – les plus riches ou les plus pauvres – ces données pourraient souffrir d'un biais, mais elles sont valides pour des études plus générales.

Par ailleurs, une autre limite tient à la multi-bancarisation. Le patrimoine, les revenus ou la consommation des ménages ne peuvent donc être évalués exactement si un ménage possède plusieurs comptes en banque.

En outre, le revenu s'est avéré difficile à appréhender car il n'est pas constitué uniquement par la somme des virements entrants. Il convient à cet égard de préciser que l'étude n'a pas porté sur les virements de compte à compte au sein d'un même ménage, mais uniquement sur les virements entre ménages différents ou en provenance d'une autre banque. Finalement, les personnes possédant deux comptes bancaires effectuent des virements de régularisation, ce qui a pour conséquence de bruiter le message sur l'évolution des revenus.

L'accès aux données

Tom OLIVIA, Insee

Dans le cadre de notre partenariat, nous avons un accès distant à une machine virtuelle (VM) de la banque, à la fois avec un identifiant personnel et avec une partie mobile du mot de passe en provenance d'une clé RSA : le code change toutes les minutes. La traçabilité et la sécurité sont donc garanties.

De plus avec la VM Windows mise à disposition à distance, nous accédons à un Data Lab, qui se trouve être un environnement de travail incluant tous les logiciels de statistiques nécessaires à l'étude *via* un navigateur web interne à la banque. Après une nouvelle série d'authentification, les données sont appelées avec les procédures SQL dans Python ou R.

Un contrôle des entrées et sorties dans la VM est pratiqué par les salariés du Crédit Mutuel car cette machine n'a pas accès à Internet. Concrètement, les agents du Crédit Mutuel mettent à notre disposition des fichiers de sortie contrôlés en amont sur une plateforme de partage de fichiers accessible depuis l'extérieur. Un système sécurisé est également appliqué aux entrées, ainsi qu'un contrôle de l'utilisation de *packages* pour éviter d'éventuels problèmes de sécurité.

Quelques exemples d'exploitations

Consommations

Le graphique étudie la consommation par dixième pondéré de revenu de février à décembre 2020, c'est-à-dire pendant les deux confinements. Pendant le premier confinement, le choc est plus fort pour les hauts déciles de revenu et plus mesuré pour les bas déciles. Ce même constat est transposé toute l'année, même si le choc du deuxième confinement est moindre. Il faut souligner que la série présentée est désaisonnalisée par un système de ratio et *dé-trendée* : la tendance préexistant à la crise Covid est nettoyée et lavée de ces données.

Patrimoine financier

Odran BONNET

Nous avons étudié l'évolution de l'épargne, plus précisément du patrimoine financier constitué de l'ensemble des soldes des différents comptes des ménages au Crédit Mutuel-CIC. Les graphiques rendent compte de l'évolution des déciles de patrimoine financier, tant en pourcentage qu'en montant.

En pourcentage, on observe que les plus bas déciles sont ceux ayant connu la plus grosse augmentation de leur patrimoine financier au moment du confinement. En réalité, ces ménages possédant peu d'argent sur leur compte en banque (100 euros) ont vu ce patrimoine augmenter de 50 ou 60 euros, mais baisser de nouveau pendant l'été. Il s'est donc produit un effet de rattrapage et de consommation, les ménages concernés ayant puisé dans l'épargne accumulée pendant le confinement.

En absolu, ce sont les plus gros déciles qui ont le plus augmenté leur épargne. Cette augmentation ne se retrouve pas nécessairement en pourcentage, car pour ces ménages possédant déjà des revenus élevés, une augmentation même importante en euros sur une année donnée, représente peu par rapport au patrimoine accumulé tout au long de la vie.

En définitive, l'un des messages principaux tient au constat de l'évolution du patrimoine financier pour tous les déciles de patrimoine, l'épargne agrégée ayant toutefois été tirée par les plus hauts patrimoines et les plus hauts revenus.

Revenus

Tom OLIVIA

Les graphiques comparent des statistiques de quartiles avec la moyenne et la médiane de revenus mensuels avant et après crise. Cette série est également désaisonnalisée et *dé-trendée*, en effectuant des ratios de déciles et non plus des moyennes comme pour la consommation.

Il apparaît que le choc des revenus est différent de celui de la consommation. Certes, une baisse relative des revenus est constatée pendant le premier confinement, mais un rebond a lieu en juin ainsi qu'une stabilisation à des niveaux inférieurs à ce que prédisait la tendance pré-Covid.

A noter que le terme « revenus » se réfère à un *proxy* nettoyé du mieux possible. De ce fait, les sujets de multi-bancarisation et de virements ponctuels polluent la série de revenus. A ce jour, il s'agit de l'un des défauts les plus prégnants de ces données.

Conclusions

La mise à disposition des données bancaires a permis de dresser un panorama de l'impact de la crise sur les ménages, en les différenciant selon leur catégorie. C'est donc l'apport principal de ces données, outre l'aspect de temps réel et *nowcasting*.

Au-delà, des études de plus long terme seront susceptibles d'être réalisées. La *Note de conjoncture* peut être consultée afin de prendre connaissance plus avant du travail mené.

