

Enrichissement de l'Enquête Mensuelle de Conjoncture de la Banque de France : enseignements de l'analyse textuelle des commentaires des chefs d'entreprise

Mathilde Gerardin¹ & Martial Ranvier²

Juillet 2021, WP #821

RÉSUMÉ

Dans le cadre de l'enquête mensuelle de conjoncture de la Banque de France, ce document tire les principaux enseignements de l'analyse textuelle des commentaires des chefs d'entreprise. Tout d'abord, la richesse de ces données est illustrée via un indice de sentiment élémentaire et l'identification des principaux mouvements sociaux depuis 2009 par listes de mots-clés. L'article présente ensuite deux applications statistiques dont la reproductibilité est discutée. La première, appliquée aux gilets jaunes de 2018 et aux grèves de 2019, a pour objectif d'estimer l'impact sur le PIB d'un événement dont l'effet est sans équivoque. La seconde, adossée à l'étude du Brexit, vise à caractériser à partir d'un modèle d'apprentissage supervisé et de *word vectors* les effets d'un événement complexe avec des incidences multiples.

Mots-clés : analyse textuelle, enquête de conjoncture, indice de sentiment, mots-clés, *word vectors*, Brexit, mouvement social

Classification JEL : C21, C45, C52, E32, D22

Les Documents de travail reflètent les idées personnelles de leurs auteurs et n'expriment pas nécessairement la position de la Banque de France. Ils sont disponibles sur publications.banque-france.fr

¹ Banque de France, mathilde.gerardin@banque-france.fr

² Banque de France, martial.ranvier@banque-france.fr

Nous remercions Laurent Ferrara pour sa relecture et ses remarques sur le document. Merci également à Valérie Chauvin, Bertrand Pluyaud et Hélène Tanguy pour leurs conseils avisés tout au long de ces travaux et de la rédaction de ce document. L'estimation de l'impact des grèves basée sur les deux variantes du modèle ISMA a été réalisée par l'équipe DiaConj. Les travaux sur la partie Brexit ont quant à eux bénéficié de l'expertise de Sophie Haincourt et Maeva Silvestrini, et d'échanges avec Antoine Berthou et Caroline Jardet.

NON-TECHNICAL SUMMARY

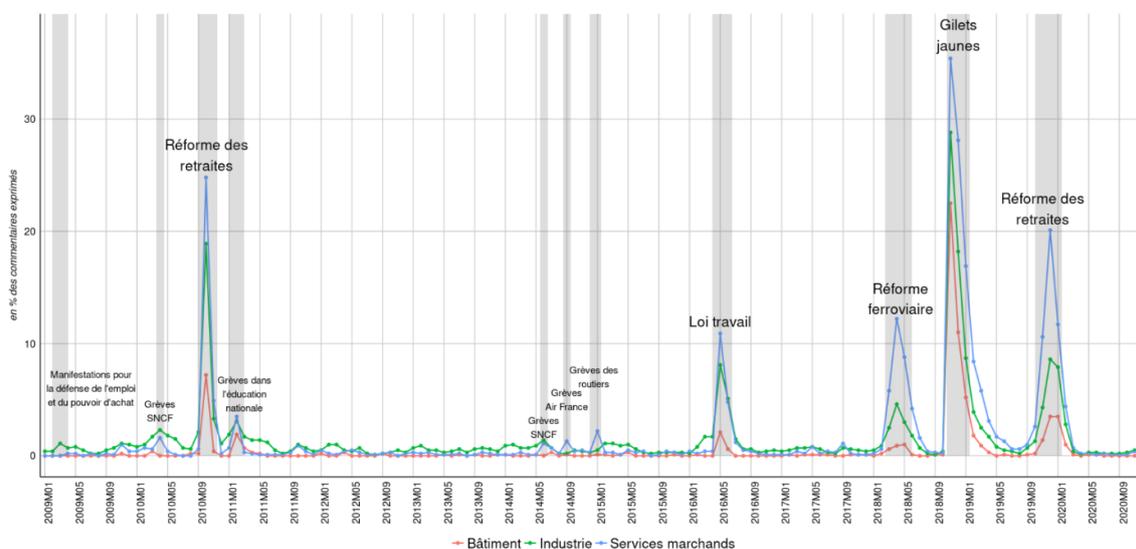
Réalisée par entretien téléphonique semi-dirigé, la collecte de l'enquête mensuelle de conjoncture est retranscrite à la fois sous forme d'échelles d'opinion et d'un résumé synthétique reprenant sous forme textuelle le contenu des informations communiquées lors de l'entretien. Ces commentaires permettent de préciser les réponses par des éléments de contexte utiles à l'analyse conjoncturelle (données chiffrées, informations factuelles sur l'actualité de l'entreprise et sur ses marchés). Ce document de travail fait le point sur les principaux résultats obtenus via l'exploitation de ces données textuelles complémentaires, depuis le développement de cet axe d'analyse en septembre 2018. Un corpus de plus de 500 000 documents a ainsi été constitué, correspondant aux déclarations des entreprises aux enquêtes industrie, services marchands et bâtiment, recueillies entre 2009 et 2021.

La richesse et le contenu informationnel de ces données est en premier lieu illustré grâce à la construction d'un indice textuel de sentiment (IS). Construit sur la base d'un dictionnaire de mots polarisés (positifs, neutres, négatifs), l'IS présente des corrélations avec l'indice du climat des affaires (ICA) comprises entre 0,6 et 0,9 selon les enquêtes et le mode de calcul. Ce résultat dénote une capacité forte des données textuelles à recomposer les informations quantitatives des soldes d'opinion, et suggère une utilité possible de l'IS pour le *nowcasting* du PIB. Ensuite, l'article présente une méthode d'identification par mots-clés permettant de retracer les occurrences dans les commentaires de certaines thématiques au cours du temps. Sa pertinence est démontrée dans le cadre de l'étude des différents mouvements sociaux qui ont émaillé les dix dernières années (voir graphique), et dont la méthode permet de comparer l'ampleur relative dans les préoccupations des entreprises interrogées.

Un tel comptage peut dans un deuxième temps être dérivé en un calcul économétrique permettant d'en estimer l'impact sur le PIB. À titre d'illustration, 20% des entreprises auraient ainsi vu leur activité perturbée par le mouvement des gilets jaunes en novembre et décembre 2018, conduisant à une perte estimée à -0,15 points de PIB au T4 2018. Cette méthode d'évaluation peut être aisément mobilisée pour toute autre application se ramenant à l'étude de deux groupes d'entreprises, affectées et non affectées par un événement.

Un cas plus complexe ne correspondant pas à cette partition dichotomique est développé dans la dernière partie de l'article et mis en application pour apprécier les effets du Brexit. À partir de la restriction du corpus aux documents évoquant le Brexit, nous identifions 6 catégories de réactions non exclusives des entreprises. Après labellisation et projection des documents dans un espace de plongement lexical de type *word2vec*, un modèle logit-lasso est entraîné. Les prévisions de ce dernier permettent de retracer de manière quantifiée la chronologie et la nature des préoccupations des entreprises françaises vis-à-vis du Brexit depuis 2016. Cette analyse met notamment en évidence un premier pic des préoccupations immédiatement après le référendum à l'été 2016, caractérisé par des effets monétaires liés à la dépréciation de la livre. Par la suite, deux autres pics dans les préoccupations, au printemps et à l'automne 2019, correspondent aux dates butoirs successivement repoussées pour la sortie du Royaume-Uni de l'Union Européenne : par crainte d'une fermeture des frontières, ces périodes sont caractérisées par un niveau élevé d'incertitude et par des ajustements d'offre et de demande des deux côtés de la Manche. Avec un travail d'adaptation de la typologie et de la labellisation, cette méthodologie a vocation à être reproduite pour l'analyse d'un événement futur où la nature de l'impact est multiple et où il y a un besoin de suivre l'évolution et la magnitude des effets au cours du temps.

Graphique : Mention des mouvements sociaux dans les commentaires de l'Enquête Mensuelle de Conjoncture



Source : Enquête Mensuelle de Conjoncture, Banque de France.

Enrichment of the Banque de France's Monthly Business Survey: Lessons from Textual Analysis of Business Leaders' Comments

ABSTRACT

In the context of the Banque de France's monthly business survey, this document presents the main findings of the textual analysis of business leaders' comments. First, the richness of these data is illustrated via an elementary sentiment index and the identification of the main social movements since 2009 by means of keywords. Then, the article presents two statistical applications whose reproducibility is discussed. The first one, applied to the 2018 yellow vests and the 2019 strikes, aims to estimate the impact on GDP of an event whose effect is unequivocal. The second, backed by the study of Brexit, aims to characterize, using a supervised learning model and word vectors, the effects of a complex event with multiple impacts.

Keywords: Textual Analysis, Business Survey, Sentiment Index, Keywords, Word Vectors, Brexit, Social Movements

Working Papers reflect the opinions of the authors and do not necessarily express the views of the Banque de France. This document is available on publications.banque-france.fr/en

Introduction

Alors que la première estimation de la croissance du PIB de la France du trimestre T est rendue publique le 30^{ème} jour du trimestre $T + 1$, les enquêtes mensuelles de conjoncture (EMC) de la Banque de France pour le mois M sont publiées le 7^{ème} jour ouvrable du mois $M + 1$. Ces données d'enquête sont ainsi le premier ensemble statistique disponible pour le suivi de la conjoncture et la prévision du PIB du trimestre en cours (*nowcasting*).

La spécificité de l'EMC – celle d'une enquête par entretien téléphonique retranscrite à la fois sous la forme d'échelles d'opinion et d'un résumé synthétique textuel, offre une source de données unique pour concevoir des outils complémentaires dédiés à l'analyse conjoncturelle.

L'originalité de ces données tient au fait qu'il ne s'agit pas d'un champ textuel ouvert comme il peut en exister dans d'autres enquêtes, pour lesquelles le contenu peut s'éloigner sensiblement de l'objet de l'enquête (réticences du déclarant à participer, griefs contre le portail de déclaration...) et dont le traitement par des méthodes d'analyse textuelle se trouverait fragilisé. Au contraire, il s'agit bien de documents reprenant sous forme textuelle le contenu des échelles d'opinion communiquées lors de l'enquête et précisées par des éléments de contexte tels que des données chiffrées (volumes produits, croissance du CA), des informations factuelles sur l'actualité de l'entreprise (survenue d'incidents de production, prise de commandes exceptionnelles) et sur ses marchés (locaux, nationaux et internationaux).

Pendant longtemps ces données n'ont été utilisées que de manière discrétionnaire, le plus souvent pour apprécier les notations individuelles (justification des variations les plus importantes, vérification des déclarations influençant fortement les résultats de l'enquête), ou semi-automatique (comptage des occurrences de mots clés sur une période sous revue, sans mise en perspective temporelle), sans que celles-ci ne conduisent à une approche systématique. Ce document de travail vise à combler cette lacune et fait le point sur les principaux résultats obtenus depuis le développement de cet axe d'étude en septembre 2018. À notre connaissance il n'existe pas à ce jour dans la littérature de travaux en analyse textuelle publiés sur ce type de données, à l'exception de ceux présentés dans le récent article du blog de la Banque d'Angleterre d'avril 2021 (Yotzov *et al.* (2021)) à partir des réponses au *Decision Maker Panel* et qui s'appuie sur des techniques très proches de celles utilisées dans cet article.

Pour ces travaux, un corpus de plus de 500 000 documents, correspondant aux retranscriptions (désignées sous le nom de « faits saillants » ou « commentaires ») des déclarations des entreprises aux enquêtes industrie, services marchands et bâtiments, recueillies entre 2009 et 2021, a été constitué. Ces données sont présentées en détail dans la première partie de l'article.

La valeur du contenu informationnel des faits saillants est examinée dans la deuxième partie. Dans un premier temps, la correspondance entre ces faits saillants et les données quantitatives macroéconomiques pouvant être extraites de l'enquête à partir des soldes d'opinion est établie par la construction d'un indice textuel de sentiments comparé à l'indice du climat des affaires, avec lequel il présente une forte corrélation. Ensuite, la présence dans les faits saillants d'une information conjoncturelle relative à des événements spécifiques est illustrée sur le cas d'événements aisément repérables dans le temps tels que la survenue de mouvements sociaux (grèves, manifestations...). Une simple identi-

fication par mots-clés permet ainsi de confirmer la présence dans les commentaires de citations des différents mouvements qui ont émaillé les dix dernières années, et d'en évaluer l'ampleur dans les préoccupations des entreprises interrogées.

Les deux dernières parties de l'article examinent, par l'intermédiaire de deux applications, l'exploitation statistique qui peut être tirée de ce contenu informationnel spécifique. Dans la première application, le comptage précédent des mentions de mouvement sociaux est ainsi dérivé en un calcul économétrique simple permettant d'estimer l'impact sur le PIB des deux plus récents d'entre eux : le mouvement des gilets jaunes à fin 2018, et les grèves en réaction au projet de réforme des retraites à fin 2019. Cette méthode d'évaluation se base sur l'identification d'un groupe de contrôle et d'un groupe de traitement ; sa pertinence repose sur l'hypothèse que l'évènement est univoque dans sa nature (impacte principalement l'activité) et sa direction (l'impact est négatif).

La seconde application vise à caractériser les effets d'un évènement complexe, avec des incidences à court terme qui peuvent être en partie positives et en partie négatives, et concerner différents grandeurs économiques. Cette démarche a en particulier été mise en oeuvre pour apprécier l'impact du Brexit sur l'activité des entreprises françaises depuis le référendum britannique de 2016. La méthode s'appuie sur la préparation d'un corpus de documents ciblés, dont une partie est étudiée et labellisée manuellement. Cet extrait du corpus est ensuite utilisé pour entraîner un modèle de type logit-lasso, après projection des documents dans un espace de plongement lexical de type *word2vec*. Le modèle ainsi appris est ensuite utilisé chaque mois pour prédire les occurrences des différents effets. Dans le cas particulier présenté ici, il permet ainsi de retracer de manière quantifiée la chronologie des préoccupations des entreprises françaises vis-à-vis du Brexit depuis le vote en faveur de la sortie du Royaume-Uni de l'Union Européenne en juin 2016.

1 Les données

1.1 Présentation de l'Enquête Mensuelle de Conjoncture (EMC)

Depuis 1976, la Banque de France interroge mensuellement des chefs d'entreprise sur leur activité récente et leurs prévisions pour le mois suivant. Aujourd'hui, pas moins de 8 500 entreprises ou établissements dans les secteurs de l'industrie, des services marchands et de la construction sont interrogés chaque mois, grâce au support des unités du réseau de la Banque de France. À l'heure du *Big Data* et de l'intelligence artificielle, ces données d'enquête demeurent à la pointe des indicateurs avancés de la croissance du PIB. Blanchet et Coueffe (2020) montrent notamment que les modèles de prévision sont d'autant plus performants avec des données récentes, notamment en début de trimestre, d'où l'importance de la fraîcheur de ces données : pour l'enquête mensuelle de conjoncture de la Banque de France, la publication statistique pour le mois M , ainsi que le détail sectoriel, est disponible au 8^{ème} jour ouvré du mois $M+1$.

La collecte des données prend la forme d'un entretien semi-dirigé par téléphone, ce qui permet le développement d'une relation de long terme entre le testeur (= l'agent de la Banque de France en charge de l'enquête) et l'informateur (= l'interlocuteur dans l'entreprise répondant mensuellement à l'enquête). En pratique, une fois l'entretien terminé, le testeur retranscrit la discussion par des notes qualitatives sur différentes variables exprimées dans une échelle de notation commune.

Certaines de ces variables sont exprimées en évolution (production, livraisons, commandes, effectifs...) tandis que d'autres sont exprimées en niveau (stocks, taux d'utilisation des capacités...). L'évolution ou le niveau peuvent être demandés en comparaison du mois précédent, parfois du mois $M-12$, ou en anticipation à l'horizon du mois $M+1$. Chaque variable est notée sur une échelle à 7 niveaux, allant de *forte diminution* (respectivement *niveau très bas*) à *forte augmentation* (respectivement *niveau très élevé*). Pour agréger ces réponses individuelles, ces dernières sont synthétisées en soldes d'opinion (allant de -200 à +200) représentant la proportion pondérée d'entreprises déclarant une hausse moins la proportion d'entreprises déclarant une baisse. Enfin, les séries temporelles publiées sont corrigées des variations saisonnières et des jours ouvrés par une méthode additive de type X12-Reg-ARIMA.

En plus de ces variables, il existe dans le questionnaire un champ textuel dit des « faits saillants », que nous présentons dans la sous-section suivante et dont l'exploitation est au cœur de cet article.

1.2 La rubrique faits saillants

La rubrique des faits saillants consiste en un champ libre de texte pouvant le cas échéant contenir des valeurs chiffrées, mais surtout de nombreuses informations qualitatives sur le contexte de la déclaration à même d'éclairer les notations des variables décrites précédemment. L'information peut ainsi indiquer un événement particulier affectant l'entreprise, un changement dans sa politique, ou tout autre fait relatif à l'environnement économique ou l'évolution de son secteur d'activité. Comme pour les variables notées, ce champ est rempli par le testeur après l'entretien téléphonique : il contient

donc ce que l’informateur a spontanément déclaré, sans cadre ou question spécifique posé¹. Un exemple de formulaire de collecte est présenté sur le graphique 7 en annexe.

Les données individuelles pour ces déclarations (appelées commentaires) existent depuis 2008, date de la mise en place du système de collecte SIRCE (Système d’Information Renforcé sur les Centralisations et les Enquêtes). En 2020, la section faits saillants avait été complétée pour 88 % des tests de conjoncture. En effet, il faut préciser que remplir ce champ n’est pas obligatoire : si aucune information particulière (autre que celle déjà retranscrite dans les variables notées) n’est donnée par l’informateur, le testeur n’a pas d’obligation à inscrire de commentaire.

Ces données textuelles représentent une source de données très riche. Jusqu’en 2018, elles étaient utilisées occasionnellement et lues manuellement pour certaines entreprises dans le cas de tests de conjonctures incohérents, pour expliquer une dynamique inattendue dans certains secteurs, ou pour apprécier l’impact éventuel d’évènements spécifiques affectant l’environnement économique des entreprises françaises. Pour illustrer le potentiel et la richesse de ces données, nous proposons dans cet article plusieurs exemples d’exploitations statistiques systématiques de ces commentaires à l’aide de techniques d’analyse textuelle.

1.3 Pré-traitement des données textuelles

Le format particulier et non structuré des données textuelles implique d’y appliquer un certain nombre de traitements pour les rendre exploitables. Nous appliquons un premier filtre pour définir notre échantillon de travail, à savoir uniquement les commentaires longs d’au moins 4 mots² et rattachés à un test de conjoncture non estimé (c’est-à-dire que l’agent de la Banque de France a bien été en contact avec l’informateur dans l’entreprise³). Nous appliquons ensuite des opérations sur le texte brut de manière à supprimer la casse, les chiffres et les accents.

Dans toutes les applications présentées dans la suite de ce document, nous adoptons une approche classique en analyse textuelle dite de « *sac de mots* » (ou *bag of words*), c’est-à-dire qu’un texte est considéré comme un vecteur de mots indépendants les uns des autres, sans considération de leur ordre mais avec conservation des répétitions éventuelles. Pour cela, nous avons utilisé le package *R tidytext* développé par Silge et Robinson (2016) et la fonction dite de « *tokenization* » qui consiste à découper le texte selon un séparateur donné, ici le caractère *espace* pour diviser le commentaire en la suite des mots qui le composent.

1. Exceptionnellement, la crise économique et sanitaire de 2020 a bousculé un peu ce cadre puisque le champ des faits saillants a été utilisé pour collecter des réponses à des questions supplémentaires (notamment, le niveau d’activité en % par rapport à la normale) en s’affranchissant des contraintes informatiques. Toutefois ces questions additionnelles n’ont pas empêché de continuer à collecter les remarques spontanées des chefs d’entreprise puisqu’il n’y a pas de contrainte forte sur le nombre de caractères autorisés (2 000 caractères).

2. Cette règle vise à éliminer les commentaires non informatifs du type « *fermé pour travaux* », « *congés habituels* », « *entreprise de bâtiment* »...

3. Lorsqu’un informateur n’a pas pu être interrogé un mois donné et qu’il s’agit d’une entreprise ou d’un établissement employant au moins 100 salariés ou occupant une place importante dans sa branche, il est possible d’extrapoler avec précaution les notations de l’enquête précédente à partir des dernières informations reçues.

Selon l'usage qu'il est fait des données, d'autres opérations classiques sont susceptibles d'être pertinentes pour en réduire le bruit (suppression des mots non informatifs) et/ou la dimension (lemmatisation⁴ du lexique). Le cas échéant, ces traitements seront précisés et explicités dans le corps du texte.

