

Un an de crise Covid : comment évaluer l'impact de la pandémie sur l'activité économique des entreprises françaises ? Construction de contrefactuels individuels et diagnostics de l'année 2020

One Year of COVID: What Impact Did the Pandemic Have on the Economic Activity of French Companies? Construction of Individual Counterfactuals and Diagnoses for 2020

Benjamin Bureau*, **Anne Duquerroy***, **Julien Giorgi****, **Mathias Lé***, **Suzanne Scott**** et **Frédéric Vinas***

Résumé – Nous étudions l'impact de la crise sanitaire sur l'activité de plus de 645 000 entreprises françaises, à partir de données individuelles permettant d'estimer leur chiffre d'affaires mensuellement. Notre modèle de microsimulation est triplement innovant. Premièrement, nous quantifions la perte d'activité par rapport à une situation contrefactuelle sans crise afin de repenser les conséquences de la crise en tenant compte des trajectoires de développement que suivaient les entreprises avant la pandémie. Deuxièmement, nous estimons ce choc au niveau individuel, afin d'étudier l'hétérogénéité des pertes d'activité. Nous mettons en lumière la dispersion du choc à la fois entre secteurs et au sein des secteurs. Le secteur explique jusqu'à 48 % de la variance des chocs d'activité mensuels observés en 2020, une part beaucoup plus importante que lors d'une année normale. Enfin, nous identifions quatre profils de trajectoires d'activité en 2020. Le secteur est le principal déterminant d'appartenance à ces profils. Conditionnellement au secteur, ces profils-types sont également corrélés à l'adaptation organisationnelle des entreprises.

Abstract – *We study the impact of the health crisis on the activity of more than 645,000 French companies using individual data to estimate their monthly turnover. Our microsimulation model is innovative in three ways. First, we quantify the loss of activity compared with a non-crisis counterfactual situation to rethink the consequences of the crisis taking into account companies' growth trajectories before the pandemic. Second, we estimate this shock at the individual level to study the heterogeneity of loss of business. We highlight the disparities of the shock both between and within sectors. The sector explains up to 48% of the variance of the monthly activity shocks observed in 2020, a much larger proportion than in a normal year. Finally, we identify four profiles of business trajectories in 2020. The industry is the primary determinant of belonging to these profiles. Conditionally to the sector, these profiles also correlate with the organisational adaptation of companies.*

Codes JEL / JEL Classification : D22, G38, H32

Mots-clés : Covid-19, activité économique, sociétés non financières, microsimulation

Keywords: COVID-19, economic activity, non-financial corporations, microsimulation

*Banque de France ; **Insee. Correspondance : julien.giorgi@insee.fr

Les auteurs remercient Jean Boissinot, Nicolas Carnot, Gilbert Cette, Olivier Gamier, Alain Gerbier, Dominique Goux, Loriane Py, Sébastien Roux et Jean-Pierre Villetelle pour les nombreux échanges et conseils qui ont nourri l'analyse présentée ici, ainsi que les participants aux réunions techniques du Comité de suivi et d'évaluation des mesures de soutien financier aux entreprises confrontées à l'épidémie de Covid-19, présidé par Benoit Cœuré, dont les retours et les exposés ont permis d'améliorer sensiblement notre analyse. Nous remercions aussi Tanguy Aubert, Julien Demuynck, Olivier Hubert et Martial Ranvier pour leur aide.

Reçu en octobre 2021, accepté en avril 2022.

Les jugements et opinions exprimés par les auteurs n'engagent qu'eux-mêmes, et non les institutions auxquelles ils appartiennent, ni a fortiori l'Insee.

Citation: Bureau, B., Duquerroy, A., Giorgi, J., Lé, M., Scott, S. & Vinas, F. (2022). One Year of COVID: What Impact Did the Pandemic Have on the Economic Activity of French Companies? Construction of Individual Counterfactuals and Diagnoses for 2020. *Economie et Statistique / Economics and Statistics*, 532-33, 3–23. doi: 10.24187/ecostat.2022.532.2069

Au début de l'année 2020, la pandémie de Covid-19 et les mesures de restrictions sanitaires ont provoqué un choc économique d'une ampleur inédite depuis la Grande Récession (Bartik *et al.*, 2020 ; Forsythe *et al.*, 2020), affectant à la fois l'offre et la demande de biens et services (Baldwin & Di Mauro, 2020). En France, ces mesures, concrétisées par plusieurs périodes de confinement et déconfinement progressif, et par un ensemble de mesures évolutives comme les couvre-feux et les restrictions de déplacement, ont provoqué un ralentissement brutal de l'activité économique. Entre 2019 et 2020 le PIB français a baissé de 7.9 % et le revenu national de 6.3 % (Amoureux *et al.*, 2021). Au mois d'avril 2020, la chute de la valeur ajoutée a dépassé les 30 %, plaçant la France parmi les pays les plus touchés de la zone euro (Heyer & Timbeau, 2020). Sur l'ensemble de l'année 2020, la valeur ajoutée des entreprises françaises a connu une baisse de 8.1 % et de 8.3 % pour les seules sociétés non financières (SNF).

Au-delà de ces constats macroéconomiques, nous cherchons à évaluer plus précisément l'impact de la crise sanitaire sur l'activité des entreprises. Cela consiste à estimer l'écart entre les niveaux d'activité observés pendant la crise et les niveaux qui auraient été observés si la crise n'avait pas eu lieu. Cette approche par le « contrefactuel » est au fondement des méthodes classiques de microéconométrie d'évaluation des politiques publiques¹. Avec la pandémie de Covid-19, l'estimation de ces montants contrefactuels d'activité pose des problèmes méthodologiques inédits. En effet, la pandémie a affecté l'intégralité des entreprises, rendant caduque les estimations reposant sur la constitution de groupes de contrôle. De surcroît, même si la pandémie a affecté l'ensemble des entreprises françaises, ses conséquences ont pu être extrêmement hétérogènes et dépendre d'une multitude de facteurs complexes, pouvant avoir des effets différents ou être inobservables. En conséquence, les exercices de modélisation de l'activité des entreprises sur cette période se sont révélés excessivement périlleux ou simplificateurs. Beaucoup de travaux ont ainsi approché l'impact de la crise en utilisant les taux de variations observés de leur activité entre 2019 et 2020 (Hadjibeyli *et al.*, 2021 ; Bourlès & Nicolas, 2021), biaisant nécessairement les estimations de l'ampleur des chocs d'activité. D'autres approches, plus structurelles, ont modélisé *ex ante* différents scénarii d'évolution de la pandémie et des mesures de restriction sanitaire afin de pouvoir estimer l'ampleur du

choc économique (Schivardi & Guido, 2020 ; Gourinchas *et al.*, 2021 ; OECD, 2020)². Ces travaux reposent sur des hypothèses théoriques importantes, dont la pertinence souffre d'un manque de vérification *a posteriori*, dans un contexte inédit où les hypothèses académiques éprouvées peuvent rencontrer des limites. De surcroît, des études sur données américaines montrent que les comportements d'auto-isollement n'ont pas toujours suivi le calendrier des restrictions sanitaires (Glaeser *et al.*, 2021 ; Gupta *et al.*, 2021 ; Sears *et al.*, 2020) et que les décisions de réouvertures des commerces n'ont pas toujours coïncidé avec la levée des restrictions (Balla-Elliott *et al.*, 2020) limitant la pertinence du recours au calendrier de ces dernières dans les modélisations d'activité. Une autre piste exploitée a été l'utilisation des données d'enquêtes (Bloom *et al.*, 2021 ; Bignon & Garnier, 2020), pouvant cependant rencontrer de faibles taux de couverture ou des risques de biais déclaratifs.

Cet article aspire à dépasser ces limites en proposant une méthode novatrice d'évaluation de l'impact de la pandémie de Covid-19 s'appuyant sur un ensemble limité d'hypothèses. L'analyse repose sur un positionnement a-théorique visant à modéliser le développement de l'activité de l'ensemble des entreprises françaises si leur dynamique d'activité n'avait pas été altérée par le début de la crise. Les dynamiques individuelles d'activité mensuelle postérieures au mois de février 2020 sont prédites à partir de mécanismes autorégressifs avant d'être comparées aux montants observés *ex post*, leur écart permettant d'approcher individuellement l'impact de la pandémie sur l'activité. Ces prévisions sont effectuées au niveau des entreprises et ne reposent pas sur une application uniforme des chocs sectoriels. Ce travail se distingue sur ce point de ceux appliquant à des données individuelles des chocs estimés entièrement ou partiellement au niveau sectoriel, ce qui en limite l'hétérogénéité (Anayi *et al.*, 2020 ; Blanco *et al.*, 2020 ; Hadjibeyli *et al.*, 2021).

La dimension sectorielle a toutefois bien été importante dans la crise, tous les secteurs n'ayant pas été affectés avec la même intensité (Danieli & Olmstead-Rumsey, 2020 ; Brinca *et al.*, 2020). En France, des différences ont été observées en fonction de l'intensité sectorielle

1. Voir par exemple Angrist & Pischke (2008).

2. La plupart de ces études ont mobilisé les estimations de l'impact de l'activité dans des modèles financiers d'évaluation de la liquidité ou du risque de défaut des entreprises.

des mesures de restriction (Baleyte *et al.*, 2021 ; Dauvin & Sampognaro, 2021)^{3,4}, de la dépendance de certains secteurs au tourisme (Škare *et al.*, 2021) ou aux chaînes de valeur internationales (Gerschel *et al.*, 2020 ; Baldwin & Tomiura, 2020). De même, la dégradation inédite des anticipations du fait de la crise (Insee, 2020) a pu contribuer à une augmentation de l'épargne de précaution des ménages et à un recentrage de leur consommation sur les biens de première nécessité (Bonnet *et al.*, 2021). Pour autant, le secteur ne semble pas expliquer complètement la diversité des situations vécues par les entreprises puisque même au sein d'un secteur donné, le degré de dépendance à l'égard des marchés étrangers (Brancati & Brancati, 2020) et les effets des mesures de distanciation sociale (Blanchard *et al.*, 2020) ont eu des effets différenciés, conduisant parfois à une réallocation de l'activité et de l'emploi entre entreprises « gagnantes » et « perdantes » (Barrero *et al.*, 2020 ; 2021). La méthode développée vise donc à mesurer l'hétérogénéité des chocs d'activité individuels, éventuellement au sein d'un même secteur et à proposer une quantification de la contribution du secteur à la variabilité des situations individuelles observées en 2020.

L'utilisation de données infra-annuelles permet d'évaluer l'impact de la crise aussi bien annuellement que mois par mois. L'exploitation des séries mensuelles de chocs d'activité permet de compléter l'analyse de l'hétérogénéité « en coupe » par une analyse dynamique de la diversité des trajectoires d'activité au fil de la pandémie. L'effet de court terme de la pandémie sur l'emploi (Barrero *et al.*, 2020), les fermetures d'entreprises (Gourinchas *et al.*, 2020) et l'activité (Fairlie, 2020 ; Bloom *et al.*, 2021) a régulièrement été mis en avant, mais certains travaux soulignent également l'inégale persistance des chocs initiaux sur l'activité (Bloom *et al.*, 2021) ou l'emploi (Chetty *et al.*, 2020 ; Cajner *et al.*, 2020). Un objectif de l'article est donc de caractériser la diversité des trajectoires d'activité des entreprises françaises au cours de l'année 2020 afin de pouvoir en proposer une typologie. La compréhension de cette typologie et du rôle joué par le secteur ou d'autres caractéristiques démographiques ou organisationnelles des entreprises visera à affiner la compréhension de l'hétérogénéité de l'impact de la crise sanitaire sur l'activité des entreprises françaises.

La suite de l'article présente les données utilisées (section 1), le détail des méthodes d'évaluation mises en place (section 2) et les principaux résultats relatifs à l'étude de l'économie française

sur l'année 2020 (section 3). Ces résultats sont ensuite discutés en conclusion.

1. Données et construction de l'échantillon

1.1. Construction de la base de données

La mesure d'activité retenue est le chiffre d'affaires, offrant une mesure brute de l'activité économique dont l'estimation est peu affectée par des hypothèses de reconstruction. Il permet d'approcher l'impact de la crise sur l'activité, indépendamment des ajustements effectués ensuite dans la gestion financière et opérationnelle des entreprises et des mesures publiques de soutien. Les données mobilisées sont issues des déclarations mensuelles faites par les entreprises à la Direction générale des finances publiques (DGFiP) dans le cadre du paiement de la taxe sur la valeur ajoutée (TVA). Elles permettent de reconstituer le chiffre d'affaires de chaque entreprise, en sommant l'ensemble de ses opérations, imposables ou non, sur le territoire français ou à l'étranger (annexe A1). Les secteurs financier et de l'administration publique, ainsi que les travailleurs indépendants et les entreprises individuelles sont exclus de l'échantillon.

Les séries de chiffre d'affaires construites à partir des déclarations nécessitent certaines corrections⁵. Des reports de déclaration d'un mois sur l'autre, se traduisant par une déclaration nulle un mois donné suivie d'une déclaration égale à l'activité de deux mois le mois suivant, ont été corrigés en répartissant l'activité du second mois entre le mois nul et le mois de rattrapage. Des valeurs aberrantes (en niveau ou en évolution) ont été corrigées en les ramenant dans la tendance de la série. Enfin, les entreprises déclarant leur chiffre d'affaires de manière trop irrégulière, pour lesquelles des simulations robustes ne pouvaient être effectuées, ont été exclues de l'échantillon de travail. Cela concerne surtout des micro-entreprises avec un faible chiffre d'affaires annuel et n'affecte donc que marginalement (0.2 point de pourcentage) le taux de couverture de l'étude en termes de chiffre d'affaires.

