

2023

00

Bases
statistiques
et généralités

Neuchâtel 2023

Méthodes de correction des effets calendrier en usage à l'OFS, en 2023

Construction, choix et application des régresseurs mis à disposition

Rapports de méthodes de la section Méthodes statistiques

Les rapports de méthodes décrivent les méthodes mathématiques et statistiques à la base des résultats et des analyses de la statistique publique. Ils présentent également l'évaluation et le développement de nouvelles méthodes en vue d'une application future. Ces publications visent d'une part à documenter les méthodes utilisées ou envisagées dans un souci de transparence et de rigueur scientifique, et d'autre part à favoriser la collaboration avec le monde scientifique et universitaire.

Les résultats numériques présentés dans les rapports de méthodes illustrent les concepts mathématiques décrits, mais ne sont pas des résultats officiels des enquêtes concernées. De même, les méthodes réellement appliquées peuvent différer légèrement de celles décrites dans ces rapports.

Les rapports de méthodes sont disponibles sous forme électronique sur le site Internet de l'OFS.

Méthodes de correction des effets calendrier en usage à l'OFS, en 2023

Construction, choix et application des régresseurs mis à disposition

Rédaction Clément Chevalier, OFS
Éditeur Office fédéral de la statistique (OFS)

Neuchâtel 2023

Éditeur: Office fédéral de la statistique (OFS)

Renseignements: Clément Chevalier, OFS, tél. +41 58 481 00 19
clement.chevalier@bfs.admin.ch

Rédaction: Clément Chevalier, OFS

Contenu: Section méthodes statistiques, OFS

Série: Statistique de la Suisse

Domaine: 00 Bases statistiques et généralités

Langue du texte original: français

Mise en page: Publishing et diffusion PUB, OFS
Ce document a été produit automatiquement à partir de banques de données. Il ne répond donc pas aux normes typographiques des publications de l'OFS.

En ligne: www.statistique.ch

Imprimés: www.statistique.ch
Office fédéral de la statistique, CH-2010 Neuchâtel,
order@bfs.admin.ch, tél. +41 58 463 60 60
Impression réalisée en Suisse

Copyright: OFS, Neuchâtel 2023
La reproduction est autorisée, sauf à des fins commerciales, si la source est mentionnée.

Numéro OFS: 338-0085

ISBN: 978-3-303-00720-4

Table des matières

Préambule et Résumé	5
1 Introduction	7
2 Le calendrier suisse	8
2.1 Jours fériés	8
2.2 Règle de calcul des dates des jours fériés variables	9
2.3 Utilisation et modification des programmes R de construction du calendrier	10
3 Calcul des régresseurs selon plusieurs hypothèses	11
3.1 Modèle de base avec jours fériés	11
3.2 Désaisonnalisation des régresseurs : modèle de référence à 14 régresseurs	12
3.3 Réduction du nombre de régresseurs	13
3.3.1 Option 1 : effet des dimanches et dimanches fériés	13
3.3.2 Option 2 : effet des dimanches et de tous les jours fériés	14
3.3.3 Option 3 : effet des lundis à vendredis non fériés	14
3.3.4 Option 4 : effet des week-ends et jours fériés	15
3.4 Calcul des régresseurs en R et fichiers à disposition	15
4 Choix des régresseurs	17
4.1 Outliers et stabilité du modèle	17
4.2 Critères de type AIC ou BIC	18
4.3 Tests statistiques	18
4.4 Spectre de la série corrigée	20
4.4.1 Définition et estimation du spectre	20
4.4.2 Fréquences de jours ouvrables et tests visuels sur le spectre	22
4.5 Interactions entre ces outils	23
5 Exemple et résultats	24
5.1 Série d'étude, code R et décomposition	24
5.2 Résultats avec l'option numéro 2	25
5.3 Résultats avec les options 1 et 3	26
6 Conclusions	29
Bibliographie	30
Annexe	31
A Code R d'appel à X13-ARIMA-SEATS	31
Rapports de méthodes de la section méthodes statistiques de l'OFS	33

Préambule

L'Office Fédéral de la Statistique (OFS) publie des statistiques mensuelles et trimestrielles dont les valeurs peuvent partiellement dépendre de causes connues et interprétables, telles que la saisonnalité ou la présence de jours fériés. Ce rapport se focalise sur la quantification et l'estimation des effets calendrier des séries temporelles. Il s'agit ici de prendre en compte la part des variations de la série qui est purement liée à la variation du calendrier. Pour la relecture attentive de ce rapport, l'auteur remercie D. Assoulin, D. Kilchmann, D. Meskaldji et J.-P. Renfer de la section méthodes statistiques (METH) de l'OFS, ainsi que Christian Furger de la section KE.

Résumé

La correction des effets calendrier d'une série temporelle consiste à prendre en compte la part des variations de la série qui est purement liée à la variation du calendrier. Par exemple, le nombre de jours d'un mois donné peut avoir une influence sur les valeurs d'une série temporelle représentant un chiffre d'affaires cumulé au cours du mois. Pour certaines activités, le nombre de jours ouvrables au cours du mois, le nombre de samedis et dimanches ou encore le nombre de jours fériés peuvent avoir une importance non négligeable. Ce rapport présente les méthodes utilisées à l'Office Fédéral de la Statistique (OFS) pour effectuer ces ajustements. Cette démarche s'inscrit dans le contexte plus vaste de la désaisonnalisation des séries temporelles à l'aide d'outils tels que X13-ARIMA-SEATS [Sax & Eddelbuettel, 2018, U.S. Census Bureau, 2020]. Ces outils sont aujourd'hui bien établis et largement répandus en statistique publique, mais leur application n'est pas aisée, surtout si l'utilisateur souhaite prendre en compte les spécificités du calendrier suisse et paramétrer son influence sur la série analysée. Ce rapport vise justement à faciliter cette démarche et se focalise uniquement sur l'influence du calendrier sur la série étudiée, en vue de produire une série corrigée de ses effets calendrier. Après une brève introduction au contexte général de la désaisonnalisation et aux principaux défis liés à la correction d'effets calendrier, une construction détaillée du calendrier suisse, utilisé par l'OFS, est présentée. Il s'agit de calculer, pour chaque mois d'une période de temps donnée, le nombre de lundis, mardis, . . . , dimanches au cours du mois et le nombre de lundis, mardis, . . . , dimanches *fériés*. Le calendrier suisse permet de calculer des séries temporelles nommées "régresseurs" selon plusieurs jeux d'hypothèses concernant l'influence des différents jours de la semaine (fériés ou non) sur les valeurs de la série temporelle d'étude. Les régresseurs permettent finalement de produire la série corrigée de ses effets calendrier. Cette démarche et les différents jeux d'hypothèses implémentés sont détaillés ainsi que des références et des outils d'aide à la décision pour choisir le jeu d'hypothèses le mieux adapté à la série temporelle étudiée. Finalement, des résultats sur une série utilisée à l'OFS sont présentés et discutés à titre d'exemple. Des fichiers et du code implémenté en langage R sont mis à disposition du public sur le site <https://www.bfs.admin.ch/bfs/fr/home/dscc/outils/methodes-statistiques/2023-06-calendrier.html>, en vue de permettre aux utilisateurs d'effectuer de façon autonome des corrections d'effets calendrier sur leurs propres séries temporelles.

1 Introduction

Les séries temporelles présentent souvent des variations qui peuvent partiellement s'expliquer à l'aide de causes connues et interprétables. Pour certaines séries économiques, ces variations peuvent dépendre de la période de l'année considérée (saisonnalité) et des spécificités calendaires de la période de mesure (effets calendrier). En statistique publique, corriger ces variations saisonnières et ces effets calendrier est une étape importante pour étudier la série et interpréter ses variations à court et long terme. Des outils tels que X13-ARIMA-SEATS [Sax & Eddelbuettel, 2018] et JDemetra+ [Grudkowska et al., 2013] permettent d'estimer les effets saisonniers et les effets calendrier d'une série, en vue de produire une nouvelle série corrigée de ses effets saisonniers et/ou calendrier.

Les méthodes utilisées pour l'estimation et la correction de la composante saisonnière sont aujourd'hui assez bien établies et sont utilisables quelle que soit la région géographique de la série étudiée [Eurostat, 2018]. Ce n'est hélas pas le cas pour l'estimation et la correction des effets calendrier. Celles-ci nécessitent en effet de fixer un calendrier, avec des jours fériés qui lui sont spécifiques. Il faut par ailleurs aussi faire des choix de modélisation, notamment sur l'influence des différents jours de la semaine lorsqu'ils sont fériés et lorsqu'ils ne le sont pas. De tels choix ne sont pas simples et il n'y a à ce jour pas de consensus clair sur une procédure automatique de choix [Attal-Toubert, 2012].

Avec les options par défaut, la correction d'effets calendriers de X13-ARIMA-SEATS ou de JDemetra+ ne tient ainsi pas compte des spécificités du calendrier suisse. Elle ne permet pas non plus d'injecter d'éventuelles informations spécifiques sur l'influence des jours de la semaine.

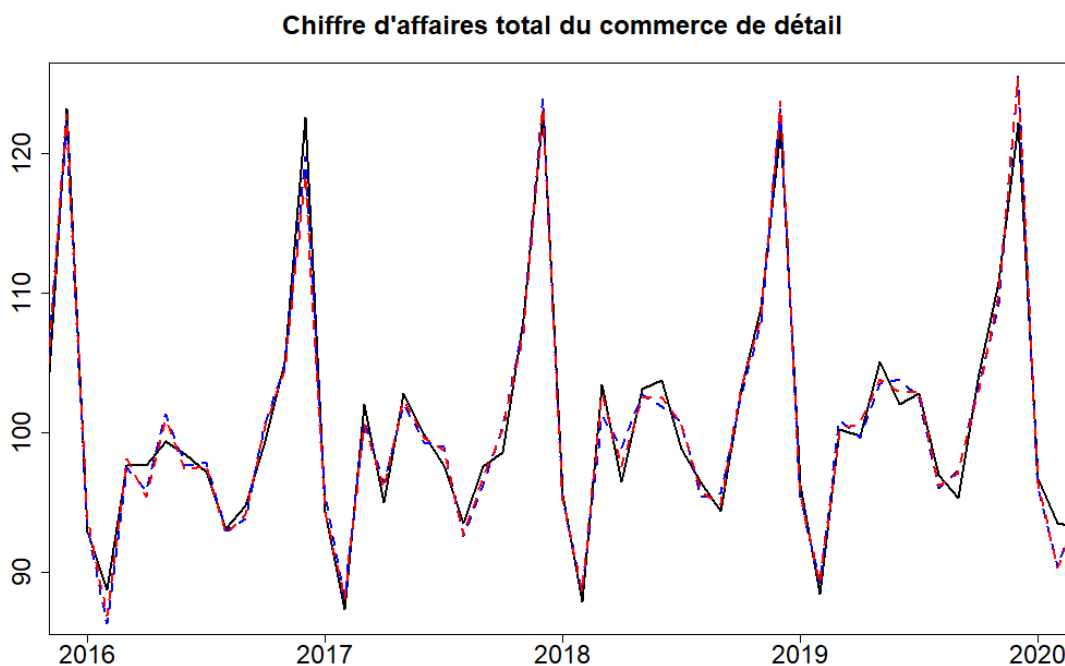


FIGURE 1

Ce rapport se focalise sur l'estimation et la correction des effets calendrier. Un exemple de correction d'effets calendrier est donné en figure 1. Le chiffre d'affaires mensuel total du com-

merce de détail au niveau national¹ (ligne noire) y est représenté sur la période 2016-2020. Les courbes pointillées rouges et bleues correspondent à la série corrigée de ses effets calendrier en utilisant le logiciel X13-ARIMA-SEATS, un calendrier spécifique à la Suisse et deux choix de modélisation. Mathématiquement, la correction d'effets calendrier d'une série temporelle mensuelle ou trimestrielle $X_t, t \geq 1$, consiste à soustraire à X_t (ou, pour un modèle multiplicatif, à diviser X_t par) la quantité :

$$S_t := \sum_{i=1}^k \xi_i Z_{it} \quad (1)$$

où les séries temporelles $Z_{1t}, \dots, Z_{kt}$ sont définies par l'utilisateur via un choix de modèle et sont appelées "régresseurs". Les coefficients $\xi_1, \dots, \xi_k$ sont des paramètres à estimer et k est le nombre de régresseurs.

L'équation (1) pose la question du choix des régresseurs $Z_{1t}, \dots, Z_{kt}$ ainsi que du nombre k de régresseurs. Dans le cadre de l'estimation d'effets calendrier, c'est-à-dire d'effets qui prennent en compte le fait que les jours de la semaine n'ont pas tous la même importance sur le phénomène observé, il peut sembler naturel de choisir comme premier régresseur Z_{1t} le nombre de lundis au cours du mois (ou du trimestre, pour une série trimestrielle) t considéré, et de procéder de même pour le nombre de mardis, ..., dimanches, de sorte à définir 7 régresseurs $Z_{1t}, \dots, Z_{7t}$. De plus, pour de nombreuses variables économiques (par exemple les ventes de détail), le fait qu'un jour soit férié peut avoir une influence importante, si bien qu'il serait pertinent de définir 7 régresseurs supplémentaires $Z_{8t}, \dots, Z_{14t}$ qui comptent le nombre de lundis, ..., dimanches fériés au cours du mois ou trimestre t . Ces idées seront utilisées par la suite, bien que les régresseurs utilisés in-fine subissent encore quelques traitements.

Si la notion de jour de la semaine (lundi, mardi, etc.) ne dépend pas de la zone géographique considérée, il faut remarquer que le calcul de régresseurs liés aux jours fériés dépend d'un choix de calendrier, lequel est spécifique à un pays, voire à une zone géographique plus petite. Ce rapport de méthode est organisé comme suit. La notion de calendrier suisse, utilisé à l'OFS, est détaillée en section 2. À partir d'un calendrier et d'un jeu d'hypothèses bien défini concernant l'influence des différents jours de la semaine sur la série, il est possible de construire les régresseurs. Ce point ainsi que plusieurs jeux d'hypothèses sont présentés en section 3. Des fichiers contenant les régresseurs calculés selon les différentes hypothèses sont publiquement accessibles ainsi que les programmes, implémentés en R, qui génèrent ces fichiers². La section 4 présente des outils d'aide à la décision quant au choix de modélisation, c'est-à-dire le choix des régresseurs. Un exemple complet d'utilisation en R est présenté et discuté en section 5.

