Rapport de méthodes

Passage du concept établissement au concept entreprise

0

Statistische Grundlagen und Übersichten
Bases statistiques et vues d'ensemble
Basi statistiche e presentazioni generali

OFS BFS UST

Office fédéral de la statistique
Bundesamt für Statistik
Ufficio federale di statistica
Uffizi federal da statistica

Neuchâtel, 2002

Die vom Bundesamt für Statistik (BFS) herausgegebene Reihe «Statistik der Schweiz» gliedert sich in folgende Fachbereiche:

0 Statistische Grundlagen und Übersichten
1 Bevölkerung
2 Raum und Umwelt
3 Arbeit und Erwerb
4 Volkswirtschaft
5 Preise
6 Industrie und Dienstleistungen
7 Land- und Forstwirtschaft
8 Energie
9 Bau- und Wohnungswesen
10 Tourismus
11 Verkehr und Nachrichtenwesen
12 Geldpolitik, Finanzmärkte und -akteure
13 Soziale Sicherheit
14 Gesundheit
15 Bildung und Wissenschaft
16 Kultur, Medien, Zeitverwendung
17 Politik
18 Öffentliche Finanzen
19 Rechtspflege
20 Gesellschaft in Bewegung (Querschnittsthemen)

La série «Statistique de la Suisse» publiée par l'Office fédéral de la statistique (OFS) couvre les domaines suivants:

0 Bases statistiques et vues d'ensemble
1 Population
2 Espace et environnement
3 Travail et vie active
4 Economie nationale
5 Prix
6 Industrie et services
7 Agriculture et sylviculture
8 Energie
9 Construction et logement
10 Tourisme
11 Transports et communications
12 Politique monétaire, marchés et acteurs financiers
13 Protection sociale
14 Santé
15 Education et science
16 Culture, médias, emploi du temps
17 Politique
18 Finances publiques
19 Droit et justice
20 Société en mutation (thèmes interdisciplinaires)

Statistik der Schweiz

Methodenbericht

Statistique de la Suisse

Rapport de méthodes

Passage du concept établissement au concept entreprise

Auteur **Monique Graf**
Office fédéral de la statistique

Editeur **Office fédéral de la statistique**

OFS BFS UST

Office fédéral de la statistique
Bundesamt für Statistik
Ufficio federale di statistica
Uffizi federal da statistica

Neuchâtel, 2002

Préambule et remerciements

Le problème du passage du concept établissement au concept entreprise a été posé en juin 2001 par Alain Vuille de la section EWL. La solution à ce problème a consisté à placer l'analyse dans un contexte matriciel, ce qui est le sujet de ce rapport. A partir de là, une collaboration efficace s'est instaurée, puisqu'Alain Vuille a pu reprendre les programmes S-plus développés pour la période 1995-2000, et les étendre au cas différent de 1991-1995. Jean-Pierre Renfer de Meth m'a aidée à comprendre dans quel contexte l'emploi en équivalents plein-temps était utilisé pour la statistique de la valeur ajoutée. Il a aussi relu ce rapport, et grâce à lui la lisibilité en est nettement améliorée. Gregory Rais de la section VG m'a éclairée sur les groupes de formes juridiques. Qu'ils soient tous chaleureusement remerciés.

Résumé

La statistique de l'emploi (BESTA) est basée sur un échantillon d'établissements. Elle est utilisée pour calibrer d'autres statistiques fédérales, comme la statistique de la valeur ajoutée qui, elle, est établie à partir d'un échantillon d'entreprises. De plus, le cadre d'échantillonnage n'est pas exactement le même dans les deux cas.

Le présent rapport décrit une méthode d'harmonisation ou de passage entre le concept établissement et le concept entreprise. Le recensement des entreprises fournit une matrice donnant le nombre d'emplois par combinaison des classes d'activité des entreprises et des établissements. C'est à partir de cette matrice qu'une matrice de passage d'un concept à l'autre est déduite, et que l'univers de sondage est adapté. La statistique de l'emploi fournit l'évolution conjoncturelle.