Avant d'aller plus loin, il nous semble important dans un premier temps d'examiner la qualité de nos données, notamment en étudiant leur contenu informationnel. Pour cela, nous développons dans la section suivante deux types d'arguments. Le premier (2.1) est à l'échelle macroéconomique et établit le lien entre un indice de sentiment construit à partir du texte des commentaires et l'indice synthétique de l'enquête dit *Indice du Climat des Affaires* (ICA). Le second (2.2) met en évidence la présence d'informations liées à certains événements, notamment dans le cas des mouvements sociaux.

4. La lemmatisation est un traitement lexical qui consiste à renvoyer chaque mot à sa forme canonique, le lemme. Par exemple, un verbe conjugué est renvoyé à son infinitif, et un adjectif à son masculin singulier.

2 Validation du contenu informationnel

2.1 Information macroéconomique : construction d'un indice de sentiment

Du texte brut à l'indice de sentiment macro-économique

L'idée de convertir des données textuelles en indicateurs synthétiques quantifiés n'est pas nouvelle. Ce type d'approche s'est considérablement développé dans la littérature depuis le début des années 2000 dans le marketing et l'analyse d'opinion (Turney (2002)) avec l'essor du web et des réseaux sociaux, occasionnant d'importants volumes de données textuelles en ligne : revues de produits (Morinaga *et al.* (2002)), débats d'idée sur les forums (Tong (2001))...

Peu après, cette approche a également été appliquée avec succès sur les questions économiques et financières. La première utilisation semble pouvoir être imputée à Antweiler et Frank (2004) puis Tetlock (2007), qui met en évidence l'apport d'articles du *Wall Street Journal* pour la prévision des fluctuations de cours de bourse.

Plus récemment, Gentzkow *et al.* (2019) proposent une revue de l'utilisation de texte comme intrant de la recherche économique. Observant que le trait caractéristique majeur de ce type de données est le grand nombre de dimensions permettant de les décrire (« *high-dimensional data* »), les auteurs énumèrent les méthodes statistiques adaptées à leur manipulation, et en décrivent l'apport en termes de recherche économique au travers de huit principaux cas d'usage⁵. Deux d'entre eux, le *nowcasting* et la mesure d'incertitude politique, présentent un intérêt particulier pour l'analyse conjoncturelle. Un trait commun de ces deux applications est le recours fréquent à la construction d'indices de sentiment.

Centrés sur ces derniers, Algaba *et al.* (2020) proposent une synthèse des travaux d'intérêt pour l'économétrie de l'analyse de données de sentiments qualitatives (principalement textuelles). Ils relèvent ainsi que ces données peuvent être utilisées comme instrument pour l'évaluation de chocs économiques de diverses natures, avec notamment Larsen et Thorsrud (2019), ou bien comme variables additionnelles aux variables explicatives usuelles de modèles de *nowcasting* et de *forecasting* dont elles permettent d'améliorer les performances prédictives, à l'instar de Ardia *et al.* (2019).

À cette fin, le sentiment latent doit être extrait du texte et mesuré quantitativement, soit au moyen d'un dictionnaire (« *lexicon-based* »), soit par *machine learning*. Algaba *et al.* (2020) affirment que l'utilisation du *machine learning* conduit à de meilleures performances *out-of-sample*, au prix de calculs plus intensifs et d'une moindre transparence du modèle. Après avoir étudié de manière approfondie 24 variantes de construction d'indices de sentiment sur 18 jeux de données labellisés, Ribeiro *et al.* (2016) soulignent les larges différences de performance entre méthodes et entre jeux de données. Ils concluent que la méthode de quantification du sentiment utilisée doit être sélectionnée avec soin en fonction de l'objectif retenu.

5. Paternité de documents, cours de la bourse, communication des banques centrales, *nowcasting*, incertitude politique, biais médiatique (« *media slant* »), définition de marché et impact de l'innovation, identification des thèmes (« *topics* ») en recherche, politique et droit.

L'utilisation d'un dictionnaire pour la construction d'un indice de sentiment est cependant considérée comme la méthode la plus directe, la plus efficace et la plus parcimonieuse. Turney (2002) la définit comme le calcul d'un sentiment pour un document, à partir du sentiment des mots contenus dans le document. Elle requiert l'utilisation d'un dictionnaire de sentiments, c'est-à-dire d'une liste de mots (ou d'expressions) – quelques centaines à quelques milliers selon les auteurs – chacun associé à un score de sentiment.

Ces dictionnaires peuvent soit avoir un caractère généraliste, soit être élaborés spécifiquement pour l'analyse d'un domaine particulier. Henry (2008) et Loughran et McDonald (2011), après avoir démontré le caractère sous-optimal de dictionnaires généralistes pour l'analyse du domaine financier, élaborent chacun leur dictionnaire propre en langue anglaise. Bortoli *et al.* (2017) reprennent ce principe sur un corpus en langue française composé d'articles du journal *Le Monde*, afin d'en extraire des indices de sentiments relatifs à l'emploi. C'est ce dictionnaire, mis à notre disposition par les auteurs, que nous utilisons pour la construction de notre indice de sentiment. Il est décrit plus amplement dans le paragraphe qui suit.

Un premier essai concluant d'indice de sentiment macro-économique

Comme évoqué précédemment, les « faits saillants » ont longtemps été utilisés à des fins d'approfondissement du contrôle de qualité des données individuelles. À titre d'exemple, pour confirmer la validité d'un recul brutal de l'activité dans une entreprise faisant jusque-là état d'une forte dynamique de croissance, la mention, dans les commentaires, d'une panne sur les chaînes de production ayant conduit à une fermeture exceptionnelle prolongée du site est particulièrement éclairante.

Au-delà de cet apport indiscutable sur le plan micro-économique, il est utile de s'interroger sur la possibilité de construire - sur la seule base des informations textuelles ainsi collectées - des indicateurs macroéconomiques spécifiques, et de pouvoir apprécier dans quelle mesure ces nouveaux indicateurs corroborent ceux obtenus à partir des soldes d'opinion collectés par l'enquête. Pour atteindre cet objectif, nous produisons un indice de sentiments de facture volontairement simple. Des développements méthodologiques plus conséquents seraient nécessaires pour employer un tel indice à des fins plus ambitieuses : cette discussion est ébauchée dans le paragraphe suivant.

La première étape du calcul consiste à transformer le corpus en « sac de mots ». Les mots sont ensuite lemmatisés⁶ afin de normaliser et réduire la taille du lexique, défini comme le nombre de mot différents présents au moins une fois dans le corpus de commentaires. Le dictionnaire rassemble, quant à lui, un total de 2 252 mots (926 lemmes), dont 529 mots (214 lemmes) positifs et 1 723 mots (712 lemmes) négatifs⁷. Tout lemme du lexique absent du dictionnaire sera donc par défaut considéré comme neutre.

Des tests de robustesse de ce dictionnaire de sentiments ont été menés sur les données de l'enquête Industrie. Le principe consiste à recouper le signe du sentiment associé à chaque lemme avec le signe d'un score basé sur les résultats de l'enquête : pour chaque mot du dictionnaire, une note corres-

6. La lemmatisation est réalisée à partir du dictionnaire Morphalou 3 développé par le laboratoire ATILF (2019).

7. La sur-représentation des termes négatifs par rapport aux termes positifs est une caractéristique usuelle de ce type de dictionnaire.

pondant à la moyenne des principaux soldes d’opinion de l’enquête déclarés dans les commentaires qui mentionnent ce mot est calculée. La comparaison des distributions de notes ainsi obtenues pour chacun des trois groupes de mots du dictionnaire (négatif / neutre / positif) est conforme à l’attendu, sauf pour un petit nombre de mots qui apparaissent, en moyenne, mal classés⁸. Un nouvel indice de sentiments est recalculé après reclassement des mots identifiés à l’étape précédente dans le dictionnaire⁹. Ce nouvel indice varie peu par rapport à l’indice initial.

L’indicateur de sentiment (IS) micro-économique est dans un deuxième temps obtenu pour chaque document de chaque période sous revue mensuelle (PSR) du corpus, par sommation algébrique des scores du dictionnaire, ramenés au nombre total de lemme du document, soit :

$$IS_{d,t} = \frac{\sum_l S_{l,d,t}}{n_{d,t}} \quad (1)$$

où $S_{l,d,t}$ désigne le score signé du lemme l du document d pour la PSR t , et $n_{d,t}$ désigne le nombre total d’occurrences de tous les lemmes (y compris les répétitions) du document d pour la PSR t . L’IS micro est donc compris, pour chaque commentaire, entre -1 et 1. L’indice de sentiment macro-économique est obtenu lors de la troisième étape par somme pondérée des indices de sentiment micro :

$$IS_t = \sum_d p_{d,t} IS_{d,t} \quad (2)$$

où $p_{d,t}$ désigne la pondération individuelle du document d pour la PSR t .

Deux variantes de l’indicateur sont calculées. La première est obtenue en fixant $p_{d,t} = \frac{1}{n_t}$, pour l’ensemble des documents et des PSR (avec n_t nombre de documents de la PSR), ce qui revient à ne pas pondérer les IS micro. La seconde est obtenue en prenant pour $p_{d,t}$ la valeur de la pondération individuelle de l’enquête mensuelle de conjoncture, qui tient compte des effectifs et de la valeur ajoutée du secteur de l’entreprise.

Le calcul de l’indice de sentiment est réalisé pour chacune des trois enquêtes mensuelles de conjoncture (industrie, services marchands, bâtiment), à partir de l’intégralité des commentaires valides, pour la période s’étalant de janvier 2010 à février 2021. Sur cette période de plus de 11 ans, le corpus ainsi constitué s’élève à près de 300 000 documents et un lexique de près de 17 000 lemmes pour l’enquête industrie, 200 000 documents et environ 14 000 lemmes pour les services marchands, 100 000 documents et plus de 10 000 lemmes pour l’enquête bâtiment.

La taille des lexiques de notre corpus est très largement supérieure à celle du dictionnaire mis à notre disposition, suggérant qu’un nombre important de mot polarisés (positivement ou négativement) pourraient à tort être labellisés comme neutres, ce qui créerait alors un aplatissement potentiel de l’indice (sous-estimation de l’impact des mots positifs en phase ascendante du cycle économique,

8. Par exemple, le mot « stable » est considéré comme « positif » dans ce dictionnaire, mais sa note moyenne au sens des soldes d’opinion est voisine de 0. De fait, parler de « stabilité de l’emploi » renvoie à une connotation positive, alors que dans l’EMC, la « stabilité » renvoie plutôt à une conjoncture en demi-teinte, ce qui conduit à reclasser le mot « stable » de « positif » à « neutre ».

9. Dans l’exemple précédent, le score du mot « stable » passe donc après reclassement de « +1 » à « 0 ». Lors du test de robustesse, 60 mots ont ainsi été détectés comme étant à reclasser (soit environ 1 % des lemmes présents dans plus de 100 documents).

sous-estimation de celui des mots négatifs en phase descendante) . Pour limiter cet effet, nous faisons le choix de limiter le calcul de l’indice de sentiments aux lemmes du lexique qui apparaissent dans au moins une centaine de documents différents, pour chacune des trois enquêtes. Ce choix permet de réduire la taille du lexique, d’un facteur de 4 à 5 selon les enquêtes, en éliminant essentiellement du « bruit » (noms propres, acronymes, abréviations, fautes de frappe. . .). Le tableau 1 résume la composition du corpus et du dictionnaire final employé pour chacune des trois enquêtes.

Tableau 1. Description du corpus et des dictionnaires utilisés

	Corpus	Lexique	Lemmes présents dans plus de 100 documents			
			Négatifs	Neutres	Positifs	Total
Industrie	297 317	16 883	246	3 910	128	4 284
Services Marchands	195 103	14 005	193	2 829	107	3 129
Bâtiment	95 289	10 193	143	1 714	93	1 950

Source : Enquête Mensuelle de Conjoncture, Banque de France, d’après Bortoli *et al.* (2017).

L’indice de sentiment ainsi obtenu est ensuite corrigé des variations saisonnières et des jours ouvrables (CVS-CJO) selon une méthode analogue à celle utilisée pour les soldes d’opinion agrégés des trois enquêtes (méthode de type X12-Reg-ARIMA)¹⁰. L’IS CVS-CJO ainsi obtenu est ensuite comparé à l’Indice du Climat des Affaires (ICA) de cette même enquête. L’ICA est une variable synthétique obtenue par l’extraction de la première composante principale d’une Analyse en Composantes Principales (ACP) menée sur les principaux soldes d’opinion de chaque enquête. L’ICA étant publié après application d’un lissage sur les 5 derniers mois, le même lissage est appliqué sur l’indice de sentiment.

Le graphique 1 résume la comparaison des deux indices. La colonne de gauche présente les comparaisons réalisées sur les versions non lissées de l’indice, elle permet de mesurer à quel point l’IS retrace la dynamique mensuelle de l’information conjoncturelle. La colonne de droite compare les versions lissées des deux indices, et permet d’apprécier comment le « cycle » filtré des soubresauts mensuels est retracé par chacun des deux indicateurs.

En dépit de la simplicité du mode de calcul, et l’utilisation d’un dictionnaire prêt à l’emploi - et sans doute perfectible par un travail ad hoc - l’IS retrace plus que convenablement le cycle matérialisé par l’ICA. En particulier sur la période récente, l’IS capte le choc des deux confinements de mars/avril et novembre 2020. Sur la période complète, les corrélations entre ICA et IS (tableau 2) sont comprises entre 0,7 (industrie) et 0,9 (services marchands) pour l’indice non pondéré ; et entre 0,6 (industrie) et 0,9 (bâtiment) pour l’indice pondéré. La prise en compte des pondérations dégrade significativement la corrélation avec l’ICA, plus particulièrement dans les services et l’industrie. Cet écart semble dû à la moins bonne prise en compte des périodes de confinement en 2020, qui représentent une grande partie de la covariance des indices, par les versions pondérées des IS.

Il est à ce titre dommage de ne pas disposer d’un historique plus long des commentaires, qui aurait notamment pu permettre d’évaluer comment l’IS se serait comporté durant les crises de 1993 et 2008.

10. L’observation mois par mois du lexique de mots utilisés montre bien une saisonnalité des phénomènes rapportés en commentaire : fermetures de sites en août et réouvertures en septembre, mention des jours fériés en mai, juin et décembre. . . Cette saisonnalité, classique dans les enquêtes de conjoncture, fait écho à celle observée dans les variables classiques de l’EMC, qui sont d’ailleurs toujours analysées en CVS-CJO.

Tableau 2. Corrélation des indices de sentiment et de climat des affaires

	Non pondéré		Pondéré	
	Non lissé	Lissé sur 5 mois	Non lissé	Lissé sur 5 mois
Industrie	0,75	0,82	0,61	0,73
Services Marchands	0,89	0,92	0,63	0,66
Bâtiment	0,86	0,89	0,79	0,87

Source : Enquête Mensuelle de Conjoncture, Banque de France.

Cette petite analyse semble toutefois suffisante pour valider la qualité des informations présentes dans l’EMC au niveau macro-économique.

Quelques pistes d’approfondissement possibles

S’agissant de l’amélioration de l’indice, le premier approfondissement qui vient à l’esprit est celui de l’amélioration du dictionnaire utilisé pour le calculer. Il paraît notamment possible d’augmenter la proportion de mots du lexique polarisés dans le dictionnaire, ce qui devrait conduire à une plus grande finesse de l’indicateur final. De même, plutôt que d’utiliser des valeurs discrètes (-1 / 0 / +1), le score de sentiment pourrait prendre une valeur continue dans l’intervalle [-1 ; 1], obtenue par apprentissage à partir des données individuelles des soldes d’opinion de l’enquête. Une telle approche, dite « *fine grained* » du dictionnaire, est proposée par Barbaglia *et al.* (2021).

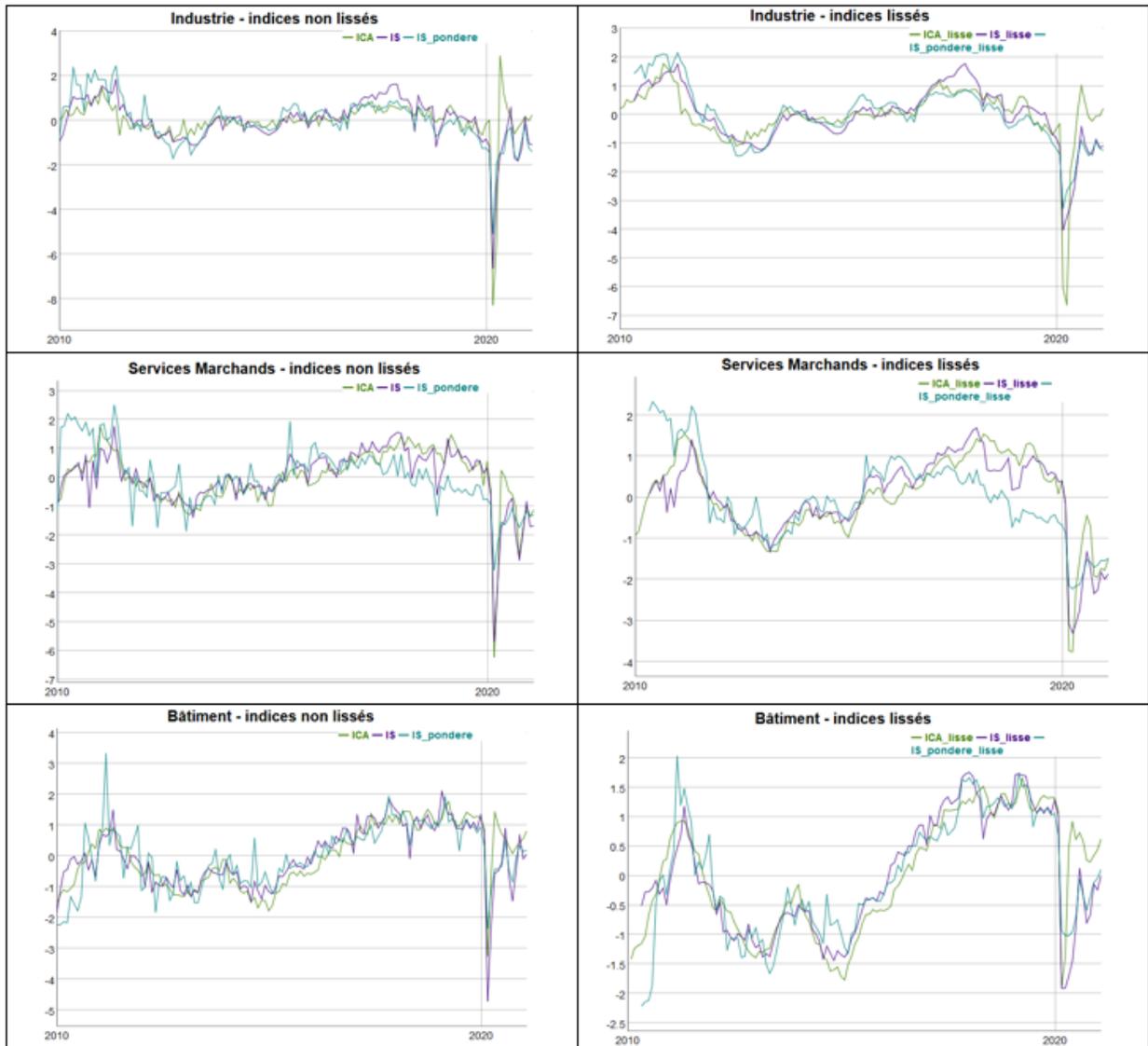
Un autre axe de développement, plus délicat sur le plan technique, pourrait être la prise en compte de certains aspects grammaticaux comme les négations (« pas mauvais » devrait avoir un effet positif sur l’IS), et les modification d’amplitude (« plutôt mauvais » devrait avoir un négatif moins fort sur l’IS que « très mauvais »). Aprigliano *et al.* (2021) fournissent un exemple de ce type de développements sur un dictionnaire spécifiquement conçu pour l’analyse de la presse italienne à fins de prévision économique.

De même, il pourrait être utile d’isoler les différents aspects de la conjoncture sur lesquels portent tel ou tel passage du commentaire afin de mieux cibler l’objet du « sentiment » mesuré par l’indice (« *aspect-based sentiment analysis* »). Pour une analyse de la presse US, Barbaglia *et al.* (2021) définissent ainsi 6 thèmes d’intérêt, pour chacun desquels ils isolent les passages d’articles afférents, en décomposent la structure grammaticale au moyen de techniques de linguistique computationnelle (*part of speech tagging, dependency parsing, tense detection...*) afin de raffiner au maximum l’intrant textuel utilisé pour le calcul des indices de sentiment.