Anne-Laure DELATTE

Merci pour cette présentation passionnante. Nous ouvrons à présent les questions pendant une dizaine de minutes.

Yvon SERIEYX

Votre démarche est très intéressante. Elle permet d'avoir des analyses en termes de revenus disponibles des individus et des ménages, à la multi-bancarité près. Il faut en effet rappeler le taux actuel de 25 % de multi-bancarité pour les individus, ce qui signifie que ce taux est encore plus élevé pour les ménages. Les conjoints n'ont pas nécessairement le même compte en banque, ou ont une banque différente. Il est donc

difficile d'acquérir une vision exacte du niveau de vie. De plus, votre décile dans l'échelle des revenus ne correspond que rarement au décile dans le niveau de vie.

Par conséquent, cette vision concerne davantage les individus que les ménages, ce qui est normal puisque les banques traitent avant tout avec des individus. Dès lors, je m'interroge sur l'utilité de la démarche, dès lors qu'elle tente de mesurer les impacts sur les ménages dans toute leur diversité, c'est-à-dire selon leur taille et leur structure.

Odran BONNET

Je partage absolument vos réserves. Ces données sont originales mais présentent en effet des défauts car elles ne se substituent pas aux données fiscales ou d'enquête, qui sont disponibles plus tard. Le taux de multi-bancarisation de 25 % que vous citez est très important, mais une grande partie de la multi-bancarisation porte aussi sur les comptes dormants. Or ces derniers ne sont pas inclus dans notre étude.

Sur le fait que les ménages soient mal captés et que les flux ne portent pas sur les revenus mais sur les virements entrants, je partage votre constat. Cependant, nous avons comparé les niveaux de vie médians à la distribution effectuée par l'Insee annuellement. Il en ressort que dans le milieu de la distribution, les revenus apparaissent à un niveau comparable. En revanche, il existe une tendance à surestimer les hauts revenus par rapport aux prévisions de l'Insee. Ce point s'explique par le fait qu'une partie des virements entrants est liée à la multi-bancarisation, donc à des transferts entre ménages.

En définitive, ces données sur la situation financière des ménages sont partielles. Pour certaines études, utilisées en évolution, elles permettront toutefois d'acquérir une première idée de la situation en temps réel des ménages.

Tom OLIVIA

Nous avons effectué une comparaison avec la série agrégée diffusée par la BdF, et les ordres de grandeur sont comparables en tendance. Ceci complète en partie la réponse sur les questions de représentativité.

Raphaëlle BERTHOLON

Je souhaite revenir sur les questions posées par le *web-scraping* en matière d'accès aux données. Peut-on imaginer une réglementation et une régulation européenne pour rendre ces données plus accessibles, au-delà de l'initiative mentionnée ? Est également en jeu la qualité des données produites, dont on peut supposer qu'elles pourraient générer les appétits d'organismes privés peu sérieux, qui pourraient en faire le commerce. Je suppose que le comité du label du Cnis s'est interrogé à ce sujet.

Jean-Charles BRICONGNE

J'ai discuté avec une juriste de la BdF sur la question de la légalité du recours au *scraping*. En France, il convient de différencier la réglementation européenne du RGPD et la réglementation générale sur la protection des données. Dans le RGPD, une exception est spécifiquement prévue pour les travaux d'étude et de recherche. Nous sommes donc couverts de ce point de vue. De plus, les données sur l'immobilier comportent peu d'aspects confidentiels. En revanche suivant les pays, les législations pénales peuvent être plus ou moins restrictives en matière de téléchargement de données, même si elles sont disponibles sur des sites publics. En l'occurrence, une entreprise privée telle que SeLoger.fr *scrapant* les données d'AVendre.fr sera susceptible d'être poursuivie. Au contraire, même s'il est peu probable qu'un site incrimine un chercheur utilisant ses données, il reste la question pénale à résoudre. Cette insécurité juridique nécessite par conséquent d'être réglée.

Anne Laure DELATTE

Nous pouvons retenir une tension entre la nécessité d'obtenir des données très rapidement - ce que permettent les données massives issues du *scraping* – pour adapter en temps réel des politiques économiques, et la qualité de ces données par rapport à la statistique publique classique. En la matière, on constate que la qualité des données est sujette à vérification. Finalement, les données fiscales disponibles trois ou six mois plus tard peuvent apparaître suffisantes pour confirmer des constats apparus pendant une crise.

Dans le point suivant, nous allons réaborder la question de l'épargne, mais à travers des données publiques administratives.

En conclusion, nous reviendrons sur la question de la réglementation discutée à l'occasion des présentations. Il nous semble que les finances publiques sont en jeu, notamment dans l'aspect de transfert des données commerciales privées vers la statistique publique. Un encadrement est en cours de réflexion pour minimiser au maximum les coûts de ce transfert. L'initiative globale présentée par Nicolas de se réunir pour sécuriser l'accès et les coûts devra par conséquent être suivie.

Nous allons à présent étudier la mesure de l'épargne des ménages, mais cette fois grâce à la collecte de données effectuée de façon plus classique.