Mots-clé

rapport de méthodes; statistique de l'emploi; BESTA; STATEM; entreprise; établissement; recensement des entreprises; RE; BZ; statistique de la valeur ajoutée; équivalent plein temps; EPT; BESTA adaptée; évolution conjoncturelle; NOGA; modèle d'adaptation; interpolation; matrice de passage

---

Complément d'information: Monique Graf, tél. 032 713 66 15
monique.graf@bfs.admin.ch

Réalisation: Service de méthodes statistiques, OFS

Diffusion: Office fédéral de la statistique
CH-2010 Neuchâtel
Tél. 032 713 60 60 / Fax 032 713 60 61
order@bfs.admin.ch

Internet: http://www.statistik.admin.ch

Numéro de commande: 338-0007

Prix: gratuit

Série: Statistique de la Suisse

Domaine: 0 Bases statistiques et vues d'ensemble

Langue du texte original: Français

Graphisme/Layout: OFS

Copyright: OFS, Neuchâtel 2002
La reproduction est autorisée, sauf à des fins commerciales,
si la source est mentionnée.

ISBN: 3-303-00232-0

Table des matières

1 Introduction **5**

2 Matrices de passage des EPT du concept établissement au concept entreprise **5**

3 Interpolation au niveau NOGA2 entre deux RE dans le concept entreprise **7**

4 Passage à une sous-catégorie **8**

5 Interpolation non linéaire des matrices de passage **8**

6 Matrices de passage des EPT du concept établissement au concept entreprise entre 1991 et 1995 **11**

7 Analyse des résultats sur le calcul des EPT au niveau NOGA5 x forme juridique **14**

Annexes **17**

Bibliographie **28**

1 Introduction

La statistique de l'emploi (BESTA) est une statistique trimestrielle basée sur un échantillon aléatoire d'établissements conçu au niveau NOGA2. Cette statistique est adaptée à l'évolution mesurée d'un recensement des entreprises (RE) à l'autre, ces recensements ayant lieu tous les 3 ans, de telle sorte qu'aux trimestres concernés la BESTA adaptée coïncide avec le RE, voir [1], [2], [3] et [4]. Il y a donc tous les 3 ans des révisions de la statistique de l'emploi. Pour d'autres statistiques, comme celle de la valeur ajoutée, l'univers ne correspond pas exactement au cadre d'échantillonnage de la BESTA, et de plus se base sur un échantillon d'entreprises et non d'établissements. Le but de cette étude est d'une part d'établir un passage du concept entreprise au concept établissement, et d'autre part de donner des estimations de l'emploi au niveau de sous-classes économiques, qui permettent de tenir compte d'univers différents pour d'autres statistiques. Le passage au concept entreprise se base sur les RE et la BESTA adaptée. Les estimateurs obtenus pour les sous-classes sont des estimateurs synthétiques utilisant pour le niveau de détail souhaité l'information fournie par les RE, et pour l'évolution conjoncturelle la BESTA au niveau NOGA2. Ils ne seront utilisés que pour ajuster l'univers. Les pondérations et les estimations publiées se feront au niveau NOGA2 uniquement.

L'emploi est ici mesuré en équivalents plein-temps (EPT).

Contraintes :

- Les estimations obtenues doivent coïncider avec les EPT des entreprises donnés par les RE aux trimestres concernés. Les RE en 1991, 95 et 2000 ont eu lieu à chaque fois au 3e trimestre. Les contraintes sont donc toujours fixées au même trimestre.
- Les estimations doivent aussi permettre de descendre à des niveaux d'activité plus détaillés, tout en garantissant l'additivité au niveau NOGA2. Dans la pratique, les sous-classes sont données par la NOGA5 croisée avec le groupe de formes juridiques (voir Annexe).

2 Matrices de passage des EPT du concept établissement au concept entreprise

Considérons, pour fixer les idées, le RE98, et posons :

$\mathbf{b}_{98} =$ vecteur-colonne des EPT du RE98 pour les 49 classes NOGA2[1] ($\mathbf{b}$ pour Betrieb, établissement).

$\mathbf{u}_{98} =$ vecteur-colonne des EPT du RE98 pour les 49 classes NOGA2 ($\mathbf{u}$ pour Unternehmen, entreprise).

$\mathbf{A}_{98}(i,j) =$ EPT agrégé des établissements de classe NOGA2 j appartenant à des entreprises de classe NOGA2 i.

$\mathbf{1} =$ vecteur-colonne de "1" de longueur 49.

