

## Campagne de l'IPI de mars 2020

### Contexte

La crise sanitaire et les mesures de confinement en lien avec le Covid-19 ont des impacts très importants sur la mesure de statistiques conjoncturelles comme l'indice de la production industrielle (IPI). Ces statistiques sont d'autant plus importantes pour comprendre les phénomènes en jeu dans l'économie sur la période.

Pour rappel, le calcul des indices de la production industrielle répond à un impératif national et européen. (règlement européen sur les statistiques de court terme). L'IPI a pour objet de suivre l'évolution mensuelle de l'activité industrielle et de la construction, à la fois à un niveau agrégé et à un niveau fin (jusqu'au niveau sous-classe de la NAF en données brutes). Il représente à ce titre une information primordiale pour le suivi de l'état de l'économie en France et est également un intrant essentiel pour l'élaboration des comptes trimestriels français.

Cette fiche complète les documents méthodologiques habituels (fiche méthodologique, Insee méthode n°133), sur certains points particuliers. Les enjeux spécifiques auxquels il a fallu prêter une attention particulière ce mois sont ainsi : la gestion de la non-réponse, nettement plus importante qu'habituellement, l'expertise des imputations automatiques / leur ajustement si nécessaire, le contrôle des modèles CVS-CJO.

### Taux de réponse

Au moment de la diffusion de l'indice, le taux de réponse brut s'établit à un peu moins de 65 % sur l'industrie manufacturière (contre plus de 90 % habituellement). En données pondérées – c'est-à-dire en tenant compte de la taille des entreprises répondantes et du poids des séries – et après intégration des séries collectées par d'autres organismes (ministère de l'agriculture, organismes professionnels agréés), le taux de réponses sur l'industrie manufacturière s'établit à 76 %, soit un niveau sensiblement plus élevé que le taux brut, et acceptable pour estimer de façon satisfaisante le niveau de l'activité industrielle en mars 2020, même si des révisions sur cette estimation apparaissent probables lors des prochaines publications (*cf. infra*).

Dans le détail, près de 90 % des séries au niveau classe de la NAF ont un taux de réponse supérieur à 50 % et 65 % des séries ont un taux de réponse supérieur à 70 %. Si l'on tient compte de la taille des séries (leur poids dans la valeur ajoutée industrielle), plus de 95 % de la valeur ajoutée est couverte par des taux de réponses supérieur à 50 % et près des trois quarts est couverte avec un taux de réponse supérieur à 70 %.

Afin de favoriser l'atteinte de taux de réponses suffisants, des relances très ciblées - sur quelques grandes entreprises (environ 150) réparties sur les branches les plus stratégiques (du fait de leur poids et/ou d'un faible taux de réponse initial observé) - ont été réalisées, en conformité avec la politique de forte limitation des relances mise en place par l'Insee dans la crise actuelle<sup>1</sup>.

### Contrôle et ajustement des imputations

Les entreprises non-répondantes pour une campagne d'IPI donnée font habituellement l'objet d'une estimation de leur production à l'aide d'une méthode d'imputation automatique (*cf. Insee méthodes n°133 pour une description de la formule*). Cette méthode donne de bons résultats en temps normal mais est moins bien adaptée en présence de taux de réponses réduits, ce qui était le cas pour un certain nombre de séries (*cf. supra*), d'autant plus dans la mesure où cela était couplé à des évolutions atypiques. C'est pourquoi les imputations ont fait l'objet d'une attention particulière.

Pour ce faire, l'Insee s'est appuyé sur d'autres données permettant de contrôler la qualité des imputations et de les ajuster si nécessaire. Il s'agit notamment des données de chômage partiel et d'arrêt de travail pour motif de maladie ou de garde d'enfant issues de la déclaration sociale nominatives (DSN) qui ont permis d'estimer l'importance du choc subi par chaque entreprise non-répondante. Dans une moindre

1 L'échantillon des enquêtes mensuelles de branches destiné à calculer l'IPI comporte un peu moins de 5 000 entreprises.

mesure, l'Insee s'est également appuyé sur des données de consommation d'électricité pour certains secteurs très consommateurs d'énergie fournies par RTE.

Au total, la quasi-totalité des imputations ont fait l'objet d'une confrontation avec l'une ou l'autre de ces sources externes, environ 250 identifiées à l'aide d'une méthode de priorisation (fonction du poids de la branche étudiée et d'un score de plausibilité de l'imputation) ont fait l'objet d'un contrôle manuel approfondi, et un peu moins d'une centaine ont fait l'objet d'un ajustement.

## Indice de production dans la construction

Les données sur lesquelles s'appuie habituellement l'Insee pour élaborer l'indice de production dans la construction (enquêtes de branches menées par la fédération française du bâtiment et la fédération nationale des travaux publics) n'ont pas pu être traitées ou collectées, à la fois sur février et mars. Afin de pouvoir produire une estimation, l'Insee et le service statistique du ministère du Développement durable (SDES) se sont appuyés sur la DSN (*cf. supra*) et une enquête des fédérations ciblée sur le fonctionnement des entreprises du secteur de la construction pendant la période de confinement.

## Corrections des variations saisonnières et pour jours ouvrables

Les séries sont corrigées des variations saisonnières (CVS) et des effets de calendrier (CJO). L'estimation de ces effets est effectuée avec la méthode X13-Arima sous JDemetra+ (disponible sur le site d'Eurostat) au niveau des classes de la NAF rév. 2. Les indices CVS-CJO des postes des niveaux supérieurs sont obtenus par agrégation des indices CVS-CJO des classes qui les composent.

La forte baisse de la plupart des séries au mois de mars ne doit cependant pas être interprétée par les modèles de CVS comme une baisse du niveau d'activité moyen d'un mois de mars et ne doit donc pas modifier les coefficients saisonniers estimés sur le passé.

Une façon de résoudre ce problème sans remettre en question les modèles habituels est l'introduction d'« outlier » dans les modèles CVS-CJO pour les séries présentant un profil saisonnier atypique un mois donné (*cf. Insee méthodes n°133* explicitant différents types d'outlier et note d'Eurostat édictant des préconisations dans la crise actuelle). Cette méthode permet alors de faire en sorte que l'évolution particulière de ce mois de mars ne soit pas intégrée à l'estimation des coefficients saisonniers et ne biaise ainsi pas à la hausse l'estimation de la chute d'activité liée à la crise actuelle.

X13-Arima permet de détecter une bonne partie de ces points atypiques de façon automatique (en s'appuyant sur des statistiques de test et des seuils de détection adaptés). Cependant, pour certaines séries, la détection peut être mise en défaut, ce qui a nécessité une vérification manuelle systématique des séries restantes pour savoir si un « additive outlier » devait être ajouté.

## Révisions

Comme évoqué plus haut, les données non disponibles au moment de la publication de l'indice ont fait l'objet d'une estimation. Avec l'arrivée et l'intégration de réponses retardataires dans l'indice au cours des mois suivants, et en dépit des mesures particulières mises en place pour affiner l'estimation des données manquantes, l'indice du mois de mars pourrait être révisé plus fortement qu'habituellement. En effet, si ce mécanisme de révision est normal (le taux de réponse du dernier mois n'est jamais à 100 % au moment de la publication), il prend ce mois une importance particulière avec la nette baisse du taux de réponse. Par ailleurs, en dépit des précautions prises sur les modèles utilisés pour corriger des variations saisonnières et des effets de calendrier, ceux-ci pourraient également être source de révisions de façon plus importante qu'habituellement.