Ces pistes d’approfondissements, dont la liste n’est évidemment pas exhaustive, ont pour point commun un alourdissement sensible de la complexité de confection de l’indice, même si des outils informatiques récents, disponibles sous forme de package *R* (*Sentometrics* de Ardia *et al.* (2020)), ou Python (notamment *Spacy* de Honnibal et Montani (2017), *TextBlob* de Loria *et al.* (2019) et *Vader* de Hutto et Gilbert (2014)) contribuent à les rendre plus accessibles. Leur mise en oeuvre ne se justifie donc que pour des objectifs plus ambitieux que la simple illustration proposée ici.

En particulier, on peut se demander dans quelle mesure un indice comme l'IS pourrait fournir une information utile pour améliorer les modèles existants de *nowcasting* du PIB trimestriel, sur le modèle du travail récemment réalisé par Kalamara *et al.* (2020) sur données britanniques.

Graphique 1. Synthèse des comparaisons des IS et ICA CVS-CJO (centrés réduits)



Source : Enquête Mensuelle de Conjoncture, Banque de France.

2.2 Information liée à un évènement : citation des mouvements sociaux

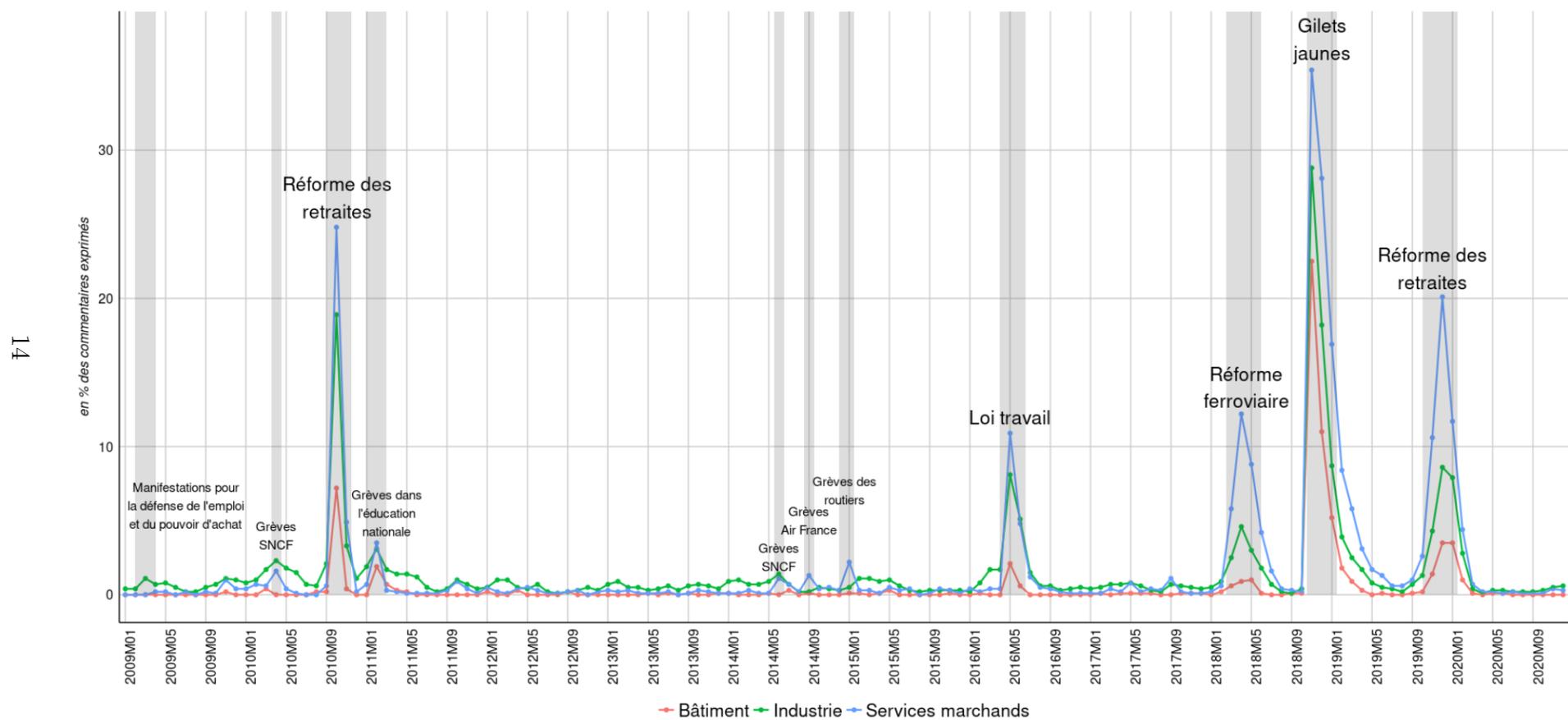
Au delà de retracer fidèlement les indicateurs macroéconomiques de l'enquête, les commentaires offrent également une source d'information supplémentaire dans le cas d'évènements affectant la conjoncture. Nous présentons dans cette partie une application à l'étude de l'impact des mouvements sociaux. Il est en effet possible d'identifier, à partir de recherches de mots-clés, quelles sont les entreprises concernées ; la série temporelle construite en rapportant, pour chaque période sous revue, la proportion d'entreprises concernées à l'ensemble des réponses, identifie très clairement les grands mouvements sociaux des dix dernières années.

Après lecture de nombreux commentaires, une liste de termes évoquant les mouvements sociaux est constituée. La difficulté est double : d'une part élaborer la liste de mots-clés la plus exhaustive possible, et d'autre part limiter les possibles cas d'homonymies. Par exemple, dans notre corpus le mot *manifestation* est très souvent utilisé dans le sens de *manifestation culturelle* et a donc été exclu. Le champ lexical des mouvements sociaux est finalement relativement restreint et les mots-clés recherchés dans les commentaires sont les suivants : « *grève* », « *mouvement social* », « *débrayage* », « *conflit social* », « *gilet jaune* », ainsi que les pluriels et formes dérivées¹¹ associés.

Une fois identifiés l'ensemble des commentaires contenant au moins un de ces mots-clés, un comptage en nombre puis en proportion du total des commentaires exprimés est réalisé par mois x enquête. Les trois séries historiques obtenues sont représentées sur le graphique 2, où chaque rectangle gris foncé correspond à un évènement majeur de l'actualité décrit plus en détails dans le tableau 7 en annexe.

11. On entend par forme dérivée un terme écrit un peu différemment mais strictement identique du point de vue du sens : par exemple « *gillet jaune* », « *giletjaune* », « *gilet-jaune* » ou « *gj* » sont des formes dérivées de « *gilet jaune* ».

Graphique 2. Mention des mouvements sociaux dans les commentaires de l'Enquête Mensuelle de Conjoncture



Source : Enquête Mensuelle de Conjoncture, Banque de France.

Cette approche permet des premières comparaisons d'ampleur dans les perceptions des chefs d'entreprise entre les différents mouvements sociaux et d'un secteur à l'autre. Ainsi, le mouvement dits des *gilets jaunes* en novembre et décembre 2018 s'avère avoir particulièrement préoccupé les entreprises, notamment dans les services marchands. Par ailleurs, le premier épisode de grèves contre la réforme des retraites en octobre 2010 semble avoir été de plus large ampleur que celui de décembre 2019. De manière générale, le graphique 2 montre aussi que les entreprises des services marchands (courbe bleue) paraissent toujours plus concernées par les mouvements sociaux que celles de l'industrie (courbe verte), et *a fortiori* que celles du bâtiment (courbe rose).

Deux types de difficultés sont susceptibles de fragiliser l'interprétation de ce graphique. La première est liée au fait que, les commentaires étant libres, il se peut que certaines entreprises soient affectées sans que ce soit évoqué lors de l'entretien et/ou retranscrit dans les faits saillants. La seconde difficulté vient du fait qu'on ne parle pas encore à ce stade d'entreprises affectées ou non affectées, mais seulement d'entreprises dont le commentaire du test de conjoncture fait mention d'un mouvement social ; or parmi ces déclarations, certaines spécifient que le mouvement social en question n'a pas eu d'effet pour l'entreprise.

Les deux hypothèses suivantes permettent néanmoins de valider l'interprétation ci-dessus qui apprécie l'impact relatif des mouvements sociaux les uns par rapport aux autres (dans le temps ou par secteur d'activité) :

1. la probabilité pour un chef d'entreprise de mentionner un mouvement social lors de l'entretien sachant son impact est stable au cours du temps,
2. la probabilité pour un testeur de faire référence à un mouvement social dans les faits saillants sachant qu'il a été évoqué par le chef d'entreprise est stable au cours du temps.

Le biais de sous-déclaration peut en revanche difficilement être résolu.

Pour aller plus loin et qualifier finement l'impact en niveau de chaque mouvement social pris isolément, l'enjeu est de distinguer, au sein des entreprises faisant mention de l'évènement, celles affectées de celles non affectées. C'est ce que nous développons dans la section 3 suivante.

3 Mesure d'impact d'un évènement à l'aide de recherches par mots-clés

Cette application des méthodes d'analyse textuelle s'inscrit dans le prolongement de la sous-section précédente : comment évaluer précisément l'impact d'un évènement à partir des informations le concernant identifiées dans les commentaires ?

La méthodologie que nous proposons n'est pas uniquement applicable à l'étude des mouvements sociaux : nous revenons plus loin sur les conditions qui nous semblent être nécessaires à la reproductibilité des étapes décrites ci-après.

L'idée sur laquelle repose l'estimation d'impact est simple : la clé est d'identifier, dans notre échantillon, les entreprises affectées par l'évènement E (= groupe de traitement) et les entreprises non affectées (= groupe de contrôle).

L'identification de ces deux groupes se fait en deux étapes. Tout d'abord, à l'aide d'une première recherche de mots-clés désignant l'évènement E , sont identifiées les entreprises dont les commentaires mentionnent l'évènement en question, quelle qu'en soit la raison. Dans l'exemple des mouvements sociaux, il s'agit des entreprises représentées sur le graphique 2. La difficulté de cette première recherche de mots-clés varie selon le champ lexical de l'évènement E : relativement simple dans le cas des mouvements sociaux, la tâche se complique par exemple dans le cas des évènements météorologiques du fait d'un grand nombre d'homonymes (le mot « climat » fait fréquemment référence au « climat *des affaires* » par exemple).

Au sein de ce premier groupe d'entreprises, il s'agit ensuite de repérer celles affectées par l'évènement E et celles non affectées. Pour cela, la partie du commentaire mentionnant l'évènement est extraite en utilisant la ponctuation et la mise en forme (restriction à la phrase ou paragraphe qui mentionne E en utilisant les points et les retours à la ligne). Sur cette portion de texte, une deuxième recherche de mots-clés regroupant des termes synonymes de « sans effet » est menée. Contrairement à la première recherche de mots-clés qui est spécifique à chaque évènement, la liste des mots-clés qualifiant les « sans effet » est quasiment identique pour tous les évènements E . Elle comprend par exemple les termes et expressions suivantes : *sans effet, pas d'effet, non concerné, pas d'impact, ras, pas touché* ... La proportion de commentaires qui mentionnent un évènement tout en précisant qu'il n'a pas eu d'impact pour l'entreprise est relativement faible dans le cadre de commentaires spontanés¹². Cependant elle est susceptible d'être plus élevée lorsqu'un évènement est un sujet d'actualité tel qu'il vient naturellement à être abordé lors de la discussion entre le chef d'entreprise et le testeur¹³. Il semble donc essentiel d'identifier et de retirer ces cas là pour ne pas biaiser l'analyse.

Les groupes dit *de contrôle* (= entreprises non affectées par E) et de *traitement* (= entreprises affectées par E) ainsi constitués, une régression économétrique sur la variable de traitement T permet d'estimer l'impact de l'évènement E sur une variable Y . La grandeur économique Y sur laquelle on

12. Cette proportion est de 7 % dans le cas des gilets jaunes en novembre-décembre 2018 et de 2 % dans le cas des grèves contre la réforme des retraites en décembre 2019.

13. Le cas extrême étant celui où il n'y a pas de groupe de contrôle, comme en mars 2020 où l'intégralité des commentaires ont fait allusion au confinement et à la Covid.

veut estimer l'impact de l'évènement est susceptible de changer¹⁴ selon l'exercice et la nature de E . La régression suivante est estimée par les moindres carrés ordinaires sur données empilées :

$$Y_{i,m} = \alpha + \beta T_{i,m} + \theta Z_{i,m} + \varepsilon_{i,m} \quad (3)$$

où $T_{i,m}$ désigne l'indicatrice qui vaut 1 si l'entreprise i est affectée par l'évènement E le mois m et 0 sinon, et $Z_{i,m}$ un ensemble de variables de contrôle pour le couple (i, m) : interaction secteur d'activité de $i \times$ mois m , taille en effectifs de i le mois m , région de i le mois m .

L'estimation est réalisée sur les données empilées du mois M de l'évènement E et du mois $M - 1$ afin de tenir compte des différences de saisonnalité sectorielles¹⁵. Le coefficient estimé pour β désigne l'impact individuel moyen de l'évènement E sur la grandeur Y d'une entreprise affectée. En multipliant cet impact par la part d'entreprises affectées, on obtient un impact consolidé sur le solde d'opinion agrégé de la variable d'intérêt Y .

La dernière étape consiste, à l'aide de l'équation qui suit, à traduire cet impact en points de PIB par rapport à une situation contrefactuelle caractérisée par l'absence de l'évènement E :

$$\Delta PIB_T = \alpha + \delta \text{solde opinion } \frac{Y}{T} + \tau \Delta PIB_{T-1} + \varepsilon \quad (4)$$

où ΔPIB_T désigne la variation trimestrielle du PIB pour le trimestre T , $\text{solde opinion } \frac{Y}{T}$ le solde d'opinion¹⁶ pour la variable Y au trimestre T , et ΔPIB_{T-1} la variation trimestrielle du PIB pour le trimestre $T - 1$, afin de tenir compte de la nature autorégressive du processus de croissance du PIB. La relation est estimée sur la période allant du deuxième trimestre 2009¹⁷ au trimestre le plus récent pour lesquelles les données sont disponibles à la date de l'évènement (en général, T-2).

Le tableau 3 ci-dessous décompose les différentes étapes du calcul de l'impact pour deux grands évènements ayant affecté la conjoncture depuis 2018, les gilets jaunes et les grèves contre la réforme des retraites. Il ne doit pas être interprété comme une volonté de proposer une hiérarchie des évènements visés par l'estimation, pas plus qu'à émettre un jugement sur leur impact structurel, mais comme une simple décomposition économétrique de leur effet à court terme tel que retranscrit dans l'enquête. Les résultats des régressions estimées pour le coefficient β sont détaillés dans le tableau 8 en annexe, celles pour le coefficient δ le sont dans le tableau 9 en annexe.

La lecture du tableau 3 montre que la proportion d'entreprises signalant être affectées par le mouvement des gilets jaunes en novembre et décembre 2018 est relativement plus importante que celle mentionnant un impact des grèves de décembre 2019. Par ailleurs, on constate que l'impact indivi-

14. Dans la plupart des cas, on s'intéresse à l'activité récente (du mois dernier) de l'entreprise. Cependant, on peut aussi s'intéresser à d'autres variables comme l'activité prévue, les livraisons, les commandes...

15. Dans le cas d'un évènement s'étalant sur 2 mois, la variable de traitement T peut prendre les valeurs $\{0, 1\}$ sur les deux mois M et $M - 1$; dans le cas d'un évènement survenant le mois M , la variable de traitement vaut 0 le mois $M - 1$ et prend les valeurs $\{0, 1\}$ le mois M .

16. Obtenu en effectuant la moyenne des soldes d'opinion CVS-CJO pour les trois enquêtes industrie, services marchands et bâtiment, pondérée par la part de chaque secteur agrégé dans la valeur ajoutée brute correspondant au champ total des trois enquêtes : 0,24 pour l'industrie, 0,64 pour les services et 0,12 pour le bâtiment. (Source : compte de production par branche, Tableau des entrées-sorties (TES) 2018 de l'Insee)

17. Le point du premier trimestre 2009 est retiré car atypique $\Delta PIB_{T12009} = -1,6$ et l'idée est d'estimer la relation entre le solde d'opinion et le PIB dans une période « normale ».

Tableau 3. Mesure d'impact d'un évènement : décomposition des résultats pour les gilets jaunes et les grèves

	Grèves retraites (déc. 2019)	Gilets jaunes (nov.-déc. 2018)
	Act. passée	Act. passée
Part d'entreprises impactées (a)	10 %	19 %
Impact individuel (β)	-24	-17
Impact solde d'opinion agrégé (c) = (a) \times (β)	-2,4	-3,2
Impact du solde d'opinion sur le PIB : δ	0,067	0,069
Impact sur le PIB trimestriel (d) = (c) \times nb mois \times ($\delta/3$)	-0,05	-0,15 <i>(sur 2 mois)</i>

Source : Enquête Mensuelle de Conjoncture, Banque de France.

duel moyen des grèves de décembre 2019 est relativement plus important que celui des gilets jaunes, mais du fait de la proportion moindre d'entreprises touchées l'effet agrégé sur le PIB est relativement plus faible, à -0,05 pt de PIB trimestriel contre -0,15 pt pour les gilets jaunes.

La même démarche est déclinable par secteurs d'activité. En particulier, la ventilation sectorielle de la part des entreprises affectées est intéressante pour mettre en évidence l'hétérogénéité des effets sur les différentes activités économiques. Par exemple, près de 70 % des entreprises du secteur du transport et de l'entreposage déclaraient subir les conséquences du mouvement des gilets jaunes en 2018 ; lors des grèves contre la réforme des retraites en 2019, c'était majoritairement les entreprises de l'hébergement restauration qui indiquaient être affectées, avec près de 40 % d'entreprises touchées.

Concernant l'extrapolation de cette méthodologie à l'étude d'autres évènements, son caractère répliquable repose en premier lieu sur la capacité et la facilité à distinguer le groupe de contrôle du groupe de traitement, et donc sur les caractéristiques du champ lexical décrivant l'évènement E . Au plus le vocabulaire est circonscrit et sans homonyme, au plus il sera facile d'identifier l'évènement E dans le texte des commentaires. Cependant, quel que soit l'évènement E considéré, il y a un risque que le groupe de contrôle soit mal défini : en effet et comme évoqué dans la section 2.2, les commentaires étant libres et spontanés, il se peut qu'une entreprise soit touchée par un évènement sans qu'elle le mentionne lors de l'entretien de conjoncture. Cette non-déclaration aura pour conséquence de :

1. sous-estimer la part des entreprises affectées, donc sous-estimer l'impact agrégé, tout en ayant une plus faible puissance statistique (dû à la taille plus restreinte de l'échantillon de traitement) ;
2. potentiellement modifier (à la hausse ou à la baisse) l'estimation de l'impact individuel moyen β , en fonction des caractéristiques des entreprises passant du groupe de contrôle au groupe de traitement.

Il semble cependant que plus un évènement est important et médiatisé, plus le risque de non-

déclaration est faible. Ceci étant, dans le cas extrême où un évènement touche de manière tellement évidente toutes les entreprises qu'il devient inutile de le déclarer, l'absence de groupe de contrôle rend cette méthode inopérante pour la mesure de l'impact (cas de la crise de la Covid-19 à partir de mars 2020). Par ailleurs, il est envisageable de contourner le problème de non-déclaration spontanée en demandant explicitement aux agents de la Banque de France d'évoquer l'évènement en question dans leurs entretiens et d'en indiquer, le cas échéant, l'effet dans les commentaires. Dans ce cas là, on perd néanmoins la possibilité de comparer l'ampleur de cet évènement par rapport à des évènements passés pour lesquels les déclarations étaient spontanées.

En outre, pour prendre en compte l'effet du mois auquel a lieu l'évènement au sein du trimestre (mois 1, mois 2 ou mois 3), une modélisation alternative à celle de l'équation 4 a été envisagée à partir du modèle ISMA (Mogliani *et al.* (2017)), utilisé pour le *nowcasting* du PIB trimestriel et reposant sur les soldes d'opinion de l'enquête de conjoncture. Dans ce modèle, la spécification séquentielle des équations intègre l'information progressivement disponible au cours du trimestre. L'exercice mené dans le cadre de l'estimation de l'impact des grèves à partir de deux variantes de ce modèle produisait une estimation comprise dans une fourchette de -0,02 à -0,06 pt de PIB, soit un résultat très proche de celui présenté dans ce document (-0,05 pt), conduisant ainsi à retenir l'équation 4.