3. Au niveau international, le rôle direct de l'intensité des mesures de restriction sur l'activité est illustré par la forte corrélation entre l'indice de restriction de l'université d'Oxford, synthétisant en temps réel le degré de restriction liés aux mesures sanitaires nationales (Hale *et al.*, 2021), et le taux de (dé)croissance du PIB au premier trimestre de l'année 2020.

4. Les secteurs 'Industriels', 'Construction', 'Transports', 'Hébergement-restauration' et les 'Autres services' – essentiellement les arts, spectacles et activités récréatives, services de coiffure et de soins corporel, réparation d'ordinateurs et autres biens personnels – ont été particulièrement affectés par ces mesures.

5. Elles sont détaillées dans Bureau *et al.* (2021a, annexe B, p. 40).

Les déclarations de TVA sont enrichies d'informations sur les caractéristiques des entreprises issues de FARE 2018 – Fichier agrégé de résultats d'Esane (Élaboration des statistiques annuelles d'entreprise) – dernier millésime disponible. L'échantillon est ainsi restreint aux entreprises présentes dans FARE 2018 et déclarant leur TVA mensuellement depuis janvier 2018. Cet appariement permet de vérifier la cohérence des chiffres d'affaires reconstruits à partir des données de TVA. Pour assurer cette cohérence, les entreprises dont le chiffre d'affaires issu de FARE diffère de plus de 35 % du chiffre d'affaires annuel reconstitué à partir des déclarations TVA en 2018 sont exclues de l'échantillon de travail. Cette vérification a été effectuée à la fois au niveau des unités légales et des groupes profilés. Lorsque la cohérence n'était pas vérifiée au niveau de l'unité légale, mais l'était au niveau du groupe profilé, ce dernier était conservé dans l'échantillon en agrégeant les chiffres d'affaires des unités légales le composant⁶. Cette condition élimine de l'échantillon certaines grosses entreprises françaises pour lesquelles les déséquilibres entre les données de bilan et les déclarations TVA sont prononcés.

1.2. Description de l'échantillon

L'échantillon final est constitué de plus de 740 000 unités légales, regroupées en 645 000 unités d'observation : 578 000 unités légales analysées en tant que telles et 68 000 groupes profilés. Il représente 85 % de la valeur ajoutée des SNF des secteurs conservés dans l'étude, hors travailleurs indépendants. Sur l'ensemble des SNF, l'échantillon couvre 71 % de la valeur ajoutée dont 81 % de la valeur ajoutée des entreprises de taille intermédiaire et des grandes entreprises (ETI-GE), 72 % de la valeur ajoutée des petites et moyennes entreprises (PME) et 38 % de la valeur ajoutée des très petites entreprises (TPE), souscrivant majoritairement aux régimes de déclaration trimestriel et annuel de la TVA.

La répartition des effectifs salariés par secteur au sein de cet échantillon est proche de celle de l'ensemble des entreprises du champ d'étude. Par rapport à la base FARE restreinte au champ de travail de l'étude, le secteur du commerce est légèrement surreprésenté, et les secteurs de l'énergie et des activités scientifiques et techniques sous-représentés. Par taille d'entreprise, la structure des effectifs est comparable à la structure générale, mais les ETI-GE sont sous-représentées dans l'échantillon au profit des PME et TPE (figure I ; pour les chiffres voir Bureau *et al.*, 2021a, annexe C, p. 41).

Les ajustements effectués sur les déclarations permettent donc de fournir une image peu déformée de la population des SNF françaises et de leurs montants d'activité.

1.3. Recours aux données d'enquête

L'étude des déterminants de la situation des entreprises pendant la crise s'enrichit de l'enquête Impact de la crise sanitaire sur l'organisation de l'activité des entreprises de l'Insee (Duc & Souquet, 2020). Cette enquête documente le comportement des entreprises durant la crise, en particulier leur stratégie d'adaptation de leur activité : proportion de salariés en télétravail, réorganisation de la logistique commerciale pendant les confinements (développement de systèmes de vente en ligne, de vente directe ou de nouveaux systèmes de livraison), adaptation de l'offre à travers le développement de nouveaux produits, activités ou services, réalisation d'investissements spécifiques notamment dans les nouvelles technologies, ainsi que la réorganisation de l'activité *via* une modification des fournisseurs et partenaires commerciaux ou encore *via* la mise en commun de ressources avec d'autres entreprises. L'appariement avec ces données restreint l'échantillon à 13 500 entreprises. Afin de conserver la même représentativité, les observations sont pondérées par calage sur marge. Cet appariement n'est effectué qu'en dernière étape de l'analyse, et n'est utilisé qu'au sein du modèle paramétrique.

2. Méthode

2.1. Estimer les chocs d'activité attribuables à la crise

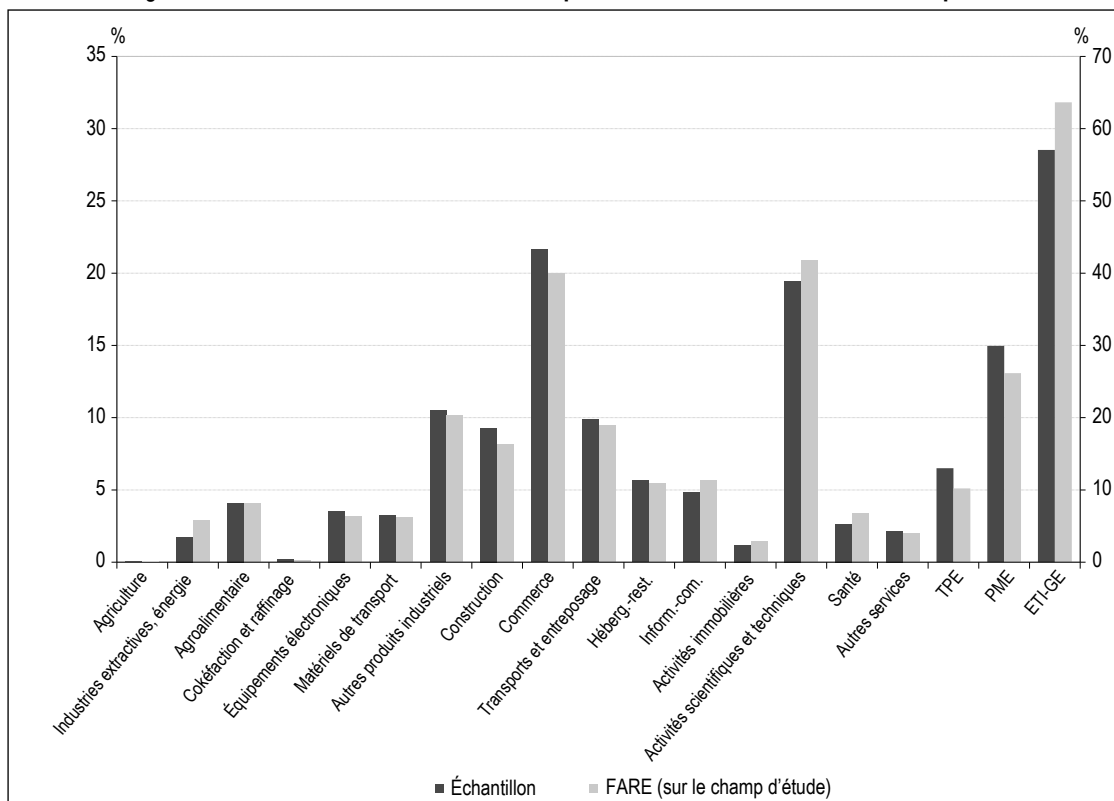
La méthode consiste à estimer des chocs d'activité attribuables à la crise sanitaire pour chacune des entreprises de l'échantillon, tout en s'assurant que l'agrégation des prévisions individuelles est cohérente avec des prévisions sectorielles robustes.

2.1.1. Estimation des dynamiques hors-crise au niveau méso-économique

Un chiffre d'affaires contrefactuel total est d'abord estimé par groupe *secteur* × *taille*. Pour cela, 16 secteurs de la nomenclature

6. En comptabilité d'entreprises, les chiffres d'affaires des différentes unités légales ne se somment pas exactement. La comparaison du chiffre d'affaires issu des comptes profilés de FARE avec le proxy obtenu par somme des chiffres d'affaires depuis les données TVA permet de conserver des unités légales dont la cohérence n'est pas vérifiée, mais dont une approximation au niveau profilé est cohérente avec les données de bilan. Cela permet notamment d'augmenter la taille de l'échantillon et le taux de couverture.

Figure I – Ventilation des effectifs salariés par secteur d'activité et taille d'entreprises



Note : les données issues de FARE comportent l'ensemble des entreprises françaises sur le champ d'étude. % secteur axe de gauche, % taille axe de droite.
Source : DGFiP, déclarations TVA ; Insee, FARE 2018. Calculs des auteurs.

agrégée A17⁷ et trois tailles d'entreprise (TPE, PME et ETI-GE) sont mobilisés, pour un total de 44 séries⁸. Le croisement du secteur et de la taille permet de conserver une bonne finesse d'analyse, même au niveau le plus agrégé des simulations, de prendre en compte la saisonnalité particulière des TPE dans certains secteurs, et d'obtenir des prédictions plus robustes des montants d'activité générés par les TPE au sein de chaque secteur.

Le chiffre d'affaires total des groupes *secteur × taille* est d'abord reconstitué mensuellement entre janvier 2015 et janvier 2020. Cette période est utilisée pour modéliser la dynamique hors-crise des 44 séries *secteur × taille*. Chacune d'elles est stationnarisée⁹, puis modélisée par modèle SARIMA en sélectionnant le couple (p_s, q_s) de paramètres d'ordres autorégressif et de moyenne mobile minimisant le critère AIC¹⁰ parmi 64 combinaisons possibles de paramètres allant de $(p_s = 1, q_s = 1)$ à $(p_s = 8, q_s = 8)$ ¹¹. Cette procédure permet d'obtenir une modélisation robuste des séries *secteur × taille* transformées et stationnarisées correspondant à l'équation (1). En notant $Y_{t,s}$ le chiffre d'affaires du groupe *secteur × taille* s à la date t , B l'opérateur de retard et $X_{t,s} = (1 - B^{12}) \log(Y_{t,s})$ ¹², les équations correspondant à chaque série s'écrivent :

$$X_{t,s} = (\varphi_{1,s} X_{t-1,s} + \dots + \varphi_{p_s,s} X_{t-p_s,s}) + \varepsilon_{t,s} - (\psi_{1,s} \varepsilon_{t-1,s} + \dots + \psi_{q_s,s} \varepsilon_{t-q_s,s}), \forall s \quad (1)$$

où $(\varepsilon_{t,s})_{t=1,\dots,T}$ désigne un bruit blanc gaussien de variance σ^2 . Ces équations sont ensuite mobilisées pour calculer la prévision linéaire optimale mensuelle d'horizon h de chaque série *secteur × taille*. Dans le cadre de l'étude,

7. Le découpage en 17 secteurs a été privilégié car il permettait des prévisions de meilleure qualité que celles obtenues avec un découpage plus fin.
8. Parmi les 48 groupes résultant du croisement secteur × taille, ceux comptant peu d'entreprises sont fusionnés par secteur. Dans les secteurs 'Agriculture' et 'Santé', les entreprises de plus de dix salariés sont regroupées. Le secteur de la 'Cokéfaction-raffinage' constitue un seul groupe.
9. La stationnarisation des séries transformées est vérifiée par des tests de Dickey-Fuller et de Dickey-Fuller augmentés (Dickey & Fuller, 1979).
10. Le critère AIC vaut $2k - 2\log(L)$, avec L la vraisemblance du modèle estimé, et k le nombre de paramètres libres du modèle. Il repose sur un compromis entre la qualité de l'ajustement et la complexité du modèle, en pénalisant les modèles ayant un grand nombre de paramètres pour limiter le surajustement (Akaike, 1998).
11. Une fois ce couple de paramètres sélectionné, l'absence d'autocorrélation sérielle des résidus, leur normalité et leur blancheur sont testées (Box & Pierce, 1970 ; Ljung & Box, 1978). La significativité des coefficients associés aux paramètres du couple est testée par un z-test. Lorsque plus d'un de ces critères n'est pas vérifié, le couple de paramètres donnant la seconde valeur la plus faible pour l'AIC est sélectionné et la procédure est répétée.
12. La différence par rapport au même mois de l'année précédente est une approche classique de stationnarisation des séries temporelles. Une décomposition mensuelle des 44 séries secteur × taille a par ailleurs permis d'identifier une tendance saisonnière, justifiant ainsi le recours à des retards d'ordre 12.

$h \in [1,11]$, la prévision étant effectuée entre février et décembre 2020 :

$$X_{T+h,s} = EL[X_{T+h,s} | X_{1,s}, \dots, X_{T,s}], \forall h \in [1,11] \quad (2)$$

Ces prévisions prennent la période janvier 2015-janvier 2020 comme période d'apprentissage. La transformation de la série des prévisions obtenues avec l'équation (2) permet d'obtenir la série $(Y_{T+h,s})_{h \in [1,11]}$ correspondant à l'estimation du chiffre d'affaires contrefactuel au cours de chaque mois de 2020 pour chaque groupe $secteur \times taille$ ^{13,14}.