2 Le calendrier suisse

2.1 Jours fériés

Les traitements d'ajustements calendrier d'une série temporelle nécessitent de connaître le calendrier, c'est-à-dire d'être capable de compter le nombre de lundis, mardis, ..., dimanches sur un intervalle de temps donné. De plus, il est nécessaire de compter le nombre de lundis, mardis, ..., dimanches fériés, ce qui fait intervenir le calendrier des jours fériés, qui dépend du pays ou de la région considérée. Cette section présente le calendrier suisse utilisé à l'OFS et les jours fériés de ce calendrier. Comme les jours fériés varient en fonction des cantons, les calendriers

1. <https://www.bfs.admin.ch/bfs/fr/home/statistiques/industrie-services/enquetes/dhu.html>

2. <https://www.bfs.admin.ch/bfs/fr/home/dscc/ouils/methodes-statistiques/2023-06-calendrier.html>

cantonaux peuvent différer du calendrier suisse. Les calendriers des cantons ne sont pas détaillés ici mais peuvent être utilisés à partir des programmes fournis en langage R, moyennant un effort de programmation. Ce point est expliqué en section 2.3.

Le calendrier suisse utilisé à l'OFS a un certain nombre de jours fériés fixes, liés à une date précise. Ces jours fériés sont, par ordre chronologique

- le 1er janvier : jour férié à 100%,
- le 2 janvier : jour férié à 60%,
- le 1er mai : jour férié à 40%,
- le 1er août : jour férié à 100%,
- le 25 décembre : jour férié à 100%,
- le 26 décembre : jour férié à 100%.

Le fait qu'un jour puisse être partiellement férié (fériel à moins de 100%) signifie que le jour est compté à la fois comme férié et non férié, avec des poids qui somment à 100%. Par exemple, comme le 2 janvier 2020 tombe un jeudi et qu'il s'agit du seul jeudi férié de ce mois alors pour la date $t = \text{janvier 2020}$ le régresseur Z_{11t} (nombre de jeudis fériés, voir section 1) prend la valeur 0.6 et le régresseur Z_{4t} (nombre de jeudis non fériés) prend la valeur 0.4.

En plus de jours fériés fixes, le calendrier suisse contient plusieurs jours fériés dont les dates varient d'une année à l'autre. Ces jours fériés sont dans l'ordre

- le vendredi Saint (ou "vendredi de Pâques") : férié à 100%,
- le dimanche de Pâques : férié à 100%,
- le lundi de Pâques : férié à 100%,
- le jeudi de l'Ascension : férié à 100%,
- le lundi de Pentecôte : férié à 100%.

Dans le calendrier suisse de l'OFS, les jours fériés liés à Pâques sont tous considérés comme fériés à 100% bien qu'ils ne soient pas forcément fériés dans tous les cantons.

Certaines années, il peut arriver qu'un jour partiellement férié coïncide avec un jour férié variable. C'est le cas par exemple du 1er mai 2008 (fériel à 40%) qui coïncidait avec le jeudi de l'Ascension de l'année 2008 (fériel à 100%). Dans ce cas-là, c'est le plus haut pourcentage qui prime, c'est-à-dire que ce jour est bien férié à 100%. La section suivante détaille le mode de calcul des dates des jours fériés variables.

2.2 Règle de calcul des dates des jours fériés variables

La totalité des jours fériés variables mentionnés en section 2.1 ont une date qui dépend de la date du dimanche de Pâques. C'est donc le calcul de cette date qu'il faut détailler en premier. Même si le calendrier Grégorien est périodique de période 400 ans, la date de Pâques n'est pas périodique de période 400 ans car elle dépend des cycles de la lune. Un algorithme est ainsi nécessaire au calcul de cette date. La date du dimanche de Pâques est fixée au premier dimanche après la pleine lune qui suit ou coïncide avec le 21 mars. Selon cette définition, le dimanche de Pâques a lieu au plus tôt le 22 mars (si le 21 mars est un jour de pleine lune et que c'est en plus un samedi) et au plus tard le 25 avril (si la pleine lune tombe au 20 mars, de sorte à ce que la pleine lune suivante tombe au 18 avril, et si le 18 avril est un dimanche).

Il existe plusieurs algorithmes pour calculer la date exacte du dimanche de Pâques d'une année donnée, comme les algorithmes de Butcher-Meeus, de Gauss, de Conway ou de Oudin [Bien, 2004]. Tous ces algorithmes sont rapides car ils ne nécessitent que des opérations élémentaires

(additions, soustractions, multiplications, divisions). En pratique, à l'OFS, l'algorithme de Oudin est implémenté en langage R. Une fois la date du dimanche de Pâques calculée, les autres jours fériés variables du calendrier se calculent facilement. En effet, le vendredi Saint a lieu 2 jours avant le dimanche de Pâques, le lundi de Pâques a lieu 1 jour après le dimanche de Pâques, le jeudi de l'Ascension a lieu 39 jours après le dimanche de Pâques, et le lundi de Pentecôte a lieu 50 jours après le dimanche de Pâques.

2.3 Utilisation et modification des programmes R de construction du calendrier

Cette section décrit l'utilisation des programmes R de construction du calendrier³. Comme mentionné précédemment, la version actuelle des programmes construit le calendrier suisse utilisé à l'OFS, mais le code a été pensé de sorte à permettre des modifications simples aux utilisateurs qui souhaiteraient utiliser un calendrier différent, comme par exemple le calendrier d'un canton. Un exemple est donné en fin de section.

Le fichier `BuildSwissCalendar.R` contient le code d'une fonction `BuildSwissCalendar` qui calcule le nombre de lundis, mardis, ..., dimanches au cours de chaque mois d'une période considérée, ainsi que le nombre de lundis, mardis, ..., dimanches fériés. Ce calcul se fait sur une période de `numberOFYear = 400` ans (valeur par défaut) à partir du premier janvier d'une année `beginYear` passée en argument de la fonction. La sortie (output) du programme est une matrice avec $12 * \text{numberOFYear}$ lignes et 18 colonnes, chaque ligne correspondant à un mois donné. Si le calcul se fait sur une période de 400 ans, alors il y a 4'800 lignes au total. Les 18 colonnes sont les suivantes :

- colonne 1 : date du premier jour du mois. Par exemple 01.05.2020.
- colonne 2 : numéro du mois considéré. Par exemple 5 pour le mois de mai.
- colonne 3 : numéro du trimestre considéré. Les valeurs possibles sont 1, 2, 3 ou 4.
- colonne 4 : nombre de jours du mois considéré. Les valeurs possibles sont 28, 29, 30 ou 31.
- colonnes 5 à 11 : nombre de dimanches, lundis, ..., samedis du mois considéré.
- colonnes 12 à 18 : nombre de dimanches, lundis, ..., samedis fériés du mois considéré.

Le calendrier obtenu sert d'input au calcul des régresseurs, voir section 3. Il est important de noter que, même si la série étudiée (dans le cadre d'un ajustement calendrier) est une série trimestrielle, le calendrier suisse contiendra toujours des informations mois par mois et non trimestre par trimestre. Ceci n'est pas un souci en pratique et n'empêche pas le bon déroulement des calculs de régresseurs qui eux seront bien des séries trimestrielles, voir section 3.

Les jours fériés du calendrier suisse utilisé à l'OFS sont généralement différents de ceux d'autres pays ou même des différents cantons de Suisse. Si l'utilisateur souhaite utiliser les programmes R avec un autre calendrier, des adaptations au code doivent être faites. La partie du code dédiée au calcul du nombre de lundis, ..., dimanches (féries ou pas) ne dépend pas du choix de calendrier et n'a pas besoin d'être modifiée. Seule la partie du code dédiée au calcul des jours fériés doit être modifiée. Cette partie du code est bien délimitée à l'aide de commentaires détaillés en anglais, et se situe dans la version actuelle entre les lignes de code 160 et 230 environ. Le code calcule le nombre de jours *féries* mois par mois, avec un bloc de code dédié pour chaque mois de l'année. Ce bloc est potentiellement vide, pour les mois de l'année qui ne contiennent aucun jour férié. Si des jours fériés fixes doivent être ajoutés pour un mois donné alors il faudra modifier le bloc de code relatif au calcul des jours fériés de ce mois. Par exemple, dans le canton de

3. Ces programmes peuvent être téléchargés sur la page
<https://www.bfs.admin.ch/bfs/fr/home/dscc/outils/methodes-statistiques/2023-06-calendrier.html>.

Neuchâtel, le 1er mars est férié. Pour ajouter ce jour comme férié, le bloc de code dédié à mars (vide pour le calendrier suisse) serait remplacé par le bloc suivant :

```
if(j == 3){
if(FirstDayOfMonth_j == 1) Off1 = 1
if(FirstDayOfMonth_j == 2) Off2 = 1
if(FirstDayOfMonth_j == 3) Off3 = 1
if(FirstDayOfMonth_j == 4) Off4 = 1
if(FirstDayOfMonth_j == 5) Off5 = 1
if(FirstDayOfMonth_j == 6) Off6 = 1
if(FirstDayOfMonth_j == 7) Off7 = 1
}
```

Le code ci-dessus indique ici que si le premier jour du mois $j=3$ (mars) est un dimanche, alors le mois considéré contient un dimanche férié ($Off1 = 1$: le chiffre 1 se rapportant aux dimanches, le 2 aux lundis, ..., le 6 aux samedis). L'instruction est répétée de façon similaire pour tous les jours de la semaine du lundi au samedi. Il faut ici noter que les éventuels jours fériés supplémentaires pour le mois de mars liés à Pâques sont traités de façon générique plus loin dans le code. L'ajout d'un jour férié comme le 1er mars ne pose pas de soucis dans la mesure où celui-ci ne tombe jamais à la même date que l'un des jours fériés variables mentionnés ci-dessus. Les utilisateurs qui souhaiteraient rajouter des jours fériés qui peuvent potentiellement tomber à la même date qu'un jour férié variable doivent prendre un peu plus de précaution dans leur programmation et s'inspirer, dans le code existant, du traitement qui est réservé au 1er mai.

3 Calcul des régresseurs selon plusieurs hypothèses

Cette section détaille la construction des régresseurs à l'aide de différentes modélisations possibles. Les équations et calculs sont présentés dans le cas où la série temporelle étudiée est une série mensuelle. Si la série d'étude est trimestrielle, les calculs présentés restent valables et ceux-ci continuent bien d'être effectués au niveau mensuel. Seule une opération supplémentaire pour adapter les résultats aux séries trimestrielles est nécessaire. Celle-ci est détaillée en section 3.4.

3.1 Modèle de base avec jours fériés

Soient $N_{1t}, \dots, N_{7t}$ le nombre de lundis, ..., dimanches *non fériés* au cours d'un mois t et $N_{8t}, \dots, N_{14t}$ le nombre de lundis, ..., dimanches *féries*. Soit X_t la série temporelle d'étude, sur laquelle des effets calendrier doivent être estimés.

Le modèle additif de base avec jours fériés décompose X_t comme la somme d'effets calendrier et d'une série ε_t qui n'a pas d'effets calendrier :

$$X_t = \varepsilon_t + \sum_{i=1}^{14} \alpha_i N_{it} \quad (2)$$

Dans le modèle multiplicatif de base, l'équation devient simplement

$$X_t = \varepsilon_t \sum_{i=1}^{14} \alpha_i N_{it} \quad (3)$$

Dans le modèle de base, l'effet de chaque jour non férié de la semaine et de chaque jour férié de la semaine peut varier, ce qui donne un modèle à 14 paramètres. Comme indiqué dans [Bell & Hillmer \[1983\]](#) section 4, dans un tel modèle les estimateurs des coefficients α_i sont fortement corrélés. Pour réduire la corrélation entre les estimateurs des coefficients du modèle de base, un modèle dit "avec contraste" est introduit. Celui-ci est obtenu en notant que la somme des jours fériés et non fériés d'un mois t donné, $N_{1t} + \dots + N_{14t}$, est égale au nombre de jours du mois t , noté ici N_t . Il est ainsi possible d'utiliser le modèle de base avec le régresseur N_t et les 13 régresseurs $N_{it} - N_{14t}$, $1 \leq i \leq 13$. Pour un modèle additif, l'équation (2) devient

$$X_t = \varepsilon_t + \tilde{\alpha}N_t + \sum_{i=1}^{13} \beta_i (N_{it} - N_{14t}) \quad (4)$$

Dans ce modèle, les 14 régresseurs sont le nombre de jours du mois N_t et les *contrastes* entre N_{it} ($1 \leq i \leq 13$) et le nombre de dimanches fériés N_{14t} . Les jours qui sont contrastés avec N_{14t} (dimanches fériés) sont les lundis, ..., dimanches non fériés, ainsi que les lundis, ..., samedis fériés.

3.2 Désaisonnalisation des régresseurs : modèle de référence à 14 régresseurs

Les régresseurs définis dans l'équation (4) sont des séries temporelles qui présentent une saisonnalité. Cela signifie qu'il existe des mois de l'année où ces régresseurs ont, en moyenne, une valeur plus forte ou plus faible qu'à d'autres mois de l'année. Ce point est facilement vérifiable en prenant l'exemple du régresseur N_t qui prend des valeurs toujours plus fortes en juillet (31 jours) qu'en février (28 ou 29 jours). De même, les mois de mars et avril ont tendance à contenir plus de lundis fériés que les autres mois en raison du lundi de Pâques. La saisonnalité dans les régresseurs n'est pas quelque chose de souhaitable pour des raisons détaillées dans [Attal-Toubert \[2012\]](#), section 1.3.3 et [Eurostat \[2015\]](#), section 2.3. Dans l'équation (4) l'estimation des coefficients de régression se fait à l'aide d'un modèle REG-ARIMA sur une série temporelle X_t qui peut avoir subi des transformations de type "log" et/ou des opérations de différenciation (voir [U.S. Census Bureau \[2020\]](#), section 4).

L'opération de désaisonnalisation des régresseurs transforme ces derniers en de nouveaux régresseurs qui ont la même moyenne empirique en janvier, février, etc. Soit Z_t , un régresseur du modèle de l'équation (4). La désaisonnalisation du régresseur consiste à remplacer Z_t par $Z_t - Z_T$ où Z_T est une série temporelle périodique de période 12 dont les valeurs sont les moyennes empiriques des Z_t sur chaque mois de l'année. Par exemple, pour le régresseur N_t (nombre de jours dans le mois), la désaisonnalisation du régresseur remplace N_t par $N_t - N_T$ et ce régresseur désaisonnalisé est traditionnellement nommé "Leap Year" ("année bissextile", en anglais). Pour tous les mois de l'année à l'exception de février, Leap Year prend la valeur 0 car N_t prend la même valeur que N_T . En effet, la longueur de chaque mois (excepté février) ne dépend pas de l'année considérée. Pour le mois de février N_t prend la valeur 28 ou 29 selon que l'année est bissextile ou non, d'où le nom du régresseur. Le mois de février a une longueur moyenne de 28.2425 jours car un tout petit peu moins d'une année sur quatre est bissextile dans le calendrier Grégorien. Ainsi, au mois de février, le régresseur Leap Year peut prendre la valeur $N_t - N_T = 28 - 28.2425 = -0.2425$ pour les mois de février des années non bissextiles, ou bien $N_t - N_T = 29 - 28.2425 = 0.7575$ pour les mois de février des années bissextiles. Les régresseurs désaisonnalisés sont de moyenne nulle pour chaque mois de l'année.