.2 La mesure de l'épargne des ménages et les comptes des entreprises

.a Dispositif de suivi de l'épargne financière des ménages en temps de crise et concept de surplus d'épargne

Yann WICKY, Banque de France

Je travaille au service des statistiques européennes de la BdF. Je présenterai dans un premier temps la manière dont la BdF a pu suivre la situation du point de vue du financement des ménages, mais également de leur épargne dès le début de la crise de la Covid et de la crise économique. L'existant nous permet d'avoir un suivi assez pointu du financement et des dépôts des ménages. Nous étudierons ensuite les mesures spécifiques mises en place. François Guinouard présentera enfin les éléments concernant plus spécifiquement l'épargne des ménages.

L'existant

Dispositif de suivi de l'épargne financière des ménages

A la veille de la crise, nous disposions déjà d'un ensemble de collectes de données nous permettant de pratiquer un suivi pertinent des dépôts et des crédits. Ces collectes sont encadrées par des règlements européens tels que le règlement BCE 2013/33 sur les encours de bilan, permettant une comparaison avec le reste de l'Eurosystème et le règlement BCE 2013/34 sur les taux d'intérêt. De plus, la banque a mis au point le dispositif Surfi, qui est un dispositif de collecte auprès des 600 établissements de crédit français.

L'environnement Surfi est composé de tableaux, dont le tableau « M_CLIENRE » sur les encours de dépôts et de crédits aux agents non financiers (ANF). Ce tableau, remis de façon mensuelle par les 200 principales banques, couvre environ 97 % du bilan. Les banques représentant le reste (3 %) remettent le tableau de façon trimestrielle.

Le tableau « M_CLIENRE » fournit un détail intéressant par poste. Il est donc possible, pour le financement, de déterminer précisément le type de crédit accordé pour les entreprises et les ménages. Pour la partie « passif », le détail est tout aussi intéressant sur le type de dépôts pratiqués par les ménages auprès des banques françaises.

Finalement, le suivi mensuel des dépôts est très précis, en particulier au moment où la crise se déclenche pour les dépôts nets (différence entre dépôt et crédit des ménages).

Mise en place d'une collecte hebdomadaire

Dès le mois de mars 2020, nous avons lancé une collecte *ad hoc* à vocation temporaire auprès des dix principales banques françaises, pour obtenir un résumé hebdomadaire du tableau « M_CLIENRE ». Nous ne demandons que quelques postes pour les sociétés non financières et les ménages. Les banques ont répondu favorablement, dans la mesure où le coût était relativement limité pour elles.

Nous avons constaté que les remises étaient de qualité, ce qui a permis une captation rapide des signaux d'une constitution importante d'épargne par les ménages. En revanche à ce niveau de fréquence infra-mensuelle, certaines informations, telles que la prise en compte des amortissements, apparaissaient moins exploitables. La même limite est apparue sur les dépôts, alors que sur le long terme, l'exploitation immédiate

permet d'étudier comment se comportent, à un instant T, les dépôts des ménages. De surcroît, la fréquence infra-mensuelle n'a pas permis la prise en compte des autres éléments d'épargne tels que l'assurance-vie et les organismes de placements collectifs (OPC).

Le tableau « M_CLIENRE » hebdomadaire a été récolté pendant trente semaines au moment de la crise, mais la communication externe a été maintenue à une fréquence mensuelle plus classique.

Dispositif de suivi de l'épargne financière des ménages : note mensuelle

François GUINOUARD, Banque de France

Les statistiques monétaires permettent de bien connaître les dépôts bancaires à une fréquence mensuelle, avec des délais d'un peu moins d'un mois après la fin d'un mois donné. En revanche, afin d'acquérir une vision plus exhaustive de l'épargne des ménages, le suivi trimestriel se base sur la comptabilité nationale et les comptes financiers. Sont ainsi disponibles tous les placements (notamment l'assurance-vie, les actions, et les obligations). Pour autant, le suivi trimestriel présente l'inconvénient de la temporalité, dans le contexte de la crise.

Nous avons donc fait en sorte d'adapter l'existant dans la manière de produire les données, afin d'élaborer une note de résultats s'appuyant sur des sources différentes de celles habituellement utilisées dans la comptabilité nationale. D'un certain point de vue, cette démarche est similaire à celle de l'Insee dans le cadre des comptes non financiers, dans la mesure où d'autres sources de données sont utilisées. Des points de vigilance ont été nécessaires quant à la qualité de ces autres sources et à leur périmètre. En particulier, nous avons cherché à contrôler la cohérence entre les données de placements financiers dont disposait la BdF et les données de l'épargne produites par l'Insee. Nous avons été attentifs à ne pas diffuser de messages divergents sur les comportements des ménages et sur leurs placements financiers.

Dans le tableau que nous avons publié, la dernière ligne correspond à la publication de l'Insee sur l'épargne financière des ménages. Les autres lignes ont été établies avec les statistiques de la BdF.

Le principal message vérifié tient au fait que les ménages ont beaucoup épargné du fait de la restriction de la consommation, cette épargne s'étant matérialisée essentiellement sur les dépôts bancaires au cours de l'année 2020. Ces constats ont été confirmés par la publication des comptes trimestriels officiels.