[1] Classes NOGA2 avec regroupement des classes 10-14, 36-37,40-41

Donc, $\mathbf{u}_{98}(i) = \sum_j \mathbf{A}_{98}(i,j)$ et $\mathbf{b}_{98}(j) = \sum_i \mathbf{A}_{98}(i,j)$, ou, sous forme plus compacte :

$$\begin{aligned} \mathbf{u}_{98} &= \mathbf{A}_{98} \cdot \mathbf{1} \\ \mathbf{b}_{98} &= \mathbf{A}_{98}^t \cdot \mathbf{1} \end{aligned} \tag{1}$$

Après vérification que $\mathbf{A}_{98}$ est bien inversible, on obtient :

$$\mathbf{A}_{98}^{-1} \cdot \mathbf{u}_{98} = \mathbf{1} = \left(\mathbf{A}_{98}^t\right)^{-1} \cdot \mathbf{b}_{98}$$

donc, en prémultipliant par $\mathbf{A}_{98}$, on a

$$\mathbf{u}_{98} = \mathbf{A}_{98} \cdot \left(\mathbf{A}_{98}^t\right)^{-1} \cdot \mathbf{b}_{98}$$

Posons

$$\mathbf{P}_{98} = \mathbf{A}_{98} \cdot \left(\mathbf{A}_{98}^t\right)^{-1} \tag{2}$$

$\mathbf{P}_{98}$ est une matrice de passage de $\mathbf{u}_{98}$ à $\mathbf{b}_{98}$:

$$\mathbf{u}_{98} = \mathbf{P}_{98} \cdot \mathbf{b}_{98} \tag{3}$$

Remarques :

- $\mathbf{P}_{98}$ est la seule matrice de passage du concept établissement au concept entreprise qui conserve toute l'information contenue dans $\mathbf{A}_{98}$. En effet

  $$\mathbf{A}_{98} = \mathbf{P}_{98} \cdot \mathbf{A}_{98}^t$$

  Dans ce sens, c'est la transformation optimale qui vérifie (3).
- Il existe une infinité de matrices de passage possibles de $\mathbf{b}_{98}$ à $\mathbf{u}_{98}$.
  La plus simple est de rang 1 ; c'est le projecteur orthogonal de $\mathbf{b}_{98}$ sur $\mathbf{u}_{98}$. Appelons-la $\mathbf{P}_{1,98.}$

  $$\mathbf{P}_{1,98} = \frac{1}{\mathbf{b}_{98}^t \cdot \mathbf{b}_{98}} \mathbf{u}_{98} \cdot \mathbf{b}_{98}^t \tag{4}$$

  On voit facilement que $\mathbf{P}_{1,98} \cdot \mathbf{b}_{98} = \mathbf{u}_{98}$, mais que $\mathbf{P}_{1,98} \cdot \mathbf{A}_{98}^t$ est une matrice dont toutes les colonnes sont proportionnelles à $\mathbf{u}_{98}$. On choisit donc la matrice $\mathbf{P}_{98}$ (2).

On constate (tableau ci-dessous) que les termes diagonaux des matrices de passage $\mathbf{P}_{95}$ et $\mathbf{P}_{98}$ sont très proches de 1, et que les termes hors diagonale sont généralement proches de 0. Cela vient du fait que l'essentiel des emplois se situent sur la diagonale de $\mathbf{A}_{95}$ et $\mathbf{A}_{98}$. En comparant la distribution des termes des matrices de passage, on voit que la classe économique des établissements est un peu plus souvent différente de celle de leur entreprise (en termes d'EPT) en 1995 qu'en 1998. Voir aussi en Annexe l'analyse des différences entre $\mathbf{u}_{95}$ (resp. $\mathbf{u}_{98}$) et $\mathbf{b}_{95}$ (resp. $\mathbf{b}_{98}$), ainsi que l'évolution de l'emploi dans les classes NOGA2 66 et 67.