3.3 Réduction du nombre de régresseurs

Le modèle de référence contient les 14 régresseurs du modèle avec contraste (4) qui sont chacun désaisonnalisés. L'utilisation de ce modèle nécessite donc d'estimer 14 paramètres, ce qui peut être beaucoup selon le cas, le nombre de paramètres devant rester faible face au nombre d'observations (longueur de la série temporelle étudiée). En outre, le modèle de référence fait l'hypothèse que les lundis, . . . , dimanches non fériés et les lundis, . . . , dimanches fériés ont tous une influence différente sur la valeur de la série X_t étudiée. Cette hypothèse n'est pas forcément nécessaire et dans certains cas il peut être réaliste de considérer que certains jours fériés ou certains jours non fériés ont la même influence. Pour diminuer le nombre de paramètres à estimer, il est ainsi possible de faire un certain nombre d'hypothèses, plus ou moins fortes, qui mènent à un modèle simplifié, avec moins de 14 régresseurs.

Dans ce rapport et dans les programmes implémentés en R, l'hypothèse du modèle de référence à 14 régresseurs est notée "option 0", et peut être utilisée pour effectuer un ajustement calendrier (voir, section 5 pour des exemples). Cette option est cependant peu utilisée en pratique [Attal-Toubert, 2012, Eurostat, 2018, Quartier-La Tente & Ladiray, 2018] pour plusieurs raisons. D'une part, d'un point de vue économique il est souvent peu réaliste de supposer que les dimanches non fériés et les dimanches fériés ont un effet différent sur la série étudiée. D'autre part, comme expliqué en section 4.1, il est souhaitable d'observer une certaine stabilité du modèle au fil du temps, lorsque les nouvelles données arrivent chaque mois ou chaque trimestre. Une modélisation trop complexe de chaque jour férié peut créer des instabilités au niveau du modèle du fait des occurrences limitées de chaque type de jours fériés sur la période d'observation.

La suite de ce rapport décrit 4 jeux d'hypothèses qui mènent à des modèles qui sont progressivement de plus en plus simplifiés.

3.3.1 Option 1 : effet des dimanches et dimanches fériés

Dans cette variante du modèle de référence, les dimanches non fériés et les dimanches fériés ont les mêmes effets calendrier sur la série temporelle étudiée. Selon l'équation (4), cela revient à supposer l'égalité de coefficients $\beta_7 = \beta_{14}$, où β_{14} est le coefficient associé aux dimanches fériés, N_{14t} , dans l'équation (4). Ce coefficient est $\beta_{14} = -\sum_{i=1}^{13} \beta_i$. L'égalité entre 2 coefficients permet ainsi d'avoir un modèle avec un paramètre de moins à estimer. Le modèle avec contraste à 14 paramètres peut se réécrire de la manière suivante (voir Attal-Toubert [2012] pour les détails du calcul) :

$$X_t = \varepsilon_t + \tilde{\alpha}N_t + \sum_{i=1}^6 \beta_i \left(N_{it} - \frac{1}{2}D_t \right) + \sum_{i=8}^{13} \beta_i \left(N_{it} - \frac{1}{2}D_t \right) \quad (5)$$

où $D_t := N_{7t} + N_{14t}$ est la nouvelle variable de contraste, égale au nombre de dimanches (fériés ou non) du mois t . Le modèle de l'équation (5) possède 13 régresseurs qui sont tous désaisonnalisés. Avant désaisonnalisation, ces régresseurs sont le nombre de jours du mois N_t et les contraste entre N_{it} , $i \notin \{7, 14\}$ et le nombre de dimanches (fériés ou non) D_t . Les jours qui sont contrastés avec D_t sont les lundis, . . . , samedis non fériés, ainsi que les lundis, . . . , samedis fériés.

L'utilisation de l'option 1 de simplification est intéressante lorsque les jours fériés (hors dimanches) ont une influence potentiellement différente sur la série étudiée, c'est-à-dire si l'impact d'un jour férié dépend de sa position dans la semaine. Par exemple, dans certains secteurs, les

jours fériés peuvent avoir un impact différent du fait de la présence de week-end prolongés (jour férié le vendredi ou lundi), voire de ponts (jour férié le jeudi ou mardi).

Pour des raisons de stabilité des révisions du modèle, [Attal-Toubert \[2012\]](#) recommande cependant d'utiliser un modèle qui suppose une influence identique des jours fériés, lesquels sont alors assimilés aux dimanches. Ce modèle correspond à l'option 2 et est détaillé en section 3.3.2. Donner une influence identique pour tous les jours fériés semble en effet être la méthode la plus courante en pratique [\[Eurostat, 2018\]](#). Dans ses "guidelines", Eurostat [\[Eurostat, 2015\]](#) insiste également sur la stabilité du modèle et l'analyse des révisions. En section 4.5, il est recommandé de n'utiliser le modèle de l'option 1 que si celui-ci est économiquement "réaliste" et si le modèle de l'option 2 (qui suppose une égalité des coefficients de tous les jours fériés) ne parvient pas à supprimer correctement les effets calendrier de la série. Il est également important d'être capable d'observer suffisamment de jours fériés sur la période d'étude pour quantifier précisément leur effet [\[Quartier-La Tente & Ladiray, 2018\]](#). Pour aider à la décision, l'hypothèse d'égalité ou non de l'influence des jours fériés peut aussi être testée statistiquement, voir section 4.3.

3.3.2 Option 2 : effet des dimanches et de tous les jours fériés

Dans cette variante du modèle de référence, les dimanches (fériés ou non) et tous les jours fériés ont le même effet sur la série étudiée. Selon l'équation (4) cela revient à supposer l'égalité de coefficients

$$\beta_7 = \beta_8 = \dots = \beta_{14}. \quad (6)$$

L'égalité entre 8 coefficients permet un modèle simplifié avec 7 paramètres de moins à estimer. Le modèle obtenu s'écrit de la manière suivante (voir [Attal-Toubert \[2012\]](#) pour les détails du calcul) :

$$X_t = \varepsilon_t + \tilde{\alpha}N_t + \sum_{i=1}^6 \beta_i \left(N_{it} - \frac{1}{8}F_t \right) \quad (7)$$

où $F_t := N_{7t} + N_{8t} + \dots + N_{14t}$ est la nouvelle variable de contraste, égale au nombre de dimanches et jours fériés du mois t . Ce modèle contient 7 régresseurs qui sont tous désaisonnalisés. Avant désaisonnalisation, ces régresseurs sont le nombre de jours du mois, et les contrastes entre N_{it} , $i \leq 6$ et F_t . Les jours qui sont contrastés avec F_t sont les lundis, ..., samedis non fériés.

Comme expliqué en section 3.3.1, l'option 2 est souvent présentée comme le modèle le plus complexe (donc, avec le plus de régresseurs) utilisé en pratique [\[Attal-Toubert, 2012\]](#), ce qui signifie que les praticiens utilisent ce modèle ou bien un modèle encore plus simplifié. En section 4.5, on défend l'idée que l'option 2 de modélisation devrait être testée en premier avant d'être éventuellement simplifiée (si des modèles plus simples parviennent à correctement éliminer les effets calendrier) ou complexifiée (si l'option 2 ne parvient pas à éliminer les effets calendrier). L'option 2 de modélisation est testée sur un exemple, et comparée aux options 1 et 3, en section 5.

3.3.3 Option 3 : effet des lundis à vendredis non fériés

Dans cette variante du modèle de référence, les hypothèses de l'option 2 (c'est-à-dire que les dimanches et jours fériés ont la même influence sur la série étudiée) sont reprises. De plus, on suppose que les jours non fériés du lundi au vendredi ont la même influence sur la série étudiée,

les samedis non fériés pouvant avoir leur influence propre, différente de celle des autres jours non fériés. En plus de l'équation (6), il y a ainsi l'hypothèse suivante :

$$\beta_1 = \beta_2 = \dots = \beta_5. \quad (8)$$

Le modèle obtenu s'écrit

$$X_t = \varepsilon_t + \tilde{\alpha}N_t + \beta_1 \left(LV_t - \frac{5}{8}F_t \right) + \beta_6 \left(N_{6t} - \frac{1}{8}F_t \right) \quad (9)$$

où F_t est la variable de contraste définie dans l'équation (7) et $LV_t := N_{1t} + N_{2t} + \dots + N_{5t}$ est le nombre de lundis, ..., vendredis non fériés du mois t . Ce modèle contient 3 régresseurs qui sont tous désaisonnalisés. Avant désaisonnalisation, ces régresseurs sont le nombre de jours du mois, le contraste entre LV_t et F_t , et le contraste entre N_{6t} (nombre de samedis non fériés) et F_t .

L'option 3 de modélisation permet une modélisation d'un effet de jours "ouvrables" (jours de la semaine non fériés du lundi au vendredi) tout en laissant la possibilité d'une influence différente pour les samedis non fériés. Il est recommandé en section 4.5 de tester les performances de ce modèle dans les cas où l'option 2 donne des résultats satisfaisants, en vue d'obtenir un modèle plus simple.

3.3.4 Option 4 : effet des week-ends et jours fériés

Dans cette dernière variante, la plus simplifiée, du modèle de référence, les hypothèses de l'option 3 sont reprises et de plus les samedis non fériés ont la même influence que les dimanches et jours fériés. Le modèle obtenu ne contient plus que 2 régresseurs et s'écrit

$$X_t = \varepsilon_t + \tilde{\alpha}N_t + \beta_1 \left(LV_t - \frac{5}{9}(F_t + N_{6t}) \right) \quad (10)$$

où LV_t et F_t sont définis dans les modèles des options précédentes. Avant désaisonnalisation, les 2 régresseurs sont le nombre de jours du mois et le contraste entre les jours de la semaine non fériés du lundi au vendredi LV_t et les jours du week-end plus jours fériés.

L'option 4 de modélisation propose un modèle simple qui peut fonctionner pour l'étude de séries temporelles dans de nombreux domaines tels que l'industrie ou les services, voir Attal-Toubert [2012] section 2.3. En pratique, cette option devrait être testée si les modèles plus complexes donnent des résultats satisfaisants.

3.4 Calcul des régresseurs en R et fichiers à disposition

Les options décrites précédemment sont implémentées en langage R. L'objectif est de calculer les régresseurs servant d'input aux programmes d'ajustement calendrier tels que X13-ARIMA-SEATS. Le code R de calcul des régresseurs ainsi les fichiers obtenus avec ce code sont publiquement disponibles⁴, et un exemple d'utilisation de ces fichiers pour estimer les effets calendrier avec X13-ARIMA-SEATS est décrit en section 5. Les utilisateurs qui ne veulent pas utiliser R peuvent télécharger directement les régresseurs produits selon les différentes options possibles, voir sections 3.3.1 à 3.3.4.

Le fichier `CalendarAdjustment.R` propose un exemple d'appel à la fonction `BuildRegressors`. Les arguments de cette fonction sont :

4. Ces fichiers peuvent être téléchargés sur la page <https://www.bfs.admin.ch/bfs/fr/home/dscg/outils/methodes-statistiques/2023-06-calendrier.html>.

- `startingYear` (valeur par défaut : 1980) : les régresseurs sont calculés à partir du 1er janvier de l'année `startingYear`. Cette année doit être antérieure ou égale à l'année de départ de la série temporelle étudiée.
- `nYear` (valeur par défaut : 50) : indique la longueur, en nombre d'année, sur laquelle les régresseurs sont calculés. Les régresseurs sont ainsi des séries chronologiques de longueur `nYear`. Les utilisateurs qui travaillent avec des séries temporelles de moins de 50 ans peuvent laisser l'argument `nYear` à sa valeur par défaut.
- `option` (valeur par défaut : 2) : indique le degré de simplification du modèle de référence. Les valeurs possibles sont 0, 1, 2, 3 ou 4. Le choix `option=0` mène au modèle de référence à 14 paramètres. Les options 1, 2, 3 et 4 sont décrites dans les sections 3.3.1 à 3.3.4.
- `monthly` (valeur par défaut : TRUE) : indique si les régresseurs doivent être calculés pour l'ajustement calendrier d'une série mensuelle (`monthly = TRUE`) ou trimestrielle (`monthly = FALSE`).
- `printCalendar` (valeur par défaut : TRUE) : indique si le calendrier suisse, qui est calculé avant de calculer les régresseurs, doit être exporté dans un fichier au format texte.
- `printRegressors` (valeur par défaut : TRUE) : indique si les régresseurs qui sont calculés (output de la fonction) doivent être exportés dans un fichier au format texte.
- `calendarPath` (valeur par défaut : NULL) : lorsque `printCalendar=TRUE`, indique le chemin et le nom du fichier contenant le calendrier suisse à exporter. Si `printCalendar=TRUE` et qu'aucun chemin/nom n'est indiqué, un fichier `SwissCalendar.dat` (nom par défaut) est créé dans le répertoire de travail de R.
- `regressorPath` (valeur par défaut : NULL) : lorsque `printRegressors=TRUE`, indique le chemin et le nom du fichier contenant les régresseurs à exporter. Si `printRegressors=TRUE` et qu'aucun chemin/nom n'est indiqué, un fichier avec un nom par défaut est créé dans le répertoire de travail de R.

La fonction `BuildRegressors` calcule, et éventuellement exporte, une matrice avec l'année en première colonne et le numéro du mois – ou trimestre, pour les séries trimestrielles – en seconde colonne. Les colonnes suivantes correspondent aux régresseurs (désaisonnalisés), dont le nombre dépend de l'option choisie, et la dernière colonne contient le régresseur `Leap Year` :

- si `option=0`, les colonnes 3 à 15 contiennent les contrastes entre les lundis, ..., dimanches non fériés, lundis, ..., samedis fériés et les dimanches fériés,
- si `option=1`, les colonnes 3 à 14 contiennent les contrastes entre les lundis, ..., samedis non fériés, lundis, ..., samedis fériés et les dimanches (fériés ou non),
- si `option=2`, les colonnes 3 à 8 contiennent les contrastes entre les lundis, ..., samedis non fériés et les dimanches plus jours fériés,
- si `option=3`, la colonne 3 contient les contrastes entre les jours non fériés de la semaine du lundi au vendredi et les dimanches plus jours fériés. La colonne 4 contient le contraste entre les samedis non fériés et les dimanches plus jours fériés,
- si `option=4`, la colonne 3 contient les contrastes entre les jours non fériés de la semaine du lundi au vendredi et les jours du week-end plus jours fériés.