En termes de communication, nous avons rapidement constaté que les ménages ont beaucoup épargné pendant la crise, et qu'il était parlant d'effectuer une comparaison par rapport à la situation normale. La communication a donc été axée sur la notion de surplus d'épargne financière, c'est-à-dire l'écart entre l'épargne constatée des ménages et l'épargne qu'ils auraient effectuée en l'absence de crise en 2019.

Toutefois, le chiffre absolu de 180 milliards d'euros d'épargne financière sur une année n'est pas nécessairement évocateur, dans la mesure où le grand public n'a pas conscience de l'ampleur de l'épargne financière en temps normal. C'est pourquoi le chiffre de surplus d'épargne de 110 milliards d'euros sur l'année 2020 a été plus généralement repris dans la communication. Ce même chiffre continue d'être abondamment repris et commenté par les décideurs politiques. Nous avons effectué plusieurs simulations pour constater l'important surplus d'épargne, essentiellement concentré sur les dépôts bancaires, sachant que la collecte hebdomadaire nous avait renseignés dès le trimestre sur l'augmentation massive de ces dépôts bancaires.

Anne-Laure DELATTE

Je retiens de votre présentation que l'excès d'épargne se confirme, mais que nous ne disposons pas d'information sur la granularité. En d'autres termes, la distribution par revenu de cette accumulation d'épargne n'est pas connue.

Yann WICKY

Pour obtenir la granularité de la donnée, il conviendrait de pratiquer le ligne à ligne, ce qui est un tout autre exercice en particulier sur la partie « dépôt ». Sur la partie « financement », la granularité forte est possible. Sur la partie dépôts, les éléments du Crédit Mutuel ont permis d'aller assez loin dans les statistiques de

dispersion, notamment. Il aurait toutefois été beaucoup trop lourd pour les banques de collecter cette information auprès d'elles.

Anne-Laure DELATTE

Ce n'est pas non plus dans le mandat de la Banque de France, qui s'intéresse plus à de l'agrégation.

Yann WICKY

Le mandat n'est pas limitant, car nous pouvons nous intéresser à de la donnée granulaire, qui a un sens du point de vue de la stabilité financière. Néanmoins, nous sommes toujours attentifs à respecter la charge statistique et à éviter de noyer les banques sous cette charge. Pendant la crise, chacun est demandeur de données. La situation est également compliquée pour les déclarants, et ne voulions pas les perturber en demandant trop de données.

Yvon SERIEYX

Même en pratiquant ligne à ligne, la vision en termes de niveau de vie et de ménages selon leur composition est impossible à retrouver. Par conséquent, si ces données sont aussi utilisées par les pouvoirs publics pour une politique de mobilisation de l'épargne, elles sont très incomplètes, sauf à les faire passer par un modèle qui les différencie. Par exemple, un retraité ayant cumulé 6 000 euros pendant le confinement est dans une situation différente de celle d'une mère de famille de deux enfants, qui sait son emploi menacé et qui a eu une accumulation mécanique de son épargne. Or ces deux lignes sont identiques dans les tableaux.

Yann WICKY

Je partage tout à fait votre point de vue. En réalité, la vocation première de notre démarche était d'avoir un suivi sur la partie financement, notamment concernant la mise en œuvre des PGE. Le fait qu'une accumulation d'épargne ait eu lieu nous a conduits à employer la notion de « surplus d'épargne », pour éviter une sorte de tromperie pour le public. Il aurait en effet été totalement erroné et inconvenant de laisser supposer que les Français s'étaient enrichis pendant la crise. C'est pourquoi nous avons été très réservés sur les publications externes.

François MOURIAUX, Banque de France

Je suis directeur des statistiques monétaires à la BdF. Je souhaite revenir sur le constat selon lequel le suivi de l'évolution de l'épargne financière des ménages, en particulier des dépôts, s'effectuerait au plan macro sans granularité. Les missions de la Banque sont de collecter toutes informations qui lui sont nécessaires pour ses analyses. Il est clair que les données distributionnelles sont importantes pour ces analyses des banques centrales. En particulier, il existe une série de réflexions sur les impacts redistributifs de la politique monétaire de bas taux d'intérêt, par exemple. Nous avons bien conscience du besoin d'améliorer le dispositif de mesure de la répartition de l'épargne financière des ménages, par catégorie socio-professionnelle ou tranche de revenus.

Par conséquent à la lumière de l'expérience des besoins analytiques apparus pendant la crise pandémique, nous allons engager des travaux sur ce sujet. Ce type de travail doit être très bien articulé avec l'enquête Patrimoine, de manière concertée avec l'Insee.

Anne-Laure DELATTE

Merci pour cette précision importante. Bien évidemment, l'effet redistributif de la politique monétaire est essentiel.

.b Impact de la crise sanitaire sur l'activité des entreprises : microsimulation sur des données individuelles

Anne DUQUERROY, Banque de France

Nous allons vous donner un aperçu d'un exercice en cours de finalisation mené conjointement par l'Insee et la Banque de France sur l'impact de la crise Covid sur l'activité des entreprises d'une part, et sur leur trésorerie et besoin de liquidités d'autre part.