Enfin, il semble important de discuter les conditions qui rendent cette méthodologie pertinente ou non pour l'évaluation de l'impact d'un évènement. Adossée à une enquête de conjoncture mensuelle, signalons en premier lieu qu'elle est utile pour mesurer des effets de court terme mais qu'elle n'est pas adaptée pour évaluer des effets à moyen ou long terme. En effet les questions de l'enquête mensuelle de conjoncture se limitent essentiellement à la situation du mois passé et du mois suivant, il paraît donc naturel de supposer que les faits évoqués dans les commentaires se bornent eux aussi au même horizon temporel.

Par ailleurs, l'évaluation de l'impact est basée sur une estimation sur une unique variable d'intérêt (Y) : cette méthodologie est donc à envisager pour estimer des effets portant principalement sur une grandeur économique (par exemple, l'activité passée). Dans le cas contraire d'un évènement affectant plusieurs variables (par exemple, un choc sur l'activité et sur l'emploi), la méthodologie semble incomplète dans la mesure où elle ne permet pas de prendre en compte les potentielles interactions entre les différentes grandeurs économiques (par exemple, si la baisse de l'activité entraîne à son tour un surcroît de baisse d'emploi).

De plus, dans le cas d'un évènement ambivalent qui toucherait certaines entreprises positivement et d'autres négativement, la méthode telle qu'elle s'avèrerait là aussi imparfaite dans la mesure où elle ne proposerait qu'une estimation d'un effet moyen, résultant de la combinaison d'effets positifs et négatifs. Des raffinements seraient nécessaires pour qualifier plus précisément les deux types d'impacts en isolant deux sous-groupes de traitement, par exemple en identifiant les facteurs à l'origine du signe des effets ou en repérant cette information dans le texte des commentaires.

En conclusion, il semble que cette approche soit à privilégier dans le cas d'évènement de court terme, facilement identifiable dans les commentaires avec des recherches par mots-clés, et dont la nature et la direction de l'impact sont sans équivoques. Les estimations d'autres institutions (voir notamment les notes de conjoncture de l'Insee en mars 2019 par Pouget *et al.* (2019a) et décembre 2019 par

Pouget *et al.* (2019b)) confirment les estimations d'impact présentées dans le tableau 3 et valident *a posteriori* la méthodologie. L'analyse textuelle se révèle donc ici être un outil d'enrichissement précieux pour l'enquête mensuelle de conjoncture : elle permet de produire des estimations d'impact très rapidement, avec le même calendrier que la production des indicateurs classiques, et ce sans ajouter d'item supplémentaire dans le questionnaire habituel.

4 Évaluation d'un évènement à effets complexes avec un modèle d'apprentissage supervisé : application au cas du Brexit

Certains évènements comme le processus de sortie du Royaume-Uni de l'Union Européenne sont persistants dans le temps et affectent les entreprises de manière plus complexe qu'un seul impact négatif sur leur activité. Dans le cas du Brexit par exemple, depuis le référendum du 23 juin 2016, certaines entreprises ont connu une baisse de la demande, tandis que d'autres ont constaté, plus ou moins ponctuellement, une hausse des commandes qui leur étaient adressées. Dans ce cadre, l'outil de recherche par mots clés atteint ses limites pour rendre compte précisément des effets rapportés par les chefs d'entreprises dans les commentaires. En effet, un même mot peut être utilisé pour décrire différentes conséquences du même évènement : c'est la combinaison de plusieurs mots qui va permettre la distinction entre un impact et un autre.

De nombreux travaux académiques ont documentés les canaux par lesquels le Brexit est susceptible d'impacter la croissance économique, à la fois pour le Royaume-Uni et pour ses partenaires. À court et moyen terme notamment, le choc de demande négatif pour les exportateurs européens serait principalement lié au canal commercial (hausse des droits de douanes et renforcement des barrières non tarifaires). Si les estimations d'ampleur varient d'un travail à l'autre en fonction du scénario et des hypothèses retenus, un net consensus se dégage néanmoins de la littérature : le Brexit est un choc largement asymétrique, l'Union Européenne (UE) représentant environ 50 % du commerce de biens britanniques mais le Royaume-Uni moins de 7 % du commerce de l'UE ; les disparités régionales sont toutefois importantes, et dans ce cadre on ne peut extrapoler l'impact moyen à l'ensemble des pays européens (voir les travaux de la Brexit Task Force (2020) pour une revue détaillée des conséquences du Brexit).

Dans ce cadre, les commentaires de l'enquête mensuelle de conjoncture fournissent un éclairage intéressant sur les impacts pour les entreprises françaises, qu'il s'agisse d'effets déjà ressentis lors de la déclaration ou anticipés à un horizon de quelques mois. Pour les exploiter, nous proposons dans cette section une méthode reposant sur de l'apprentissage supervisé qui permet d'extraire des commentaires de l'information sur un évènement complexe et diffus tel que le Brexit. Cette approche donne la possibilité de caractériser individuellement quelles sont les entreprises affectées et, le cas échéant, par quel type d'effet, nous permettant ainsi de décrire au niveau macroéconomique les principaux effets à l'œuvre et leur dynamique au cours du temps.

La présentation de cette méthode et de ses résultats se décompose en quatre sous-sections. La première (4.1) a pour objectif la construction d'une base d'apprentissage en associant un ou plusieurs effets ressentis à chaque commentaire. La seconde (4.2) décrit la construction du modèle à partir du corpus précédent. La troisième (4.3) définit une règle de classification originale compte tenu des cas d'effets multiples. Enfin, la dernière sous-section (4.4) résume les résultats obtenus et discute de la reproductibilité de la méthode à de futurs évènements de ce type.

4.1 Élaboration de la base d'apprentissage

Sélection du champ d'analyse et préparation des données

Notre période d'analyse court de janvier 2016, quelques mois avant le référendum sur l'appartenance du Royaume-Uni à l'Union européenne, à décembre 2020, date de la fin de la période de transition au cours de laquelle les entreprises britanniques conservaient leur accès au marché européen. Le champ d'étude est restreint aux secteurs de l'industrie manufacturière.

En premier lieu, nous identifions les commentaires qui mentionnent le terme « Brexit », signifiant par là que ce sujet est d'intérêt pour l'entreprise¹⁸. Nous obtenons un échantillon de 1 381 commentaires qui traitent du Brexit, sur un total de 174 579 commentaires valides d'entreprises dans l'enquête industrie sur la période. Le premier constat est donc que peu d'entreprises françaises semblent préoccupées par les conséquences du Brexit (en moyenne 0,8 % des commentaires sur la période).

S'en tenir à l'identification des commentaires qui mentionnent le Brexit est incomplet car ces commentaires incluent d'autres types d'informations non pertinentes pour l'exercice et susceptibles de brouiller l'analyse d'un potentiel impact du Brexit. Pour les distinguer, nous divisons le commentaire en différentes unités de texte (portions de phrases ou phrases entières, portions de paragraphes ou paragraphes entiers) en utilisant la ponctuation (point, point-virgule, retour à la ligne) comme délimiteur dans la *tokenization*. L'idée est de sélectionner, parmi ces unités de texte, celles se rapportant au Brexit : il y en a au moins une dans chaque commentaire (celle contenant le terme « Brexit », justifiant la présence du commentaire dans l'échantillon), et potentiellement plusieurs. Pour les identifier, on utilise une recherche de mots-clés relatifs au Royaume-Uni : « *livre sterling* », « *UK* », « *RU* », « *Angleterre* », « *Grande Bretagne* »... Cette recherche élargie permet de bien capter les cas où, après avoir utilisé le terme Brexit dans une première partie du commentaire, le testeur y fait à nouveau référence mais sans répéter le mot Brexit.

À ce stade, l'échantillon est donc composé des commentaires mentionnant au moins une fois le Brexit et restreints à la partie du texte pertinente. Cette unité d'analyse est appelée « *document* » dans ce qui suit, pour la distinguer du commentaire initial. Dans l'exemple de commentaire ci-dessous, le *document* est constitué par les portions de texte en bleu :

« *Fortes livraisons en ce mois de janvier grâce à une importante demande venant d'angleterre (brexit), marché intérieur stable. Dans le même temps, la production est volontairement limitée, s'ensuit un déstockage important. Prix restant stables dans l'ensemble. Hausse attendue de la production, la forte demande venant de l'export et spécialement de l'angleterre devrait tirer l'activité sur février/mars.* »

18. Il s'avère que le champ lexical pour désigner le Brexit est principalement centré autour de ce terme, avec très peu d'expressions désignant le Brexit par analogie, ce qui rend cette phase d'identification relativement aisée.

Typologie des différents effets

Une lecture attentive du corpus de documents sélectionnés à l'étape précédente permet d'identifier quels sont les différents impacts du Brexit cités par les entreprises¹⁹, et parmi eux, ceux qui apparaissent de manière récurrente. À l'issue de cette lecture, nous construisons une typologie de six effets du Brexit, décrits dans l'encadré ci-dessous.

Typologie des effets du Brexit

Catégorie n°1 : pas d'effet ^a

Cette catégorie regroupe les documents mentionnant le Brexit pour souligner que celui-ci n'a pas d'effet pour l'entreprise. Les expressions utilisées sont très similaires d'un document à l'autre : « *brexit : pas d'impact* », « *brexit : ras* », « *entreprise non impactée par le Brexit* »...

Catégorie n°2 : baisse de la demande

La seconde catégorie fait référence aux entreprises rencontrant une baisse de leur demande à cause du Brexit : morosité des marchés mondiaux, perte de clients anglais, report de projets, attentisme...

Catégorie n°3 : hausse de la production

Ce troisième type d'effet décrit une hausse de la production pour les entreprises dont les clients au Royaume-Uni passent des commandes additionnelles dans l'objectif de sécuriser leurs stocks, notamment par crainte d'une fermeture des frontières.

Catégorie n°4 : augmentation des stocks

La quatrième thématique est la symétrique de la précédente : elle décrit le cas d'entreprises françaises augmentant leurs stocks de matières premières ou de produits finis en provenance du Royaume-Uni.

Catégorie n°5 : incertitude

La cinquième catégorie regroupe les documents indiquant de l'incertitude quant à la manière dont la sortie du Royaume-Uni impactera l'entreprise, en lien avec le flou entourant les conditions d'échanges post-Brexit (droits de douanes, barrières non tarifaires...).

Catégorie n°6 : coûts additionnels

Le sixième groupe rassemble les documents déclarant des coûts additionnels liés au Brexit, qu'ils soient monétaires ou non : dépréciation de la livre sterling, hausse des prix des matières premières, changements organisationnels voire relocalisations, effets de concurrence...

^a. Comme dans le cas des mouvements sociaux, il se peut que le chef d'entreprise mentionne le Brexit durant l'entretien de conjoncture, mais pour préciser que son entreprise n'est pas concernée.

19. Comme évoqué précédemment, on peut retrouver dans un même document plusieurs effets différents.

Dès lors qu'un document sera identifié comme « sans effet », par définition il ne pourra pas être associé conjointement à l'un des autres effets.

Labellisation de l'échantillon d'apprentissage

L'étape suivante consiste à labelliser certains documents de chaque catégorie d'effet afin de construire une base d'apprentissage pour le modèle. En pratique et comme illustré sur le tableau 4, la labellisation prend la forme de 6 variables indicatrices, chacune codant pour une catégorie d'effet, renseignées par des 0 (= absence de l'effet considéré) ou des 1 (= présence de l'effet considéré). Suffisamment de documents ont été étiquetés pour chaque catégorie, de sorte que le modèle ait assez de matière pour apprendre à caractériser chaque effet. Les labels ont par ailleurs été distribués relativement uniformément au cours du temps (plus précisément, d'un trimestre à l'autre) de manière à ne pas confondre les conditions économiques d'un mois donné avec un effet particulier. Cela permet également de mieux capter l'hétérogénéité de chaque effet.

Tableau 4. Illustration de la labellisation de l'échantillon d'apprentissage

	1. Pas d'effet	2. Baisse de l'activité	3. Hausse de la production	4. Augmentation des stocks	5. Incertitude	6. Coûts additionnels
Exemple A	0	0	1	1	0	0
Exemple B	1	0	0	0	0	0
Exemple C	0	1	0	0	0	0
Exemple D	0	0	0	0	1	0
Exemple E	0	0	0	0	0	1

Exemple A : « *augmentation des stocks* = effet brexit très belle année 2019 : *brexit profite à cette filiale de groupe anglais* tant en terme de production que de gain de change »

Exemple B : « carnet en baisse suite à la perte d'un client important en gb (*non lié au brexit*) »

Exemple C : « en revanche *baisse des exports* vers nissan gb du fait du brexit »

Exemple D : « sur le brexit : *beaucoup d'interrogations* du monde de la pêche : l'accès aux eaux irlandaises, écossaises, la fiscalité, le rapatriement de l'argent »

Exemple E : « *augmentation des coûts de transport* (effet brexit) »

Source : Enquête Mensuelle de Conjoncture, Banque de France.

Au total, nous avons ainsi labellisé 654 documents (soit 47 % des 1 381 documents mentionnant le Brexit) répartis entre les différents topics comme indiqué sur le tableau 10 en annexe.

Nous faisons ici implicitement l'hypothèse que le vocabulaire utilisé pour caractériser chaque type d'effet du Brexit est stable au cours du temps. Cette hypothèse est nécessaire, autrement l'apprentissage du modèle serait vain pour classifier des déclarations futures. Au vu des nombreux commentaires que nous avons lu et, *a posteriori*, des performances du modèle en matière de prévision (cf. infra), cette hypothèse semble très raisonnable.

4.2 Construction du modèle

Dans cette deuxième étape, nous construisons le modèle qui a pour objectif de classer dans la ou les bonne(s) catégorie(s) d'effet(s) chaque document évoquant le Brexit. Nous recourons à une classification binaire, c'est-à-dire que chaque catégorie est testée contre toutes les autres (approche dite *one vs all*), ce qui implique qu'il y a en fait un modèle pour chaque catégorie d'effet. Chacun de ces six modèles de classification binaire est indépendant et construit de la même manière.

Les variables explicatives

Les variables explicatives injectées dans le modèle sont de deux types :

- (1) des variables construites à partir des *word vectors*
- (2) des variables détectant la présence de certains mots-clés

(1) Les *word vectors* sont construits à partir d'un modèle de type *word2vec* (voir Mikolov *et al.* (2013)) entraîné sur un large corpus de texte. Les modèles *word2vec* sont des réseaux de neurones visant à reconstruire le contexte linguistique des mots dans un espace de grande dimension²⁰. En sortie, chaque mot du corpus est caractérisé par un vecteur de telle sorte que deux mots ayant des significations proches sont également proches géométriquement parlant. Les valeurs des coordonnées n'ont pas d'interprétation en soi, si ce n'est de positionner les mots les uns par rapport aux autres dans l'espace sémantique (voir graphique 8 en annexe pour une illustration).

Notre corpus de texte ayant une taille limitée, nous exploitons des vecteurs déjà entraînés sur un autre corpus de texte bien plus grand (approche dite de *transfer learning*, voir Fauconnier (2015) pour les *word vectors* réutilisés). L'inconvénient est que ces vecteurs n'ont pas été construits sur du texte s'inscrivant strictement dans le même contexte que celui de nos documents. Cependant, l'avantage décisif est lié à la taille importante du corpus sur lequel est entraîné ce modèle puisqu'il s'agit, avec 1,6 milliard de mots, d'un des corpus les plus importants en langue française (voir Baroni *et al.* (2009) pour plus de détails sur la construction du corpus par aspiration des sites internet du domaine *.fr*). En sortie, chacun des 1,6 milliard de mots est représenté par un vecteur de dimension 200.

Pour passer des vecteurs de mots à des variables explicatives construites au niveau des documents, les mots vides²¹ sont ignorés et les autres termes lemmatisés²² avant de faire la moyenne des vecteurs des lemmes constituant le document. D'autres stratégies d'agrégation décrites en annexe page 44 ont été testées, en gardant comme principe que le vecteur du document est défini comme la moyenne simple des vecteurs des mots qui le composent, mais en filtrant plus ou moins les mots à considérer dans le document.

(2) 6 variables additionnelles viennent s'ajouter aux 200 variables précédentes issues des *word vectors*. Elles indiquent, pour chaque catégorie d'effet e , la proportion de termes dans le document appartenant à la liste l_e^l des mots caractéristiques pour cet effet. Ces listes de mots clés sont définies

20. On parle de *plongement lexical*, ou *word embedding* en anglais.

21. Les mots vides, ou *stop words*, désignent les mots non significatifs pour la compréhension du texte : articles, pronoms, prépositions, auxiliaires...

22. La lemmatisation est réalisée à partir du dictionnaire Morphalou 3 produit par le laboratoire ATILF (2019)

par les auteurs et regroupent les termes les plus fréquemment utilisés pour décrire l'effet en question. Par exemple, pour la catégorie 1 « pas d'effet », la liste contient les termes : « *impact* », « *effet* », « *peu* », « *pas* », « *aucun* », « *concerné* », « *sans* », « *ras* »... Les listes complètes des termes pour les 6 effets sont détaillées en annexe page 47. Pour chaque document i et pour chaque effet e , la variable explicative suivante est ainsi créée :

$$W_{e_i} = \frac{\text{Nombre de mots du document } i \text{ appartenant à } l_e^l}{\text{Nombre total de mots dans le document } i} \quad (5)$$

D'autres variantes visant à caractériser la présence de certains termes dans le texte des documents ont été testées et sont décrites en annexe page 4.4.

L'entraînement des modèles

Les 206 variables explicatives qui caractérisent chaque document sont injectées, pour chacun des 6 modèles (1 modèle par effet), dans un algorithme logit-lasso (Tibshirani (1996)) qui s'écrit :

$$\underset{\beta_0, \beta_1, \dots, \beta_p}{\text{Min}} \frac{1}{2} \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{i,j} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad \text{s/c } \lambda \geq 0 \quad (6)$$

où y_i vaut 1 si le document i décrit l'effet considéré et 0 sinon, et les $x_{i,j}$ les variables explicatives décrites dans le paragraphe précédent.

Ce modèle de régression combine à la fois simplicité dans l'interprétation des coefficients, une étape de sélection de variables, et est compatible avec un nombre de variables élevé. Le paramètre de pénalisation $lambda$ ²³ (λ) est choisi par validation croisée, en paramétrant le nombre de blocs à 10 dans la fonction *cv.glmnet* du package *glmnet* de Friedman *et al.* (2010). Le critère de performance utilisé est celui de l'AUC (*Area Under the Curve*, aire sous la courbe ROC²⁴).

Pour l'évaluation de ces modèles, nous mettons en place une procédure de validation croisée à 10 blocs (voir graphique 3 ci-dessous), répétée 5 fois. Pour chaque itération, nous sauvegardons l'AUC, le nombre de variables sélectionnées, et le paramètre lambda : ces résultats sont ensuite moyennés sur l'ensemble des 10 blocs x 5 répétitions.

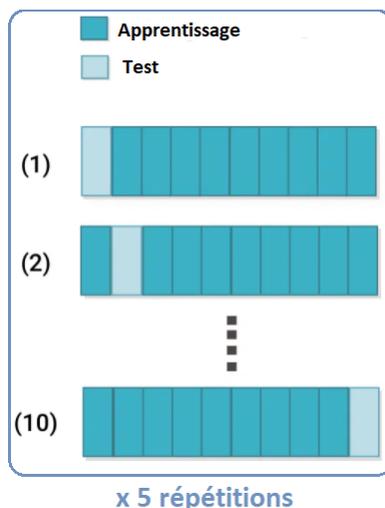
Pour chacun des 6 effets, les performances des modèles sont indiquées dans la deuxième colonne du tableau 5. En moyenne, l'AUC est de 0,87 : il est maximal pour la prédiction de la catégorie d'effet 4 (augmentation des stocks) à 0,94, et minimal pour la catégorie 6 (coûts additionnels) à 0,75. Les résultats détaillés des estimations sont présentés dans le tableau 15 en annexe.

Enfin, pour confirmer les performances de ces modèles, nous les appliquons sur un échantillon de

23. Le paramètre *lambda* est un coefficient de régularisation qui permet de contrôler l'impact de la pénalité imposée sur les coefficients estimés de la régression : $\lambda \sum_{j=1}^p |\beta_j|$

24. Ce critère est relativement courant dans le cas d'un problème de classification binaire : il a notamment l'avantage d'être indépendant du seuil de probabilité défini pour la classification, ce qui laisse la possibilité de déterminer ce dernier en toute dernière étape.

Graphique 3. Illustration de la validation croisée sur 10 blocs répétée 5 fois



Note : Les données sont partitionnées en 10 blocs. À chaque itération $i = 1, \dots, 10$, le modèle est appris sur les 9 blocs $\neq i$ puis testé sur le 10^{ème} bloc i . Cette procédure est répétée 5 fois pour faire varier la partition initiale en 10 blocs.

test, c'est-à-dire des commentaires qui n'ont jamais servi à l'estimation du modèle. Les performances obtenues en termes d'AUC sont indiquées dans la troisième colonne du tableau 5 : pour chaque catégorie, l'AUC sur l'échantillon de test est proche de celui obtenu en validation croisée, avec un écart de 0,03 en moyenne et au maximum de 0,07 sur la catégorie 2.