Remarque : lorsque la série d'étude est trimestrielle et non mensuelle, les régresseurs doivent être calculés trimestriellement et non mensuellement. La fonction `BuildRegressors` effectue cette opération si l'argument `monthly` est mis à `FALSE`. Pour calculer les régresseurs sur une base trimestrielle, il suffit de commencer à calculer les régresseurs comme si la série d'étude était une série mensuelle. Ensuite, les résultats obtenus sont simplement agrégés (sommés) trimestre par trimestre. Par exemple, pour le premier trimestre de l'année 2008, la valeur des régresseurs à ce trimestre est obtenue en sommant la valeur des régresseurs mensuels prise aux mois de janvier, février et mars 2008. Les régresseurs ainsi obtenus demeurent sans saisonnalité, et leur moyenne sur un trimestre donné est nulle.

4 Choix des régresseurs

Les différentes options présentées en section 3.3 posent naturellement la question du choix de la variante optimale. L'objectif de cette section n'est pas de définir une procédure automatique de choix mais plutôt de décrire une liste non exhaustive d'outils d'aide à la décision. Ces outils sont aussi illustrés à travers un exemple en section 5.

D'une manière générale, avant d'utiliser les outils décrits ci-dessous, il est important de garder en mémoire que les régresseurs choisis doivent correspondre à des hypothèses qui ont un sens économique ("economic rationale", voir Eurostat [2015], section 2.3). Les auteurs de Quartier-La Tente & Ladiray [2018] insistent également sur ce point et n'hésitent pas à parler de "bon sens". Les outils d'aide à la décision pour le choix des régresseurs doivent ainsi être vus comme un complément à la logique économique, et ne sauraient remplacer cette dernière.

4.1 Outliers et stabilité du modèle

Le modèle de base avec jours fériés (voir, équation (4)) est un modèle REG-ARIMA qui représente la série étudiée comme la somme (ou le produit, dans le cas d'un modèle multiplicatif) d'effets calendrier et d'une série temporelle ε_t suivant un modèle ARIMA. Une présentation détaillée du modèle REG-ARIMA et de l'estimation de ses paramètres est donnée dans le manuel de X13-ARIMA-SEATS, U.S. Census Bureau [2020], section 4. Le fichier d'output au format ".out" produit par X13-ARIMA-SEATS donne une estimation des coefficients associés à chaque régresseur, ainsi qu'une matrice de covariance pour quantifier les incertitudes.

Dans le cas où l'utilisateur définit lui-même les régresseurs à utiliser pour corriger les effets calendrier (par exemple : à l'aide de l'une des options 1 à 4 décrites en sections 3.3.1 à 3.3.4), X13-ARIMA-SEATS permet de rajouter des régresseurs supplémentaires pour prendre en compte d'éventuels outliers. Le terme "outlier" renvoie ici à un concept plus large que celui de valeurs extrêmes ou aberrantes et le lecteur peut consulter U.S. Census Bureau [2020], section 4.3 pour une présentation détaillée des différents types d'outliers et de leur traitement par X13-ARIMA-SEATS. Un point important est que les outliers sont traités en ajoutant des régresseurs supplémentaires à la partie "REG" du modèle REG-ARIMA, c'est-à-dire la partie censée traiter des ajustements calendrier. Le nombre total de régresseurs utilisés dans le modèle REG-ARIMA est ainsi égal au nombre de régresseurs correspondant au modèle choisi pour la correction des effets calendrier plus le nombre d'outliers. Les utilisateurs de X13-ARIMA-SEATS peuvent choisir de laisser le logiciel détecter automatiquement les outliers (type et date). Une autre solution est de définir manuellement les dates où les outliers surviennent, ainsi que le type d'outliers qui survient à ces dates. Ces choix méthodologiques concernant les outliers ne sont pas anodins et ont un impact important sur les résultats obtenus concernant la correction des effets calendrier. D'une manière générale, la recommandation d'Eurostat (voir Eurostat [2015], section 2.7) est de définir manuellement les dates et types d'outliers et d'avoir une interprétation économique claire pour soutenir ces choix. Par exemple, le chiffre d'affaires dans le commerce de détail diminue fortement en mars 2020 suite à la pandémie de Covid 19. La valeur de la série en mars 2020 peut ainsi être fixée à l'avance comme un outlier. L'alternative qui consiste à détecter automatiquement les outliers peut rendre la comparaison entre modèles plus compliquée, voir section 4.2.

Un exemple d'ajustement calendrier avec définition manuelle des outliers est présenté en section 5. Un argument supplémentaire en faveur de ce choix méthodologique est la stabilité du modèle dans le temps, c'est-à-dire la propension du modèle à varier au fur et à mesure que

de nouvelles données sont assimilées. Dans le contexte de l'ajustement calendrier, il est souhaitable que les coefficients de régression du modèle REG-ARIMA obtenus varient peu dans le temps car des changements dans le modèle viennent affecter les valeurs passées de la série corrigée de ses effets calendrier. Ce point est nommé "analyse des révisions" et est discuté dans [U.S. Census Bureau \[2020\]](#) sections 6.2 et 7.8. Dans l'exemple donné en section 5, l'analyse de révision est effectuée en observant l'évolution dans le temps des coefficients de régression. Ces informations sont obtenues en ajoutant l'instruction suivante dans la "spec file" de X13-ARIMA-SEATS : `history{estimates = td}`.

4.2 Critères de type AIC ou BIC

Lors d'un appel à X13-ARIMA-SEATS, le fichier .out produit donne des critères (indicateurs) de type AIC (Akaike Information Criterion) ou BIC (Bayesian Information Criterion), qui permettent sous certaines conditions de comparer les modèles REG-ARIMA obtenus. Quatre critères sont calculés, et sont décrits dans [U.S. Census Bureau \[2020\]](#), section 5.5. Parmi les options 1 à 4 décrites en sections 3.3.1 à 3.3.4, il serait ainsi tentant d'effectuer une désaisonnalisation avec chaque jeu de régresseurs puis de choisir le modèle qui donne la meilleure valeur possible (c'est-à-dire la valeur la plus basse) sur un critère donné, par exemple l'AIC. [U.S. Census Bureau \[2020\]](#), sections 5.5.1 à 5.5.3, met cependant en garde contre cette pratique. En effet, des modèles avec différents jeux de régresseurs ne peuvent être comparés entre eux à l'aide de critères de type AIC ou BIC que si l'ensemble des conditions suivantes sont satisfaites :

- les outliers (détectés automatiquement ou insérés manuellement par l'utilisateur) sont les mêmes dans tous les modèles comparés, voir section 5.5.1,
- les transformations appliquées initialement sur les données (par exemple une transformation en log) sont les mêmes dans tous les modèles comparés, voir section 5.5.2,
- les degrés de différenciation sur la série ε_t (voir, équation (4)) pour modéliser celle-ci à l'aide d'un modèle ARIMA sont les mêmes dans tous les modèles comparés, voir section 5.5.3.

Dans l'exemple utilisé en section 5 les conditions énoncées ci-dessus sont satisfaites. En effet : les outliers sont définis manuellement, une transformation en "log" est effectuée pour chaque modèle, et les degrés d et D obtenus dans les modèles $\text{ARIMA}(p, d, q)(P, D, Q)$ sont $d = 1$ et $D = 1$ pour chaque option testée.

4.3 Tests statistiques

Lors d'un appel à X13-ARIMA-SEATS avec un jeu de régresseurs fixé, un fichier .out contenant les estimations des différents coefficients associés à chaque régresseur est produit. Si l'utilisateur en fait la demande, une matrice de covariance est également produite. Le nombre de coefficients estimés est égal au nombre total de régresseurs choisis plus le nombre d'outliers, définis manuellement ou détectés automatiquement.

Comme expliqué en section 3.3, les différents jeux de régresseurs possibles sont construits sur des hypothèses liées à l'égalité de coefficients associés à certains régresseurs. Intuitivement, l'option 2 (qui regroupe les jours fériés) pourrait être préférée à l'option 1 (qui distingue les jours fériés du lundi au samedi) si les coefficients de régression obtenus associés aux jours fériés lors de l'utilisation de l'option 1 sont relativement proches. Similairement, l'option 3 (qui rassemble les jours non fériés du lundi au vendredi) pourrait être préférée à l'option 2 si l'utilisation de cette dernière mène à un modèle avec des coefficients de régression proches pour les jours non fériés du lundi au vendredi.

La notion de proximité des coefficients estimés dans un modèle REG-ARIMA peut être testée statistiquement. L'INSEE (Institut national de la statistique et des études économiques, France) propose d'ailleurs une procédure automatique de choix des régresseurs basée en partie sur des tests statistiques [Attal-Toubert, 2012]. Dans ce rapport, cette procédure n'est pas reprise telle quelle mais des tests statistiques sont décrits et ont été implémentés en langage R. Le code est mis à la disposition des utilisateurs.

Soit $\hat{\beta}$ le vecteur des coefficients estimés par le modèle REG-ARIMA. Soit p le nombre de régresseurs utilisés spécifiquement pour la correction d'effets calendrier et soit k le nombre total de régresseurs, incluant à la fois les p régresseurs de correction des effets calendrier et ceux associés aux éventuels outliers. Soit s la période de la série d'étude (par exemple, $s = 12$ pour une série mensuelle et 4 pour une série trimestrielle) et soient d et D les degrés de différenciation aux lag 1 et s respectivement, utilisés dans le modèle ARIMA pour la série sans effets calendrier. Soit n le nombre de données (longueur) de la série temporelle étudiée. Le vecteur $\hat{\beta}$ est un vecteur colonne de taille k . Soit $\hat{\Sigma} \in \mathbb{R}^{k \times k}$ la matrice de covariance estimée pour les coefficients. Tester l'égalité entre certains coefficients de régression du vecteur β revient à tester l'hypothèse $H_0 : R\beta = \mathbf{0}_q$ où R est une certaine matrice de contraintes de taille $q \times k$, q étant le nombre de contraintes testées et $\mathbf{0}_q$ un vecteur colonne de taille q dont les éléments sont égaux à 0.

Par exemple, si l'option 1 est utilisée, le vecteur $\hat{\beta}$ contient des coefficients correspondants dans l'ordre à l'influence des lundis, ..., samedis non fériés, des lundis, ..., samedis fériés, puis du régresseur "Leap Year" et enfin des éventuels outliers. Tester l'hypothèse d'égalité entre les 6 coefficients $\beta_7 = \beta_8 = \dots = \beta_{12}$ (i.e. influence identique des lundis, ..., samedis fériés) revient à tester $\beta_7 - \beta_8 = 0, \dots, \beta_{11} - \beta_{12} = 0$ et peut ainsi s'écrire sous la forme $R\beta = \mathbf{0}_q$ où $q = 5$ est le nombre de contraintes et $R \in \mathbb{R}^{q \times k}$ est la matrice

$$R = \begin{pmatrix} 0 & \dots & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & & 0 \\ 0 & & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & & 0 \\ 0 & & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & \dots & 0 \end{pmatrix}$$

Sous l'hypothèse $H_0 : R\beta = \mathbf{0}_q$, la statistique

$$\frac{n - d - sD - k}{q(n - d - sD)} (R\hat{\beta})^T (R\hat{\Sigma}R^T)^{-1} (R\hat{\beta}) \quad (11)$$

est distribuée approximativement selon une loi de Fisher à q et $n - d - sD - k$ degrés de liberté [Lytras et al., 2007]⁵.

Le fichier BuildRegressors.R contient le code d'une fonction TestRegressorsEquality qui permet d'effectuer ce type de test dans le cas particulier où les régresseurs sont choisis avec l'une des options décrites dans ce rapport. Les arguments de la fonction sont les suivants :

- `obj` : objet renvoyé par la fonction `seas` du package "seasonal"⁶. La fonction `seas` est supposée avoir été utilisée avec l'une des variantes (`option`) décrite dans ce rapport. L'utilisateur doit aussi avoir demandé à la fonction `seas` d'estimer la matrice de covariance des coefficients de régression Σ . Un exemple d'appel à `seas` est présenté en annexe A.

5. dans Lytras et al. [2007], le test statistique est effectué à l'aide d'un modèle qui utilise $D = 0$ ce qui fait que le terme " sD " présent dans ce rapport n'apparaît pas dans la référence.

6. Le package `seasonal` permet en effet d'utiliser X13-ARIMA-SEATS depuis une interface R via la fonction `seas`.

- `option` : numéro de l'option utilisée lors de l'appel à `seas`. Les valeurs possibles sont 0, 1, 2 et 3. Le test statistique effectué dépend de ce numéro et permet de déterminer si le modèle numéro `option+1` pourrait être préféré au modèle `option`.
- `s` : périodicité de la série d'étude. La valeur par défaut est 12 (série mensuelle).
- `n` : nombre d'observations de la série temporelle étudiée. Cet argument peut éventuellement être omis car il constitue l'un des (nombreux) outputs de la fonction `seas`. Dans ce cas, la valeur de `n` est récupérée dans l'un des champs de l'objet `obj`.
- `d` : degré de différenciation au lag 1 utilisé dans le modèle ARIMA. Pour les mêmes raisons que l'argument `n`, cet argument peut éventuellement être omis.
- `D` : degré de différenciation au lag `s` utilisé dans le modèle ARIMA. Cet argument peut éventuellement être omis.
- `print` : argument prenant la valeur `TRUE` (défaut) ou `FALSE` pour afficher les résultats des tests statistiques directement dans R.