Sur la partie activité, un séminaire Insee a eu lieu. Sur la partie trésorerie, la publication officielle aura lieu fin juin. Nous évoquerons par conséquent la méthodologie uniquement, car nous ne pouvons pas encore communiquer sur les résultats.

Présentation de l'exercice : *nowcasting* sur la trésorerie des entreprises

Le but de cet exercice est d'étudier comment le choc lié à la Covid-19 et aux mesures de restriction sanitaires a conduit à une contraction de l'activité mettant sous tension la liquidité des entreprises non financières françaises. Nous sommes particulièrement intéressés par l'aspect individuel et donc, par la granularité, la dispersion et l'hétérogénéité.

Cet exercice de microsimulation comptable part de données individuelles de TVA, pour quantifier le choc d'activité mois après mois en 2020, ce choc étant exprimé par rapport à une situation contrefactuelle (soit l'activité normale sans crise, estimée à partir d'un modèle). Puis les impacts du choc d'activité sur la trésorerie des entreprises seront estimés, en intégrant le recours aux différents dispositifs de soutien public tels que l'activité partielle, le fonds de solidarité et les reports de charges fiscales et sociales.

Le déficit de trésorerie, s'il existe et la demande de crédit ou financement nécessaire pour absorber ce déficit, seront également étudiés. Enfin, l'évaluation de la soutenabilité de la hausse de l'endettement nécessaire à absorber le choc au niveau individuel sera estimée, avec une approche de stabilité financière.

Au fur et à mesure de ces quatre étapes, l'hétérogénéité des situations est caractérisée par taille des entreprises, caractéristique, secteur ou cotation Banque de France.

La démarche et la mobilisation des données pour cet exercice sont uniques. Il s'agit d'un exemple de collaboration entre les deux institutions, sans passer par le Centre d'accès sécurisé aux données (CASD). L'objectif était de mettre en commun des données que nous n'avions pas l'habitude d'apparier, notamment les données du registre des crédits individuels Banque de France, les données d'activité, et les données de recours aux dispositifs de soutien.

L'ensemble de ces données ont été mises à disposition sur le serveur de l'Insee (hors données de cotation non partageables), tandis que les économistes de la Banque de France ont été accueillis à l'Insee pendant toute la durée de l'étude.

La communication sera coordonnée entre les deux institutions. Pour l'heure, seul un blog sur la partie « choc d'activité » a été publié sur les deux sites. La sortie de deux *working papers* est attendue pour la fin juin, respectivement sur le choc d'activité et sur le choc de liquidité.

Un assemblage unique de données

Suzanne SCOTT, Insee

L'exercice a nécessité un assemblage de données unique et complexe, en particulier :

- les bilans des entreprises (base Fare) ;
- les données de déclaration de TVA des entreprises, à partir desquelles leur chiffre d'affaires a été reconstitué ;
- les reports de cotisations patronales ;
- les besoins de renouvellement des crédits Banque de France (Anacredit)

- les prêts garantis par l'État ;
- les encours mobilisés et mobilisables.

D'autres données envisagées au début du projet, mais finalement non mobilisées car elles n'étaient pas disponibles, sont celles relatives au Fonds de solidarité et à l'impôt sur les sociétés.

Champ de l'étude

Pour parvenir à notre échantillon final, une série de restrictions successives ont été opérées, à commencer par des restrictions de champ. Nous avons donc retiré le secteur financier, le secteur de l'enseignement et des administrations publiques, les entrepreneurs individuels et les personnes physiques, ainsi que les entreprises présentant des données imputées. Enfin, nous avons supprimé les entreprises présentant des déclarations de TVA mensuelles non exploitables, même après correction des reports de déclaration et des valeurs aberrantes liées au mode de collecte des données.

Ce travail de mise en cohérence est nécessaire pour exploiter nos données, de sorte que nous avons dû éliminer les entreprises pour lesquelles les chiffres d'affaires n'étaient pas cohérents entre les déclarations de Fare 2018 et les déclarations de TVA.

Finalement, l'échantillon de plus de 745 000 unités légales représente 85 % de la valeur ajoutée de notre périmètre de travail.

En termes de représentativité, la répartition des entreprises entre les différents secteurs d'activité se retrouve, ce qui est moins le cas pour la répartition entre les différentes tailles d'entreprises. En effet, les micro-entreprises sont moins représentées puisqu'elles déclarent au régime annuel et non mensuel. Il existe également une légère sous-représentation des grandes entreprises, en raison de problèmes de déséquilibre entre leurs déclarations de TVA et les bilans annuels de Fare.

Estimation des chocs d'activité individuels

La logique a conduit à calculer un taux de variation entre le chiffre d'affaires reconstruit à partir des déclarations de TVA et un chiffre d'affaires contrefactuel correspondant à la situation qui aurait été observée en l'absence de crise sanitaire. Le chiffre d'affaires contrefactuel a été calculé en deux étapes. Nous avons en premier lieu estimé un contrefactuel agrégé au niveau des combinaisons de secteur et de taille, puis ventilé ce chiffre d'affaires entre les entreprises de chaque combinaison. Un volume d'activité a ainsi été attribué aux entreprises en fonction de leur dynamique récente de part de marché et de leur saisonnalité propre. Cette méthode a permis d'obtenir des chocs au niveau très fin, tant individuel que sectoriel ou sur l'ensemble de l'économie, à divers degrés de granularité.