Tableau 5. Performances des modèles en validation croisée et sur l'échantillon de validation

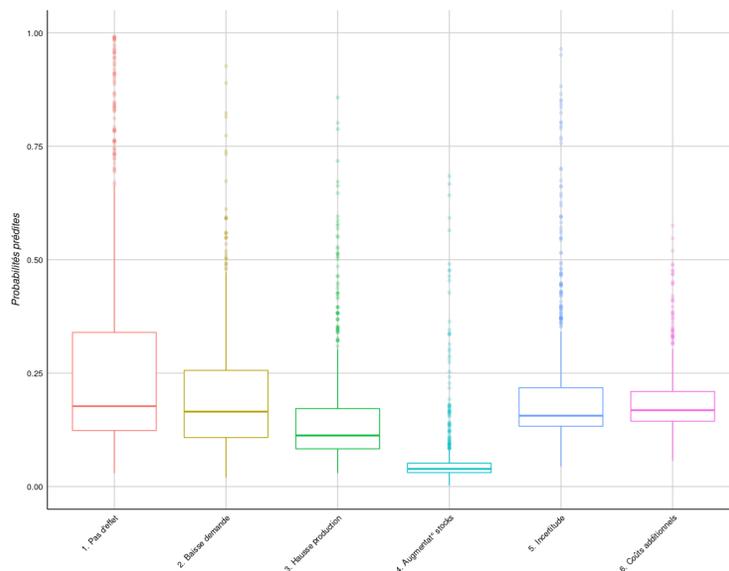
	AUC en validation croisée	AUC sur l'échantillon de test
Catégorie 1 : pas d'effet	0,92	0,92
Catégorie 2 : baisse de la demande	0,85	0,78
Catégorie 3 : hausse de la production	0,90	0,88
Catégorie 4 : augmentation des stocks	0,94	0,96
Catégorie 5 : incertitude	0,83	0,86
Catégorie 6 : coûts additionnels	0,75	0,72

Source : Enquête Mensuelle de Conjoncture, Banque de France.

4.3 Règle de classification

Le graphique 4 représente les distributions des probabilités prédites en validation croisée pour chaque catégorie d'effet. Deux constats en découlent : d'une part, les probabilités sont relativement faibles, avec des valeurs médianes autour de 0,2 ; d'autre part les distributions sont dissemblables d'un effet à l'autre : notamment les probabilités de la catégorie d'effet 4 sont très concentrées à gauche avec un troisième quartile inférieur à 0,1, alors que celles de la catégorie 1 sont un peu plus distribuées vers la droite avec un deuxième quartile supérieur à 0,3.

Graphique 4. Distributions des probabilités prédites pour chaque catégorie d’effet (échantillon de test)



Source : Enquête Mensuelle de Conjoncture, Banque de France.

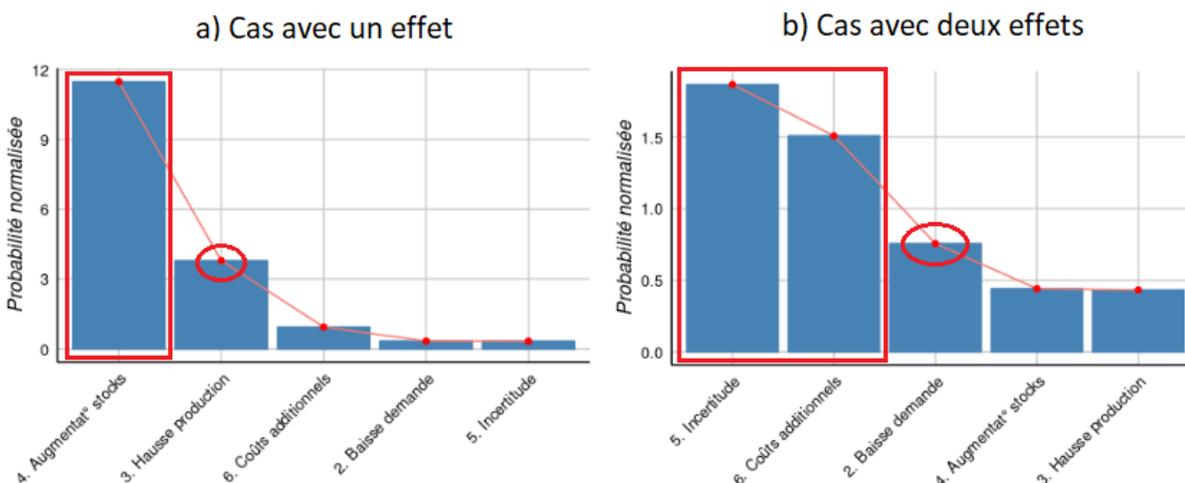
Nous proposons une règle de classification combinant plusieurs critères. En amont, les probabilités prédites pour chaque catégorie d’effet sont normalisées par la proportion de cet effet dans la base d’apprentissage. Cette normalisation permet de s’assurer de la comparabilité des probabilités d’une catégorie d’effet à l’autre. La règle de classification suivante est ensuite appliquée pour chaque document :

1. dans le cas où la probabilité de la catégorie « sans effet » est supérieure au seuil maximisant la somme *sensitivité + spécificité*²⁵ pour cette catégorie (statistique de (Youden (1950))), alors le document est classé comme « sans effet » ;
2. sinon, les probabilités des 5 autres catégories sont rangées par ordre décroissant et une règle du coude est appliquée : le document sera classifié dans les effets dont les probabilités prédites sont situées à gauche du point d’inflexion, comme illustré dans le graphique 5 ci dessous.

Cette règle de décision permet d’une part de s’assurer que l’ensemble des documents est effectivement classé dans une catégorie d’effet, et d’autre part d’autoriser une classification dans plusieurs catégories le cas échéant. Les performances obtenues en termes de sensibilités et spécificités pour chaque catégorie sont indiquées dans les colonnes 2 à 4 du tableau 6 : en moyenne, la somme *sensitivité + spécificité* est de 1,57, avec un maximum à 1,73 atteint pour la catégorie 4 (augmentation des stocks) et un minimum à 1,33 pour la catégorie 6 (coûts additionnels). Au niveau du nombre de catégories prévues, cette règle affiche un taux de bonnes prédictions de 74 % ; plus précisément, elle a tendance à classer dans un peu plus de catégories que n’en a le document en réalité.

²⁵. La sensibilité et la spécificité sont deux mesures interdépendantes et dépendantes du seuil de classification. La sensibilité désigne la capacité à détecter les vrais positifs, tandis que la spécificité désigne la capacité à détecter les vrais négatifs.

Graphique 5. Règle de classification : illustration de la règle du coude



Source : Enquête Mensuelle de Conjoncture, Banque de France.

Tableau 6. Sensitivités et spécificités pour chaque catégorie d’effet

	Échantillon de validation croisée			Échantillon de test		
	Sensitivité	Spécificité	Somme	Sensitivité	Spécificité	Somme
Catégorie 1 : pas d’effet	0,93	0,76	1,69	0,91	0,81	1,72
Catégorie 2 : baisse de la demande	0,85	0,64	1,49	0,87	0,55	1,42
Catégorie 3 : hausse de la production	0,92	0,76	1,68	0,87	0,70	1,57
Catégorie 4 : augmentation des stocks	0,91	0,82	1,73	0,95	0,70	1,65
Catégorie 5 : incertitude	0,90	0,58	1,48	0,87	0,68	1,56
Catégorie 6 : coûts additionnels	0,86	0,47	1,33	0,86	0,40	1,27

Source : Enquête Mensuelle de Conjoncture, Banque de France.

Sur l’échantillon de test (données non utilisées dans la procédure de validation croisée), cette règle de classification donne des performances de même ordre de grandeur, indiquées dans les colonnes 5 à 7 du tableau 6 : à l’exception des catégories 1 et 5, les autres catégories voient leur performance se dégrader de l’ordre de 0,1 sur la somme (*sensitivité + spécificité*), principalement à cause d’une spécificité moindre.

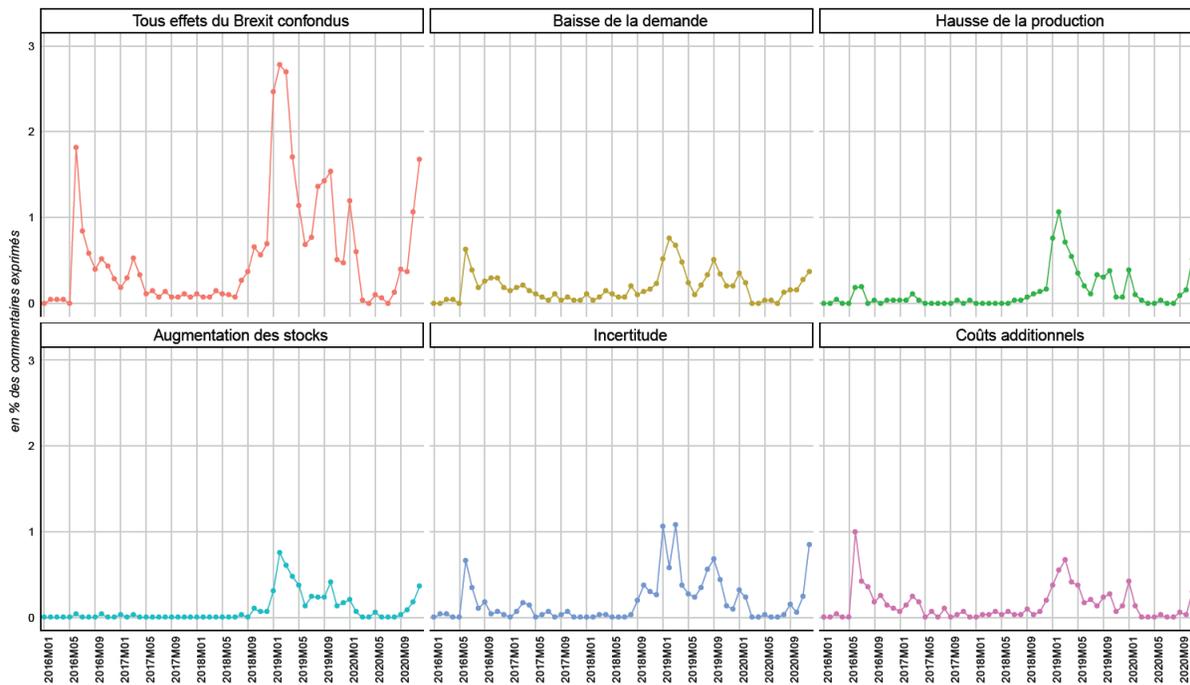
4.4 Résultats

En appliquant la règle de classification sur les probabilités prédites pour l’ensemble des documents (qu’ils soient labellisés ou non), nous sommes en mesure de suivre l’évolution et la magnitude des différents effets du Brexit en fonction du temps, comme présenté sur le graphique 6. Le premier cadre en haut à gauche indique chaque mois le pourcentage d’entreprises décrivant un effet du Brexit quel qu’il soit ; les cadrans suivants détaillent plus précisément l’évolution des effets 2 à 6 de la typologie.

Chronologiquement, un premier pic apparaît en juin 2016, mois du referendum sur l’appartenance du Royaume-Uni à l’Union Européenne : à cette date, près de 2 % des chefs d’entreprises dans l’indus-

trie mentionnent un effet du Brexit lors de l'entretien de conjoncture. Parmi ceux là, une majorité indiquent des effets monétaires²⁶ qui persisteront durant l'été 2016. Par ailleurs, toujours sur cette période de l'été 2016, autour de 0,5 % des chefs d'entreprise indiquent une baisse de la demande et/ou des incertitudes quant aux effets du Brexit pour leur entreprise.

Graphique 6. Évolution des effets du Brexit au cours du temps



Source : Enquête Mensuelle de Conjoncture, Banque de France.

S'ensuit une période relativement calme du point de vue du Brexit, sans effet notable signalé par les entreprises jusqu'à la fin de l'été 2018. Cette période correspond au temps de négociation des conditions de sortie de l'Union Européenne pour le Royaume-Uni. Mais dès l'automne 2018, des premiers signaux apparaissent, caractérisés par une inquiétude grandissante des chefs d'entreprises. Ces signaux se matérialisent ensuite très concrètement au premier trimestre 2019 avec des ajustements d'offre et de demande des deux côtés de la Manche. Les acteurs économiques sont en effet inquiets d'une fermeture brutale des frontières au 29 mars 2019, date officielle du Brexit, l'accord de retrait approuvé au Conseil Européen en novembre 2018 étant rejeté trois fois par le Parlement britannique les 15 janvier, 12 mars et 29 mars. En conséquence, des entreprises françaises connaissent une baisse de leur demande, résultat de clients anglais préférant changer de fournisseur, tandis que d'autres augmentent leur production pour satisfaire les commandes additionnelles de leurs clients britanniques. En parallèle, certaines entreprises françaises stockent elles aussi par précaution des produits importés du Royaume-Uni afin d'anticiper tout problème aux frontières à la fin du mois de mars 2019. Finalement, l'Union Européenne accordera plusieurs reports de la date du Brexit, d'abord au 12 avril, puis au 31 octobre et enfin au 31 janvier 2020.

26. Après le 23 juin, la monnaie britannique avait en effet perdu près de 15 % en quelques jours.

En lien avec les négociations très compliquées sur l'accord de sortie, la suite de l'année 2019 est caractérisée par un niveau relativement élevé d'incertitudes. À ce sujet, les travaux de Bloom *et al.* (2019) suggèrent que la baisse graduelle de l'investissement des entreprises britanniques (dont le niveau est estimé inférieur de 11 % en 2019 à ce qu'il aurait été sans le vote en faveur du Brexit) est la conséquence de l'incertitude persistante provoquée par le résultat du référendum de juin 2016. Face à ce niveau d'incertitude élevé et qui perdure dans le temps, les entreprises françaises pourraient, par le même canal, engager des ajustements qualitativement similaires.

Par ailleurs, la baisse de la demande reste en 2019 une conséquence du Brexit signalée de manière récurrente par les entreprises, mais les stocks de précaution tendent à disparaître des deux côtés de la Manche. En parallèle, une proportion non négligeable de chefs d'entreprises indiquent des coûts supplémentaires en lien avec le Brexit, par exemple en lien avec des restructurations d'entreprises pour celles ayant des établissements dans les deux pays, ou la création de nouveaux circuits d'échanges en anticipation des futures formalités douanières. Une partie de ces coûts additionnels vient également du temps passé par les dirigeants à planifier et organiser les conséquences du Brexit. Pour les entreprises outre-manche, Bloom *et al.* (2019) ont montré que la préparation de ces ajustements intra-entreprises étaient la principale explication de la baisse de la productivité des entreprises britanniques estimée entre -2 et -5 % 3 ans après le référendum.

Finalement, lors de la sortie effective du Royaume-Uni de l'Union Européenne au 31 janvier 2020, les impacts du Brexit apparaissent de bien moindre ampleur comparé à ceux de juin 2016, mars 2019 et octobre 2019. L'explication vient de la période de transition qui s'ouvre au 31 janvier 2020 et qui durera encore onze mois durant lesquels « rien ne change » pour les entreprises britanniques qui bénéficient encore de l'accès au marché européen. À l'issue de cette transition, fin décembre 2020 un regain d'incertitude gagne les chefs d'entreprises françaises, accompagné de stocks de précaution des deux côtés de la Manche : en effet, dès le 1^{er} janvier 2021 de nouvelles formalités post-Brexit entrent en vigueur pour les échanges entre les deux pays.

Les résultats présentés sur le graphique 6 sont également déclinables en un panorama sectoriel, par taille d'entreprise et par région. À ce titre, les entreprises de plus de 250 salariés sont bien plus affectées par les conséquences du Brexit que les entreprises de 249 salariés ou moins (voir la planche de résultats 11 en annexe) : au pic de mars 2019, elles sont plus de deux fois plus nombreuses à indiquer un effet du Brexit lors de l'entretien de conjoncture. La situation géographique de l'entreprise est également un facteur important : les entreprises des Hauts-de-France, entretenant en moyenne des liens plus étroits avec le Royaume-Uni, indiquent plus fréquemment un impact du Brexit par rapport aux entreprises des autres régions françaises (à titre d'exemple, la planche de résultats 12 en annexe compare les entreprises des régions Hauts-de-France et Occitanie).

Au delà de la magnitude des effets, l'analyse sectorielle permet également de mettre en évidence des temporalités différentes d'un secteur à l'autre : à titre d'exemple et comme illustré sur la planche de résultats 13 en annexe, l'industrie automobile a été impactée très tôt (dès juin 2016) relativement au secteur des produits en caoutchouc, plastique et autres produits minéraux non métalliques où le pic d'incertitude est atteint en mars 2019. Les structures d'échanges entre le France et Royaume-Uni pourraient éclairer ce décalage temporel : d'une part, l'interdépendance très forte de la production automobile entre la France et le Royaume-Uni induit une contrainte forte pour l'assemblage des véhicules et rend donc impératif des ajustements dans les chaînes de production : plusieurs grands constructeurs automobiles ont annoncé des fermetures de sites, suppressions d'emploi ou relocalisa-

tions. D'autre part, les produits en caoutchouc et plastiques font parti des produits les plus exportés vers l'Angleterre : dans ce cadre, les entreprises françaises seraient principalement inquiètes de l'application de droits de douanes qui viendraient pénaliser la compétitivité de leurs exportations²⁷.

En conclusion, la méthodologie proposée permet de classifier au sein d'une typologie des documents évoquant un sujet commun mais avec des impacts variés d'une entreprise à l'autre. En particulier, elle semble d'autant plus pertinente dans les cas où la nature de l'impact est multiple (activité, prix, demande, livraisons...avec potentiellement des effets positifs pour les uns et négatifs pour les autres) et où il y a un besoin de suivre l'évolution et la magnitude des effets au cours du temps. Au prix d'un travail d'adaptation sur la typologie et la labellisation, et sous réserve d'un nombre suffisant d'observations, il est tout fait envisageable de reproduire cette méthodologie pour l'analyse d'un évènement futur. Une telle application des méthodes d'analyse textuelle aux retours de terrains indiqués par les chefs d'entreprise apporte en effet un éclairage précieux au diagnostic conjoncturel.

La reproductibilité se fait bien entendu à un coût relativement plus élevé que l'application précédente basée sur des recherches par mots-clés. Notamment, la typologie d'effets repose avant tout sur la lecture et la compréhension des commentaires : elle mobilise donc une expertise métier sur le sujet d'intérêt et contient nécessairement une part de subjectivité. Sur ce point, un axe de développement de la méthode consisterait à objectiver la classification retenue par des méthodes d'extraction des topics (*topic modeling*) telles que la LDA (voir Blei *et al.* (2003)), méthodes pour lesquelles l'investissement peut-être relativement conséquent étant donné la sensibilité des résultats au choix des paramètres et la difficulté d'interpréter les résultats qui en découlent. Enfin, la constitution d'un échantillon d'apprentissage est une étape fastidieuse mais un passage obligé pour la mise en place d'un modèle d'apprentissage supervisé. Toujours est-il qu'avec ces ajustements, cette combinaison de méthodes d'analyse textuelle et de machine learning permet de rendre compte des effets complexes qui affectent les entreprises françaises lors d'un évènement tel que celui du Brexit.

27. Pour plus de détails sur les impacts sectoriels, voir par exemple cet article de la Chambre de Commerce et d'Industrie des Hauts-de-France.

Discussion et conclusion

Cet article a développé plusieurs types d'applications de l'analyse textuelle aux commentaires des chefs d'entreprises interrogés dans l'enquête mensuelle de conjoncture de la Banque de France. Il s'agit d'un des rares articles mettant en valeur l'utilisation de ce type de méthode, sur ce type de données.

Nous avons montré que les commentaires sont riches d'information pour compléter le diagnostic conjoncturel donné par les variables usuelles de l'enquête. Parmi les applications présentées, la méthodologie de mesure d'impact d'un évènement à l'aide de recherches par mots-clés est la plus aboutie : elle a été éprouvée dans plusieurs situations dans lesquelles les mesures d'impact ont été confirmées par les mesures concordantes produites par d'autres institutions. En plus des cas d'applications présentés ici, elle a par ailleurs récemment été mise en œuvre pour mesurer l'impact des annonces gouvernementales du 31 mars 2021 et de la fermeture des écoles pour trois semaines au mois d'avril sur les anticipations d'activité à fin mars des entreprises pour le mois d'avril.