La fonction `TestRegressorsEquality` renvoie un objet qui contient plusieurs champs dont la statistique du test de Fisher (voir équation (11)) et une p-valeur. Le test statistique effectué dépend de l'option (`option`) choisie :

- si `option=0`, l'hypothèse testée est l'égalité $\beta_7 = \beta_{14}$ (voir, équation (4)), c'est-à-dire que les dimanches non fériés et fériés ont la même influence. Dans le modèle de l'équation (4), le coefficient β_{14} est l'influence des dimanches fériés et est égal à $-\sum_{i=1}^{13} \beta_i$. Le test effectué ici est ainsi un test du type $R\beta = 0$ où la matrice de contraintes R ne contient qu'une seule contrainte. Si la p-valeur obtenue avec le test est suffisamment grande (typiquement, supérieure à 0.05), l'hypothèse d'égalité des coefficients n'est pas rejetée ce qui indique que le test ne tranche pas contre la simplification du modèle de l'option 0 en faveur du modèle de l'option 1. Il est rappelé aux lecteurs que l'option 0 n'est pas recommandée en pratique.
- si `option=1`, le test effectué se rapporte à l'équation (5). L'hypothèse testée est l'influence égale des jours fériés du lundi au samedi. Dans l'équation (5), cela correspond à l'hypothèse $H_0 : \beta_8 = \dots = \beta_{13}$. Dans le code R ces coefficients sont situés en position 7, 8, ..., 12. Au total, $q = 5$ contraintes sont testées. Si la p-valeur obtenue avec le test est suffisamment grande, l'hypothèse d'égalité des coefficients n'est pas rejetée. Cela indique que le test ne donne pas d'argument contre la simplification du modèle de l'option 1 en un modèle de l'option 2.
- si `option=2`, le test effectué se rapporte à l'équation (7). L'hypothèse testée est l'influence égale des jours non fériés du lundi au vendredi. Dans l'équation (7), cela correspond à l'hypothèse $H_0 : \beta_1 = \dots = \beta_5$. Au total, $q = 4$ contraintes sont testées. Selon la p-valeur obtenue, le test donne des arguments en faveur (ou en défaveur) du modèle de l'option 3 par rapport à celui de l'option 2.
- si `option=3`, le test effectué se rapporte à l'équation (9). L'hypothèse testée est l'influence égale des samedis non fériés et des dimanches plus jours fériés. Dans l'équation (9), cela correspond à l'hypothèse $H_0 : \beta_6 = -\frac{5}{8}\beta_1 - \frac{1}{8}\beta_6$. Au total, une seule contrainte ($q = 1$) est testée. Selon la p-valeur obtenue, le test donne des arguments en faveur (ou en défaveur) du modèle de l'option 4 par rapport à celui de l'option 3.

4.4 Spectre de la série corrigée

4.4.1 Définition et estimation du spectre

Les modèles REG-ARIMA estimés par X13-ARIMA-SEATS font l'hypothèse que la série d'étude X_t se décompose en une somme de régresseurs identifiés (i.e. la partie "REG" du modèle) et d'une série ε_t suivant un modèle ARIMA – qui est donc (éventuellement après un nombre limité

de différenciations) stationnaire. La notion de densité spectrale (ou “spectre”, voir par exemple Eurostat [2018], section 10.5) est un outil utile pour étudier les propriétés de la série ε_t (éventuellement différenciée), et en particulier pour voir si celle-ci ne comporte plus d’effets calendrier notable. Il est en effet souhaitable que les effets calendrier de la série initiale soient traités dans la partie “REG” du modèle ARIMA et ne subsistent plus dans la série ε_t .

Soit une série W_t stationnaire suivant un modèle ARMA de moyenne nulle. Soit la fonction d’autocovariance $\gamma(h) := E(W_t W_{t+h})$, qui quantifie la dépendance entre 2 observations séparées de h pas de temps. La densité spectrale de la série W_t est définie comme la transformée de Fourier discrète de sa fonction d’autocovariance. Dans le cas d’un modèle ARMA, cela donne la fonction suivante, définie sur l’intervalle $[-1/2, 1/2]$:

$$g(\lambda) = \gamma(0) + 2 \sum_{j=1}^{\infty} \gamma(j) \cos(2\pi j \lambda). \quad (12)$$

L’égalité

$$\gamma(j) = 2 \int_0^{1/2} \cos(2\pi j \lambda) g(\lambda) d\lambda \quad (13)$$

montre que $g(\lambda)$ s’interprète comme une contribution, dans le domaine des fréquences, à la fonction d’autocovariance de W_t . De plus, $g(\lambda)$ étant une fonction paire (c’est-à-dire que $g(-\lambda) = g(\lambda)$), l’analyse de cette fonction peut se faire sur l’intervalle $[0, 1/2]$ au lieu de $[-1/2, 1/2]$.

Étant donné une série discrète, centrée, de données $y_1, \dots, y_n$ provenant d’une série temporelle Y_t , plusieurs estimateurs existent pour estimer la densité spectrale, et deux d’entre eux sont implémentés dans le programme X13-ARIMA-SEATS. L’estimateur le plus connu et répandu est peut-être le “périodogramme” [Findley et al., 1998]. Celui-ci est relié au carré du module de la transformée de Fourier discrète de la série de données et est exprimé en décibels par X13-ARIMA-SEATS :

$$\tilde{g}(\lambda) := 10 \log_{10} \left(\frac{1}{n} \left| \sum_{j=1}^n y_j \exp(-2i\pi j \lambda) \right|^2 \right) \quad (14)$$

pour $0 \leq \lambda \leq 1/2$. L’estimateur $\tilde{g}(\lambda)$ de la densité spectrale n’est cependant pas celui par défaut. L’estimateur par défaut, dit “spectre-AR” (“AR-spectrum” en anglais), peut être obtenu en notant que tout processus ARMA peut s’écrire sous une représentation purement autorégressive $AR(\infty)$, avec éventuellement une infinité de termes. De plus, la densité spectrale d’un processus ARMA s’écrit dans le cas général comme une fonction simple des paramètres du modèle ARMA, voir Eurostat [2018], section 10.5. En considérant ainsi que la série de données $y_1, \dots, y_n$ est issue d’un processus $AR(m)$ où m est un “grand” nombre (ce qui est une hypothèse acceptable compte tenu de l’existence d’une représentation $AR(\infty)$), l’estimateur spectre-AR, exprimé en décibels, s’écrit

$$\hat{g}(\lambda) := 10 \log_{10} \left(\frac{\sigma_m^2}{2\pi \left| 1 - \sum_{j=1}^m \hat{\phi}_j \exp(2i\pi j \lambda) \right|^2} \right) \quad (15)$$

où $\hat{\phi}_1, \dots, \hat{\phi}_m$ sont les coefficients obtenus en régressant linéairement la série y_t (qui est centrée) à l’aide des prédicteurs $y_{t-1}, \dots, y_{t-m}$ et où σ_m^2 est la variance empirique des résidus de cette régression linéaire. Dans X13-ARIMA-SEATS, m est égal à 30 par défaut, et cette valeur peut être modifiée par l’utilisateur.

X13-ARIMA-SEATS affiche pour plusieurs séries, dans un fichier d'output .out, les valeurs de $\hat{g}(\lambda)$, ou bien $\tilde{g}(\lambda)$ si l'estimation par périodogramme est explicitement choisie. Ces valeurs sont données sur l'intervalle $[0, 1/2]$, aux fréquences $0/120, 1/120, \dots, 60/120$. Les séries pour lesquelles les graphiques sont produits sont dans l'ordre

1. les résidus du modèle REG-ARIMA,
2. la série originale ajustée des effets calendrier et outliers (aussi donnée en table B1 du fichier .out), différenciée une fois,
3. la série corrigée de ses effets saisonniers (aussi donnée en table E2 du fichier .out), différenciée une fois,
4. la composante irrégulière (aussi donnée en table E3 du fichier .out), voir [Eurostat \[2018\]](#), section 3 pour une définition.

Si les effets calendrier sont convenablement corrigés à l'aide de régresseurs bien choisis, les séries mentionnées ci-dessus sous 1. et 4. ne devraient normalement pas comporter d'effets calendrier. En effet, les résidus du modèle REG-ARIMA ne doivent plus, par définition, avoir d'effets calendrier résiduels. La composante irrégulière représente la partie des variations de la série liés à des événements imprévisibles, ce qui inclut les outliers mais exclut les effets calendrier. Pour vérifier que ces deux séries ne comportent plus d'effets calendrier, il est possible d'analyser visuellement l'allure du spectre de ces séries en se focalisant sur certaines fréquences.

4.4.2 Fréquences de jours ouvrables et tests visuels sur le spectre

La périodicité du calendrier Grégorien implique une certaine périodicité sur le nombre de lundis, ..., dimanches au cours du mois ou trimestre considéré (voir, [Bell et al. \[2012\]](#), section 11). Ainsi, le spectre d'une série dont les effets calendrier ne sont pas corrigés présente souvent des pics à des fréquences particulières, dites "fréquences de jours ouvrables" ("trading day frequency", en anglais). Pour une série mensuelle, les fréquences de jours ouvrables sont les mêmes que celles observées en traçant simplement le périodogramme d'une série mensuelle déterministe, égale au nombre de lundis (ou n'importe quel autre jour de la semaine) au cours du mois ou trimestre.

Pour une série mensuelle, les fréquences de jours ouvrables les plus importantes, qui ressortent pour de nombreuses séries temporelles avec des effets calendrier, sont 0.348125 et dans une moindre mesure 0.431458. Les pics aux autres fréquences sont généralement négligeables. La valeur 0.348125 s'obtient en notant qu'il y a en moyenne 4.348125 semaines dans un mois de l'année. La valeur 0.431458 est plus difficile à interpréter [[Bell et al., 2012](#)]. Pour une série trimestrielle, [[Bell et al., 2012](#)] cite 7 fréquences différentes qui sont importantes et/ou non négligeables. Lorsque le spectre d'une série est analysé, ce sont les valeurs à ces fréquences qui permettent de décider de la qualité de la correction des effets calendrier.

Dans le cas présent, les spectres à analyser sont les spectres 1. (résidus du modèle REG-ARIMA) et 4. (composante irrégulière) définis en section 4.4.1. Les cas possibles sont les suivants.

- Si l'une des deux séries présente des "pics significatifs" (tels que définis à la fin de cette section) à au moins l'une des fréquences de jours ouvrables, cela signifie que le modèle REG-ARIMA utilisé n'a pas soustrait la totalité des effets calendrier de la série initiale. Cela suggère un choix trop restrictif du jeu de régresseurs, c'est-à-dire des hypothèses trop simplificatrices sur les effets des différents jours de la semaine. Par exemple, si l'utilisateur a choisi l'option numéro 3 décrite en section 3.3.3, la présence d'effets calendrier résiduels suggère de faire de nouveaux tests avec un modèle plus complexe, par exemple l'option numéro 2.

- Si les deux séries ne présentent pas de pics significatifs aux fréquences de jours ouvrables, cela suggère que les régresseurs utilisés ont permis de bien capter l'ensemble des effets calendrier. Cela indique qu'il n'est à priori pas nécessaire de complexifier le modèle en choisissant un jeu de régresseurs plus riche (par exemple : en choisissant l'option numéro 1 si l'option numéro 2 permet déjà de bien corriger les effets calendrier), et que le jeu de régresseurs actuel est peut-être adapté. De manière générale, l'utilisateur devrait chercher à corriger les effets calendrier en utilisant le modèle le plus simple possible. Ainsi, si l'option numéro 2 donne des résultats satisfaisants, l'utilisateur pourrait tout de même tester si les résultats demeurent satisfaisants avec l'option numéro 3, et choisir définitivement l'option 2 seulement si ce n'est pas le cas. Un exemple est proposé en section 5.

La notion de "pic significatif" à une fréquence donnée est détaillée dans [U.S. Census Bureau \[2020\]](#), section 6.1.2 et [Bell et al. \[2012\]](#), section 11.4. Dans son fichier .out, X13-ARIMA-SEATS représente le spectre sous forme d'un graphique au format texte (voir figure 3 pour un exemple). Le graphique est mis à l'échelle sur 52 lignes de sorte qu'un écart d'une ligne soit égal à

$$\frac{1}{52} \left(\max_k \hat{g}(k) - \min_k \hat{g}(k) \right), 0 \leq k \leq 1/2.$$

Les fréquences de jours ouvrables sont représentées à l'aide de la lettre "T" (trading day). Soit λ une fréquence de jours ouvrables fixée. Un pic du spectre à la valeur λ est jugé comme visuellement significatif si la valeur $\hat{g}(\lambda)$ est suffisamment grande et si elle dépasse les valeurs adjacentes de "6 lignes" ou "6 étoiles" ("6 stars peak", en anglais). Plus précisément, les deux conditions suivantes doivent être remplies :

1. $\hat{g}(\lambda)$ doit être supérieur à la médiane des valeurs du spectre $\text{med}_{k=0/120, 1/120, \dots, 60/120} \hat{g}(k)$,
2. $\hat{g}(\lambda)$ doit dépasser les valeurs adjacentes à droite et à gauche (i.e. $\hat{g}(\lambda \pm 1/120)$) d'au moins $6/52 (\max_k \hat{g}(k) - \min_k \hat{g}(k))$.

Les conditions énoncées ci-dessus sont valables pour une série mensuelle. Dans ce cas-là les fréquences de jour ouvrable, marquées graphiquement avec un "T" par X13-ARIMA-SEATS, sont 0.348125 et 0.431458. Pour une série trimestrielle, [Bell et al. \[2012\]](#) montre que l'utilisation du spectre et du test visuel de X13-ARIMA-SEATS donne des performances incertaines pour détecter la présence ou l'absence d'effets calendrier. Ainsi, le spectre en tant qu'outil d'aide à la décision pour le choix des régresseurs est à utiliser en priorité pour des séries mensuelles.

4.5 Interactions entre ces outils

Dans le cadre d'un choix entre les jeux de régresseurs 1 à 4 détaillés en section 3.3, l'option numéro 2 qui donne une influence identique aux dimanches et jours fériés, tout en permettant aux jours non fériés du lundi au samedi d'avoir leur propre effet, peut être considérée comme une option de référence. Cela signifie que le choix de cette option devrait être envisagé en premier, et que choisir in fine un modèle différent ne se ferait que si des arguments concrets viennent étayer cette décision. Pour une série mensuelle, la recommandation de départ serait ainsi d'observer à l'aide du spectre (voir section 4.4) si les effets calendrier sont correctement éliminés en utilisant l'option 2.

En cas de résultats favorables (pas de pics aux fréquences de jours ouvrables), des modèles plus simples que celui de l'option 2 peuvent être envisagés. Pour cela, il est possible de vérifier – avec le modèle simplifié – la présence de pics aux fréquences de jours ouvrables. Il est

aussi possible de comparer les indicateurs de type AIC obtenus, si les conditions énumérées en section 4.2 pour effectuer les comparaisons sont satisfaites. Enfin, des tests statistiques d'égalité de coefficients peuvent être effectués sur le modèle de l'option 2. Selon le cas, les différents indicateurs peuvent donner des messages contradictoires. Par exemple : un rejet des tests d'égalité de coefficients avec l'option 2 mais des spectres qui demeurent satisfaisants si l'on teste l'option 3. Ce point reflète que les différents modèles peuvent dans certains cas donner des résultats proches et que plusieurs choix sont parfois acceptables. Les interprétations économiques sont alors utiles pour finaliser le choix du modèle.