Un choc d'activité important, inégalement distribué

Une perte d'activité conséquente de -15 %, sur l'ensemble de la période choquée de février à décembre 2020, est constatée. Au moment du premier confinement, la perte constatée en avril atteignait -36 %, avant un rattrapage modéré sur la fin de l'année mais sans constater d'effet de rebond.

Cette approche macroéconomique masque les situations individuelles très variées, que les calculs contrefactuels individuels permettent d'observer.

La vision en coupe permet de constater une forte dispersion intra-sectorielle des pertes d'activité. Pour autant le rôle important du secteur n'explique pas toute l'hétérogénéité du choc d'activité individuel. Une approche plus dynamique par classification des trajectoires d'entreprises aboutit à identifier quatre profils homogènes de chocs au cours de l'année, dont les profils médians sont représentés dans le graphique présenté.

Il est constaté que la trajectoire diffère au cours des trois grandes périodes de l'année : une variation de l'intensité du choc lors du premier confinement, puis une variation de l'ampleur du rattrapage pendant le déconfinement et, finalement, la présence ou non selon les profils, d'un impact du second confinement. Est notée également la présence d'un groupe de 6 % des entreprises pour lesquelles l'activité n'a jamais repris à la suite du choc du premier confinement.

L'explication de la probabilité d'appartenir à chacun de ces groupes est encore en cours, mais les résultats préliminaires vont dans le sens d'une probabilité accrue, pour les micro-entreprises les plus jeunes et pour les entreprises du secteur de la culture, du tourisme et de la restauration, d'appartenir aux deux groupes les plus choqués.

Choc de trésorerie : méthodologie

Anne DUQUERROY

A partir des données de TVA nettoyées, et donnant des chroniques de chiffre d'affaires individuelles pour chaque entreprise, l'ensemble des décaissements et encaissements des entreprises sont retracés mois par mois sur toute l'année 2020. La méthodologie adoptée est celle d'un flux de trésorerie selon la méthodologie d'analyse financière suivie par la Direction des entreprises de la Banque de France.

Le tableau de flux issu des données Fiben de la Banque de France en novembre 2020 est communiqué à titre d'exemple. A partir du chiffre d'affaires issu des données de TVA, sont retranchées un certain nombre de données intermédiaires (charges, masse salariale), pour parvenir à un niveau d'excédent brut d'exploitation (EBE). Puis sont retranchés un certain nombre de flux, dont les décaissements liés à la variation du besoin en fonds de roulement (BFR), au paiement de l'impôt sur les sociétés ou des dividendes.

Le flux de financement apparaît positif, avant la variation de trésorerie finale.

Le tableau de flux – sauf le bloc des flux de financement - est modélisé pour obtenir une variation de trésorerie. L'interrogation porte ensuite sur le besoin d'augmenter l'endettement ou les fonds propres pour absorber le choc de trésorerie lié à la crise.

Au total, l'ensemble du raisonnement est mené avant PGE, en incorporant les dispositifs de soutien pour gonfler le flux de trésorerie à un mois donné. L'estimation sera ensuite réconciliée avec les demandes effectives des entreprises au titre du PGE ou avec les données de la centrale des risques sur l'augmentation des flux des encours de crédit.

Finalement, il s'agit d'une approche dynamique intéressante à présent que les bilans 2020 disponibles. Cette approche nous permet d'avoir une image mensuelle – à l'inverse d'un bilan – et de mesurer la dynamique de la liquidité selon les secteurs et les confinements.

Choc de trésorerie : grille d'analyse

Le concept repris dans les études similaires est celui de l'impasse de trésorerie ou l'illiquidité. Face à une variation de trésorerie négative ou à un déficit de *cash-flow*, l'entreprise puise dans son *cash* disponible. Une fois celui-ci épuisé, l'entreprise est dite illiquide.

Pour notre part, nous avons privilégié le concept de besoin de financement opérationnel. Dans cette hypothèse, l'entreprise absorbe la baisse de *cash-flow* liée au choc d'activité, tout en conservant un volant de trésorerie opérationnelle pour aborder la phase de reprise de l'activité.

Ainsi, l'impact en termes de liquidité sera comparé avec et sans mesures de soutien, ce qui est intéressant pour évaluer l'efficacité des dispositifs de soutien publics. De manière plus extrapolative, il sera également intéressant d'examiner les conséquences de la cessation des mesures de soutien, en termes de besoin de liquidité pour l'entreprise.

Enfin, nous sommes attachés à obtenir une analyse granulaire pour identifier les entreprises ayant des besoins de financement, tout comme celles ayant connu des variations de trésorerie positives : l'analyse sera menée en fonction de leurs caractéristiques et de la cotation de la Banque de France. De même, l'impact des mesures de soutien et le type d'entreprises aidées seront étudiés.