2.1.2. Calcul des montants d'activité contrefactuels et des chocs individuels

La seconde étape consiste à répartir le chiffre d'affaires contrefactuel estimé entre toutes les entreprises des groupes $secteur \times taille$. Cette ventilation se fait itérativement, en commençant par février 2020 pour finir par décembre 2020. La part de marché mensuelle de chaque entreprise dans son groupe incorpore sa saisonnalité propre et sa dynamique récente de développement. Formellement, la part individuelle attribuée à chaque entreprise i du groupe $secteur \times taille s$ au premier mois t (ici février 2020) vaut :

$$S_{i,s,t} = \frac{1}{2} \left(S_{i,s,t-12} + \frac{1}{3} \sum_{j=1}^3 S_{i,s,t-j} \right) * \left(1 + \frac{1}{2} \frac{\left(\sum_{j=1}^3 Y_{i,t-j} - \sum_{j=1}^3 Y_{i,t-12-j} \right)}{\left(\sum_{j=1}^3 Y_{i,t-j} + \sum_{j=1}^3 Y_{i,t-12-j} \right)} \right) \quad (3)$$

avec $S_{i,s,t}$ la part de marché de l'entreprise i au sein du groupe $secteur \times taille s$ à la date t . La part de marché attribuée à chaque entreprise au mois de février correspond à la moyenne de sa part de marché au cours des trois mois précédents¹⁵ et de sa part de marché en février 2019¹⁶, à laquelle est ajouté un coefficient de pondération individuel visant à incorporer la tendance de croissance ou de décroissance de l'entreprise sur l'année écoulée. Ce coefficient s'appuie sur la structure des indicateurs de Haltiwanger et Davis ; il est borné par construction entre 0 et 2 et centré en 1 (Davis & Haltiwanger, 1992). Au-dessus de 1, il permet d'incorporer une tendance de croissance, et en dessous, de décroissance.

Les parts de marchés contrefactuelles des mois de mars ($t+1$) à décembre 2020 ($t+10$) sont calculées de la même façon mais en remplaçant les parts de marché des mois postérieurs à février 2020 par celles estimées aux itérations précédentes :

$$\left\{ \begin{array}{l} S_{i,s,\hat{t}+1} = \frac{1}{2} \left(S_{i,s,(t+1)-12} + \frac{1}{3} \left(S_{i,s,t} + \sum_{j=2}^3 S_{i,s,(t+1)-j} \right) \right) * \\ \left(1 + \frac{1}{2} \frac{\left(\left(Y_{i,(\hat{t}+1)-1} + \sum_{j=2}^3 Y_{i,(t+1)-j} \right) - \sum_{j=1}^3 Y_{i,(t+1)-12-j} \right)}{\left(Y_{i,(\hat{t}+1)-1} + \sum_{j=2}^3 Y_{i,(t+1)-j} + \sum_{j=1}^3 Y_{i,(t+1)-12-j} \right)} \right) \\ S_{i,s,\hat{t}+10} = \frac{1}{2} \left(S_{i,s,(t+10)-12} + \frac{1}{3} \sum_{j=2}^3 S_{i,s,(\hat{t}+10)-j} \right) * \\ \left(1 + \frac{1}{2} \frac{\left(\sum_{j=1}^3 Y_{i,(\hat{t}+10)-j} - \sum_{j=1}^3 Y_{i,(t+10)-12-j} \right)}{\left(\sum_{j=1}^3 Y_{i,(\hat{t}+10)-j} + \sum_{j=1}^3 Y_{i,(t+10)-12-j} \right)} \right) \end{array} \right. \quad (4)$$

Les parts de marchés mensuelles sont ensuite ajustées de sorte qu'elles somment à 1 au sein de chaque groupe :

$$S_{i,s,t}^{\sim} = \frac{S_{i,s,t}}{\sum_{i=1}^n S_{i,s,t}} \quad (5)$$

Le chiffre d'affaires contrefactuel individuel est le produit de la part de marché individuelle estimée et de l'activité contrefactuelle totale du groupe auquel appartient l'entreprise au mois t :

$$CA_{i,s,t} = S_{i,s,t}^{\sim} Y_{t,s} \quad (6)$$

Le choc d'activité mensuel estimé est la différence, en pourcents, entre le chiffre d'affaires observé et ce chiffre d'affaires contrefactuel :

$$Choc_{i,s,t} = \left(\frac{CA_{i,s,t} - CA_{i,s,t}^{\sim}}{CA_{i,s,t}^{\sim}} \right) * 100 \quad (7)$$

En sommant – mois par mois ou sur l'année – les chiffres d'affaires contrefactuels de tout l'échantillon ou d'un secteur donné et en les comparant à l'agrégation des chiffres d'affaires observés sur le même périmètre, il est possible de construire des chocs d'activité agrégés. Les analyses de la distribution des chocs d'activité individuels

13. La qualité de ces prévisions secteur \times taille du modèle est testée sur l'année 2019. Sur une année sans crise, on attend des prévisions contrefactuelles qu'elles correspondent aux montants de chiffre d'affaires observés. Sur l'ensemble des mois de chacune des séries, la valeur absolue de l'écart entre le montant observé et le montant simulé est de 2 % en moyenne et le montant observé est compris dans l'intervalle de confiance au seuil de 95 % du montant prédit (détails dans Bureau et al., 2021a, annexe F, figure F.1). Sur 2019, le modèle développé permet d'obtenir de meilleurs résultats qu'une modélisation naïve attribuant comme chiffre d'affaires mensuel le chiffre d'affaires du même mois de l'année précédente pour 85 % des mois de l'ensemble des 44 séries secteur \times taille.

14. Ces prévisions coïncident aussi avec l'Enquête mensuelle de conjoncture de la Banque de France (détails dans Bureau et al., 2021a, annexe F, figure F.4). Le coefficient de corrélation entre les chocs mensuels estimés dans l'étude et par l'enquête se situe autour de 0.8.

15. Cette moyenne glissante lisse les potentiels résultats exceptionnels et donne une image plus robuste du poids de l'entreprise au sein du groupe.

16. La part de marché du même mois de l'année précédente permet d'incorporer la saisonnalité mensuelle des entreprises, élément important si celle-ci diffère de la saisonnalité du groupe.

tels que calculés en (7) permettent d'affiner ces résultats en identifiant entreprises gagnantes et perdantes, y compris au sein d'un même secteur.

2.1.3. Mesure de l'impact de la crise par des indicateurs distributionnels des chocs d'activité

Les chiffres d'affaires contrefactuels individuels estimés correspondent à des scénarii robustes de ce qui aurait pu être observé pour chacune des entreprises, sur la base de l'ensemble de l'information disponible au déclenchement de la crise. Toutefois, malgré les précautions méthodologiques, il est possible qu'au niveau des entreprises, les prévisions diffèrent des montants d'activités qui auraient été observés. D'une part car les montants de chiffre d'affaires individuels déclarés par les entreprises sont beaucoup plus volatiles que les montants agrégés et ne présentent pas nécessairement la même saisonnalité. D'autre part car l'attribution des parts de marché contrefactuelles s'appuie sur les dynamiques observées l'année précédant l'exercice de prévision, ce qui rend l'exercice périlleux pour les entreprises présentant un rythme de développement non linéaire. Ainsi, même en l'absence de crise, le modèle de calcul des chocs individuels conduit à l'estimation de chocs non nécessairement nuls et pouvant fluctuer autour de zéro. En ce sens, l'analyse de la prévalence des entreprises gagnantes ou perdantes en 2020, et de l'ampleur de ces gains ou pertes, doit se concentrer sur leur distribution et sur son caractère exceptionnel durant la crise.

La comparaison de la distribution des chocs d'activité en 2020 avec celle obtenue en répliquant la simulation sur l'année 2019 permet de comparer les déviations des trajectoires attendues simulées par le modèle l'année de la crise à celles d'une année sans crise. L'intensité de la déformation de cette distribution par rapport à 2019, où l'on attend des déviations proches de zéro, permet d'illustrer l'intensité de l'impact de la crise sanitaire. Cette déformation est mesurée avec la distance de Hellinger, comprise entre 0 et 1 et mesurant la similarité entre deux distributions statistiques. En notant f et g les fonctions de densité des distributions comparées, la distance de Hellinger vaut la racine carrée de la formule suivante¹⁷ :

$$H^2(f, g) = \frac{1}{2} \int (\sqrt{f(x)} - \sqrt{g(x)})^2 dx \\ = 1 - \int \sqrt{f(x)g(x)} dx$$

La comparaison des distributions des chocs d'activité est effectuée sur les chocs annuels

et mensuels. Les densités des distributions des chocs individuels sont estimées par noyau.

Le recours aux données individuelles met en avant des chocs dispersés. La contribution du secteur à cette hétérogénéité doit être évaluée et pour ce faire, la variance mensuelle des chocs d'activité individuels est décomposée entre une part attribuable au secteur d'activité et une part résiduelle, attribuable à d'autres facteurs. Le découpage sectoriel utilisé est le niveau le plus fin de la nomenclature d'activité française, présentant 732 catégories. La méthode de décomposition, usuelle (Gibbons *et al.*, 2014 ; Helpman *et al.*, 2017), est la suivante :

$$V = Var(Choc_{i,s}) = \underbrace{\sum_s \frac{n_s}{n} \cdot Var_s(Choc_{i,s})}_{\text{Variance intra-classe}} \\ + \underbrace{\sum_s \frac{n_s}{n} \cdot (\overline{Choc_s} - \overline{Choc})^2}_{\text{Variance inter-classe}}$$

avec $\overline{Choc} = \frac{1}{n} \sum_i Choc_{i,s}$ et $\overline{Choc_s} = \frac{1}{n_s} \sum_{i \in s} Choc_{i,s}$,

$Choc_{i,s}$ le choc subi par la firme i du secteur s et n le nombre d'entreprises dans l'échantillon.

2.2. Partition des entreprises selon leur trajectoire de chocs

La constitution d'une série de chocs d'activité mensuels pour chaque entreprise de l'échantillon rend comparables les trajectoires de toutes les entreprises, indépendamment de leur montant d'activité attendu et observé, offrant la possibilité d'identifier des groupes homogènes de séries de chocs mensuels.

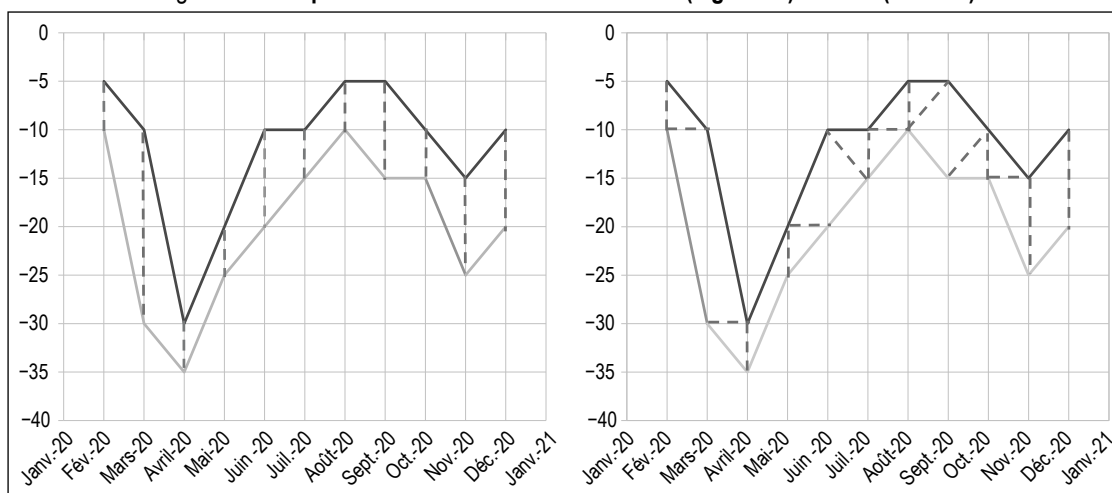
2.2.1. Construction d'une typologie par clustering de courbes

Les profils de chocs des entreprises sur l'ensemble de l'année 2020 sont identifiés par clustering de courbes. Cette méthode consiste à partitionner une population de séries en un nombre donné de classes homogènes selon la distance DTW (*Dynamic Time Warping*¹⁸ – cf. Berndt & Clifford, 1994 ; Ratanamahatana & Keogh, 2004). La figure II illustre la différence entre cette distance et une distance euclidienne : l'approche euclidienne compare simplement les séries point par point, alors que l'approche DTW compare les séries deux à deux et déforme l'ordre des points pour les aligner autant que possible. Cette déformation n'a lieu qu'au sein d'une

17. L'analyse a été reproduite avec d'autres distances statistiques (Kullback-Leibler, Bhattacharyya), pour des conclusions identiques.

18. Détails dans Bureau *et al.* (2021a, p. 15).

Figure II – Comparaison des distances euclidienne (à gauche) et DTW (à droite)



Lecture : la distance euclidienne en mars 2020 est la différence entre les deux séries ce mois-ci. La distance DTW va plutôt comparer la valeur de mars de la série noire à la valeur de février de la série grise, dont elle est plus proche, et inversement la valeur de mars de la série grise sera comparée à celle d'avril.

fenêtre de largeur égale à 10 % de la taille de la série, soit un mois (Aghabozorgi *et al.*, 2015).

Les trajectoires de chocs mensuels sont réparties en k classes de sorte à minimiser la distance DTW entre les éléments d'une même classe (Sardá-Espinosa, 2019). Pour cela k trajectoires sont tirées aléatoirement dans l'échantillon, pour constituer le centre de chacune des classes. Les autres trajectoires sont alors comparées aux différents centres et assignées à la classe dont le centre est le plus proche. Lorsque toutes les séries ont été classées, les séries médianes des classes deviennent les nouveaux centres et le processus est répété jusqu'à convergence de la partition, ou jusqu'à atteindre le nombre maximal d'itérations. La partition finale dépend à la fois du nombre de classes choisi et des centres initiaux. Une partition en quatre classes a été choisie ici afin d'optimiser la qualité de la partition tout en conservant un nombre de classes important¹⁹. Le clustering a été répété dix fois afin d'assurer la stabilité de la partition finale. La confusion, i.e. la part d'entreprises changeant de classe entre ces répétitions, reste proche de zéro dans l'ensemble de ces répétitions.