Distribution des éléments des matrices de passage

	Eléments	Min.	1er Qu.	Médiane	Moy.	3ème Qu.	Max.
$\mathbf{P}_{98}$	diagonaux	0.9935	0.9998	1	0.9997	1	1.002
$\mathbf{P}_{98}$	hors diagonale	-0.0948	-5.603e-006	-4.821e-009	0.0004	3.894e-006	0.5479
$\mathbf{P}_{95}$	diagonaux	0.9391	0.9997	1	0.9984	1	1.005
$\mathbf{P}_{95}$	hors diagonale	-0.1884	-1.805e-005	-1.028e-008	0.0008	1.650e-005	0.6962

3 Interpolation au niveau NOGA2 entre deux RE dans le concept entreprise

La BESTA est calée sur le RE aux trimestres où le recensement est effectué, et un facteur de correction est appliqué à la valeur observée de la BESTA aux trimestres intermédiaires, voir [1], [2], [3] et [4]. Ce facteur dépend de la branche économique et du nombre de trimestres séparant le trimestre t des deux RE adjacents. On parle alors de BESTA *adaptée* qui peut être vue comme une estimation du RE au trimestre t. Considérons ici les RE ayant eu lieu au 3e trimestre de 1995 et de 1998 (ces trimestres sont notés $3Q95$ et $3Q98$).

Posons

$\widehat{\mathbf{b}}_t =$ vecteur-colonne de longueur 49 des EPT de la BESTA adaptée au trimestre t.
$\widehat{\mathbf{u}}_t =$ vecteur-colonne déduit selon le concept entreprise.
$w_t = (t - 3Q95)/12$ poids de $\mathbf{P}_{98}$.

Nous proposons de calculer $\widehat{\mathbf{u}}_t$ de la manière suivante :

$$\begin{aligned} \widehat{\mathbf{u}}_t &= (w_t \mathbf{P}_{98} + (1 - w_t)\ \mathbf{P}_{95}) \cdot \widehat{\mathbf{b}}_t \qquad 3Q95 \leq t \leq 3Q98 \\ &= \widehat{\mathbf{P}}_t \cdot \widehat{\mathbf{b}}_t \end{aligned} \tag{5}$$

La matrice $\widehat{\mathbf{P}}_t$ est une interpolation linéaire des matrices de passage des deux RE, et estime la matrice de passage au trimestre t. Remarquons que la linéarité de l'interpolation est nécessaire pour respecter la contrainte d'additivité au niveau des sous-catégories.

On a alors :

$$\begin{aligned} \widehat{\mathbf{u}}_{3Q98} &= \mathbf{u}_{98} \\ \widehat{\mathbf{u}}_{3Q95} &= \mathbf{u}_{95} \end{aligned} \tag{6}$$

4 Passage à une sous-catégorie

Soit c une sous-catégorie d'une classe NOGA2, et $\mathbf{A}_{98}^{(c)}$ le vecteur-ligne des EPT correspondant, c'est-à-dire :

$\mathbf{A}_{98}^{(c)}(j) =$ EPT agrégés des établissements de classe NOGA2 j appartenant à une entreprise de la sous-classe c

De même pour 1995.

Posons encore

$$\begin{aligned} \mathbf{P}_{98}^{(c)} &= \mathbf{A}_{98}^{(c)} \cdot \left(\mathbf{A}_{98}^{t}\right)^{-1} \\ \mathbf{P}_{95}^{(c)} &= \mathbf{A}_{95}^{(c)} \cdot \left(\mathbf{A}_{95}^{t}\right)^{-1} \end{aligned} \tag{7}$$

Alors l'emploi estimé de la sous-catégorie c dans le concept entreprise est donné par

$$\begin{aligned} \widehat{u}_t^{(c)} &= \left(w_t \mathbf{P}_{98}^{(c)} + (1 - w_t)\, \mathbf{P}_{95}^{(c)}\right) \cdot \widehat{\mathbf{b}}_t \qquad && 3Q95 \leq t \leq 3Q98 \\ \widehat{u}_t^{(c)} &= \mathbf{P}_{98}^{(c)} \cdot \widehat{\mathbf{b}}_t && 3Q98 \leq t \leq 4Q00 \end{aligned} \tag{8}$$

Les valeurs aux trimestres $3Q99$ et $3Q00$ seront révisées, lorsque le RE01 sera disponible.