Si le modèle de l'option 2 ne permet pas de supprimer correctement les effets de jours ouvrables (présence de pics aux fréquences "trading day" dans le spectre), un modèle plus complexe tel que celui de l'option 1 devrait être envisagé, en gardant à l'esprit que cette option est peu utilisée en pratique. Les arguments qui pourraient pousser à préférer l'option 1 à l'option 2 sont de meilleurs indicateurs de type AIC, des tests statistiques qui rejettent l'égalité des coefficients des différents jours fériés et un spectre qui ne présente pas d'effet "trading day" avec l'option 1 alors qu'il en présente avec l'option 2. La viabilité à long terme de l'option 1 peut aussi être investiguée à l'aide des analyses de révision. Il s'agit ici de vérifier que les coefficients de régression associés aux jours fériés ne varient pas trop dans le temps, au fur et à mesure que les données sont incorporées. Un exemple est donné en section 5.

5 Exemple et résultats

5.1 Série d'étude, code R et décomposition

Dans cette section, l'outil X13-ARIMA-SEATS [Sax & Eddelbuettel, 2018] est utilisé pour estimer les effets calendrier d'une série temporelle. La série étudiée est le chiffre d'affaires mensuel total du commerce de détail en Suisse de janvier 2000 à décembre 2021, voir figure 2. En obser-

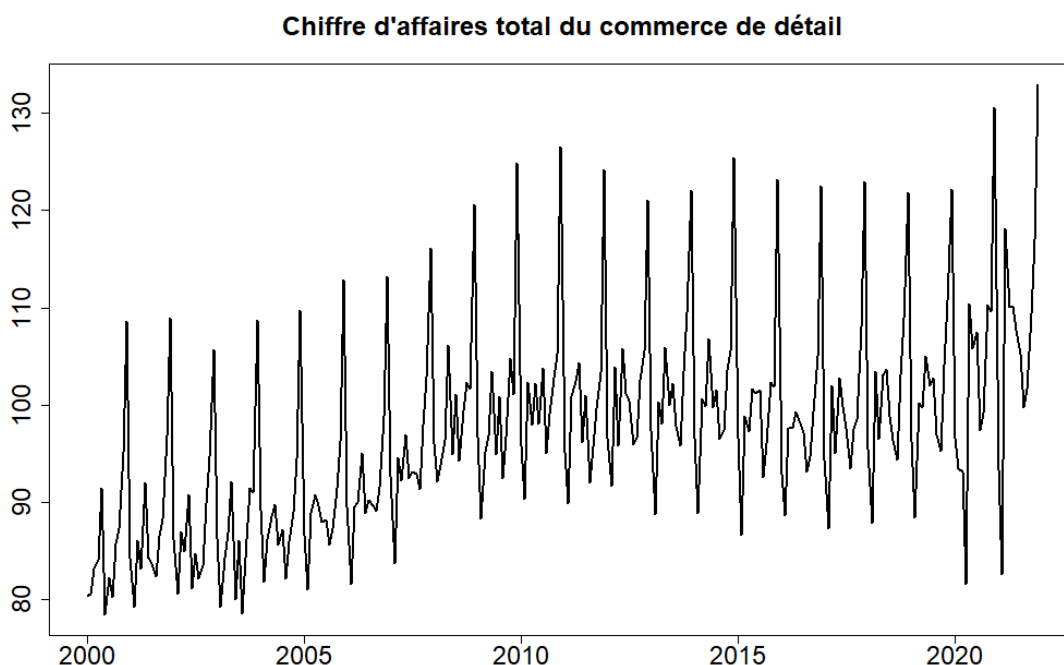


FIGURE 2

vant le graphique, il apparaît que la série présente une forte composante saisonnière avec des chiffres d'affaires élevés au mois de décembre et plus faibles en janvier et février. Les effets calendrier de la série ne sont facilement visibles en observant simplement la série et doivent être estimés. Dans cet exemple, les effets calendrier sont estimés à l'aide d'un modèle multiplicatif dans lequel la série initiale, "Detailhandelsumsatz" (DHU) en allemand, s'écrit à l'aide du produit suivant :

$$\text{DHU} = \text{TrendCycle} * \text{Seaso} * \text{Irregular} * \text{CalendarEffect} * \text{Outlier} \quad (16)$$

où TrendCycle est une estimation de la composante tendance-cycle, Seaso est la composante saisonnière, Irregular est la composante irrégulière, CalendarEffect représente les effets calendrier et Outlier les effets dus aux valeurs aberrantes détectées. Selon cette décomposition multiplicative, la série corrigée de ses effets calendrier est simplement DHU/CalendarEffect.

Pour estimer les composantes de l'équation (16), le logiciel X13-ARIMA-SEATS est ici utilisé depuis R, à l'aide du package `seasonal`. Un bloc de code R est donné en annexe A. Dans le code R, des données (série d'étude et fichiers contenant les régresseurs) sont importées et un appel à la fonction `seas` du package `seasonal` est effectué avec plusieurs choix différents de régresseurs. Lors de ces appels, les outliers sont définis manuellement et sont communs à chaque modèle testé pour faciliter la comparabilité, voir section 4.2. Les dates où les outliers sont insérées ne doivent cependant pas être vues comme une recommandation de l'OFS : celles-ci ont été obtenues à l'aide de l'algorithme de détection automatique d'outliers appliqué à l'option numéro 2 de modélisation. La fonction `out` permet de visualiser le fichier `.out` produit par X13-ARIMA-SEATS.

5.2 Résultats avec l'option numéro 2

Pour l'option 2, qui assimile les jours fériés à des dimanches, les coefficients estimés dans le modèle de l'équation (7) sont lisibles dans le fichier `.out` correspondant et reproduits dans la table 1. Sans grande surprise, la table 1 suggère un coefficient élevé associé aux samedis

TABLE 1 Modèle de régression obtenu en choisissant les régresseurs de l'option 2, tel que décrit par X13-ARIMA-SEATS.

	Parameter Estimate	Standard Error	t-value
AO2020.Mar	-0.1012	0.01667	-6.07
AO2020.Apr	-0.2550	0.01492	-17.08
AO2021.Feb	-0.1058	0.01697	-6.23
AO2021.Mar	0.1014	0.01888	5.37
User-defined Trading Day			
xreg1 (lundis non fériés)	0.0008	0.00177	0.43
xreg2 (mardis non fériés)	0.0089	0.00212	4.21
xreg3 (mercredis non fériés)	0.0090	0.00208	4.30
xreg4 (jeudis non fériés)	0.0133	0.00220	6.07
xreg5 (vendredis non fériés)	0.0066	0.00174	3.78
xreg6 (samedis non fériés)	0.0150	0.00225	6.67
User-defined Leap Year			
xreg7 (Leap Year)	0.0294	0.00628	4.68

non fériés et plus faible pour les lundis non fériés. Le fichier `.out` produit par X13-ARIMA-SEATS représente le spectre de plusieurs séries (voir section 4.4 pour des détails). Le spectre des

résidus du modèle REG-ARIMA est donné en figure 3 et ne comporte pas de pics aux fréquences de jours ouvrables, repérées à l'aide de la lettre "T". Cela indique que l'option 2 permet de prendre en compte convenablement les effets calendrier de la série. Des conclusions similaires sont obtenues en observant le spectre de la composante irrégulière, qui n'est pas reproduit ici. Le

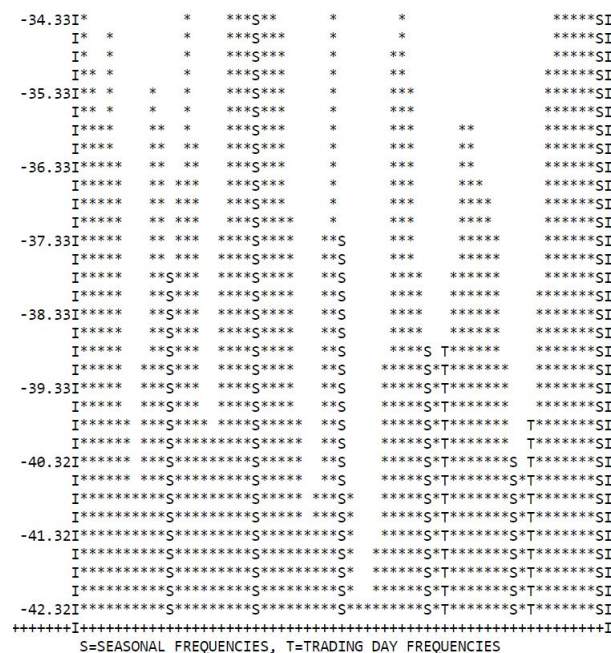


FIGURE 3 Spectre des résidus du modèle REG-ARIMA affiché dans le fichier .out de X13-ARIMA-SEATS en choisissant l'option 2 de modélisation

modèle ARIMA obtenu dans la modélisation REG-ARIMA est un $ARIMA(0,1,1)(0,1,1)$, donc un modèle avec $d = 1$ différenciation au lag 1 et $D = 1$ différenciation saisonnière (au lag 12). Le critère AICC ("F-corrected-AIC") vaut 1'017.61. A ce stade de l'analyse, l'option numéro 2 de modélisation semble acceptable et la question est à présent de savoir si le modèle peut être davantage simplifié.

5.3 Résultats avec les options 1 et 3

Lorsque l'option numéro 3 de modélisation est utilisée, le spectre des résidus du modèle REG-ARIMA (non reproduit ici) présente un pic à la fréquence de jours ouvrables 0.348125. Ce pic est signalé dans le fichier .out à l'aide du texte "regARIMA model residuals (1 Trading Day peak(s))". Cela suggère que l'option numéro 3, qui suppose un effet égal des jours non fériés du lundi au vendredi, est associée à un modèle trop simple qui ne permet pas de supprimer correctement les effets calendrier de la série. Hormis le spectre, le modèle ARIMA obtenu dans la modélisation REG-ARIMA est un $ARIMA(0,1,1)(0,1,1)$, ce qui est le même modèle que pour l'option 2 et permet une comparabilité sur la base de critères de type AIC. Le critère AICC vaut 1'047.94, une valeur moins bonne que pour l'option 2.

A ce stade, l'option 2 apparaît ainsi comme un bon choix de régresseurs pour la série étudiée. Des résultats sont néanmoins présentés pour l'option numéro 1 qui permet des influences différentes des différents jours fériés. L'idée est ici d'illustrer ce qui pourrait être fait dans le cas où l'option numéro 2 ne donnerait pas des résultats pleinement satisfaisants (par exemple, dans le cas où des pics apparaîtraient aux fréquences de jours ouvrables sur le spectre des résidus du modèle REG-ARIMA). Les coefficients de régression obtenus avec l'option 1 sont reproduits

en table 2. Ici le modèle ARIMA obtenu est encore un ARIMA(0,1,1)(0,1,1) et le critère AICC vaut 988.86, une valeur meilleure que celle obtenue avec l'option 2. Le test d'égalité des coefficients

TABLE 2 Modèle de régression obtenu en choisissant les régresseurs de l'option 1, tel que décrit par X13-ARIMA-SEATS.

	Parameter Estimate	Standard Error	t-value
AO2020.Mar	-0.0940	0.01535	-6.12
AO2020.Apr	-0.2523	0.01328	-19.00
AO2021.Feb	-0.1026	0.01570	-6.54
AO2021.Mar	0.1031	0.01799	5.73
User-defined Trading Day			
xreg1 (lundis non fériés)	0.0037	0.00217	1.71
xreg2 (mardis non fériés)	0.0060	0.00201	2.99
xreg3 (mercredis non fériés)	0.0039	0.00226	1.72
xreg4 (jeudis non fériés)	0.0077	0.00239	3.23
xreg5 (vendredis non fériés)	0.0092	0.00193	4.79
xreg6 (samedis non fériés)	0.0115	0.00220	5.21
xreg7 (lundis fériés)	-0.0049	0.00269	-1.81
xreg8 (mardis fériés)	0.0012	0.00435	0.29
xreg9 (mercredis fériés)	-0.0033	0.00401	-0.81
xreg10 (jeudis fériés)	-0.0072	0.00351	-2.06
xreg11 (vendredis fériés)	0.0114	0.00332	3.44
xreg12 (samedis fériés)	-0.0032	0.00449	-0.70
User-defined Leap Year			
xreg13 (Leap Year)	0.0327	0.00562	5.83

associés aux jours fériés donne une p-valeur très faible, ce qui n'est pas surprenant si l'on observe les coefficients associés aux jours fériés. Par exemple, le coefficient associé aux vendredis fériés est positif avec une t-valeur (ratio entre le coefficient estimé et son écart-type) de 3.44 alors que celui associé aux jeudis fériés est négatif avec une t-valeur d'environ -2 . Les jeudis fériés semblent ainsi affecter très négativement le commerce, peut-être du fait d'un "bridge effect", c'est-à-dire d'un effet de week-end prolongé du jeudi au dimanche (peu d'activité commerciale le vendredi, même si celui-ci est un jour ouvrable).

Dans ces conditions, l'option numéro 1 semble être un choix acceptable, d'autant plus que les analyses de révision (données en fin de fichier .out, non publiées dans ce rapport) qui testent la "stabilité" de ce modèle avec beaucoup de paramètres donnent des résultats satisfaisants, avec notamment des influences fortes et inverses des jeudis et vendredis fériés. Il faut cependant garder à l'esprit que le modèle de l'option 2, plus simple, permettait déjà de supprimer les effets calendrier de manière satisfaisante et pourrait ainsi constituer le choix de référence. Les séries corrigées de leurs effets calendrier à l'aide des options 1 et 2 sont représentées en figure 1 en début de rapport. Les effets calendrier expliquant une part relativement limitée des variations de la série (en comparaison aux effets de tendance ou de saisonnalité), une visualisation plus aisée est donnée en figure 4. Sur cette figure, les effets calendrier, multiplicatif, sont directement représentés. La série corrigée des effets calendrier se calcule en prenant la série de départ divisée par les effets calendrier. La figure 4 montre que la correction des effets calendrier va ici modifier les valeurs de la série initiale d'au maximum $\pm 3\%$. Il semble que l'option 1 (courbe bleue) suggère des corrections globalement proches de celles de l'option 2, et parfois un peu plus fortes en valeur absolue.