2.2.2. Expliquer la ventilation des entreprises entre les profils-types identifiés : mise en place d'un modèle de classification

L'identification des profils de trajectoire et la répartition des entreprises s'appuient exclusivement sur les chocs d'activité mensuels estimés, mais ces derniers peuvent être corrélés avec des caractéristiques des entreprises. Afin d'éclairer *a posteriori* la répartition des entreprises entre ces types de trajectoires, nous étudions les corrélations entre le profil assigné aux entreprises et leurs caractéristiques.

Les variables explicatives utilisées dans le modèle sont le secteur d'activité, la taille des entreprises, leur date de création, l'existence d'une activité exportatrice et des variables relatives au développement de la vente en ligne, de systèmes de livraison, de nouveaux produits ou services, la réorganisation de l'activité, la mise en commun de moyens avec d'autres entreprises et la réalisation d'investissements spécifiques dans les nouvelles technologies pendant la crise. Elles proviennent de FARE et de données d'enquête. L'appariement avec les données d'enquête restreint l'échantillon à 13 500 entreprises. Afin de conserver un échantillon dont la proportion d'entreprises affectée à chaque profil de trajectoire est similaire à la structure de l'échantillon, et afin d'avoir une distribution identique en termes de taille, secteur d'activité, date de création et existence d'une activité exportatrice, des poids sont attribués aux entreprises avec une méthode de calage sur marge (Deville & Särndal, 1992 ; Rebecq, 2016). Le modèle utilisé est un modèle logit multinomial non ordonné estimé par réseau de neurones avec la méthode Broyden-Fletcher-Goldfarb-Shanno²⁰.

3. Résultats

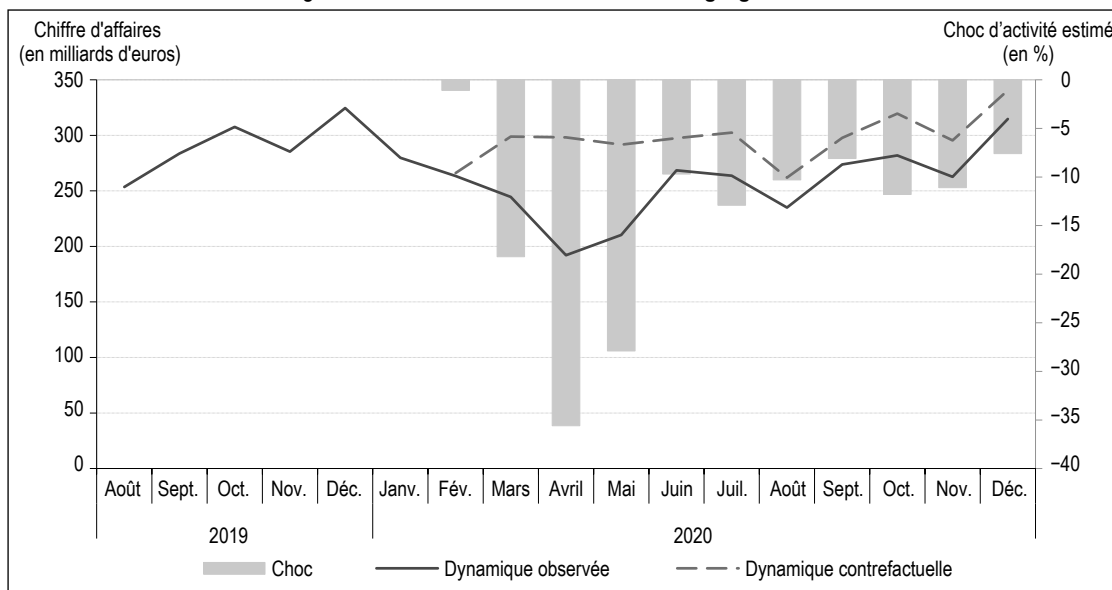
3.1. Un impact de la crise très important sur l'activité des entreprises et d'ampleur variable au cours de l'année

L'activité économique totale a été très ralentie pendant le premier confinement de 2020. Entre mars et mai, son niveau se situe 27 % en dessous de son niveau estimé en l'absence de crise (figure III). Sur le seul mois d'avril,

19. Détails dans Bureau *et al.* (2021a, annexe H, p. 86).

20. Détails dans Bureau *et al.* (2021a, annexe K, p. 91).

Figure III – Évolution du choc d'activité agrégé en 2020



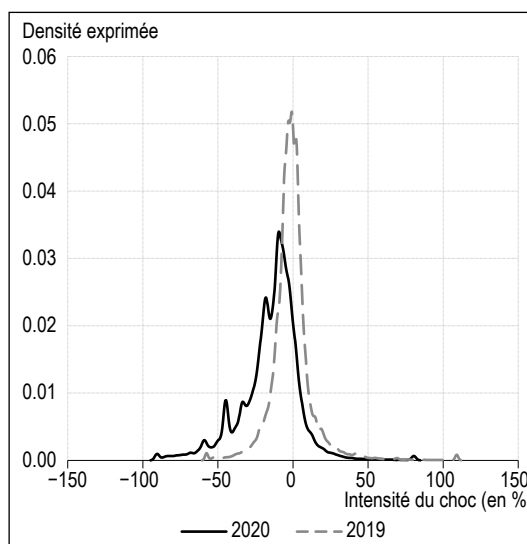
Source : DGFIP, déclarations TVA. Calculs des auteurs.

cet écart est de -35% . L'activité économique a ensuite rebondi entre juin et octobre, tout en restant 10% en dessous de son niveau attendu. La perte d'activité du printemps n'a donc pas été compensée par une activité plus élevée en été ou au début de l'automne. Au quatrième trimestre, qui inclut le deuxième confinement, la perte d'activité est estimée à environ 10% . D'une part, le deuxième confinement a été plus court et moins contraignant que le premier. D'autre part, les entreprises ont pu davantage adapter leur stratégie et leur organisation qu'au début de la crise.

Sur l'année 2020, les montants totaux de chiffre d'affaires de l'économie française se sont détachés de leur trajectoire attendue, avec plus ou moins d'intensité selon le mois étudié. Ces écarts systématiquement négatifs au niveau macro-économique sont le résultat de chocs d'activité à la fois positifs et négatifs au niveau individuel. En l'absence de crise (2019), la distribution modélisée des chocs d'activité individuels est symétrique, centrée en zéro et de variance faible. Au contraire en 2020 la distribution des chocs annuels n'est plus symétrique et connaît une forte translation vers la gauche, traduisant une plus forte proportion de chocs négatifs (figure IV). Les pertes d'activité agrégées reflètent donc la plus forte prévalence, en 2020, de chocs d'activité individuels négatifs, parfois de grande intensité.

La déformation des distributions de chocs d'activité individuels évolue mensuellement en fonction de l'intensité du choc économique. La distance de Hellinger, comparant les distributions

Figure IV – Distributions des chocs d'activités individuels en 2019 et 2020

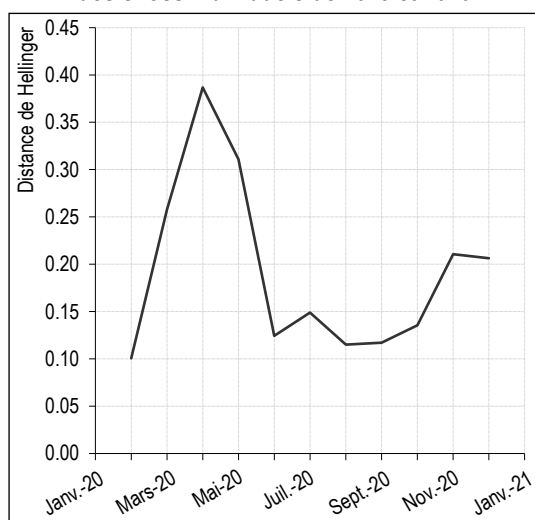


Note : la densité est estimée par noyau.
Source : DGFIP, déclarations TVA. Calculs des auteurs.

de chocs d'activité pour le même mois de 2019 et de 2020, illustre cette évolution (figure V)²¹. La dissimilarité mesurée est très faible pour le mois de février – premier mois modélisé, durant les prémices de la crise. Ensuite, la déformation des chocs individuels semble s'intensifier selon le calendrier des mesures de restriction : plus forte en avril, avant de réduire progressivement jusqu'au mois d'octobre, moment où le couvre-feu, puis le deuxième confinement furent instaurés.

21. Détails dans Bureau et al. (2021a, annexe F.1.b, figures F.2 et F.3).

Figure V – Distance de Hellinger entre les distributions mensuelles des chocs individuels de 2019 et 2020



Source : DGFiP, déclarations TVA. Calculs des auteurs.

3.2. Une hétérogénéité des situations individuelles dépassant la seule appartenance sectorielle

Même lorsque l'impact de la crise est le plus prononcé, les distributions des chocs d'activité individuels révèlent un certain nombre d'entreprises connaissant des déviations positives par rapport à leur trajectoire attendue. Au cœur du premier confinement, certaines entreprises s'en sortent donc au moins aussi bien qu'elles l'auraient pu en l'absence de crise. Cette hétérogénéité pose question, en particulier en ce qui concerne le rôle du secteur d'activité dans les différences observées, d'autant plus que les dissimilarités des distributions de chocs d'activité entre 2019 et 2020 sont plus prononcées durant les confinements.

En effet, le premier confinement a bien constitué un choc pour tous les secteurs, mais d'ampleur variable. L'« Hébergement-restauration » et la « Fabrication de matériels de transport » sont les deux secteurs dont l'activité économique a été la plus déprimée, avec une perte d'activité estimée respectivement à -71% et -54% entre mars et mai (figure VI). Les secteurs « Information et communication », « Agriculture » et « Agroalimentaire » ont mieux résisté (respectivement -13% , -11% et -9%). Au contraire, lors du deuxième confinement, seuls certains secteurs ont vu leur activité se dégrader à nouveau de manière significative après la reprise modérée de la période estivale : « Hébergement-restauration » (-54%) et les « Autres services » (-33%). Pour l'essentiel des autres secteurs, l'activité a baissé de façon plus limitée²². Dans plusieurs secteurs industriels, comme l'électronique et les autres

produits industriels, l'activité économique a rebondi entre les deux confinements et s'est presque rétablie au niveau attendu lors du second (-3% et -5% respectivement).

Si les différences inter-sectorielles sont prononcées à ce niveau de découpage, elles n'expliquent vraisemblablement pas l'ensemble de la diversité des situations individuelles connues par les entreprises françaises. La figure VII présente, pour chaque secteur, les principaux quantiles, pondérés par les effectifs salariés²³, de la distribution des chocs d'activité estimés pour les entreprises les composant en 2020 et en 2019. En 2020, la majorité des entreprises de chaque secteur connaît des pertes d'activité et les secteurs connaissent des situations hétérogènes avec des chocs médians très différents. Chaque secteur compte ainsi des dispersions très importantes avec des entreprises très pénalisées, parfois cessant leur activité, et d'autres atteignant au contraire leur niveau d'activité attendu, malgré la crise. Ces différences observées au sein d'un découpage en 17 secteurs peuvent s'expliquer par le fait que les mesures de restriction sanitaire, et en particulier les fermetures temporaires, ont touché des secteurs définis plus finement. Au niveau le plus fin de la nomenclature d'activité française (732 catégories ou « sous-secteurs »), on remarque en effet que les chocs médians annuels varient beaucoup entre les sous-secteurs d'un même secteur agrégé, y compris parmi ceux les plus touchés. Par exemple, dans le secteur « Hébergement-restauration », les établissements de restauration rapide ont mieux résisté (-34%) que les débits de boissons (-55%) ou les services de traiteurs (-70%), tous contraints à la fermeture en mars²⁴. Dans les « Autres services », la perte la plus importante concerne la gestion des salles de spectacle (-80%) alors que les services funéraires se sont maintenus (-4%). De même, les secteurs les plus préservés, comme le commerce et la fabrication alimentaire, comprennent également des sous-secteurs fortement touchés (grands magasins -52% , pâtisseries -23%) et d'autres connaissant des gains, bien que modérés, d'activité (commerce de détail d'appareils électroménagers $+8\%$, fabrication de pâtes alimentaires $+8\%$).