Les recherches par mots-clés, qui font la force de cette méthodologie, peuvent par ailleurs être utilisées plus largement pour décrire l'occurrence de certaines thématiques dans les commentaires. Le travail présenté dans le billet de blog n°160 de la Banque de France reprend par exemple cette idée pour mettre en évidence les stratégies d'adaptation des entreprises lors du confinement de mars 2020. Au même titre, l'indicateur d'incertitude publié dans le point de conjoncture de la Banque de France à fin février 2021 met en évidence les variations de l'inquiétude exprimée par les chefs d'entreprise à partir d'un comptage de mots-clés. Ce principe de recherches par mots-clés est donc à même d'être fréquemment utilisé pour identifier des entreprises affectées par une thématique donnée et en mesurer l'impact sur une grandeur économique.

S'agissant de l'indice de sentiment, la conclusion de la section 2.1 soulève une piste d'approfondissement intéressante pour de futurs travaux. En effet, une fois démontré que l'indice de sentiment construit à partir des commentaires corrobore l'indice du climat des affaires, se pose la question de l'éventuelle information supplémentaire que cet indice pourrait apporter pour la prévision du PIB au sein d'un modèle de *nowcasting* de type ISMA. L'indice de sentiment, et le dictionnaire sur lequel il repose, pourrait alors être optimisés précisément dans cette optique de prévision.

Enfin, l'application d'apprentissage supervisé présentée pour identifier des effets complexes et diffus dans le temps propose elle aussi une piste d'approfondissement à même de renforcer la méthode. Plutôt que d'utiliser des word vectors pré-entraînés sur du vocabulaire généraliste, il s'agirait d'entraîner un modèle word2vec sur notre propre corpus de données, et par là améliorer la qualité de la position relative des mots les uns par rapport aux autres dans le champ lexical des enquêtes de conjoncture. Plus largement, ces vecteurs de mots ainsi entraînés pourraient également être utilisés pour renseigner sur la proximité des mots caractéristiques de certaines périodes sous revue, l'identification de ces mots saillants faisant l'objet d'un autre projet en cours de développement.

Bibliographie

- Andres ALGABA, David ARDIA, Keven BLUTEAU, Samuel BORMS et Kris BOUDT : Econometrics meets sentiment : An overview of methodology and applications. *Journal of Economic Surveys*, 34(3):512–547, 2020.
- Werner ANTWEILER et Murray Z FRANK : Is all that talk just noise? the information content of internet stock message boards. *The Journal of finance*, 59(3):1259–1294, 2004.
- Valentina APRIGLIANO, Simone EMILIOZZI, Gabriele GUAITOLI, Andrea LUCIANI, Juri MARCUCCI et Libero MONTEFORTE : The power of text-based indicators in forecasting the italian economic activity. *Bank of Italy Temi di Discussione (Working Paper) No*, 1321, 2021.
- David ARDIA, Keven BLUTEAU, Samuel BORMS et Kris BOUDT : The r package sentometrics to compute, aggregate and predict with textual sentiment. *Journal of Statistical Software, forthcoming*, 2020.
- David ARDIA, Keven BLUTEAU et Kris BOUDT : Questioning the news about economic growth : Sparse forecasting using thousands of news-based sentiment values. *International Journal of Forecasting*, 35(4):1370–1386, 2019.
- ATILF : Morphalou [lexique]. ORTOLANG (Open Resources and TOols for LANGuage) –www.ortolang.fr, 2019. Analyse et traitement informatique de la langue française (ATILF) - UMR 7118.
- Banque de FRANCE : Point sur la conjoncture française à fin février 2021. Points de conjoncture de la Banque de France, 2021.
- Luca BARBAGLIA, Sergio CONSOLI et Sebastiano MANZAN : Forecasting with economic news. *In Available at SSRN : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3698121*, 2021.
- Marco BARONI, Silvia BERNARDINI, Adriano FERRARESI et Eros ZANCHETTA : The wacky wide web : a collection of very large linguistically processed web-crawled corpora. *Language resources and evaluation*, 43(3):209–226, 2009.
- Maël BLANCHET et Mélanie COUEFFE : Améliorer l’estimation du pib en temps réel grâce aux grands ensembles de données. Lettre Trésor-Éco n°254 - Janvier 2020, 2020.
- David M BLEI, Andrew Y NG et Michael I JORDAN : Latent dirichlet allocation. *The Journal of machine learning research*, 3:993–1022, 2003.
- Nicholas BLOOM, Philip BUNN, Scarlet CHEN, Paul MIZEN, Pawel SMIETANKA et Gregory THWAITES : The impact of brexit on uk firms. *National Bureau of Economic Research*, 2019.
- Clément BORTOLI, Stéphanie COMBES et Thomas RENAULT : Comment prévoir l’emploi en lisant le journal. *Note de conjoncture, INSEE, mars*, pages 35–43, 2017.
- BREXIT TASK FORCE : A review of economic analyses on the potential impact of brexit. *Occasional Paper Series*, (249), 2020.

- Jean-Philippe FAUCONNIER : French word embeddings. <http://fauconnier.github.io>, 2015.
- Jerome FRIEDMAN, Trevor HASTIE et Rob TIBSHIRANI : Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1):1, 2010.
- Matthew GENTZKOW, Bryan KELLY et Matt TADDY : Text as data. *Journal of Economic Literature*, 57(3):535–74, 2019.
- Mathilde GERARDIN et Martial RANVIER : L’adaptation des entreprises au confinement : l’apport de l’analyse textuelle. BlocNoteEco de la Banque de France n°160 - Mai 2020, 2020.
- Elaine HENRY : Are investors influenced by how earnings press releases are written? *The Journal of Business Communication (1973)*, 45(4):363–407, 2008.
- Matthew HONNIBAL et Ines MONTANI : spacy 2 : Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing. *To appear*, 7(1):411–420, 2017.
- Clayton HUTTO et Eric GILBERT : Vader : A parsimonious rule-based model for sentiment analysis of social media text. *In Proceedings of the International AAAI Conference on Web and Social Media*, volume 8, 2014.
- Eleni KALAMARA, Arthur TURRELL, Chris REDL, George KAPETANIOS et Sujit KAPADIA : Making text count : economic forecasting using newspaper text. *Bank of England Working Paper*, 2020.
- Vegard H LARSEN et Leif A THORSRUD : The value of news for economic developments. *Journal of Econometrics*, 210(1):203–218, 2019.
- Steven LORIA, P KEEN, M HONNIBAL, R YANKOVSKY, D KARESH, E DEMPSEY *et al.* : Textblob : simplified text processing ; 2018. *Online* : <https://textblob.readthedocs.io/en/dev/Accessed>, pages 08–02, 2019.
- Tim LOUGHRAN et Bill McDONALD : When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of finance*, 66(1):35–65, 2011.
- Tomas MIKOLOV, Kai CHEN, Greg CORRADO et Jeffrey DEAN : Efficient estimation of word representations in vector space. *arXiv preprint arXiv :1301.3781*, 2013.
- Matteo MOGLIANI, Olivier DARNÉ et Bertrand PLUYAUD : The new miba model : Real-time nowcasting of french gdp using the banque de france’s monthly business survey. *Economic Modelling*, 64:26–39, 2017.
- Satoshi MORINAGA, Kenji YAMANISHI, Kenji TATEISHI et Toshikazu FUKUSHIMA : Mining product reputations on the web. *In Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 341–349, 2002.
- Julien POUGET, Frédéric TALLET, Juliette GRANGIER et Louise VIARD-GUILLOT : Soutiens budgétaires en zone euro, sur fond de tensions - note de conjoncture - mars 2019. Note de conjoncture de l’Insee - Mars 2019, 2019a.

- Julien POUGET, Frédéric TALLET et Thomas LABOUREAU : Clair-obscur - note de conjoncture - décembre 2019. Note de conjoncture de l'Insee - Décembre 2019, 2019b.
- Filipe N RIBEIRO, Matheus ARAÚJO, Pollyanna GONÇALVES, Marcos André GONÇALVES et Fabrício BENEVENUTO : Sentibench-a benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science*, 5(1):1–29, 2016.
- Julia SILGE et David ROBINSON : tidytext : Text mining and analysis using tidy data principles in r. *Journal of Open Source Software*, 1(3):37, 2016.
- Paul C TETLOCK : Giving content to investor sentiment : The role of media in the stock market. *The Journal of finance*, 62(3):1139–1168, 2007.
- Robert TIBSHIRANI : Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society : Series B (Methodological)*, 58(1):267–288, 1996.
- Richard M TONG : An operational system for detecting and tracking opinions in on-line discussion. *In Working Notes of the ACM SIGIR 2001 Workshop on Operational Text Classification*, volume 1, 2001.
- Peter D TURNEY : Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. *arXiv preprint cs/0212032*, 2002.
- Ivan YOTZOV, Nick BLOOM, Philip BUNN, Paul MIZEN, Pawel SMIETANKA et Greg THWAITES : What matters to firms? new insights from survey text comments. Bank Underground, blog written by staff at the Bank of England- April 2021, 2021.
- William J YAUDEN : Index for rating diagnostic tests. *Cancer*, 3(1):32–35, 1950.

Annexes

Graphique 7. Extrait du formulaire de collecte rempli par l'agent du réseau après l'entretien pour l'Enquête Mensuelle de Conjoncture



BANQUE DE FRANCE
EUROSYSTEME



Système d'Information Renforcé sur les Centralisations et les Enquêtes

SIRCE

Menu

- Messages
- Requête
- Référentiel - Sources
- Référentiel élaboré
- Gestion Arborescences composées
- Collecte
- >> Génération questionnaires
- >> **Saisie des réponses**
- Calendrier
- Calcul
- Administration
- DECONNEXION

Saisie des réponses

Questionnaire

Source : 000520153-2932Z	Début coll. : 20/12/2019	Nb répondu : 35 / 40
RECORD FRANCE	Retard Nat. : 07/01/2020	1ère réponse saisie le : 03/01/2020
<input type="checkbox"/> Correspondant absent	Retard Ter. : 07/01/2020	Dernière modification : 08/01/2020

Annotation Source

[Afficher / Masquer les critères](#)

Bordereau Industrie

<p>Evolution de la Production / M-1</p> <p>PRTEM1 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 2,5 <input type="radio"/> 3 <input checked="" type="radio"/> 3,5 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> SR</p>	<input type="checkbox"/> Imputation
<p>Evolution des Livraisons / M-1</p> <p>LITEM1 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 2,5 <input type="radio"/> 3 <input type="radio"/> 3,5 <input checked="" type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> SR</p>	<input type="checkbox"/> Imputation
<p>Evolution des Commandes reçues / M-1</p> <p>CRTEM1 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 2,5 <input type="radio"/> 3 <input checked="" type="radio"/> 3,5 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> SR</p>	<input type="checkbox"/> Imputation
<p>Evolution des Commandes intérieures reçues ds le mois / M-1</p> <p>CRFEM1 <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 2,5 <input type="radio"/> 3 <input checked="" type="radio"/> 3,5 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> SR</p>	<input type="checkbox"/> Imputation
<p>Evolution des Commandes reçues de l'étranger / M-1</p> <p>CREEM1 <input type="radio"/> 1 <input type="radio"/> 2 <input checked="" type="radio"/> 2,5 <input type="radio"/> 3 <input type="radio"/> 3,5 <input type="radio"/> 4 <input type="radio"/> 5 <input type="radio"/> SR</p>	<input type="checkbox"/> Imputation

Quels sont les faits saillants relatifs à cette période ?

ZRFSM0

LES LIVRAISONS SONT FORTEMENT RaLENTIS pour l'export du fait des blocages douaniers ou géopolitiques (Turquie Liban Egypte Iran Algerie et Chili) De ce fait hausse des stocks partiellement compensés par une baisse de la

Imputation

Retour

Imprimer

Initialiser

Contrôler

Supprimer

Valider

Tableau 7. Principaux mouvements sociaux ayant affectés la conjoncture française de 2009 à 2020

Mouvement social	Période	Indications sur l'ampleur du mouvement
Manifestations pour la défense de l'emploi et du pouvoir d'achat	Mars 2009	Le 19 mars 2009, entre 1,2 million (source Intérieur) et 3 millions (source syndicale) de personnes défilent pour la défense de l'emploi et des salaires.
Grèves SNCF	Avril 2010	15 jours de grève au total, le plus long mouvement qu'ait connu la SNCF depuis le conflit sur les régimes spéciaux de retraite en 2007.
Réforme des retraites	Mars à nov. 2010	Au pic de la mobilisation le 12 octobre 2010 les syndicats revendiquent 3,5 millions de manifestants, le ministère de l'intérieur 1,23 million.
Grèves dans l'éducation nationale	Février 2011	Le 10 février 2011, le taux de grévistes dans les écoles primaires a été compris entre 20% (ministère de l'Education nationale) et 35% (syndicat FSU), et entre 15% et 50% dans les collèges et lycées
Grèves SNCF	Juin 2014	10 jours de grève, avec un pic à 28% de grévistes, le plus long conflit à la SNCF depuis 2010.
Grèves Air France	Septembre 2014	14 jours de grève des pilotes, le plus long mouvement social de l'histoire d'Air France.
Grèves des routiers	Janvier 2015	Le dimanche 18 janvier 2015, les routiers ont commencé une grève reconductible pour réclamer des augmentations de salaires : opérations escargot, blocage de sites industriels ou de dépôts pétroliers.
Loi travail	Mars à sept. 2016	Le 31 mars 2016, le mouvement atteint son apogée : entre 390 000 et 1,2 million de personnes défilent dans toute la France. A partir de la mi-mai, des dépôts pétroliers sont bloqués, ce qui provoque une ruée à la pompe et entraîne une pénurie temporaire d'essence et gazole dans 4 000 des 12 000 stations-service de France.
Réforme ferroviaire	Avril à juin 2018	A partir du 2 avril 2018, l'intersyndicale des salariés s'est engagée dans une grève perlée à l'échelle nationale, de 48 heures tous les cinq jours pour une durée de 3 mois.
Gilets jaunes	À partir de nov. 2018	A partir du 17 novembre 2018, sont organisés des blocages d'axes routiers et de carrefours giratoires, avec des manifestations nationales chaque samedi.
Réformes des retraites	Déc. 2019 à janv. 2020	Plus important mouvement social en France depuis les manifestations contre la réforme des retraites en 2010. Le 5 décembre 2019, entre 0,8 (Intérieur) et 1,5 million (CGT) de personnes manifestent.

Tableau 8. Estimations des impacts individuels des gilets jaunes et des grèves : résultats des régressions

	Grèves	Gilets jaunes
<i>Variable Y</i>	Activité passée	Activité passée
<i>Mois M</i>	Décembre 2019	Décembre 2018
Etre impacté	-23,62 (2,82)***	-16,72 (2,17)***
Etre impacté x Mois <i>M</i> (Réf : Mois <i>M</i> - 1)	-	2,19 (3,23)
Mois <i>M</i> (Réf : Mois <i>M</i> - 1)	oui	oui
Secteurs	oui	oui
Secteurs x Mois <i>M</i>	oui	oui
Régions	oui	oui
Effectifs	oui	oui
Constante	oui	oui
R^2 ajusté	10,89 %	11,90 %
<i>N</i>	12 573	12 357

Notes :

Les écarts types sont entre parenthèses. *** : p-valeur < 0.001, ** : p-valeur < 0.01, * : p-valeur < 0.05
 Pour les grèves, la variable de traitement vaut 0 en novembre 2019 et appartient à $\{0, 1\}$ en décembre 2019.
 Pour les gilets jaunes, le mouvement s'étant étalé sur 2 mois, la variable de traitement appartient à $\{0, 1\}$ en novembre et décembre 2018.

Source : Enquête Mensuelle de Conjoncture, Banque de France.

Tableau 9. Estimations de la relation entre solde d'opinion et variation du PIB : résultats des régressions

	Grèves	Gilets jaunes
<i>Période d'estimation</i>	T2 2009 - T2 2018	T2 2009 - T2 2019
<i>solde opinion</i> $\frac{Y}{T}$	0,067 (0,012)***	0,069 (0,013)***
ΔPIB_{T-1}	-0,32 (0,15)*	-0,35 (0,16)*
Constante	oui	oui
R^2 ajusté	42,97 %	44,61 %
N	41	37

Note : Les écarts types sont entre parenthèses. *** : p-valeur < 0.001, ** : p-valeur < 0.01, * : p-valeur < 0.05

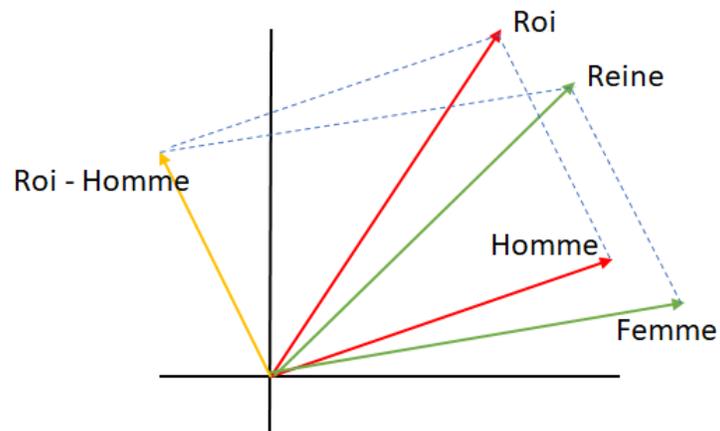
Source : Insee et Enquête Mensuelle de Conjoncture, Banque de France.

Tableau 10. Nombre de documents labellisés pour chaque catégorie d'effet

	Nombre de documents labellisés
Catégorie 1 : pas d'effet	187
Catégorie 2 : baisse de la demande	131
Catégorie 3 : hausse de la production	100
Catégorie 4 : augmentation des stocks	39
Catégorie 5 : incertitude	137
Catégorie 6 : coûts additionnels	121
Total	654

Source : Enquête Mensuelle de Conjoncture, Banque de France.

Graphique 8. Illustration du principe des *wordvectors*



Ce graphique illustre l'idée des *word vectors* : les mots « roi » et « reine » sont proches géométriquement car ils ont une signification similaire, de même que les mots « homme » et « femme ». Par ailleurs, les mots appartenant au même genre sont relativement plus proches (« femme » et « reine »), et les opérations vectorielles donnent des résultats cohérents dans l'espace sémantique : le vecteur « homme » soustrait au vecteur « roi » donne un vecteur proche de celui de « reine ».

Identification des effets du Brexit

Compléments sur les variables explicatives injectées dans les modèles

(1) Stratégies d'agrégation des word vectors

Notre unité d'analyse étant à l'échelle d'un document, il s'agit de définir une stratégie permettant de passer des vecteurs de mots à des variables explicatives construites au niveau des documents. Nous proposons plusieurs variantes pour cette étape d'agrégation, en gardant comme principe de base que le vecteur du document est défini comme la moyenne simple des vecteurs des mots qui le composent. Les différentes versions permettent de filtrer plus ou moins les mots à considérer dans le document, l'idée ici étant de réduire le bruit en ignorant certains mots inutiles à la compréhension du texte.

Le tableau 11 ci-dessous caractérise quatre possibilités²⁸ d'agrégation des *word vectors*. La première version *A*, la plus simple, conserve l'ensemble des termes du document, sans aucune transformation. La version *B* (correspondant à celle retenue pour le modèle et décrite dans le corps du document) ne tient pas compte des mots vides²⁹ dans le calcul du vecteur moyen à l'échelle du document et comprend une étape de lemmatisation³⁰ ; ainsi, les *word vectors* utilisés pour le calcul du vecteur moyen sont ceux des termes lemmatisés. La version *C* ne tient pas non plus compte des mots vides et comprend aussi une étape de lemmatisation, mais seuls les 50 % des termes les plus fréquents dans le corpus sont conservés. Enfin, la version *D* est similaire à la *C* à la différence qu'elle est plus restrictive puisque seuls les 25 % des termes les plus fréquents dans le corpus sont conservés pour le calcul du vecteur à l'échelle du document.

Tableau 11. Stratégies d'agrégation des word vectors à l'échelle des documents

	Suppression des mots vides	Lemmatisation	Termes gardés
Version A	Non	Non	Tous
Version B	Oui	Oui	Tous
Version C	Oui	Oui	Les 50 % les plus fréquents
Version D	Oui	Oui	Les 25 % les plus fréquents

Source : Enquête Mensuelle de Conjoncture, Banque de France.

(2) Variables additionnelles

Des variables additionnelles viennent s'ajouter aux 200 variables précédentes issues des *word vectors*. Elles visent à caractériser la présence de certains termes dans le texte des documents, avec différentes variantes portant à la fois sur :

- la grandeur à mesurer : absence ou présence des mots via des variables indicatrices, ou occurrences des mots via des variables de comptage (en fréquence ou en proportion),

28. De nombreuses autres méthodes d'agrégation des *word vectors* sont imaginables. Il s'agit ici d'arbitrer entre temps passé à construire ces variables, temps de calcul, temps d'analyse... et amélioration du pouvoir prédictif du modèle.