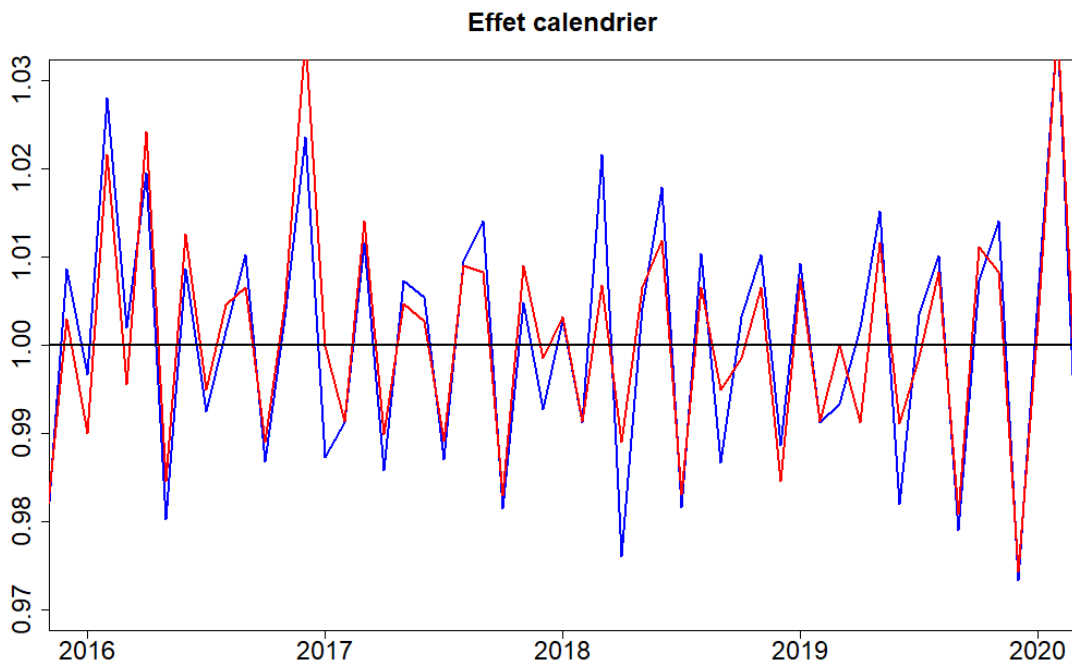


FIGURE 4 Effets calendrier (multiplicatif) obtenus en utilisant les options numéro 1 (courbe bleue) et 2 (courbe rouge).

6 Conclusions

L'estimation des effets calendrier d'une série temporelle est importante en statistique publique (voir par exemple [Eurostat \[2015\]](#), sections 2.3 et 2.4). Elle contribue à améliorer l'interprétation des variations à court terme (d'un mois, ou d'un trimestre à l'autre) d'une série temporelle. Ce rapport vise à faciliter la compréhension de cette démarche et à permettre aux utilisateurs de procéder à ces estimations de manière autonome. À cette fin, des fichiers sont mis à disposition du public. Il s'agit d'une part de fichiers au format texte qui contiennent les régresseurs calculés selon les options décrites dans ce rapport, pour des séries mensuelles ou trimestrielles. Ces fichiers peuvent être utilisés comme input lors d'un appel à X13-ARIMA-SEATS. Un exemple concret est produit dans ce rapport. D'autre part, l'OFS met aussi à disposition du code en langage R. Il s'agit ici de permettre aux spécialistes de générer eux-mêmes leurs fichiers de régresseurs voire même si nécessaire d'utiliser un calendrier différent du calendrier suisse de l'OFS. Dans tous les cas, l'utilisation d'un calendrier adapté en lieu et place du calendrier par défaut de X13-ARIMA-SEATS est recommandée par Eurostat [[Eurostat, 2015](#)].

L'estimation des effets calendrier reste une thématique sur laquelle peu de pratiques standardisées existent. Des outils d'aide à la décision et un exemple sont présentés dans ce rapport. Dans certains cas, plusieurs choix de jeux de régresseurs permettent une modélisation adaptée, ce qui donne une certaine marge de manœuvre aux utilisateurs dans leurs choix. L'OFS est attentive à la littérature scientifique sur cette thématique et pourrait à l'avenir améliorer ses outils et recommandations. Les améliorations pourraient porter sur plusieurs objets. D'une part, il est possible de proposer un éventail plus large d'options possibles pour le choix des régresseurs. Certaines options supplémentaires pourraient par exemple être mieux adaptées aux spécificités des séries temporelles d'une branche économique donnée. D'autre part, à l'instar de ce que fait l'INSEE [[Attal-Toubert, 2012](#)], une procédure automatisée du choix des régresseurs pourrait être proposée.

Bibliographie

- ATTAL-TOUBERT, K. (2012). Régresseurs pour effets de calendrier : Comment les construire, comment les choisir ? Journées de méthodologie statistique, INSEE.
URL http://www.jms-insee.fr/2012/S14_3_ACTE_ATTAL_JMS2012.PDF
- BELL, W. R. & HILLMER, S. C. (1983). Modeling Time Series With Calendar Variation. *Journal of the American Statistical Association* **78**, 526–534.
- BELL, W. R., HOLAN, S. H. & MCELROY, T. S. (2012). *Economic Time series : modeling and seasonality*. CRC Press.
- BIEN, R. (2004). Gauss and beyond : The making of Easter algorithms. *Archive for history of exact sciences* **58**, 439–452.
- EUROSTAT (2015). ESS guidelines on seasonal adjustment. Tech. rep., Eurostat.
URL <https://ec.europa.eu/eurostat/documents/3859598/6830795/KS-GQ-15-001-EN-N.pdf>.
- EUROSTAT (2018). *Handbook on Seasonal Adjustment*. Tech. rep., Eurostat.
URL <https://ec.europa.eu/eurostat/documents/3859598/8939616/KS-GQ-18-001-EN-N.pdf>.
- FINDLEY, D. F., MONSELL, B. C., BELL, W. R., OTTO, M. C. & CHEN, B.-C. (1998). New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. *Journal of Business & Economic Statistics* **16**, 127–152.
- GRUDKOWSKA, S., BUONO, D., PALATE, J. & CIEBIERA, W. (2013). Advanced Tools for Time Series Analysis and Seasonal Adjustment in the New JDemetra+. *JSM Proceedings Paper* .
- LYTRAS, D. P., FELDPAUSCH, R. M. & BELL, W. R. (2007). Determining seasonality : a comparison of diagnostics from X-12-ARIMA. *US Census Bureau* .
- QUARTIER-LA TENTE, A. & LADIRAY, D. (2018). Du bon usage des modèles REG-ARIMA en désaisonnalisation. *Journées de méthodologie statistique de l'Insee* .
- SAX, C. & EDELBUETTEL, D. (2018). Seasonal Adjustment by X-13ARIMA-SEATS in R. *Journal of Statistical Software* **87**, 1–17.
- U.S. CENSUS BUREAU (2020). Official X-13-ARIMA-SEATS manual. Tech. rep., U.S. Census Bureau.
URL <https://www2.census.gov/software/x-13arima-seats/x13as/windows/documentation/docx13as.pdf>.

Annexe

A Code R d'appel à X13-ARIMA-SEATS

Le bloc de code R suivant charge le package `seasonal`, importe les données ainsi que les régresseurs produits avec les options 1, 2 et 3, sur une période qui couvre la période de la série d'étude. La fonction "seas" est ensuite appelée avec chacune de ces 3 options. Des tests statistiques d'égalité des coefficients sont effectués et les effets calendrier ainsi que les séries corrigées des effets calendrier sont calculées. Les effets calendrier obtenus selon les options 1 et 2 sont représentés en figure 4.

```
library(seasonal)
data <- read.table(file = "filedhu_2022.dat",
                  header = FALSE, sep = "", dec = ".")
DHU <- ts(data = data, start = 2000, frequency = 12)

regressors_option1 <- import.ts("SwissRegressors1980_2029_option1.dat")
regressors_option2 <- import.ts("SwissRegressors1980_2029_option2.dat")
regressors_option3 <- import.ts("SwissRegressors1980_2029_option3.dat")

obj_option2 <- seas(x = DHU,
                  xreg=regressors_option2,
                  transform.function = "log",
                  regression.aictest=c("user"),
                  regression.usertype=c("td","td","td","td","td","td","lpyear"),
                  regression.variables=c("A02020.Mar","A02020.Apr",
                                         "A02021.Feb","A02021.Mar"),
                  regression.start="1980.1", outlier.types="none",
                  history.estimates="td", x11.mode = "mult",
                  estimate.print=c("rcm","acm"), estimate.save=c("rcm","acm"))

obj_option3 <- seas(x = DHU,
                  xreg=regressors_option3,
                  transform.function = "log",
                  regression.aictest=c("user"),
                  regression.usertype=c("td","td","lpyear"),
                  regression.variables=c("A02020.Mar","A02020.Apr",
                                         "A02021.Feb","A02021.Mar"),
                  regression.start="1980.1", outlier.types="none",
                  history.estimates="td", x11.mode = "mult",
                  estimate.print=c("rcm","acm"), estimate.save=c("rcm","acm"))

obj_option1 <- seas(x = DHU,
                  xreg=regressors_option1,
                  transform.function = "log",
                  regression.aictest=c("user"),
                  regression.usertype=c("td","td","td","td","td","td",
                                         "td","td","td","td","td","td","lpyear"),
                  regression.variables=c("A02020.Mar","A02020.Apr",
                                         "A02021.Feb","A02021.Mar"),
```

```

        regression.start="1980.1",outlier.types="none",
        history.estimate="td",x11.mode = "mult",
        estimate.print=c("rcm","acm"),estimate.save=c("rcm","acm"))

out(obj_option2)
out(obj_option3)
out(obj_option1)

TestRegressorsEquality(obj = obj_option1,option = 1,print = TRUE)
TestRegressorsEquality(obj = obj_option2,option = 2,print = TRUE)

calendar_1 <- series(x=obj_option1,series = c("x11.calendar"))
calendar_2 <- series(x=obj_option2,series = c("x11.calendar"))
calendar_3 <- series(x=obj_option3,series = c("x11.calendar"))

DHU_calendarAdj_option1 <- DHU/calendar_1
DHU_calendarAdj_option2 <- DHU/calendar_2
DHU_calendarAdj_option3 <- DHU/calendar_3

```

Rapports de méthodes de la section méthodes statistiques de l'OFS
Methodenberichte der Sektion Statistische Methoden des BFS
Methodology reports published by the FSO's Statistical Methods Section

Chevalier, C. (2023). Méthodes de correction des effets calendrier en usage à l'OFS en 2023. Construction, choix et application des régresseurs mis à disposition. Numéro OFS : 338-0085

Mariétan, R. (2022). Comportement en matière de voyages 2020. Révision de la pondération. Numéro OFS : 338-0084

Massiani, A. (2021). Enquête suisse sur la population active 2015-2020 : Révision de la pondération et de l'estimation de la variance. Numéro OFS : 338-0083

Chevalier, C., Qualité, L. (2021). Enquête suisse sur la structure des salaires 2018. Révision de la pondération. Numéro OFS : 338-0082

Qualité, L., Potterat, J. (2020). Enquête suisse sur la structure des salaires 2018. Révision du plan de sondage. Numéro OFS : 338-0081

Potterat, J., Qualité, L., Assoulin, D. (2019). Strukturhebung der eidgenössischen Volkszählung. Stichprobenplan und -ziehung, Gewichtung, Schätzverfahren und Pooling 2010-2018. BFS-Nummer : 338-0080

Potterat, J., Assoulin, D., Nicoletti, J.-M. (2019). Beschäftigungsstatistik BESTA. Revision 2015 : Stichprobenrahmen, Stichprobenplan und Hochrechnung. BFS-Nummer : 338-0079

Massiani, A. (2017). Estimation de la couverture du nouveau recensement en Suisse en 2012. Numéro OFS : 338-0078

Nedyalkova, D., Assoulin, D. (2017). Construction of full-time equivalents for the Swiss structural business statistics. FSO number : 338-0077

Ferster, M. (2016). Energieverbrauchsstatistik EVS2014 - Stichprobe, Hochrechnung und Vergleichbarkeit mit der EVS2013. BFS-Nummer : 338-0076

Panchard, C. (2014). Enquête sur la participation aux activités culturelles 2008. Plan de sondage et estimations. BFS-Nummer : 338-0073

Assoulin, D. (2014). Wertschöpfungsstatistik. Revision 2009 : Statistische Datenaufbereitung und Hochrechnung. BFS-Nummer : 338-0072

Wilhelm, M. (2014). Echantillonnage boule de neige. La méthode de sondage déterminé par les répondants. Numéro OFS : 338-0071

Assoulin, D. (2013). Wertschöpfungsstatistik. Revision 2009 : Stichprobenrahmen und Stichprobenplan. BFS-Nummer : 338-0070

Ferster, M. (2013). EVS I - Energieverbrauchsstatistik 2002 bis 2007 : Stichprobenplan und Hochrechnung. BFS-Nummer : 338-0069

Assoulin, D. (2013). Zusatzerhebung für die landwirtschaftliche Betriebszählung 2010 : Stichprobenplan und Hochrechnung. BFS-Nummer : 338-0068

Potterat, J. (2012). Use of conversion keys for NOGA2002-2008. FSO number : 338-0067-05

Andrade, B., Salamin P.-A. (2012). Enquête sur la situation sociale et économique des étudiant-e-s des hautes écoles suisses 2009. Cadre de sondage, plan d'échantillonnage et méthodes d'estimation. Numéro OFS : 338-0066

Potterat, J., Panchard, C., Kilchmann, D. (2012). Umweltschutzausgaben der Unternehmen 2009 (UWSA2009). Stichprobenplan, Einsetzungen, Gewichtung und Schätzverfahren. BFS-Nummer : 338-0065

Potterat, J. (2012). Benutzung der Umsteigeschlüssel NOGA 2002-2008. BFS-Nummer : 338-0064

Potterat, J. (2011). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2009 (KNBB09). Stichprobenplan, Gewichtung und Schätzverfahren. BFS-Nummer : 338-0063