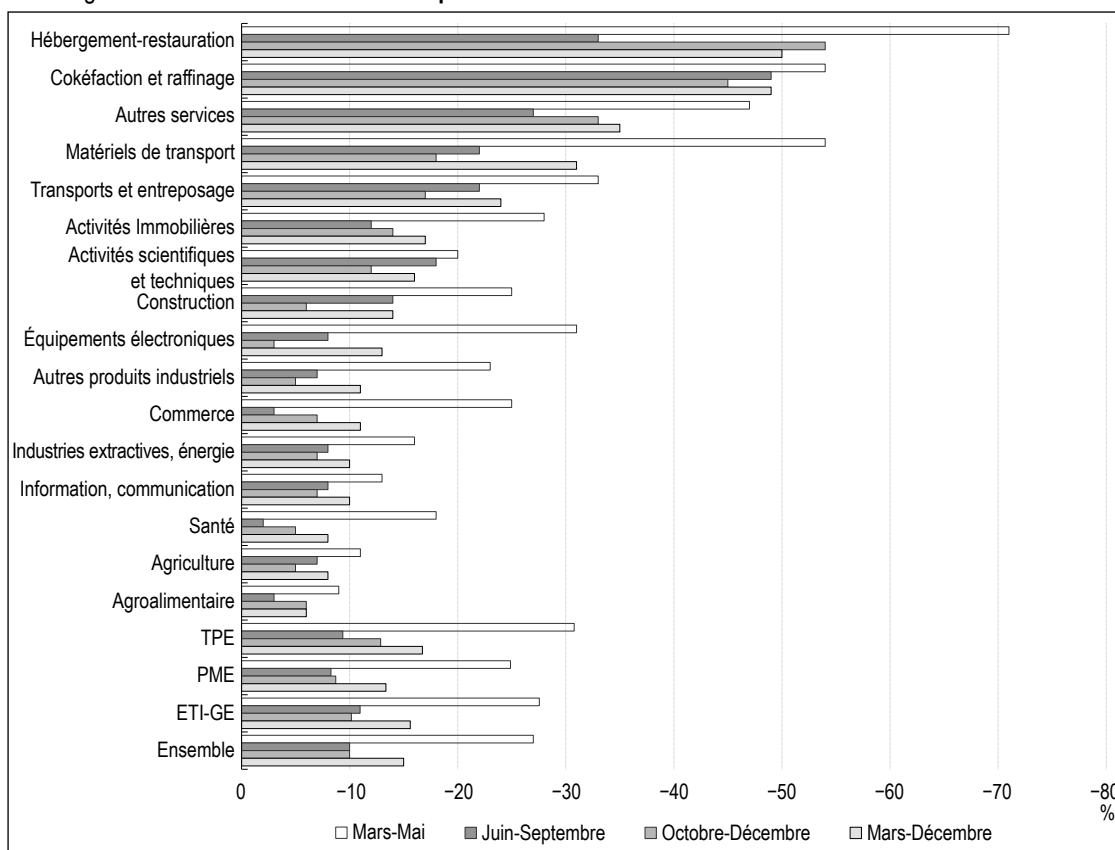
Pour autant, la diversité des situations des entreprises françaises durant la pandémie de Covid-19

22. Détails dans Bureau et al. (2021a, annexe E, p. 62).

23. La dispersion des chocs individuels d'activité pondérés par les effectifs reflète la dispersion des chocs pour les salariés appartenant à ces entreprises. La dispersion non pondérée reflète la dispersion des chocs pour les entreprises, soit pour les TPE, ces dernières étant largement majoritaires dans l'économie et dans l'échantillon.

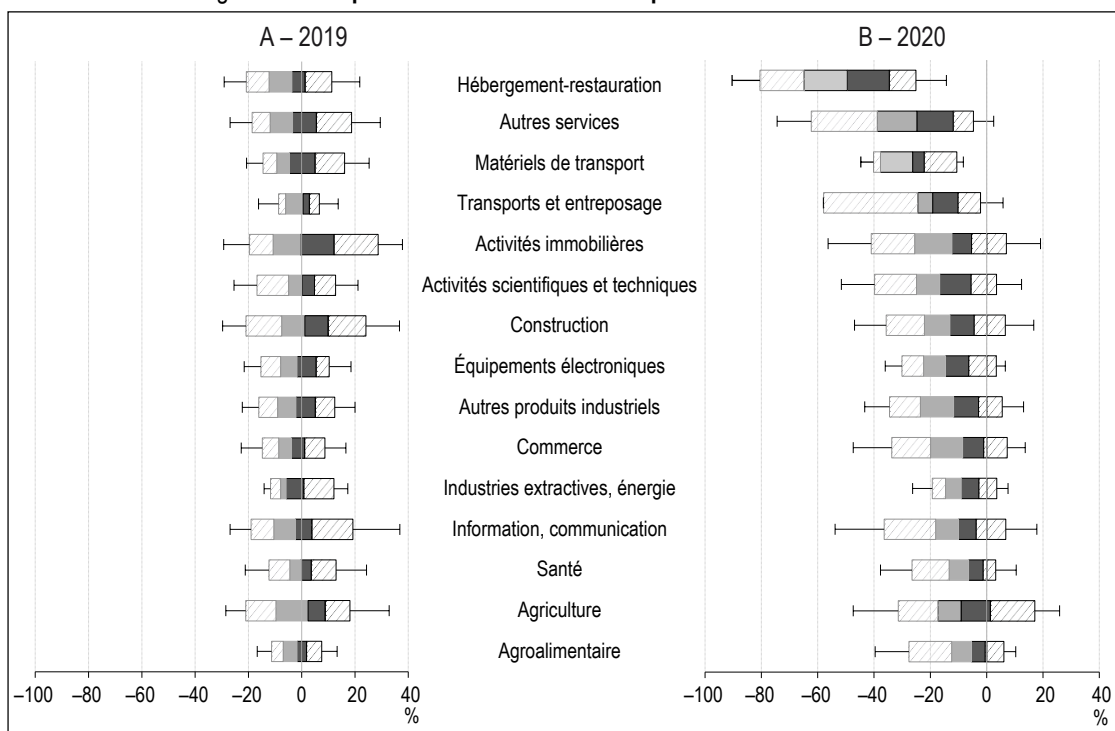
24. Voir Bureau et al. (2021b).

Figure VI – Choc d'activité économique cumulé de mars à décembre 2020 : ventilation sectorielle



Source : DGFIP, déclarations TVA. Calculs des auteurs.

Figure VII – Dispersion des chocs d'activité par secteur en 2019 et 2020



Note : chaque ligne représente la ventilation des chocs individuels d'activité au sein d'un secteur via une boîte à moustaches. Les différents segments des boîtes distinguent les quantiles à 10 %, 25 %, 50 % (médiane), 75 % et 90 %. Les extrémités (moustaches) représentent la valeur des quantiles à 5 % et 95 %. En raison de ses faibles effectifs, le secteur 'Cokéfaction-raffinage' n'est pas représenté.

Lecture : en 2020, dans le secteur dans le secteur 'Hébergement-restauration', 5 % des salariés travaillent dans une entreprise ayant connu un choc d'activité au moins inférieur ou égal à -90.4 %, et 50 % avec un choc inférieur ou égal à -50 %.

Source : DGFIP, déclarations TVA. Calculs des auteurs.

est-elle uniquement attribuable à l'activité exercée, même considérée à son niveau de découpage le plus fin ? La décomposition de la variance mensuelle des chocs d'activité individuels entre une part attribuable au secteur d'activité (732 catégories) et une part résiduelle permet d'évaluer la contribution de l'activité à la diversité des situations des entreprises. En 2020, le secteur d'activité contribue à la variance des chocs d'activité individuels à hauteur de 43 %²⁵, largement plus qu'en 2019 (figure VIII). La contribution du secteur à l'hétérogénéité des chocs est également plus importante durant les mois de confinement, ayant inégalement affecté les secteurs d'activité. En avril 2020, le secteur contribue pour 48 % à la variance des chocs avec pondération par les effectifs. Le rôle du secteur dans la dispersion des chocs est également systématiquement plus important dans les secteurs S1 et S1bis²⁶, plus concernés par des mesures de restriction sanitaires et de fermetures administratives.

Le rôle joué par le secteur d'activité dans les déviations individuelles des entreprises par rapport à la trajectoire modélisée de leur activité est largement plus important en 2020 qu'en 2019. En 2020 cette contribution du secteur à l'hétérogénéité des situations est plus importante sur les mois concernés par des mesures sanitaires marquées. Cependant, même à son pic, cette contribution ne représente que la moitié de toute l'hétérogénéité observée, d'autres facteurs doivent donc nécessairement jouer sur les chocs

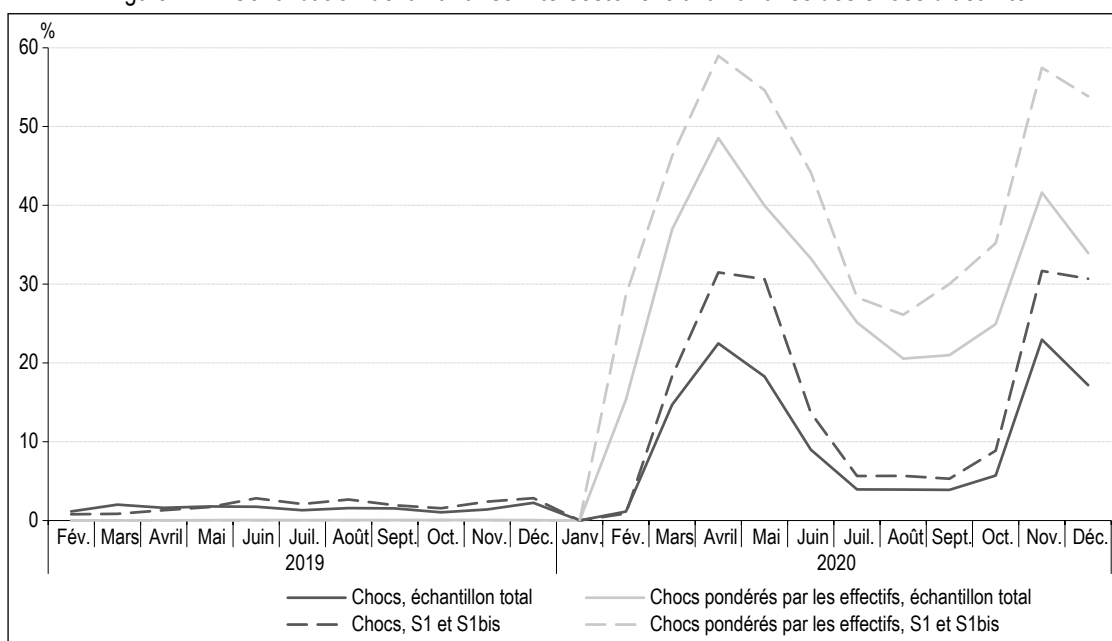
d'activité observés. Afin de traiter conjointement l'hétérogénéité des situations individuelles, au mois le mois, et les différents facteurs pouvant expliquer ces différences entre entreprises ainsi que leur évolution dans le temps, l'analyse est prolongée en deux étapes : d'abord en identifiant *ex post* une typologie pertinente des différentes trajectoires d'activité des entreprises au cours de l'année 2020 afin de regrouper les entreprises dont l'évolution des gains ou pertes d'activité ont été comparables au fil de l'année. Ensuite, en étudiant les déterminants de l'appartenance à chacun des profils de trajectoires de la typologie grâce au modèle de classification multinomial.

3.3. Quatre profils-types de trajectoire de chocs pendant la crise

La construction de la typologie des trajectoires individuelles des entreprises au cours de l'année 2020 permet de distinguer quatre trajectoires-types de chocs mensuels d'activité subis en 2020 (figure IX). Chacun de ces profils se distingue des autres à la fois par l'ampleur du choc subit au début de la pandémie et par la capacité de

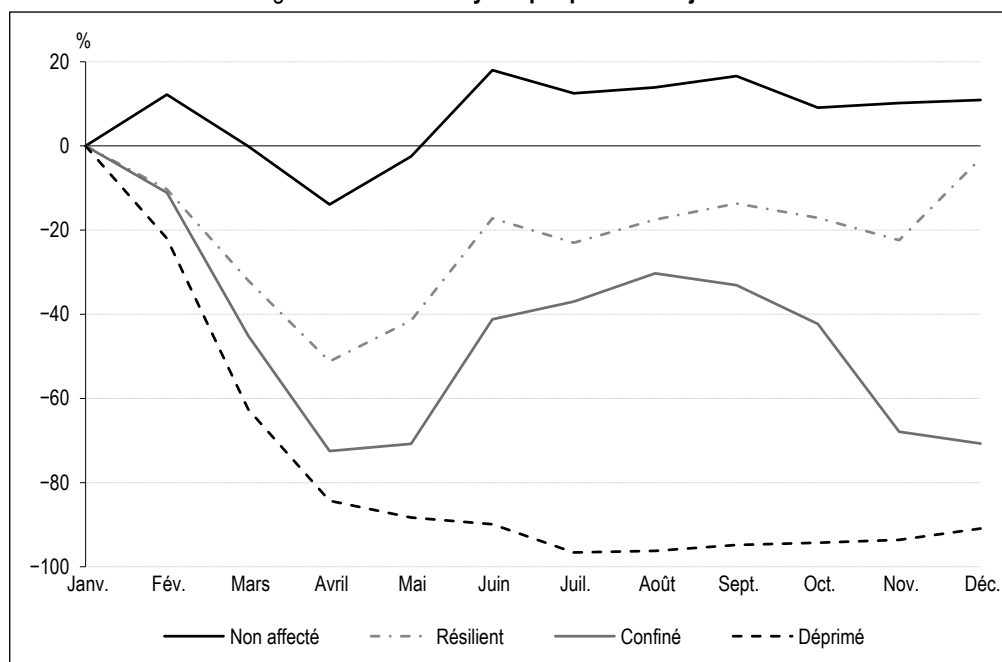
25. Décomposition avec pondération par les effectifs salariés.
26. Les listes des secteurs S1 et S1bis sont définies par les modifications successives du décret du 30 mars 2020 relatif au fonds de solidarité. L'évolution de ces listes a été reconstruite, mois par mois, sur l'ensemble de l'année 2020. La liste S1 concerne des secteurs particulièrement touchés par la crise et les fermetures administratives, en particulier dans les domaines de la restauration, du tourisme, de l'événementiel, de la culture et du sport. La liste S1bis concerne des secteurs liés, par exemple, la distribution de films ou l'édition de livres.