L'ensemble des résultats seront publiés à la fin juin.

Anne-Laure DELATTE

Sur le graphique comportant les quatre courbes, vous avez indiqué que la probabilité de se trouver dans ces courbes était liée à certains secteurs particuliers tels que le tourisme et la restauration. Les informations sont donc déjà disponibles.

Anne DUQUERROY

La dimension sectorielle joue en effet un rôle, mais d'autres facteurs ont été explorés tels que l'âge de l'entreprise ou le fait d'être ou non exportatrice. Nous avons été surpris par le fait que le secteur n'expliquait pas tous les constats. En temps normal, l'appartenance sectorielle n'explique pas l'hétérogénéité de l'activité, alors qu'au contraire pendant la crise, il s'agit d'un facteur partiel d'explication. Pour autant, une grande part de la variance a d'autres explications, de sorte qu'il conviendra d'analyser les raisons pour lesquelles les entreprises de secteur et taille équivalents se sont mieux adaptées.

Raphaëlle BERTHOLON

Les annonces des mesures de sortie de crise ont déjà été effectuées, mais on peut supposer que les résultats de vos travaux permettront éventuellement d'y apporter des correctifs.

Anne DUQUERROY

Sur ce point, je serai très prudente car l'outil n'est pas destiné à prévoir des correctifs. Il est possible de faire des constats sur l'intensité du soutien ou sur les entreprises qui ont bénéficié de telle forme de dispositif, ou encore sur celles qui n'auraient pas pu s'en sortir sans ces dispositifs. En revanche, il nous est difficile de prévoir la situation dès que les dispositifs d'aides auront été supprimés.

Raphaëlle BERTHOLON

J'ai en tête des mesures spécifiques sectorielles, calibrées de manière assez fine. Sur celles-ci, les interrogations subsistent, notamment dans les domaines de l'évènementiel ou du tourisme.

Anne DUQUERROY

Dans la simulation, nous avons eu connaissance des données de recours à l'activité partielle puisque nous observons les indemnités versées aux entreprises. Nous avons également connaissance des données de report des cotisations sociales de l'Agence centrale des organismes de sécurité sociale (Acoss), et avons simulé des reports d'IS. Nous n'avons pas eu la chance d'avoir les données du Fonds de solidarité, qui ont été simulées en fonction de la taille, du secteur et du chiffre d'affaires.

Anne-Laure DELATTE

A Bercy, un grand nombre de simulations sont effectuées. Je pense par conséquent que votre question, Raphaëlle, pourrait concerner la Direction générale du Trésor.

Merci beaucoup pour l'ensemble de ces présentations. Nous nous réjouissons de cette collaboration entre l'Insee et la Banque de France, qui a permis de fusionner des données jamais fusionnées habituellement, pour obtenir des éclairages tout à fait nouveaux.

Avis de la commission – Isabelle ANXIONNAZ, Cnis

« La commission « Système financier et financement de l'économie » a pris connaissance de quelques avancées et expériences en termes de statistiques économiques et financières, menées dans le contexte de la crise sanitaire, pour en mesurer les conséquences.

Les travaux menés à l'Insee, visant à fournir un diagnostic rapide pour que les politiques publiques puissent s'adapter au mieux à la situation économique et sociale, ont révélé des difficultés à mesurer certaines activités, notamment celles issues des services non marchands, ainsi que la nécessité de décliner un diagnostic par catégorie, aussi bien au niveau des entreprises que des ménages. Le contexte a par ailleurs confirmé l'importance de pouvoir comparer les performances nationales avec celles autres pays.

Les enquêtes mensuelles de conjoncture de la Banque de France ont été enrichies de questions supplémentaires et les commentaires libres qu'elles comportent ont été exploités en utilisant de nouvelles techniques de traitement de l'information.
Le recours à l'exploitation des données massives a été mis en œuvre pour le suivi conjoncturel dans plusieurs domaines, en France et à l'étranger.

Un suivi à haute fréquence du marché immobilier au Royaume-Uni a été appréhendé par la Banque de France à partir de données issues du webscraping d'informations disponibles sur les sites immobiliers. L'OCDE fait des prévisions économiques en temps réel à un rythme hebdomadaire pour 45 pays à partir de données de recherche Google agrégées par Google Trends. L'exploitation des données de plateformes comme Indeed ou de réseaux sociaux comme Linked In par les opérateurs propriétaires fournit une estimation en temps quasi réel des évolutions du marché du travail, au niveau de l'offre ou de la demande.

Aux États-Unis, Harvard coordonne des travaux d'exploitation menés dans différents champs affectés par la crise à partir de plusieurs sources privées, qui alimentent une base de données individuelles anonymisées accessible à tous en temps réel. L'exploitation de données d'un panel de cartes bancaires, menée par la Banque de France, a permis une observation de la situation financière des ménages en 2020.
Des exploitations nouvelles de sources existantes ont été menées pour suivre l'épargne financière des ménages à partir des données de la Banque de France, ou de manière plus originale à partir d'une mise en commun inhabituelle de données entre l'Insee et la Banque de France pour mesurer l'impact de la crise sur l'activité des entreprises.