- Kilchmann, D., Potterat, J., Genoud, S. (2011). Gütertransporterhebung 2008. Stichprobenplan, Datenaufbereitung, Gewichtung und Schätzverfahren. BFS-Nummer : 338-0062
- Graf, E. (2010). Enquête suisse sur la santé 2007. Plan d'échantillonnage, pondérations et analyses pondérées des données. Numéro OFS : 338-0061
- Eichenberger P., Hulliger B., Potterat J. (2010). Describing the Anticipated Accuracy of the Swiss Population Survey. FSO number : 338-0060
- Graf, E. (2010). Étude empirique de l'attrition du Panel Suisse de Ménages : vers une caractérisation du profil du non-répondant. Numéro OFS : 338-0059
- Graf, E. (2009). Weightings of the Swiss Household Panel : SHP_I wave 9, SHP_II wave 4, SHP_I et SHP_II combined. FSO number : 338-0058
- Graf, E. (2009). Pondérations du Panel Suisse de Ménages : PSM_I vague 9, PSM_II vague 4, PSM_I et PSM_II combinés. Numéro OFS : 338-0057-05
- Qualité, L., Tillé, Y. (2009). Estimation de la précision d'évolutions dans l'enquête sur la valeur ajoutée. Numéro OFS : 338-0056
- Renaud, A., Panchard, C. et Potterat, J. (2008). Statistique de l'emploi. Révision 2007 : méthodes d'estimation. Numéro OFS : 338-0055
- Graf, E. (2008). Pondérations du PSM. PSM_I vague 8, PSM_II vague 3, PSM_I et PSM_II combinés. Numéro OFS : 338-0054
- Andrade, B., Graf, M. (2008). Enquête suisse sur la structure des salaires 2006. Aspects méthodologiques du modèle des salaires "Salarium". Numéro OFS : 338-0053
- Renaud, A. (2008). Statistique de l'emploi. Révision 2007 : cadre de sondage et échantillonnage. Numéro OFS : 338-0052
- Graf, E. (2008). Pondérations du SILC pilote. SILC_I vague 2, SILC_II vague 1, SILC_I et SILC_II combinés. Numéro OFS : 338-0051
- Kilchmann, D. (2008). Statistik der sozialmedizinischen Institutionen 1999-2004 und Krankenhausstatistik 1999-2002. Einsetzungen für fehlende Daten. BFS-Nummer : 338-0050
- Renaud, A. (2008). Technologies de l'information et de la communication. Estimations sur la base de la statistique de la valeur ajoutée. Numéro OFS : 338-0049
- Assoulin, D. (2007). Wertschöpfungsstatistik. Einsetzungsversuche für fehlende Antworten grosser Unternehmen. BFS-Nummer : 338-0048
- Kilchmann, D. (2007). Beherbergungsstatistik Campingplätze. Stichprobenrahmen und Schätzverfahren 2005/06. BFS-Nummer : 338-0047
- Gabler, S., Häder, S. (2007). Haushalts- und Personenerhebungen. Machbarkeit von Random Digit Dialing in der Schweiz. BFS-Nummer : 338-0046
- Ferrez, J., Graf, M. (2007). Enquête suisse sur la structure des salaires. Programmes R pour l'intervalle de confiance de la médiane. Numéro OFS : 338-0045
- Renaud, A. (2007). Harmonisation de la scolarité obligatoire en Suisse (HarmoS). Design général de l'enquête et échantillon des écoles. Numéro OFS : 338-0044
- Potterat, J. (2007). Betriebszählung 2005. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. BFS-Nummer : 338-0043
- Hulliger, B. (2006). Umweltschutzausgaben der Unternehmen 2003, Stichprobenplan, Datenaufbereitung und Schätzverfahren. BFS-Nummer : 338-0042
- Renfer, J.-P. (2006). Enquête sur les chiffres d'affaires du commerce de détail. Plan d'échantillonnage et méthodes d'estimation. Numéro OFS : 338-0041
- Salamin, P.-A. (2006). Statistique de l'aide sociale dans le domaine de l'asile. Plan de sondage et extrapolations pour l'enquête pilote 2005. Numéro OFS : 338-0040

- Renaud, A. (2006). Statistique suisse des bénéficiaires de l'aide sociale. Pondération des communes 2004. Numéro OFS : 338-0039
- Graf, M. (2006). Swiss Earnings Structure Survey 2002-2004. Compositional data in a stratified two-stage sample : Analysis and precision assessment of wage components. FSO number : 338-0038
- Potterat, J. (2006). Pensionskassenstatistik 2004. Statistische Methoden zur Schätzung der provisorischen Ergebnisse. BFS-Nummer : 338-0037
- Potterat, J. (2006). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe im Jahr 2004. Stichprobenplan, Gewichtung und Schätzverfahren. BFS-Nummer : 338-0036
- Kilchmann, D. (2006). Vierteljährliche Wohnbaustatistik. Stichprobenplan, statistische Datenaufarbeitung und Schätzverfahren 2005. BFS-Nummer : 338-0035
- Kilchmann, D. (2006). Erhebung über Forschung und Entwicklung in der schweizerischen Privatwirtschaft 2004. Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. BFS-Nummer : 338-0034
- Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 2. BFS-Nummer : 338-0033
- Kilchmann, D., Eichenberger, P., Potterat, J. (2005). Volkszählung 2000. Statistische Einsetzungsverfahren Band 1. BFS-Nummer : 338-0032
- Graf, M., Matei, A. (2005). Enquête suisse sur la structure des salaires 2002. La précision du salaire brut standardisé médian. Numéro OFS : 338-0031
- Graf, E., Renfer, J.-P. (2005). Enquête suisse sur la santé 2002. Plan d'échantillonnage, pondération et estimation de la précision. Numéro OFS : 338-0030
- Potterat, J. (2005). Mietpreis-Strukturerhebung 2003. Gewichtung und Schätzverfahren. BFS-Nummer : 338-0029
- Potterat, J. (2005). Landwirtschaftliche Betriebszählung 2003. Schätzverfahren für die Zusatzerhebung. BFS-Nummer : 338-0028
- Renaud, A. (2004). Coverage estimation for the Swiss population census 2000. Estimation methodology and results. FSO number : 338-0027
- Kilchmann, D. (2004). Revision des Schweizerischen Lohnindex. Schätzmethoden der Lohnindices und deren Varianzschätzer. BFS-Nummer : 338-0026
- Graf, M. (2004). Enquête suisse sur la structure des salaires 2002. Plan d'échantillonnage et extrapolation pour le secteur privé. Numéro OFS : 338-0025
- Renaud, A. (2004). Analyse de données d'enquêtes. Quelques méthodes et illustration avec des données de l'OFS. Numéro OFS : 338-0024
- Renaud, A., Potterat, J. (2004). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sous-couverture (P-sample) et qualité du cadre de sondage des bâtiments. Numéro OFS : 338-0023
- Graf, M. (2004). Fusion de données. Etude de faisabilité. Numéro OFS : 338-0022
- Potterat, J. (2003). Mietpreis-Strukturerhebung 2003. Entwicklung des Stichprobenplans und Ziehung der Stichprobe. BFS-Nummer : 338-0021
- Potterat, J. (2003). Landwirtschaftliche Betriebszählung 2003. Stichprobenplan der Zusatzerhebung. BFS-Nummer : 338-0020.
- Renaud, A. (2003). Estimation de la couverture du recensement de la population de l'an 2000. Echantillon pour l'estimation de la sur-couverture (E-sample). Numéro OFS : 338-0019
- Hulliger, B. (2003). Bereinigung der Stichprobe, Ersatz fehlender Werte und Schätzverfahren. Erhebung über F+E in der schweizerischen Privatwirtschaft 2000. BFS-Nummer : 338-0018
- Renfer, J.-P. (2003). Enquête 2000 sur la recherche et le développement dans l'économie privée en Suisse. Plan d'échantillonnage. Numéro OFS : 338-0017

- Potterat, J. (2003). Kosten und Nutzen der Berufsbildung aus Sicht der Betriebe. Schätzverfahren. BFS-Nummer : 338-0016
- Graf, M., Matei, A. (2003). Stratégie de choix des modèles de désaisonnalisation. Application aux séries de l'emploi total. Numéro OFS : 338-0015
- Potterat, J., Salamin, P.A. (2002). Betriebszählung 2001. Methoden für die Datenbereinigung. BFS-Nummer : 338-0014
- Renaud, A. (2002). Programme international pour le suivi des acquis des élèves (PISA). Plans d'échantillonnage pour PISA 2000 en Suisse. Numéro OFS : 338-0013
- Renfer, J.-P. (2002). Enquête 2001 sur les coûts et l'utilité de la formation des apprentis du point de vue des établissements. Plan d'échantillonnage. Numéro OFS : 338-0012
- Potterat, J., Salamin, P.A. (2002). Betriebszählung 2001. Stichprobenplan und Schätzverfahren für die provisorischen Ergebnisse. BFS-Nummer : 338-0011
- Graf, M. (2002). Enquête suisse sur la structure des salaires 2000. Plan d'échantillonnage, pondération et méthode d'estimation pour le secteur privé. Numéro OFS : 338-0010
- Renaud, A., Eichenberger P. (2002). Estimation de la couverture du recensement de la population de l'an 2000. Procédure d'enquête et plan d'échantillonnage de l'enquête de couverture. Numéro OFS : 338-0009
- Kilchmann, D., Hulliger, B. (2002). Stichprobenplan für die Obstbaumzählung 2001. BFS-Nummer : 338-0008
- Graf, M. (2002). Passage du concept établissement au concept entreprise. Numéro OFS : 338-0007
- Salamin, P.A. (2001). La technique de la double enquête pour la statistique du transport routier de marchandise. Numéro OFS : 338-0006
- Peters, R., Renfer, J.-P. et Hulliger, B. (2001). Statistique de la valeur ajoutée 1997-1998. Procédure d'extrapolation des données. Numéro OFS : 338-0005
- Potterat, J., Hulliger, B. (2001). Schätzung der Sägereiproduktion mit der Sägerei-Erhebung PAUL. BFS-Nummer : 338-0004
- Graf, M. (2001). Désaisonnalisation. Aspects méthodologiques et application à la statistique de l'emploi. Numéro OFS : 338-0003
- Hüsler, J., Müller, S. (2001). Schlussbericht Betriebszählung 1995 (BZ 95), Mehrfach imputierte Umsatzzahlen. BFS-Nummer : 338-0002
- Renaud, A. (2001). Statistique suisse des bénéficiaires de l'aide sociale. Plan d'échantillonnage des communes. Numéro OFS : 338-0001
- Hulliger, B., Eichenberger, P. (2000). Stichprobenregister für Haushalterhebungen : Umstellung auf Telefonnummern ohne Namen und Adressen, Abläufe für Erstellung und Stichprobenziehung. BFS-Nummer : 338-0000

Programme des publications de l'OFS

En tant que service statistique central de la Confédération, l'Office fédéral de la statistique (OFS) a pour tâche de rendre les informations statistiques accessibles à un large public. Il utilise plusieurs moyens et canaux pour diffuser ses informations statistiques par thème.

Les domaines statistiques

- 00 Bases statistiques et généralités
- 01 Population
- 02 Espace et environnement
- 03 Travail et rémunération
- 04 Économie nationale
- 05 Prix
- 06 Industrie et services
- 07 Agriculture et sylviculture
- 08 Énergie
- 09 Construction et logement
- 10 Tourisme
- 11 Mobilité et transports
- 12 Monnaie, banques, assurances
- 13 Sécurité sociale
- 14 Santé
- 15 Éducation et science
- 16 Culture, médias, société de l'information, sport
- 17 Politique
- 18 Administration et finances publiques
- 19 Criminalité et droit pénal
- 20 Situation économique et sociale de la population
- 21 Développement durable, disparités régionales et internationales

Les principales publications générales

L'Annuaire statistique de la Suisse



L'Annuaire statistique de la Suisse de l'OFS constitue depuis 1891 l'ouvrage de référence de la statistique suisse. Il englobe les principaux résultats statistiques concernant la population, la société, l'État, l'économie et l'environnement de la Suisse.

Le Mémento statistique de la Suisse



Le mémento statistique résume de manière concise et attrayante les principaux chiffres de l'année. Cette publication gratuite de 52 pages au format A6/5 est disponible en cinq langues (français, allemand, italien, romanche et anglais).

Le site Internet de l'OFS: www.statistique.ch

Le portail «Statistique suisse» est un outil moderne et attrayant vous permettant d'accéder aux informations statistiques actuelles. Nous attirons ci-après votre attention sur les offres les plus prisées.

La banque de données des publications pour des informations détaillées

Presque tous les documents publiés par l'OFS sont disponibles gratuitement sous forme électronique sur le portail Statistique suisse (www.statistique.ch). Pour obtenir des publications imprimées, vous pouvez passer commande par téléphone (+41 58 463 60 60) ou par e-mail (order@bfs.admin.ch).
www.statistique.ch → Trouver des statistiques → Catalogues et banques de données → Publications

Vous souhaitez être parmi les premiers informés?



Abonnez-vous à un Newsmail et vous recevrez par e-mail des informations sur les résultats les plus récents et les activités actuelles concernant le thème de votre choix.
www.news-stat.admin.ch

STAT-TAB: la banque de données statistiques interactive



La banque de données statistiques interactive vous permet d'accéder simplement aux résultats statistiques dont vous avez besoin et de les télécharger dans différents formats.
www.statatab.bfs.admin.ch

Statatlas Suisse: la banque de données régionale avec ses cartes interactives



L'atlas statistique de la Suisse, qui compte plus de 4500 cartes, est un outil moderne donnant une vue d'ensemble des thématiques régionales traitées en Suisse dans les différents domaines de la statistique publique.
www.statatlas-suisse.admin.ch

Pour plus d'informations

Centre d'information statistique

+41 58 463 60 11, info@bfs.admin.ch

La correction des effets calendrier d'une série temporelle consiste à prendre en compte la part des variations de la série qui est purement liée à la variation du calendrier. Ce rapport présente les méthodes utilisées à l'Office Fédéral de la Statistique (OFS) pour effectuer ces ajustements. Cette démarche s'inscrit dans le contexte plus vaste de la désaisonnalisation des séries temporelles à l'aide d'outils tels que **X13-ARIMA-SEATS**. Ces outils sont aujourd'hui bien établis et largement répandus en statistique publique, mais leur application n'est pas aisée si l'utilisateur souhaite prendre en compte les spécificités du calendrier suisse et paramétrer son influence sur la série analysée. Ce rapport vise justement à faciliter cette démarche et se focalise uniquement sur l'influence du calendrier sur la série étudiée, en vue de produire une série corrigée de ses effets calendrier. Après une brève introduction au contexte général de la désaisonnalisation et aux principaux défis liés à la correction d'effets calendrier, une construction détaillée du calendrier suisse, utilisé par l'OFS et mis à disposition des utilisateurs, est présentée. Le calendrier suisse permet de calculer des séries temporelles nommées «régresseurs» selon plusieurs jeux d'hypothèses concernant l'influence des différents jours de la semaine (fériés ou non) sur les valeurs de la série temporelle d'étude. Les régresseurs permettent finalement aux utilisateurs de produire la série corrigée de ses effets calendrier. Cette démarche et les différents jeux d'hypothèses implémentés sont détaillés ainsi que des références et des outils d'aide à la décision pour choisir le jeu d'hypothèses le mieux adapté à la série temporelle étudiée. Des fichiers contenant les régresseurs ainsi que les programmes qui génèrent ces fichiers sont publiquement accessibles sur la page <https://www.bfs.admin.ch/bfs/fr/home/dscc/outils/methodes-statistiques/2023-06-calendrier.html>.

En ligne

www.statistique.ch

Imprimés

www.statistique.ch
Office fédéral de la statistique
CH-2010 Neuchâtel
order@bfs.admin.ch
tél. +41 58 463 60 60

Numéro OFS

338-0085

ISBN

978-3-303-00720-4

Les informations publiées ici contribuent à mesurer la réalisation des objectifs de développement durable (ODD).



Système d'indicateurs MONET 2030

www.statistique.ch → Trouver des statistiques → Développement durable → Système d'indicateurs MONET 2030

**La statistique
compte pour vous.**

www.la-statistique-compte.ch