Figure VIII – Contribution de la variance intersectorielle à la variance des chocs d'activité



Note : chaque courbe correspond à la part de variance intersectorielle dans la variance totale des chocs, chaque mois.
Lecture : en avril 2020 la variance intersectorielle représente 23 % de la variance totale des chocs sur l'échantillon, 48 % si l'on pondère par les effectifs.
Source : DGFIP, déclarations TVA. Calculs des auteurs.

Figure IX – Chocs moyens par profil de trajectoire



Lecture : les entreprises du profil « non affecté » ont connu un choc moyen de -14 % en avril 2020.
Source : DGFIP, déclarations TVA. Calculs des auteurs.

résilience, i.e. à retrouver sa trajectoire attendue hors-crise. Plus précisément, se distinguent :

- Les entreprises « non affectées » (36 % des entreprises et 42 % des salariés) : le premier confinement a eu un impact limité sur ces entreprises, avec un choc moyen²⁷ de -14 % en avril, suivi d'un rattrapage de l'activité attendue à partir du mois de juin. À l'exception du premier confinement, la distribution des chocs au sein de ce groupe est comparable à celle d'une année « normale ».

- Les entreprises « résilientes » (38 % des entreprises et 44 % des salariés) : leur perte d'activité initiale est plus substantielle, avec un choc moyen de -51 % en avril. À partir du mois de juin, les pertes sont moindres et le choc moyen reste stable autour de -20 % sur la fin d'année.

- Les entreprises « confinées » (20 % des entreprises et 12 % des salariés) : leur trajectoire moyenne se caractérise par des chocs important lors des confinements (-72 % en avril, -70 % en novembre et décembre) et reprise limitée de l'activité pendant l'été.

- Les entreprises « déprimées » (6 % des entreprises et 2 % des salariés) : leur activité s'est effondrée lors du premier confinement (-84 % en moyenne en avril), sans reprise à l'été. Les chocs médians de ces entreprises sont proches de -100 % d'avril à décembre 2020, et un tiers d'entre elles déclare un chiffre d'affaires nul sur l'ensemble de cette période.

3.4. Caractériser la trajectoire d'activité des entreprises : au-delà du secteur d'activité, leur adaptation organisationnelle

La partition des entreprises entre les profils de trajectoire est « non supervisée », et ne dépend donc que des chocs d'activité estimés pour chacune des entreprises. L'exploration des corrélations entre les caractéristiques des entreprises et leur profil de trajectoire permet d'éclairer *ex post* les logiques ayant présidé aux difficultés qu'elles ont pu rencontrer.

Les coefficients issus du modèle de classification permettant d'étudier ces corrélations sont statistiquement significatifs²⁸ (annexe A2). Le secteur d'activité est le facteur prépondérant de la répartition des entreprises entre trajectoires-types. Il représente près de 85 % de la part de la ventilation des entreprises expliquée par le modèle²⁹. Cette part est attribuable à la très forte dépendance sectorielle des profils les plus touchés, composés quasi intégralement d'entreprises de secteurs fermés administrativement pendant les confinements. Conditionnellement aux autres variables, les secteurs ayant la plus

27. Toutes les moyennes sont calculées sur des séries winsorisées à droite : les chocs supérieurs au 95^e centile de choc sont ramenés à la valeur de ce quantile.

28. Les observations sont pondérées par les coefficients issus du calage sur marge lors de la régression. Cette pondération peut avoir un impact positif sur la significativité des effets affichés.

29. Estimation par mobilisation des R^2 ajustés de Cox & Snell (1989).

forte probabilité d'appartenir au profil « non affecté » sont ceux de la fabrication de produits électroniques grand public, les sous-secteurs de l'industrie alimentaire, les activités vétérinaires et le secteur médical. Dans le profil « résilient » se trouvent majoritairement les secteurs de fabrication d'articles de joaillerie, d'ordinateurs et d'équipements périphériques et d'équipements automobiles. Les secteurs ayant la plus forte probabilité d'appartenir au profil « confiné » sont ceux des transports ferroviaires et des bibliothèques et musées. Enfin, dans le profil « déprimé », les sous-secteurs de la culture, de l'hébergement et de la restauration et du tourisme et des transports de voyageurs sont les plus surreprésentés.

Conditionnellement au secteur, l'effet des autres variables sur la probabilité de se retrouver dans les différentes classes est significatif, mais de plus faible ampleur. Autrement dit, la différence absolue de probabilité d'appartenir à un profil est beaucoup plus forte entre deux secteurs différents qu'entre deux modalités d'une autre variable du modèle. Toutefois, en exprimant les effets de chacune de ces variables en pourcentage

d'évolution de la probabilité d'être affecté à chaque profil³⁰, plusieurs éléments se dégagent (figure X).

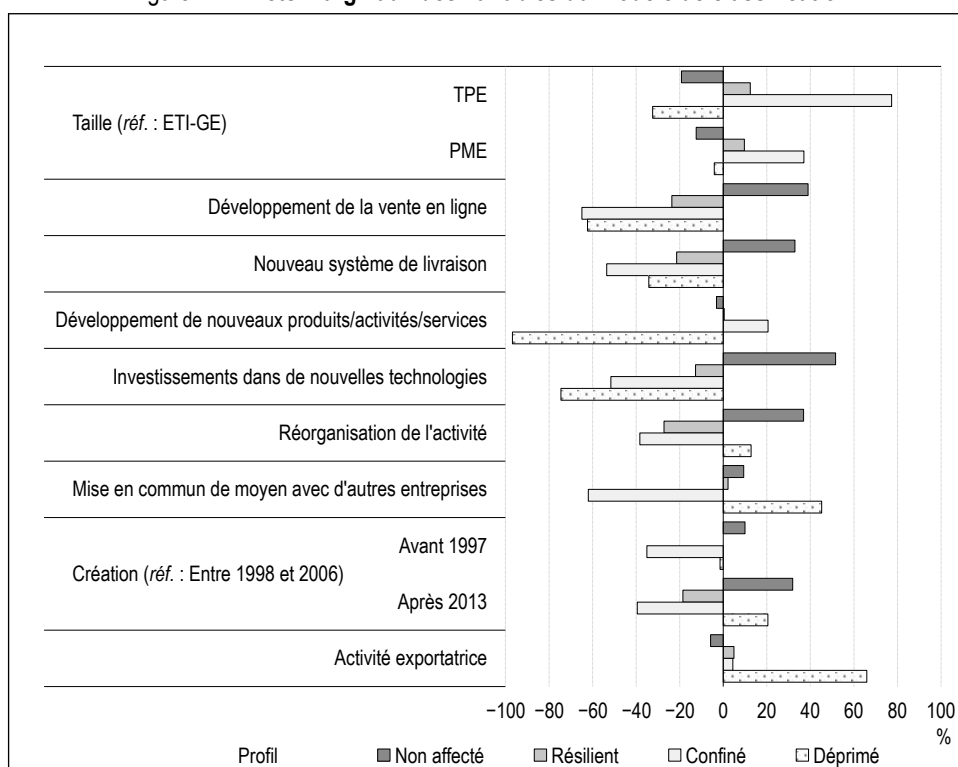
Les PME, particulièrement les TPE, plus touchées en moyenne pendant les confinements ont, toutes choses égales par ailleurs, plus de chance d'appartenir au profil « confiné », illustrant les difficultés spécifiques rencontrées par les TPE dans un certain nombre de secteurs.

30. Ces effets s'appuient sur le calcul des probabilités prédites à la moyenne d'appartenir à chaque profil pour toutes les modalités des variables explicatives catégorielles. En comparant ces probabilités en faisant varier uniquement la modalité d'une même variable catégorielle, il est possible de calculer l'effet relatif du passage d'une modalité à une autre sur la probabilité d'appartenir à chaque profil. Formellement, l'effet d'une variable binaire j sur la probabilité d'appartenir au profil c vaut ici :

$$\text{Effet}_{j,c} = \frac{\mathbb{P}(\text{Profil}_c | X_j = 1, X_{-j}) - \mathbb{P}(\text{Profil}_c | X_j = 0, X_{-j})}{\mathbb{P}(\text{Profil}_c | X_j = 0, X_{-j})}, \forall c \in [1,4]$$

Ces effets ont aussi été calculés en prenant la moyenne des variations des probabilités prédites individuelles, sans impact sur les tendances des résultats. Cette mesure rend commensurable les variations de probabilités prédites attribuables à chaque variable explicative et pour chaque trajectoire d'activité, indépendamment de la taille de ces groupes. Sur l'exploitation des probabilités prédites pour les modèles logit voir Long, 1997 ; Pryanishnikov & Zigova, 2003 ; Stratton et al., 2008 ; Peng & Nichols, 2003 ; Wulff, 2015.

Figure X – Effets marginaux des variables du modèle de classification



Lecture : les entreprises ayant développé la vente en ligne depuis le début de la crise ont 1.38 fois plus de chances d'appartenir au profil « non affecté » que les autres entreprises. Autrement dit, l'effet marginal du développement de la vente en ligne sur l'appartenance au profil « non affecté » est de +38 %. Formellement :

$$\text{Effet}_{\text{Venteenligne, Profil}_1} = \frac{\mathbb{P}(\text{Profil}_1 | X_{\text{Venteenligne}} = 1, X_{-\text{Venteenligne}}) - \mathbb{P}(\text{Profil}_1 | X_{\text{Venteenligne}} = 0, X_{-\text{Venteenligne}})}{\mathbb{P}(\text{Profil}_1 | X_{\text{Venteenligne}} = 0, X_{-\text{Venteenligne}})} = 0.38$$

Source : DGFIP, déclarations TVA ; Insee, enquête Impact de la crise sanitaire sur l'organisation de l'activité des entreprises. Calculs des auteurs.

Le fait d'exporter est quant à lui associé à une probabilité plus élevée d'appartenir au profil « déprimé », probablement du fait de la dépendance aux débouchés étrangers et de la chute de la demande extérieure. Le développement de nouveaux produits et systèmes de ventes suite à la crise est associé à une probabilité plus élevée d'appartenir au profil « non affecté », et plus faible d'appartenir aux profils les plus touchés. Il en va de même pour la réalisation d'investissements spécifiques dans de nouvelles technologies, en particulier numériques. La capacité à s'adapter aux restrictions sanitaires, en particulier celles concernant la manière d'organiser l'offre et la distribution de ses produits, a donc été importante.

La réorganisation de l'activité et la mise en commun de ressources avec d'autres entreprises sont liées à une plus forte probabilité d'appartenir à la fois aux profils « non affecté » et « déprimé ». Les entreprises ayant rapidement adapté leur activité ont pu maintenir leur niveau de chiffre d'affaires. Au contraire, la mise en commun de ressources a pu être contrainte *a posteriori* pour les entreprises les plus touchées, expliquant un effet marginal positif dans le profil « déprimé » par un mécanisme de causalité inverse.

Ces résultats permettent une exploration plus détaillée des variables corrélées à l'hétérogénéité des situations observées. En particulier, si le secteur est bien le facteur explicatif principal des trajectoires de chocs des entreprises, les corrélations observées avec d'autres de leurs caractéristiques, notamment leur stratégie d'adaptation pendant la crise, permettent de mieux comprendre la dispersion observée.

* *
*

Un résultat saillant de cette analyse est la forte prévalence des entreprises ayant traversé la crise sans s'éloigner du niveau de développement qu'elles auraient pu connaître hors crise. Les pertes agrégées d'activité sont importantes mais masquent deux dimensions de la crise. D'une part, toutes les entreprises n'ont pas connu de pertes d'activité, et d'autre part, même si la plupart des entreprises n'a pas su compenser le choc initial, une part substantielle d'entre elles a su redresser sa trajectoire d'activité pour approcher, voire dépasser, le scénario contrefactuel. Il est ainsi particulièrement notable que le groupe des entreprises dites « non affectées » rassemble plus d'un tiers des entreprises et des salariés,

soit plus que les entreprises « confinées » et « déprimées ». Affiner la compréhension des conséquences de la crise nécessite d'identifier les entreprises s'en étant mieux sorties, à l'autre extrémité du spectre comptant des entreprises ayant pratiquement cessé leur activité à partir du mois de mars. De ce point de vue, les adaptations organisationnelles, particulièrement les investissements dans de nouvelles technologies, sont importantes tant elles sont corrélées aux trajectoires d'activité préservées et semblent avoir permis d'atténuer partiellement les difficultés liées à certaines mesures de restriction sanitaire définies au niveau sectoriel. Le fait que la capacité à mettre en pratique des stratégies d'adaptation organisationnelles après le déclenchement de la crise ait pu être inégale entre les entreprises soulève la question du rôle de cette dernière dans l'exacerbation ou l'atténuation des situations observées avant la crise. Autrement dit, les pertes d'activité ont-elles été plus prononcées pour les entreprises qui étaient déjà en difficulté lorsque la crise s'est déclenchée ?