La commission salue les initiatives de la statistique publique et encourage la consolidation des travaux en cours et la poursuite des investigations.

Concernant l'utilisation des données massives, la commission note que ces données sont granulaires, disponibles rapidement, « homogènes » d'un pays à l'autre, avec un coût marginal proche de zéro et une charge de collecte faible. Mais, en tant que traces numériques des activités humaines, elles ne sont pas conçues à des fins scientifiques et ne présentent pas de garantie en termes de représentativité, d'historique, de stabilité de définition, voire de pérennité. C'est pourquoi la commission retient que l'utilisation de ces données peut être un appui utile aux données traditionnelles, lorsque la priorité est la rapidité. Mais elle recommande que ces données soient systématiquement rapprochées des statistiques publiques et que des études de validation soient menées, pour celles qui présentent un potentiel plus pérenne.

Elle préconise de plus qu'une réflexion soit menée sur les possibilités d'actualisation du cadre juridique de la transmission des bases détenues par des organismes privés vers les services de statistique publique, afin que celui-ci puisse couvrir des travaux tels que ceux réalisés en 2020.
La commission encourage par ailleurs le rapprochement entre producteurs, la mise en commun de leurs sources respectives et la poursuite des études croisées permettant d'enrichir les analyses des données individuelles, à l'instar des expériences menées conjointement par l'Insee et la Banque de France. Elle souhaite que ces études et leurs résultats, une fois consolidés, soient rendus publics et facilement accessibles.

Enfin, la commission souhaite que les comparaisons internationales des statistiques macro-économiques puissent bénéficier d'une coordination des méthodes et d'une convergence des pratiques, au niveau européen, mais aussi mondial. »

Yvon SERIEYX

Est-il possible, dans l'avant-dernier paragraphe, d'ajouter sur les ménages et les limites *« notamment sur leur type et leur composition »* ? Je pense que la disposition aurait été plus indiquée dans les recommandations, mais je ne vois pas où ajouter la phrase.

Après la phrase commençant par « *C'est pourquoi la commission retient que l'utilisation de ces données peut être un appui utile aux données traditionnelles »,* il pourrait être ajouté :

« Une vigilance particulière est à accorder à la connaissance du niveau de vie et à la composition des ménages. »

Jean-Charles BRICONGNE

Après *« un suivi à haute fréquence du marché immobilier au Royaume Uni »*, il serait intéressant d'ajouter *« permettent d'obtenir des prix en niveau et de richesse immobilière comparables aux données officielles d'un certain nombre de pays de l'OCDE*. »

Françoise MAUREL, Cnis

L'ajout me convient, mais il pourrait être utile de recommander que les données soient systématiquement rapprochées des statistiques publiques « *et de leurs concepts ».* Cette disposition interviendrait à place de la remarque sur le niveau de vie.

Odran BONNET

Ce n'est pas un « *panel de cartes bancaires* » mais plutôt un panel de *« comptes »* bancaires.

Je me joins à la dernière remarque sur l'importance de respecter les concepts.

CONCLUSION

Anne-Laure DELATTE

Merci pour la richesse des présentations, qui nous ont permis d'apprendre beaucoup.

Je remercie très chaleureusement Isabelle Anxionnaz, dont ce sera la dernière séance en commun avant qu'elle prenne sa retraite. J'ai beaucoup apprécié ton patronage pour cette commission.

Merci également à Stéphane Tagnani pour l'organisation, ainsi qu'à Cécile Golfier et Guillaume Houriez pour leur mobilisation. La prochaine commission se tiendra le 18 novembre 2021.

La séance est levée à 17 heures 45.

***Tous les documents sont disponibles sur le site* cnis.fr *à la page de la commission*.**

DOCUMENTS PRÉPARATOIRES

- Bilan 2020 de la statistique publique - Commission Système financier et financement de l'économie (pdf - 446 Ko)
- Programme de travail 2021 de la statistique publique - Commission Système financier et financement de l'économie (pdf - 332 Ko)

DOCUMENTS COMPLÉMENTARES

- Programme annuel de travail 2021 - Diaporama (pdf - 403 Ko)
- Les enjeux de la crise sanitaire dans le champ des statistiques macroéconomiques - Insee - Diaporama (pdf - 533 Ko
- Les innovations apportées à l'enquête mensuelle de conjoncture - Banque de France - Diaporama (pdf - 1 Mo)
- Apports du webscraping pour le suivi du marché immobilier du Royaume-Uni - Banque de France - Diaporama (pdf - 3 Mo)
- Utilisation des données massives : état des lieux en France et expériences étrangères réussies - OCDE - Diaporama (pdf - 1 Mo)
- L'exploitation de données bancaires pour observer la situation financière des ménages en 2020 - Insee - Diaporama (pdf - 577 Ko)
- Suivi de l'épargne financière des ménages en temps de crise et concept de surplus d'épargne - Banque de France - Diaporama (pdf - 924 Ko)
- Microsimuation sur les données individuelles d'entreprises - Banque de France - Insee - Diaporama (pdf - 1 Mo)