29. Les mots vides, ou *stop words*, désignent les mots non significatifs pour la compréhension du texte : articles, pronoms, prépositions, auxiliaires...

30. La lemmatisation est réalisée à partir du dictionnaire Morphalou 3 produit par le laboratoire ATILF (2019)

- le champ des termes à caractériser : ensemble des mots sans transformation, uniquement des termes décrivant les différents effets du Brexit, uniquement les termes les plus fréquents, avec ou sans lemmatisation, avec ou sans mots vides...

La combinaison de ces différents paramètres nous conduit à proposer 15 jeux de variables explicatives additionnelles, soit, au total en croisant avec les 4 versions des variables *word vectors*, 60 matrices de variables explicatives à tester. Les caractéristiques de l'ensemble des versions des variables additionnelles sont synthétisées dans le tableau 12 page suivante ; un exemple détaillé de la construction de deux d'entre-elles est également présenté page 47.

Dans ce qui suit, les différents jeux de variables explicatives sont identifiés par un code sur deux positions : une lettre allant de *A* à *D* désignant la stratégie d'agrégation utilisée pour la construction des variables *word vectors*, suivie d'un chiffre allant de 1 à 15 caractérisant les variables additionnelles ajoutées. Le modèle *B2* désigne le modèle présenté dans le corps de l'article.

Tableau 12. Caractéristiques des différentes variantes pour la construction des variables additionnelles

Variante	Mots vides	Réduction du lexique	Principe général	Mode de calcul	Précisions
1	conservés	aucune	Une variable pour chaque catégorie d'effet	en proportion du nb de mots du document	Chaque liste l_e est une liste de mots-clés caractéristiques de l'effet e . Pour chaque catégorie d'effet e , la variable X_e indique la présence de mots-clés caractéristiques de e : $X_e = (n \text{ mots de la liste } l_e) / (n \text{ mots du commentaire})$
2	supprimés	lemmatisation	Une variable pour chaque catégorie d'effet	en proportion du nb de mots du document	Chaque liste l'_e est une liste de mots-clés lémmatisés caractéristiques de l'effet e . Pour chaque catégorie d'effet e , la variable X_e indique la présence de mots-clés caractéristiques de e : $X_e = (n \text{ mots de la liste } l'_e) / (n \text{ mots du commentaire})$
3	supprimés	lemmatisation	Matrice documents-termes	fréquences brutes des mots	Matrice documents-termes complète
4	supprimés	racinisation	Matrice documents-termes	fréquences brutes des mots	Matrice documents-termes complète
5	supprimés	lemmatisation	Matrice documents-termes	fréquences brutes des mots	Matrice documents-termes uniquement pour les termes qui apparaissent plus de 10 fois dans le corpus
6	supprimés	racinisation	Matrice documents-termes	fréquences brutes des mots	Matrice documents-termes uniquement pour les termes qui apparaissent plus de 10 fois dans le corpus
7	supprimés	lemmatisation	Matrice documents-termes	fréquences brutes des mots	Matrice documents-termes uniquement pour les termes appartenant à la liste L^*
8	supprimés	racinisation	Matrice documents-termes	fréquences brutes des mots	Matrice documents-termes uniquement pour les termes appartenant à la liste L^*
9	supprimés	lemmatisation	Matrice documents-termes	en proportion du nb de mots du document	Matrice documents-termes uniquement pour les termes appartenant à la liste L^*
10	supprimés	racinisation	Matrice documents-termes	en proportion du nb de mots du document	Matrice documents-termes uniquement pour les termes appartenant à la liste L^*
11	supprimés	lemmatisation	Une variable pour chaque catégorie d'effet	en proportion du nb de mots du document	Chaque liste L'_e est une liste de mots-clés lémmatisés caractéristiques de l'effet e . Pour chaque catégorie d'effet e , la variable X_e indique la présence de mots-clés caractéristiques de e : $X_e = (n \text{ mots de la liste } L'_e) / (n \text{ mots du commentaire})$
12	supprimés	lemmatisation	Matrice documents-termes	fréquences brutes des mots	Matrice documents-termes uniquement pour les termes appartenant à la liste L^* Ajout de toutes les variables d'interactions 2x2
13	supprimés	lemmatisation	Matrice documents-termes	en proportion du nb de mots du document	Matrice documents-termes uniquement pour les termes appartenant à la liste L^* Ajout de toutes les variables d'interactions 2x2
14	supprimés	lemmatisation	Matrice documents-termes	fréquences brutes des mots	Matrice documents-termes uniquement pour les termes appartenant à la liste L^* Ajout de toutes les variables d'interactions 2x2 Les interactions sont discrétisées à 0 (les 2 mots sont absents), 1 (les 2 mots sont présents 1 fois), ou 2 (les 2 mots sont présents avec l'un présent plus d'une fois)
15	supprimés	lemmatisation	Matrice documents-termes	en proportion du nb de mots du document	Matrice documents-termes uniquement pour les termes appartenant à la liste L^* Ajout de toutes les variables d'interactions 2x2 Les interactions sont discrétisées à 0 (les 2 mots sont absents), 1 (les 2 mots sont présents 1 fois), ou 2 (les 2 mots sont présents avec l'un présent plus d'une fois)

Listes de mots-clés utilisées :

- $l_{pas \text{ d'effet}} = \{impact, effet, peu, pas, aucun, concerné, sans, ressenti, ras, effets, concernée, impacté, incidence, impactée\}$
 $l_{baisse \text{ demande}} = \{baisse, export, marché, activité, demande, ventes, baisser, conséquences, difficultés, diminution, ralentissement, retrait, morose, défavorable\}$
 $l_{hausse \text{ production}} = \{commandes, stocks, production, anticipation, demande, hausse, prévision, précaution, soutenue, commande, anticipe, anticiper, augmentation, stockage\}$
 $l_{augmentation \text{ stocks}} = \{stocks, matières, stock, fournisseurs, premières, prévision, finis, produits, pièces, anticiper, frontières, précaution, retards, risques, sécurité, semi, sur, surstockage, anticipation, anticipé\}$
 $l_{incertitude} = \{incertitude, incertitudes, risque, conséquences, douane, inquiétude, inquiétudes, interrogations, inquiète\}$
 $l_{coûts \text{ additionnels}} = \{prix, livre, change, cours, client, clients, sterling, marges\}$
 $l'_{pas \text{ d'effet}} = \{impacter, effet, peu, pas, aucun, concerner, sans, ressentir, ras, incidence, perturber, conséquence\}$
 $l'_{baisse \text{ demande}} = \{baisser, exporter, marché, activité, demander, vendre, conséquence, difficulté, diminuer, ralentir, retrait, morose, défavorable\}$
 $l'_{hausse \text{ production}} = \{commander, stocker, produire, anticiper, demander, hausser, prévoir, précaution, soutenir, augmenter\}$
 $l'_{augmentation \text{ stocks}} = \{stocker, matière, fournir, premier, prévoir, finir, produire, pièce, anticiper, frontière, précaution, retarder, risquer, sécuriser, semi, sur, surstocker\}$
 $l'_{incertitude} = \{incertain, risquer, conséquence, douanier, inquiéter, interroger\}$
 $l'_{coûts \text{ additionnels}} = \{prix, livrer, changer, courir, client, sterling, marger, compétitif, organiser, réorganiser, douanier, gpb\}$
 $L'_{pas \text{ d'effet}} = \{pas, impacter, commander, concerner, client, effet, marché, aucun, société, activité, entreprise, peu\}$
 $L'_{baisse \text{ demande}} = \{impacter, commander, baisser, client, effet, marché, exporter, activité, demander, français, vendre\}$
 $L'_{hausse \text{ production}} = \{commander, stocker, client, effet, anticiper, produire, activité, demander, hausser, important, fort, prévoir\}$
 $L'_{augmentation \text{ stocks}} = \{stocker, anticiper, entreprise, prévoir, matière, premier, peu, fournir, dur, lier, inquiéter, risquer, pièce, toujours, semi, finir, sécuriser, retarder, produire, précaution, importer, frontière, encore, élever, constituer, augmenter\}$
 $L'_{incertitude} = \{pas, impacter, client, incertain, marché, prix, exporter, pouvoir, lier, inquiéter\}$
 $L'_{coûts \text{ additionnels}} = \{impacter, stocker, baisser, client, effet, marché, prix, livrer, hausser, changer, entreprise, courir, matière, premier, suite, marger, lier\}$
 $L^* = \{baisser, commander, marché, exporter, activité, demander, vendre, client, stocker, produire, anticiper, hausser, important, fort, prévoir, matière, premier, fournir, peu, dur, risquer, inquiéter, marger, finir, livrer, pièce, semi, concerner, retarder, impacter, effet, incertain, aucun, précaution, prix, changer, importer, frontière, constituer\}$

Exemples détaillés de la construction de 2 jeux de variables additionnelles

Une première solution pour l’ajout de variables additionnelles consiste à définir, avec une expertise métier et pour chaque catégorie e d’effet à modéliser, une liste de mots-clés caractéristiques lemmatisés l_e^l . Après suppression des mots vides et lemmatisation du lexique, pour chaque catégorie e et document i est ensuite créée une variable W_{e_i} telle que :

$$W_{e_i} = \frac{\text{Nombre de mots du document } i \text{ appartenant à } l_e^l}{\text{Nombre total de mots dans le document } i} \quad (7)$$

Cette stratégie correspond à la variante n°2 dans le tableau 12 et à celle utilisée pour la construction du jeu de variables $B2$, celui du modèle présenté dans le corps de l’article.

Dans le cas du Brexit, voici un extrait des listes l_e^l pour chaque catégorie :

- $l_{pas\ d'effet}^l$: {*impacter, effet, peu, pas, aucun, concerner, sans, ressentir, ras, incidence, perturber, conséquence* }
- $l_{baisse\ demande}^l$: {*baisser, exporter, marché, activité, demander, vendre, conséquence, difficulté, diminuer, ralentir, retrait, morose, défavorable* }
- $l_{hausse\ production}^l$: {*commander, stocker, produire, anticiper, demander, hausser, prévoir, précaution, soutenir, augmenter* }
- $l_{augmentation\ stocks}^l$: {*stocker, matière, fournir, premier, prévoir, finir, produire, pièce, anticiper, frontière, précaution, retarder, risquer, sécuriser, semi, sur, surstocker* }
- $l_{incertitude}^l$: {*incertain, risquer, conséquence, douanier, inquiéter, interroger* }
- $l_{coûts\ additionnels}^l$: {*prix, livrer, changer, courir, client, sterling, mager, compétitif, organiser, réorganiser, douanier, gpb* }

Dans ce cas, chaque document est *in fine* caractérisé par 200 variables *word vectors* + 6 variables additionnelles soit 206 variables explicatives au total.

Une autre possibilité (correspondant à la variante n°3 dans le tableau 12) consiste à construire une matrice documents-termes après avoir supprimé les mots vides et lemmatisé le vocabulaire. Alors que dans l’exemple précédent on créait 6 variables, une matrice documents-termes créée autant de variables additionnelles que de termes distincts dans le corpus de document. Par ailleurs, alors que précédemment toutes les variables additionnelles étaient comprises entre 0 et 1 (normalisées par le nombre de mots du documents), ici pour chaque document i , la valeur de la variable pour le terme t est égale au nombre d’occurrences de t dans i : c’est donc une matrice creuse, dont les coefficients valent au maximum le nombre de mots du document. Dans notre cas, cela nous conduit à ajouter 2 175 variables supplémentaires ayant des valeurs comprises entre 0 et 86 ; cette matrice documents-termes a 99 % de coefficients nuls. Notons que dans ce cas la construction de ces variables ne nécessite ni expertise métier ni compréhension fine des termes utilisés pour désigner chaque catégorie d’effet.

Résultats pour l'ensemble des jeux de variables explicatives testés

Le tableau 13 indique, pour chaque catégorie d'effet, le jeu de variables explicatives et l'AUC du modèle le plus performant. Ce dernier varie entre 0,69 et 0,98 selon l'effet à caractériser et les variables explicatives injectées. Le premier constat est que certaines catégories semblent plus faciles à identifier que d'autres : alors que pour la prédiction de l'effet «augmentation des stocks» (catégorie 4) les AUC peuvent être très élevés avec un maximum à 0,98, la meilleure performance atteinte pour l'identification d'effets monétaires négatifs est inférieure de 0,20 avec un AUC maximum à 0,78.

Le deuxième constat est que, selon l'effet à prédire, ce n'est pas toujours le même jeu de variables explicatives qui donne la meilleure performance. Selon le seul critère de l'AUC, il faudrait donc utiliser le jeu de variables explicatives *C4* pour prédire la catégorie des « sans effet » tandis qu'il faudrait utiliser le jeu de variables *D6* pour prédire la catégorie d'effet lié à l'incertitude. Remarquons tout de même que les stratégies d'agrégation des *word vectors* sans filtrer sur les termes les plus fréquents (stratégies *A* et *B*) font partie des jeux de variables explicatives permettant d'obtenir les meilleures performances dans 4 cas sur 6 ; concernant les variables additionnelles, elles sont construites dans 3 cas sur 6 à partir de matrices documents-termes³¹ (stratégies numérotées 4 et 6) et dans 3 cas sur 6 à partir de listes de mots-clés³² (stratégies numérotées 2, 11 et 14).

Tableau 13. Performances maximales pour prédire chacune des catégories d'effets

	AUC maximal	Jeu de variables explicatives correspondant
Catégorie 1 : pas d'effet	0,93	<i>C4</i>
Catégorie 2 : baisse de la demande	0,85	<i>B2</i>
Catégorie 3 : hausse de la production	0,93	<i>A4</i>
Catégorie 4 : augmentation des stocks	0,98	<i>A14</i>
Catégorie 5 : incertitude	0,85	<i>D6</i>
Catégorie 5 : coûts additionnels	0,78	<i>B11</i>

Source : Enquête Mensuelle de Conjoncture, Banque de France.

Le fait que ce ne soit pas toujours le même jeu de variables explicatives qui donne les meilleures performances est probablement lié à la définition même des effets. Ces derniers sont décrits avec un champ lexical plus ou moins vaste, des termes plus ou moins discriminants : en fonction, ils seront relativement plus facilement identifiables par des mots-clés ciblés, ou, au contraire, par une matrice document-termes complète.

Un modèle unique, performant et parcimonieux

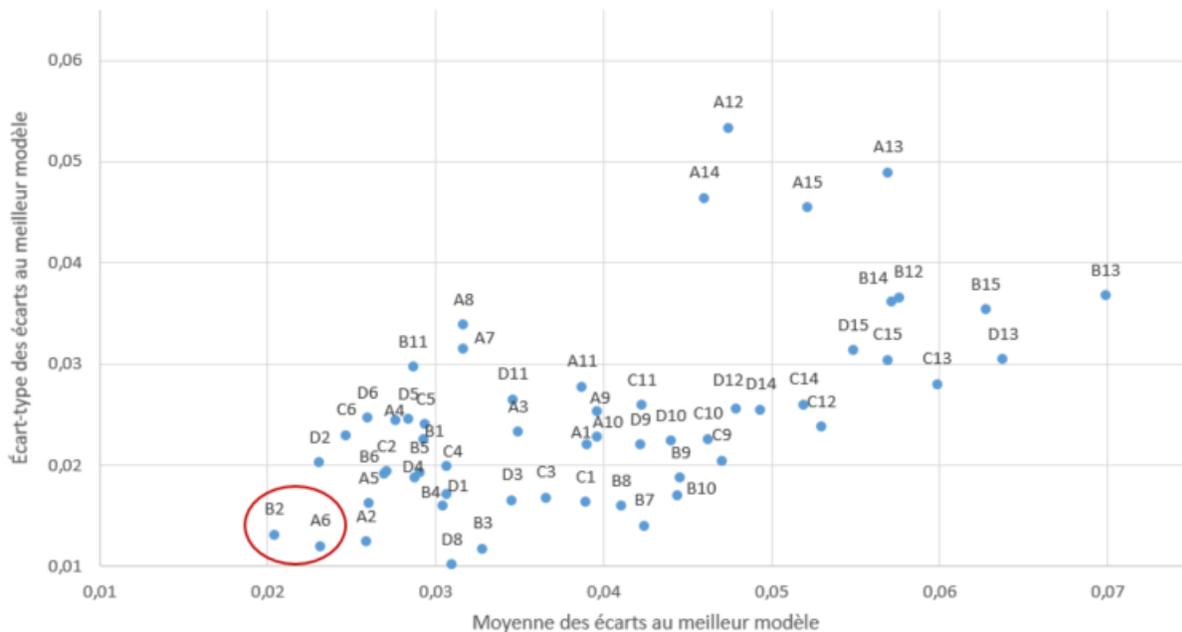
Pour faciliter la comparaison des modèles d'un effet à l'autre, nous choisissons de retenir un unique jeu de variables explicatives pour prédire les 6 catégories d'effets. Ce modèle doit à la fois dégrader le moins possible la performance par rapport à celle du meilleur modèle de chaque effet, tout en étant le plus parcimonieux possible. Pour cela, l'idée est de minimiser conjointement la moyenne

31. Une matrice documents-termes décrit la fréquence d'un ensemble de mots (en colonnes) dans une collection de documents (en lignes).

32. Ces listes de mots-clés, définies par les auteurs, sont caractéristiques des différents effets de la typologie des effets (voir l'annexe page 47 pour un exemple).

des écarts de performance au meilleur modèle ainsi que l'écart-type de ces écarts, ce qui revient à rechercher le point le plus proche de l'origine sur le graphique 9.

Graphique 9. Moyennes et écart-types des écarts de performance au meilleur modèle



Source : Enquête Mensuelle de Conjoncture, Banque de France.

Aucun modèle ne se distingue à la fois du point de vue de la moyenne et de l'écart-type. Les deux modèles *B2* et *A6* sont les meilleurs candidats, avec des écarts au meilleur modèle faibles (proches de 0,02 point en moyenne), et ce quel que soit l'effet à prédire (écarts-types des écarts proches de 0,01). Ces deux modèles ont les caractéristiques décrites dans le tableau 14 ci-dessous, à comparer avec les caractéristiques des modèles donnant, pour chaque effet, la performance maximale (tableau 13).

Tableau 14. AUC et nombre de variables sélectionnées pour les modèles candidats à un modèle unique

Modèle :	AUC			Nombre de variables sélectionnées		
	Meilleur modèle	B2	A6	Meilleur modèle	B2	A6
Catégorie 1 : pas d'effet	0,93	0,92	0,93	131	41	75
Catégorie 2 : baisse de la demande	0,85	0,85	0,82	39	40	114
Catégorie 3 : hausse de la production	0,93	0,90	0,92	142	28	90
Catégorie 4 : augmentation des stocks	0,98	0,94	0,95	34	23	25
Catégorie 5 : incertitude	0,85	0,83	0,84	75	28	73
Catégorie 6 : coûts additionnels	0,78	0,75	0,74	47	41	58
Moyenne	0,89	0,87	0,86	78	33	72
Écart-type	0,07	0,07	0,08	48	8	30

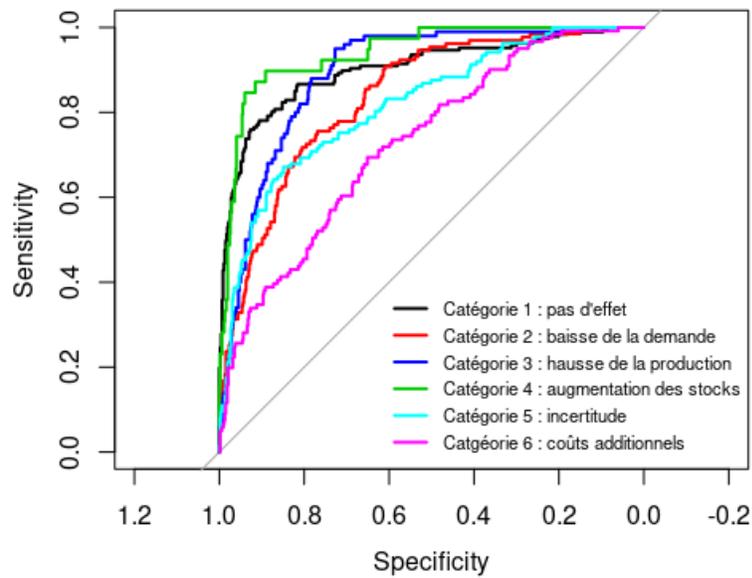
Source : Enquête Mensuelle de Conjoncture, Banque de France.