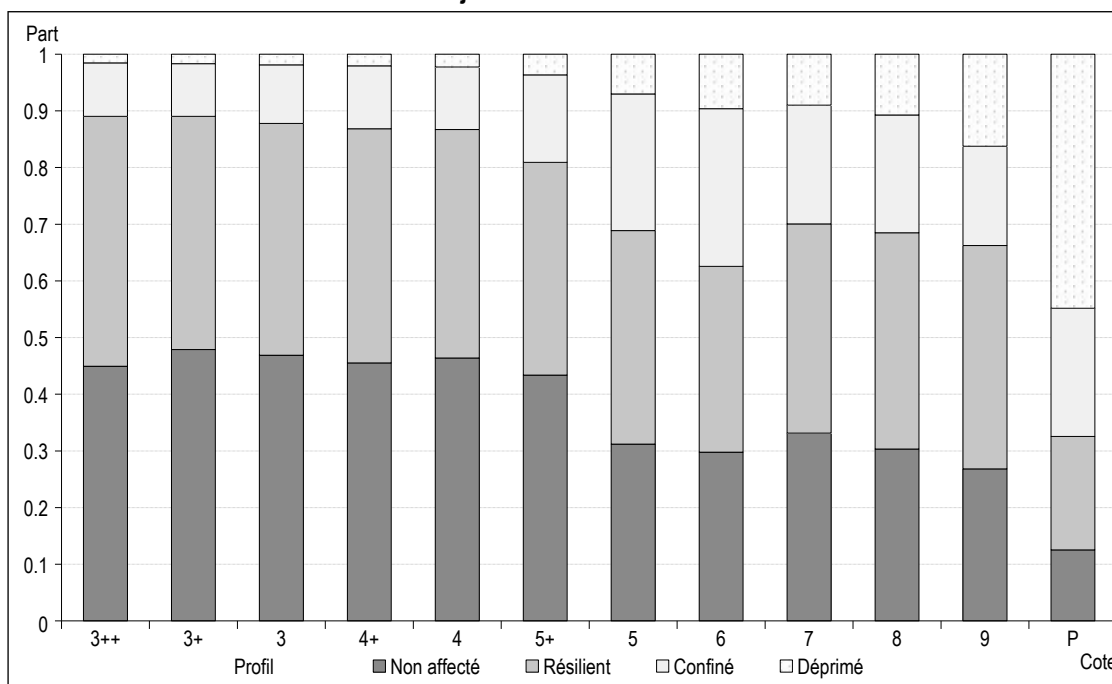
La cotation Banque de France des entreprises évalue les risques associés aux prêts accordés aux entreprises en estimant la capacité de ces dernières à honorer leurs engagements financiers à un horizon de trois ans³¹ et offre donc un indicateur de la santé financière des entreprises avant la crise³². En croisant cet indicateur avec la catégorie de trajectoire d'activité connue par l'entreprise en 2020, on remarque que les entreprises les mieux cotées sont plus souvent classées parmi les trajectoires préservées (non affectées et résilientes) alors que les entreprises considérées comme fragiles avant la crise ont plus souvent connu des trajectoires fortement affectées (confinées et déprimées). Les cotes vont de 3++, pour les entreprises dont la capacité à faire face à leurs engagements est jugée excellente, à P pour les entreprises en procédure collective (i.e. en redressement ou liquidation judiciaire)³³. Parmi les entreprises les mieux notées (3++) à fin 2019, 45 % ont été « non affectées » (figure XI).

31. La cotation est celle du 31 décembre 2019. Pour les groupes profilés, la cotation de la tête de groupe, renseignée dans FARE, est retenue. Si le Siren de la tête de groupe n'est pas connu, on retient la cotation de l'unité légale présentant la plus forte valeur ajoutée au sein de l'entreprise profilée.

32. La cotation est effectuée par la Banque de France sur la base de l'analyse des informations comptables, financières et judiciaires des entreprises, de leurs éventuels incidents de paiements sur effets de commerce et d'informations qualitatives remontées par les chefs d'entreprises.

33. Un certain nombre d'entreprises ne sont pas cotées et se voient attribuer la cote 0. Ce sont les entreprises pour lesquelles la Banque de France ne dispose pas de documentation comptable récente ou n'a recueilli aucune information défavorable d'incident de paiement sur effets, de décision ou d'information judiciaire. Ces cotes sont exclues des ventilations présentées, mais représentent une part assez stable entre les différentes catégories de trajectoire.

Figure XI – Ventilation des entreprises par cotation Banque de France au 31 décembre 2019 et trajectoire d'activité en 2020



Lecture : parmi les entreprises cotées 3++ au 31 décembre 2019, 45 % appartiennent au profil « non affecté ».
 Source : DGFIP, déclarations TVA ; cotations Banque de France. Calcul des auteurs.

Cette part décroît à mesure que l'on descend dans les niveaux de cotation pour atteindre 30 % pour les entreprises cotées 5 et 12.5 % pour celles cotées P. Ce gradient s'inverse lorsque l'on considère les entreprises « déprimées », puisqu'elles représentent 1.5 % des entreprises cotées 3++, valeur la plus faible parmi toutes les cotes. Cette part augmente à mesure que les cotations diminuent : elle atteint 7 % pour les entreprises cotées 5, 16 % de celles cotées 9 et 45 % de celles cotées P.

Ces cotes peuvent être corrélées au secteur d'activité ou à d'autres caractéristiques des entreprises comme leur taille ou leur âge, voire leur capacité à adapter leurs comportements et leur organisation pendant la crise (Bureau *et al.*, 2021a). Les statistiques présentées sont descriptives et ne doivent pas être analysées indépendamment des résultats du modèle de classification présenté³⁴ mais elles portent toutefois un enseignement complémentaire : les entreprises dont la trajectoire s'est le plus éloignée du niveau de développement qui aurait été attendu en 2020 sont les entreprises qui étaient déjà vulnérables avant le début de la pandémie. Autrement dit, l'impact strictement attribuable à la crise a été plus important pour les entreprises qui étaient *ex ante* fragiles. Il est donc possible que la crise ait exacerbé les différences préexistantes en pesant plus lourdement sur les entreprises déjà en difficulté³⁵.

L'approche développée dans cet article vise à dépasser les débats théoriques relatifs à la crise pour étudier l'impact qui lui est réellement attribuable. En construisant des scénarii de référence individuels, ce travail invite à la fois à repenser les conséquences de la crise en tenant compte des trajectoires de développement que suivaient les entreprises avant la pandémie, mais également à prolonger la réflexion proposée par une modélisation des situations financières des entreprises durant la crise, incorporant à la fois les aides publiques et les adaptations des comportements des entreprises (versement de dividendes, consommations intermédiaires, investissements). Cette modélisation financière permettrait d'évaluer les besoins financiers en incorporant les montants (et pertes) d'activité réalisés en 2020 et de quantifier les montants de trésorerie nécessaires à la reprise d'un niveau d'activité cohérent avec la dynamique connue avant la crise, qui pourrait être estimée grâce aux contrefactuels de cette étude. Ces développements font l'objet de travaux supplémentaires. □

34. L'accès restreint à ces données n'a permis de travailler que sur la ventilation agrégée des entreprises par cotation et par profil de trajectoire établi par notre étude. Une inclusion dans le modèle multinomial aurait pu apporter des éléments supplémentaires.

35. Ces constats font écho aux évaluations de l'Institut des politiques publiques selon lesquelles la crise a frappé plus durement des entreprises peu productives, avec un effet sectoriel marqué (Bach *et al.*, 2020). Ici nous montrons que cet impact est plus prononcé, y compris par rapport à la trajectoire que les entreprises auraient connue sans crise. Bureau *et al.* (2022, ce numéro) montrent également que les mesures publiques de soutien n'ont pas davantage bénéficié aux entreprises les plus fragiles *ex ante*.

BIBLIOGRAPHIE

- Aghabozorgi, S., Shirkhorshidi, A. S. & Wah, T. Y. (2015).** Time-series clustering – A decade review. *Information Systems*, 53, 16–38. <http://dx.doi.org/10.1016/j.is.2015.04.007>
- Akaike, H. (1998).** Information theory and an extension of the maximum likelihood principle. In: Parzen, E., Tanabe, K., Kitagawa, G. (eds), *Selected Papers of Hirotugu Akaike*. Springer Series in Statistics. New York, NY: Springer. https://doi.org/10.1007/978-1-4612-1694-0_15
- Amoureux, V., Héam, J.-C. & Laurent, T. (2020).** Les comptes de la Nation en 2020. *Insee Première* N° 1860, mai 2021. <https://www.insee.fr/fr/statistiques/5387891>
- Anayi, L., Button, R., Walker, D., ... & Wise, T. (2020).** Financial Stability Report – Technical annex: updated estimates of the cash-flow deficit of UK companies in a COVID-19 scenario. *Bank of England*, August. <https://www.bankofengland.co.uk/financial-stability-report/2020/updated-estimates-of-the-cash-flow-deficit-of-uk-companies-in-a-covid-19-scenario-technical-annex>
- Angrist, J. D. & Pischke, J. S. (2008).** *Mostly harmless econometrics*. Princeton University Press. <https://doi.org/10.1515/9781400829828>
- Bach, L., Ghio, N., Guillouzouic, A. & Malgouyres, C. (2020).** Impact de la crise et des mesures budgétaires 2020-2021 sur les entreprises. *Conférence sur l'impact de la crise et des mesures budgétaires 2020-2021*. Institut des Politiques Publiques, 16 novembre 2020. <https://www.ipp.eu/wp-content/uploads/2020/11/slides-ENTREPRISES-cepremap-ipp-16nov2020.pdf>
- Baldwin, R. & Di Mauro, B. W. (2020).** Economics in the time of COVID-19: A new eBook. *VOX CEPR Policy Portal*. <https://cepr.org/sites/default/files/news/COVID-19.pdf>
- Baldwin, R. & Tomiura, E. (2020).** Thinking ahead about the trade impact of COVID-19. *Economics in the Time of Covid-19*, 59. <https://cepr.org/sites/default/files/news/COVID-19.pdf>
- Balla-Elliott, D., Cullen, Z. B., Glaeser, E. L., Luca, M. & Stanton, C. T. (2020).** Business Re-Opening During the COVID-19 Pandemic. NBER, *Working paper* N° 27362. <https://www.nber.org/papers/w27362>
- Baleyte, J., Bourgeois, A., Favetto, B., Heam, J.-C., Lequien, M. & Ralle, P. (2021).** L'économie française en 2020 : une année de bouleversements. *Insee Analyses* N° 64, mai 2021. <https://www.insee.fr/fr/statistiques/5389038>
- Barrero, J. M., Bloom, N. & Davis, S. J. (2020).** COVID-19 is also a reallocation shock. NBER, *Working paper* N° 27137. <https://www.nber.org/papers/w27137>
- Barrero, J. M., Bloom, N., Davis, S. J. & Meyer, B. H. (2021).** COVID-19 Is a Persistent Reallocation Shock. *AEA Papers and Proceedings*, Vol. 111, 287–291. <http://dx.doi.org/10.29338/wp2021-03>
- Bartik, A. W., Bertrand, M., ..., Luca, M. & Stanton, C. (2020).** The impact of COVID-19 on small business outcomes and expectations. *Proceedings of the national academy of sciences*, 117(30), 17656–17666. <https://doi.org/10.1073/pnas.2006991117>
- Berndt, D. J. & Clifford, J. (1994).** Using Dynamic Time Warping to Find Patterns in Time Series. *KDD Workshop*, Vol. 10, 359–370. <https://www.aaai.org/Papers/Workshops/1994/WS-94-03/WS94-03-031.pdf>
- Bignon, V. & Garnier, O. (2020).** Mesurer l'impact de la crise Covid-19. *Revue de l'OFCE*, 166(2), 45–57. <https://doi.org/10.3917/reof.166.0045>
- Blanchard, O., Philippon, T. & Pisani-Ferry, J. (2020).** *A new policy toolkit is needed as countries exit COVID-19 lockdowns*. Bruegel Policy Contribution 12. <https://www.econstor.eu/bitstream/10419/237647/1/1726773043.pdf>
- Blanco, R., Mayordomo, S., Menéndez, Á. & Mulino, M. (2020).** *Spanish Non-financial Corporations' Liquidity Needs and Solvency After the COVID-19 Shock*. Banco de España. <https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSerias/DocumentosOcasiones/20/Files/do2020e.pdf>
- Bloom, N., Fletcher, R. S. & Yeh, E. (2021).** The impact of COVID-19 on US firms. NBER, *Working Paper* N° 28314. <http://www.nber.org/papers/w28314>
- Bonnet, O., Olivia, T. & Roudil-Valentin, T. (2021).** En 2020, la chute de la consommation a alimenté l'épargne, faisant progresser notamment les hauts patrimoines financiers : quelques résultats de l'exploitation de données bancaires. *Insee, Note de conjoncture*, mars 2021. <https://www.insee.fr/fr/statistiques/5232043?sommaire=5232077>
- Bourlès, L. & Nicolas, Y. (2021).** Quel est l'impact de la crise sanitaire sur les secteurs culturels au cours du premier trimestre 2021 ? Ministère de la Culture, DEPS, *Note de conjoncture*, N° 2021-3, juillet 2021. <https://www.culture.gouv.fr/content/download/292629/3354238?version=10>

- Box, G. E. & Pierce, D. A. (1970).** Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, 65(332), 1509–1526. <https://www.jstor.org/stable/2284333>
- Brancati, E. & Brancati, R. (2020).** Heterogeneous Shocks in the COVID-19 Pandemic: Panel Evidence from Italian Firms. *SSRN Electronic Journal*. <http://dx.doi.org/10.2139/ssrn.3597650>
- Brinca, P., Duarte, J. B. & Faria-e-Castro, M. (2020).** Measuring sectoral supply and demand shocks during COVID-19. *European Economic Review*, 139. <https://doi.org/10.1016/j.euroecorev.2021.103901>
- Bureau, B., Duquerroy, A., Giorgi, J., Lé, M., Scott, S. & Vinas, F. (2021a).** Une année de crise Covid : impact sur la dynamique de l'activité des entreprises en France. Une évaluation sur données individuelles. *Document de travail N°2021-002*. Insee, juillet 2021. <https://www.insee.fr/fr/statistiques/5411702>
- Bureau, B., Duquerroy, A., Giorgi, J., Lé, M., Scott, S. & Vinas, F. (2021b).** Le secteur de l'hébergement-restauration à travers la crise sanitaire de 2020. *Les entreprises en France*, coll. « Insee Références », novembre 2021. <https://www.insee.fr/fr/statistiques/5758810?sommaire=5759063>
- Bureau, B., Duquerroy, A., Giorgi, J., Lé, M., Scott, S. & Vinas, F. (2022).** Un examen au microscope de l'impact de la crise sanitaire et des mesures de soutien public sur la situation financière des entreprises. *Economie et Statistique / Economics and Statistics*, 532-33, 25–45 (ce numéro).
- Cajner, T., Crane, L. D., Decker, R. A., Grigsby, J., ... & Yildirmaz, A. (2020).** The US labor market during the beginning of the pandemic recession. NBER, *Working Paper N° 27159*. <https://doi.org/10.3386/w27159>
- Chetty, R., Friedman, J. N., Hendren, N., Stepner, M. & The Opportunity Insights Team. (2020).** *How did COVID-19 and stabilization policies affect spending and employment? A new real-time economic tracker based on private sector data*. Vol. 27431. Cambridge, MA: National Bureau of Economic Research. <http://www.nber.org/papers/w27431>
- Cox, D. R. & Snell, E. J. (1989).** *The Analysis of Binary Data*. 2nd ed. London: Chapman and Hall. <https://doi.org/10.1201/9781315137391>
- Danieli, A. & Olmstead-Rumsey, J. (2020).** Sector-specific shocks and the expenditure elasticity channel during the COVID-19 crisis. *Available at SSRN*. <http://dx.doi.org/10.2139/ssrn.3593514>
- Dauvin, M. & Sampognaro, R. (2021).** Dans les coulisses du confinement : modélisation de chocs simultanés d'offre et de demande. Sciences Po OFCE. *Working Paper N° 05/2021*. <https://hal-sciencespo.archives-ouvertes.fr/hal-03374280/document>
- Davis, S. J. & Haltiwanger, J. (1992).** Gross job creation, gross job destruction, and employment reallocation. *The Quarterly Journal of Economics*, 107(3), 819–863. <https://doi.org/10.2307/2118365>
- Deville, J. C. & Särndal, C. E. (1992).** Calibration estimators in survey sampling. *Journal of the American statistical Association*, 87(418), 376–382. <https://doi.org/10.1080/01621459.1992.10475217>
- Dickey, D. A. & Fuller, W. A. (1979).** Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), 427–431. <https://doi.org/10.2307/2286348>
- Duc, C. & Souquet, C. (2020).** L'impact de la crise sanitaire sur l'organisation et l'activité des sociétés. *Insee Première N° 1830*, décembre 2020. <https://www.insee.fr/fr/statistiques/4994488>
- Fairlie, R. (2020).** The impact of COVID-19 on small business owners: Evidence from the first three months after widespread social-distancing restrictions. *Journal of Economics & Management Strategy*, 29(4), 727–740. <https://dx.doi.org/10.1111%2Fjems.12400>
- Forsythe, E., Kahn, L. B., Lange, F. & Wiczer, D. (2020).** Labor demand in the time of COVID-19: Evidence from vacancy postings and UI claims. *Journal of Public Economics*, 189, 104238. <https://doi.org/10.1016/j.jpubeco.2020.104238>
- Insee (2020).** En avril 2020, chute inédite de la confiance des ménages. *Informations Rapides N° 2020-106*, 28 avril 2021. <https://www.insee.fr/fr/statistiques/4483357>
- Gerschel, E., Martinez, A. & Mejean, I. (2020).** Propagation of shocks in global value chains: the coronavirus case. *Notes IPP*, (53). <https://www.ipp.eu/wp-content/uploads/2020/03/n53-notesIPP-march2020.pdf>
- Gibbons, S., Overman, H. G. & Pelkonen, P. (2014).** Area disparities in Britain: Understanding the contribution of people vs. place through variance decompositions. *Oxford Bulletin of Economics and Statistics*, 76(5), 745–763. <https://doi.org/10.1111/obes.12043>
- Glaeser, E. L., Jin, G. Z., Leyden, B. T. & Luca, M. (2021).** Learning from deregulation: The asymmetric impact of lockdown and reopening on risky behavior during COVID-19. *Journal of Regional Science*, 61(4), 696–709. <https://doi.org/10.1111/jors.12539>
- Gourinchas, P. O., Kalemli-Özcan, Ş., Penciakova, V. & Sander, N. (2020).** *Estimating SME failures in real time: an application to the COVID-19 crisis*. NBER, *Working Paper N° 27877*. <https://doi.org/10.3386/w27877>

- Gourinchas, P. O., Kalemli-Özcan, Ş., Penciakova, V. & Sander, N. (2021).** COVID-19 and SMEs: A 2021 “time bomb”? NBER, *Working paper* N° 28418.
<https://doi.org/10.3386/w28418>
- Gupta, S., Nguyen, T., Raman, S., ... & Wing, C. (2021).** Tracking public and private responses to the COVID-19 epidemic: evidence from state and local government actions. *American Journal of Health Economics*, 7(4), 361–404. <https://doi.org/10.1086/716197>
- Hadjibeyli, B., Roulleau, G. & Bauer, A. (2021).** Live and (don't) let die : the impact of COVID-19 and public support on french firms. Direction générale du Trésor, *Document de travail* N° 2021-2.
<https://www.tresor.economie.gouv.fr/Articles/2021/04/20/live-and-don-t-let-die-the-impact-of-covid-19-and-public-support-on-french-firms-1>
- Hale, T., Angrist, N., Goldszmidt, R., ... & Tatlow, H. (2021).** A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, 5(4), 529–538.
<https://doi.org/10.1038/s41562-021-01079-8>
- Helpman, E., Itzhoki, O., Muendler, M. A. & Redding, S. J. (2017).** Trade and inequality: From theory to estimation. *The Review of Economic Studies*, 84(1), 357–405. <https://doi.org/10.1093/restud/rdw025>
- Heyer, É. & Timbeau, X. (2020).** Évaluation de la pandémie de Covid-19 sur l'économie mondiale. *Revue de l'OFCE*, 166(2), 59–110. <https://doi.org/10.3917/reof.166.0059>
- Ljung, G. M. & Box, G. E. (1978).** On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303 <https://doi.org/10.2307/2335207>
- Long, J. S. (1997).** Regression models for categorical and limited dependent variables. *Advanced Quantitative Techniques in the Social Sciences*, 7.
- OECD (2020).** *Corporate Sector Vulnerabilities During the COVID-19 Outbreak: Assessment and Policy Responses*. OECD Publishing.
<https://www.oecd.org/coronavirus/policy-responses/corporate-sector-vulnerabilities-during-the-covid-19-outbreak-assessment-and-policy-responses-a6e670ea/>
- Peng, C. Y. J. & Nichols, R. N. (2003).** Using multinomial logistic models to predict adolescent behavioral risk. *Journal of Modern Applied Statistical Methods*, 2(1), 16. <https://doi.org/10.22237/jmasm/1051748160>
- Pryanishnikov, I. & Zigova, K. (2003).** Multinomial logit models for the Austrian labor market. *Austrian Journal of Statistics*, 32(4), 267–282. <https://doi.org/10.17713/ajs.v32i4.461>
- Ratanamahatana, C. A. & Keogh, E. (2004).** Everything you know about dynamic time warping is wrong. *Third workshop on mining temporal and sequential data*.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.432.4253&rep=rep1&type=pdf#page=53>
- Rebecq, A. (2016).** Icarus: un package R pour le calage sur marges et ses variantes. *9^e colloque francophone sur les sondages*, Gatineau (Canada). http://paperssondages16.sfds.asso.fr/submission_54.pdf
- Sardá-Espinosa, A. (2019).** Time-series clustering in r using the dtwclust package. *The R Journal*, RJ-2019-023. <https://doi.org/10.32614/RJ-2019-023>
- Schivardi, F. & Guido, R. (2020).** A simple method to estimate firms' liquidity needs during the COVID-19 crisis with an application to Italy.
<https://iris.luiss.it/retrieve/handle/11385/198776/104683/SchivardiRomano.pdf>
- Sears, J., Villas-Boas, J. M., Villas-Boas, V. & Villas-Boas, S. B. (2020).** Are We #StayingHome to Flatten the Curve? *Department of Agricultural and Resource Economics*, first version April, 5.
<http://dx.doi.org/10.2139/ssrn.3569791>
- Škare, M., Soriano, D. R. & Porada-Rochoń, M. (2021).** Impact of Covid-19 on the travel and tourism industry. *Technological Forecasting and Social Change*, 163, 120469. <https://doi.org/10.1016/j.techfore.2020.120469>
- Stratton, L. S., O'Toole, D. M. & Wetzel, J. N. (2008).** A multinomial logit model of college stopout and dropout behavior. *Economics of Education Review*, 27(3), 319–331. <http://dx.doi.org/10.1016/j.econedurev.2007.04.003>
- Wulff, J. N. (2015).** Interpreting results from the multinomial logit model: Demonstrated by foreign market entry. *Organizational Research Methods*, 18(2), 300–325. <https://doi.org/10.1177/1094428114560024>
-

CONSTRUCTION DES SÉRIES DE CHIFFRE D'AFFAIRES

La formule d'estimation du chiffre d'affaires à partir des données de déclaration de la TVA est la suivante :

$$CA_{i,t} = CAF_{i,t} + CAE_{i,t}$$

$$\Leftrightarrow CA_{i,t} = (BI_{i,t} - AA_{i,t} - AOI_{i,t} + a * b * AONI_{i,t}) + (UE_{i,t} + HUE_{i,t} + a * (1 - b) * AONI_{i,t})$$

avec a et b fixés par défaut à 1.

Tableau A1 – Dictionnaire des variables issues des déclarations de TVA

| | |
|---|---|
| Base imposable hors taxe France (BI) | Opérations réalisées en France à taux normal 20 % + en France métropolitaine à taux réduit 5.5 % + en France métropolitaine à taux réduit 10 % + dans les DOM à taux normal 8.5 % + dans les DOM à taux réduit 2.1 % + anciens taux + opérations imposables à un taux particulier |
| Achats autoliquidés (AA) | Achats de prestations de services intracommunautaires + Importations + Acquisitions intracommunautaires + Livraisons d'électricité, de gaz naturel, de chaleur ou de froid imposables en France + Achats de biens ou de prestations de services réalisés auprès d'un assujetti non établi en France |
| Autres opérations imposables (AOI) | Autres opérations imposables |
| Autres opérations non imposables (AONI) | Autres opérations non imposables |
| Exportations vers l'Union européenne (UE) | Livraisons intracommunautaires à destination d'une personne assujettie – Ventes BtoB + Livraisons d'électricité, de gaz naturel, de chaleur ou de froid non imposables en France |
| Exportations hors de l'Union européenne (HUE) | Exportation hors UE |

ANNEXE 2

RÉSULTATS DU MODÈLE DE CLASSIFICATION

Tableau A2 – Résultats de la régression du modèle mlogit

| Variables indépendantes | Variable dépendante | | | | | |
|---|---------------------|----------|----------------------|----------|------------------------|----------|
| | Profil « confiné » | | Profil « résilient » | | Profil « non affecté » | |
| Taille : TPE | 0.330*** | (0.027) | 0.786*** | (0.044) | -0.178*** | (0.054) |
| Taille : PME | 0.224*** | (0.027) | 0.447*** | (0.044) | 0.090* | (0.054) |
| Date de création : avant 1997 | 0.385*** | (0.009) | 0.251*** | (0.013) | -0.020 | (0.022) |
| Date de création : entre 1998 et 2006 | 0.481*** | (0.009) | 0.780*** | (0.012) | 0.091*** | (0.020) |
| Date de création : entre 2007 et 2012 | 0.343*** | (0.009) | -0.098*** | (0.012) | 0.229*** | (0.019) |
| Activité exportatrice | 0.108*** | (0.010) | 0.103*** | (0.014) | 0.566*** | (0.027) |
| Dév. de la vente en ligne | -0.598*** | (0.013) | -1.375*** | (0.022) | -1.304*** | (0.037) |
| Dév. de nouveaux systèmes de livraison | -0.525*** | (0.013) | -1.049*** | (0.022) | -0.703*** | (0.036) |
| Dév. de nouveaux produits/activités/services | 0.035*** | (0.011) | 0.219*** | (0.016) | -3.410*** | (0.063) |
| Investissement dans de nouvelles technologies | -0.553*** | (0.020) | -1.141*** | (0.029) | -1.784*** | (0.042) |
| Réorganisation de l'activité | -0.632*** | (0.013) | -0.797*** | (0.021) | -0.194*** | (0.035) |
| Mise en commun de moyens | -0.068*** | (0.013) | -1.057*** | (0.022) | 0.283*** | (0.026) |
| Effectifs en télétravail | -0.004*** | (0.0002) | -0.008*** | (0.0002) | -0.009*** | (0.0004) |
| Constante | -7.878*** | (0.038) | -7.120*** | (0.057) | -7.748*** | (0.072) |
| AIC | 1 199 764.000 | | 1 199 764.000 | | 1 199 764.000 | |
| N | | | 13 426 | | | |

*p<0.1 ; **p<0.05 ; ***p<0.01.

Source : DGFIP, déclarations TVA ; Insee, enquête Impact de la crise sanitaire sur l'organisation de l'activité des entreprises. Calculs des auteurs.