À la lecture du tableau 14, nous constatons que les modèles *B2* et *A6* ont des performances relativement homogènes en terme d'AUC. En revanche, ils sont plus hétérogènes du point de vue du nombre

de variables sélectionnées, le modèle *A6* ayant en moyenne plus de deux fois plus de variables que le modèle *B2*. À performance proche, nous préférons privilégier un modèle parcimonieux pour éviter tout risque de sur-apprentissage et retenons donc le modèle *B2* qui semble le meilleur compromis avec (i) des bonnes performances par rapport aux meilleurs AUC, (ii) des performances homogènes d'une catégorie d'effet à l'autre, et (iii) un nombre de variables limité. Les courbes ROC pour ce modèle sont représentées sur le graphique 10.

Ce modèle est donc retenu comme le modèle unique pour prédire les probabilités d'appartenir respectivement à chacune des six catégories et correspond à celui présenté dans le corps de l'article.

Graphique 10. Courbes ROC du modèle *B2* pour chaque catégorie d'effet



Source : Enquête Mensuelle de Conjoncture, Banque de France.

Tableau 15. Résultats détaillés des estimations pour les modèles de prédiction des différents effets du Brexit

	Catégorie 1	Catégorie 2	Catégorie 3	Catégorie 4	Catégorie 5	Catégorie 6
	Pas d'effet	Baisse demande	Hausse produc ^o	Augmenta ^o stocks	Incertitude	Coûts additionnels
<i>Cste</i>	-2,12	-1,84	-3,09	-3,93	-1,27	-1,53
X1	0,00	0,00	0,18	0,00	0,00	0,00
X2	0,00	0,00	0,00	0,00	-0,22	0,00
X3	2,07	0,00	0,00	0,00	0,00	0,00
X5	0,20	0,00	0,00	0,00	0,00	0,00
X7	0,00	-0,24	0,00	0,00	0,00	0,00
X9	0,00	0,00	0,10	0,00	0,00	-0,28
X10	0,00	0,00	-0,16	0,00	0,00	0,00
X11	0,00	0,00	0,00	0,58	0,00	0,00
X14	0,00	0,00	0,00	0,00	0,00	-0,18
X22	0,00	-1,57	0,00	0,00	0,07	0,00
X23	0,00	0,00	-0,15	0,00	0,00	0,00
X25	0,00	0,00	0,00	0,00	0,00	-0,02
X26	0,00	0,00	0,00	0,00	0,00	-0,47
X27	1,17	0,00	-0,43	0,00	0,00	-0,44
X31	0,79	0,00	-0,24	-0,42	0,00	0,00
X32	0,00	0,08	0,00	0,00	0,00	0,00
X33	0,00	0,00	-0,33	0,00	0,00	0,00
X34	-0,03	0,00	0,00	0,00	0,00	0,00
X36	0,00	0,00	-0,80	-1,22	0,00	0,00
X37	-0,16	0,00	1,23	-0,43	0,00	0,00
X40	0,18	0,00	-0,02	0,00	0,00	0,00
X41	0,00	0,00	0,00	0,00	-0,05	0,00
X43	0,00	0,00	-1,11	0,00	0,00	0,00
X44	0,00	0,00	1,31	0,00	0,00	0,00
X47	-0,25	0,00	0,00	0,00	0,32	0,00
X48	-0,11	0,00	0,00	0,00	0,00	0,00
X49	0,83	0,00	-1,45	0,00	0,00	0,00
X50	-0,03	0,00	0,00	0,00	0,00	0,00
X51	0,00	0,00	0,43	0,00	0,00	0,00
X52	-0,14	0,00	0,00	0,00	0,00	0,00
X54	0,00	0,00	0,69	0,00	0,00	0,00
X55	0,00	0,00	0,62	0,00	0,00	0,00
X57	0,00	0,00	0,00	0,98	0,00	0,16
X59	0,00	0,00	-0,48	0,00	0,00	-0,41
X64	0,00	0,00	0,00	0,00	0,46	0,00
X65	0,00	0,00	0,41	0,00	0,00	0,00
X66	0,00	0,00	0,00	0,00	0,00	0,38
X69	0,39	-0,17	0,00	-0,20	0,00	0,00
X70	0,00	0,00	-0,07	0,00	0,00	0,00
X71	0,00	0,00	-0,66	0,00	0,00	0,00
X72	0,00	0,00	0,00	0,00	0,00	-0,38
X73	0,52	0,00	0,00	0,00	0,00	0,00
X74	0,74	-0,05	0,00	0,00	0,00	0,00
X77	1,23	-0,64	0,00	0,00	0,00	0,00
X78	0,00	0,00	-0,33	0,00	0,00	0,00
X79	0,00	0,00	0,82	0,00	0,00	0,00
X80	0,00	0,00	1,59	0,00	0,00	-1,06
X82	0,08	0,00	0,00	0,00	0,00	-0,03
X84	0,00	0,00	-0,24	0,00	0,00	0,00
X85	0,00	0,00	0,00	0,16	0,00	0,00
X86	0,30	0,00	0,00	0,30	0,00	0,00
X87	-1,09	0,00	0,25	0,00	0,00	0,00
X89	0,00	0,00	-0,56	0,00	0,00	0,00
X91	0,00	0,00	0,29	0,00	0,00	0,00
X92	0,00	-0,04	0,00	0,00	0,00	0,00
X93	-0,14	1,18	-1,72	0,00	0,00	0,26
X95	0,00	0,81	0,00	0,00	0,00	0,00
X96	0,00	0,00	-0,07	0,00	0,00	0,00

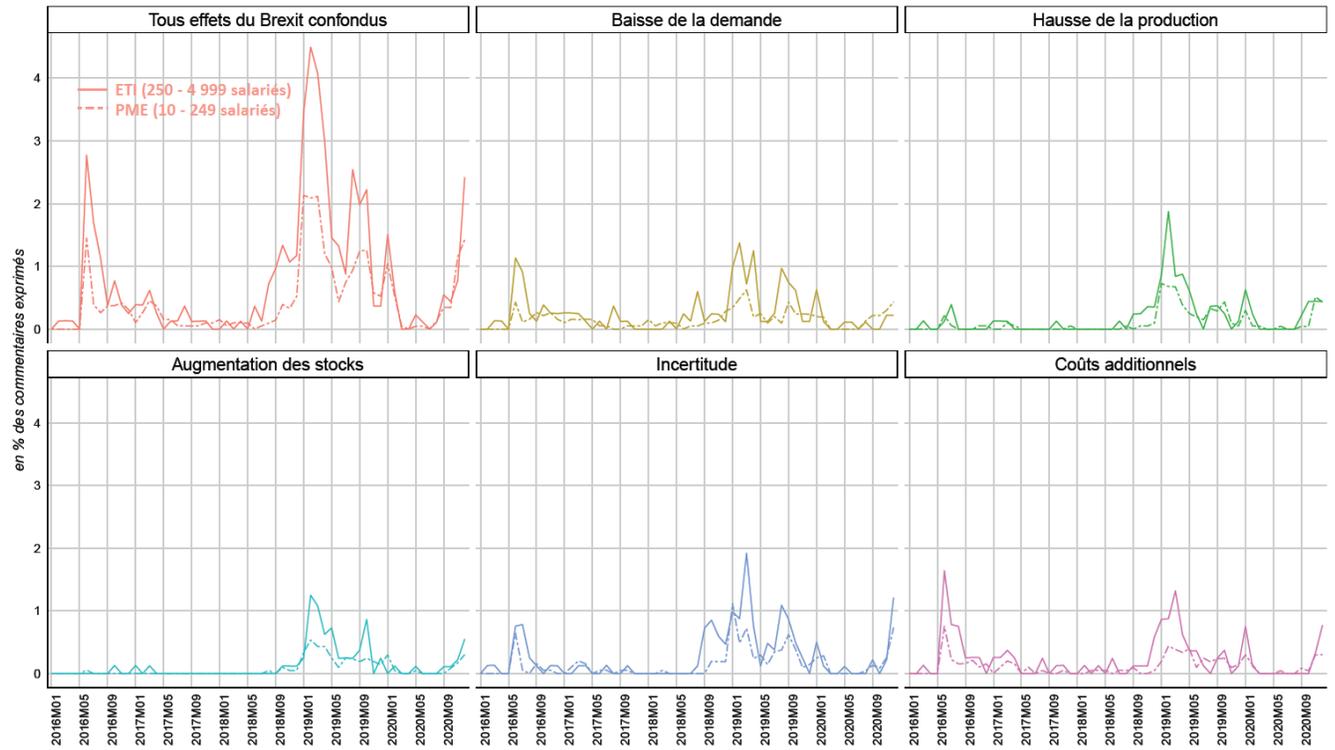
X98	0,00	0,00	-0,46	0,00	0,00	0,00
X99	0,00	0,00	0,00	-0,50	0,00	0,00
X100	0,19	0,00	-0,52	0,00	-0,31	0,00
X103	0,32	0,00	0,00	0,00	-0,01	0,43
X104	0,00	0,00	0,00	0,00	0,00	0,11
X106	0,00	-0,71	0,00	0,00	0,00	0,00
X107	0,00	-0,14	-0,89	0,19	0,00	0,00
X109	0,00	0,00	0,00	-1,17	0,00	0,00
X110	0,00	0,00	0,05	0,00	0,00	0,00
X111	0,00	0,68	0,00	-0,25	0,00	0,00
X114	0,22	0,00	0,00	0,00	0,00	0,00
X115	0,00	-0,03	0,77	0,00	0,00	0,00
X116	-0,82	0,00	0,51	0,00	0,00	0,00
X117	0,00	0,00	-0,33	0,00	0,00	-0,68
X118	0,00	0,00	0,02	0,00	0,00	0,00
X119	0,24	0,00	0,00	0,00	0,00	0,00
X120	0,00	0,00	0,00	0,00	0,00	0,03
X122	0,00	0,00	-0,23	0,77	0,00	0,00
X123	0,00	0,00	0,47	0,00	0,00	0,00
X124	0,00	0,00	-0,60	0,00	0,00	0,00
X127	0,00	0,00	1,25	0,00	0,00	0,00
X128	0,12	0,00	0,00	0,00	0,00	0,00
X130	0,78	0,00	0,00	0,00	0,00	0,00
X131	0,00	0,00	-1,59	-0,20	0,00	0,00
X132	0,13	0,00	-0,11	0,00	0,00	0,00
X133	0,00	0,00	1,49	0,00	0,00	0,00
X134	0,00	0,41	-0,35	0,00	0,00	0,00
X135	-0,16	0,00	0,00	0,00	0,00	0,00
X138	0,00	0,00	0,00	-2,15	0,00	0,00
X140	0,00	0,00	-0,85	0,33	0,00	0,00
X142	0,00	0,00	0,00	0,00	0,00	0,24
X146	-0,06	0,00	0,00	0,38	0,00	0,00
X148	-0,97	0,00	0,00	0,00	0,00	0,00
X149	0,00	0,00	0,98	0,00	0,00	0,00
X150	0,00	0,00	-0,67	0,00	0,00	0,00
X151	0,00	0,00	0,00	0,00	0,00	-0,02
X152	0,80	0,00	0,00	0,00	0,00	0,00
X153	0,00	0,47	-0,36	0,00	0,00	0,00
X155	0,00	1,37	0,00	0,00	0,00	-0,11
X157	-0,68	-0,50	0,00	0,00	0,55	0,00
X160	0,00	0,00	0,00	0,00	0,00	0,23
X163	0,00	0,00	0,00	-0,27	0,00	0,00
X165	-0,46	0,00	0,35	0,00	0,00	0,00
X166	0,42	0,00	0,00	0,00	0,00	0,00
X168	0,00	0,00	0,00	0,00	0,00	0,01
X169	0,00	0,00	0,20	0,00	0,00	0,00
X170	0,36	0,00	0,00	-0,21	0,00	0,00
X172	0,00	0,00	0,00	0,25	0,00	-1,04
X173	0,00	0,00	0,00	0,00	0,00	-0,28
X175	-0,90	0,00	0,00	0,07	0,00	0,00
X177	0,00	0,00	0,00	0,99	0,00	0,00
X179	0,97	-1,33	0,35	0,00	0,00	-0,60
X180	0,00	-0,13	0,00	0,00	0,00	0,00
X181	0,00	0,00	0,00	0,00	-0,27	0,00
X183	-0,75	0,75	0,42	0,00	0,00	-0,91
X186	0,00	0,00	0,00	-0,85	0,00	0,00
X188	-0,15	0,00	0,00	0,92	0,00	0,00
X191	0,00	0,00	0,68	0,00	0,00	0,00
X192	0,00	0,00	0,02	-0,15	0,00	0,00
X193	-0,09	0,00	-1,17	0,00	0,00	0,00
X197	0,08	0,00	0,00	0,00	0,00	0,00
X198	0,00	0,00	-0,78	0,00	0,00	0,00
X199	0,00	0,15	0,00	0,00	0,00	0,00
X200	0,00	0,00	0,00	-0,28	0,00	0,00
W1	12,85	-1,19	-2,57	0,00	-2,01	-1,62

W2	-1,34	7,11	-0,91	0,00	0,00	0,00
W3	-0,05	0,00	15,51	0,00	-6,09	0,00
W4	0,00	0,00	-5,78	13,05	0,00	0,00
W5	-12,78	-1,62	0,00	0,00	16,38	-0,08
W6	0,00	0,00	0,00	-2,41	0,00	6,80
N var	47	24	61	28	12	28
λ	0,00979	0,0213	0,00717	0,00859	0,0248	0,0207
AUC	0,92	0,86	0,91	0,96	0,85	0,75
N obs	654	654	654	654	654	654

Note : Les variables X1 à X200 sont les variables construites à partir des *wordvectors*. Les variables X absentes du tableau sont celles non sélectionnées (= coefficient nul) pour les 6 modèles. Les variables W1 à W6 sont les variables caractérisant la présence de certains termes caractéristiques de chaque effet.

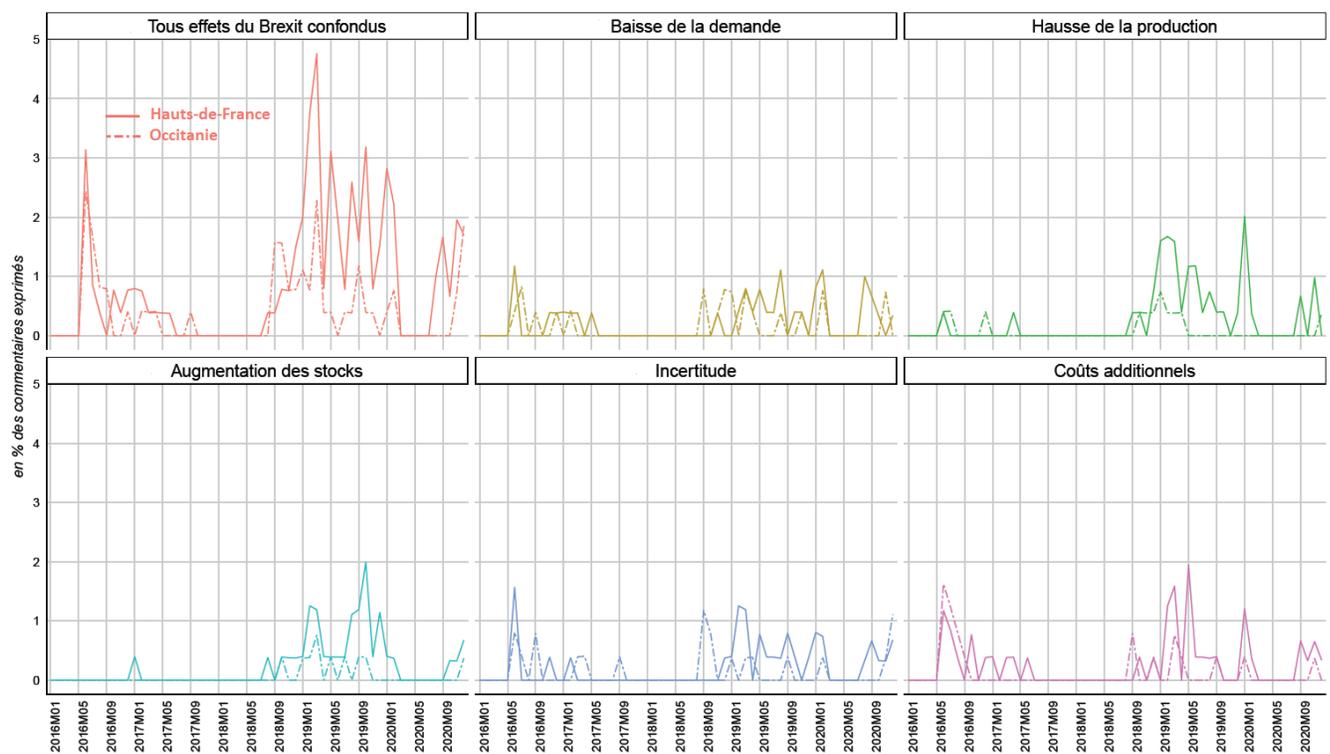
Source : Enquête Mensuelle de Conjoncture, Banque de France.

Graphique 11. Évolution des effets du Brexit au cours du temps : comparaison entre ETI et PME



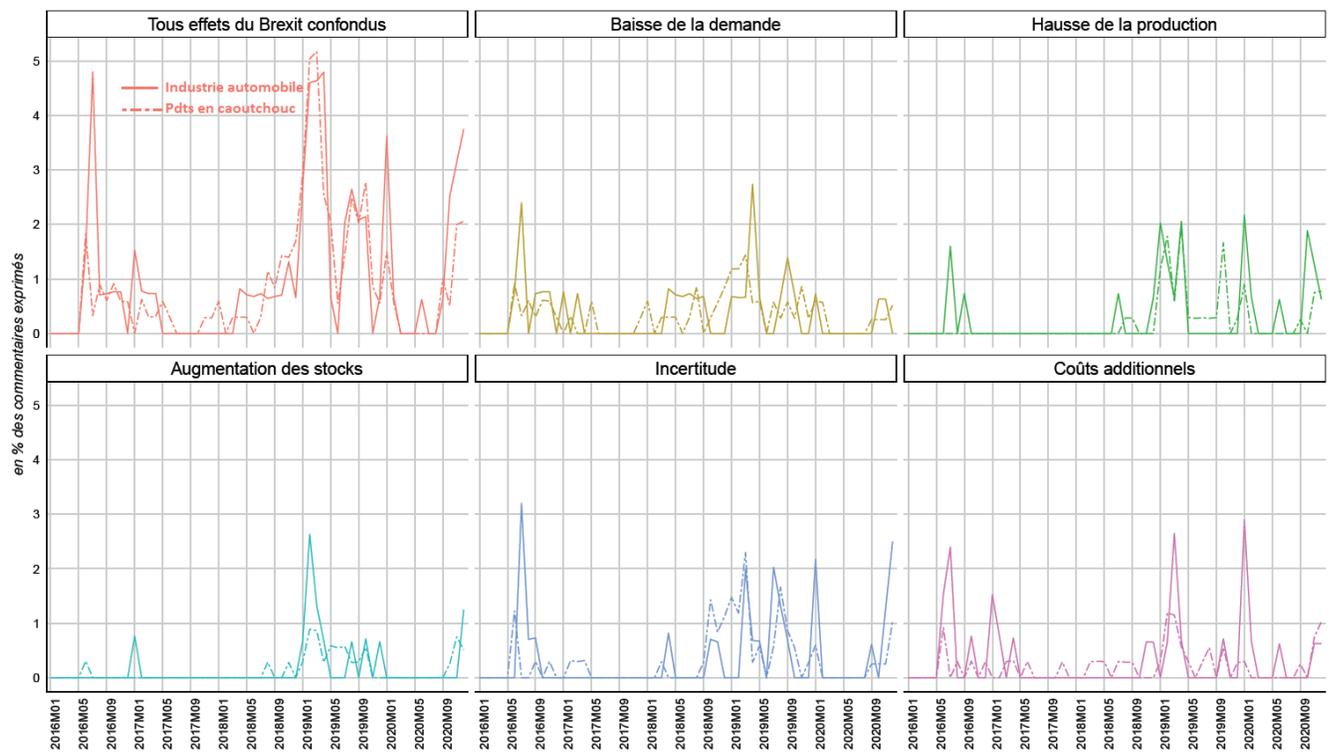
Source : Enquête Mensuelle de Conjoncture, Banque de France.

Graphique 12. Évolution des effets du Brexit au cours du temps : comparaison entre Hauts-de-France et Occitanie



Source : Enquête Mensuelle de Conjoncture, Banque de France.

Graphique 13. Évolution des effets du Brexit au cours du temps : comparaison entre industrie automobile et fabrication de produits en caoutchouc et plastique



Source : Enquête Mensuelle de Conjoncture, Banque de France.