On a bien l'additivité des sous-classes

$$\sum_{c \in i} \widehat{u}_t^{(c)} = \widehat{\mathbf{u}}_t(i)$$

car, par exemple,

$$\sum_{c \in i} \mathbf{A}_{98}^{(c)}(j) = \mathbf{A}_{98}(i, j)$$

5 Interpolation non linéaire des matrices de passage

La matrice de passage $\widehat{\mathbf{P}}_t$ au trimestre t (5) a été estimée par interpolation linéaire de $\mathbf{P}_{95}$ et $\mathbf{P}_{98}$.

On pourrait aussi penser à une interpolation non linéaire, analogue au modèle développé pour l'adaptation de la BESTA aux RE. Nous allons d'abord généraliser ce modèle à la formulation matricielle utilisée pour le passage du concept établissement au concept entreprise, puis démontrer que $\widehat{\mathbf{P}}_t$ est l'approximation du 1er ordre de la matrice de passage obtenue par analogie avec ce modèle d'adaptation.

Modèle d'adaptation

Rappelons que dans ce modèle, la correction de l'accroissement total de la statistique de l'emploi est effectuée dans le but de faire coïncider cet accroissement avec celui des RE sur la même période, de telle sorte que cette correction soit répartie uniformément sur les accroissements relatifs des trimestres concernés. Cela mène à la formule de correction que nous rappelons ici :

Pour la classe NOGA2 $i, i = 1, ..., 49$, posons

$b_t(i) =$ EPT de la BESTA brute, mais corrigée des effets non dûs à l'adaptation aux RE (essentiellement le changement d'échantillon au $1Q97$).

$\widehat{b}_t(i) =$ EPT de la BESTA adaptée (i-ème composante de $\widehat{\mathbf{b}}_t$, cf §2).

$z_{95}(i) =$ rapport des EPT du RE95 et de l'échantillon au $3Q95$.

$z_{98}(i) =$ idem pour 1998.

La formule d'adaptation s'écrit alors

$$\widehat{b}_t(i) = (z_{95}(i))^{1-w_t} (z_{98}(i))^{w_t} b_t(i) \tag{9}$$

En introduisant $\mathbf{Z}_{95}$ et $\mathbf{Z}_{98}$, matrices diagonales ayant $z_{95}(i)$ et $z_{98}(i)$ comme i-ème terme diagonal, (9) s'écrit directement sous forme matricielle :

$$\widehat{\mathbf{b}}_t = \mathbf{Z}_{95}^{1-w_t} \cdot \mathbf{Z}_{98}^{w_t} \cdot \mathbf{b}_t \tag{10}$$

Remarque Appliquer une fonction (comme l'élévation à la puissance w_t) à une matrice diagonale ne pose aucun problème : il suffit d'appliquer cette fonction à chaque élément diagonal.

Ecrivons encore (10) sous une forme équivalente, utile par la suite :

$$\widehat{\mathbf{b}}_t = \exp\left[(1 - w_t) \ln(\mathbf{Z}_{95}) + w_t \ln(\mathbf{Z}_{98})\right] \cdot \mathbf{b}_t \tag{11}$$

Transposition aux matrices de passage

Revenons aux matrices de passage du concept établissement au concept entreprise. Il serait dans la logique du modèle d'adaptation aux RE de faire une interpolation non linéaire de $\mathbf{P}_{95}$ et $\mathbf{P}_{98}$. Comme $\mathbf{P}_{95}$ et $\mathbf{P}_{98}$ ne sont pas symétriques, on a que :

$$\mathbf{P}_{95} \cdot \mathbf{P}_{98} \neq \mathbf{P}_{98} \cdot \mathbf{P}_{95}$$

La forme (10) n'est donc pas directement transposable aux matrices de passage.

Ecrivons formellement la formule d'interpolation analogue à (11) :

$$\widetilde{\mathbf{P}}_t = \exp\left[(1 - w_t) \ln(\mathbf{P}_{95}) + w_t \ln(\mathbf{P}_{98})\right] \tag{12}$$

Le calcul d'une fonction d'une matrice $\mathbf{P}$ - dans le cas le plus simple où $\mathbf{P}$ est diagonalisable - s'effectue en 2 étapes :

– Trouver la matrice diagonale $\mathbf{D}$ des valeurs propres et une matrice inversible $\mathbf{Q}$ de vecteurs propres à droite, telles que $\mathbf{P} = \mathbf{Q}^{-1} \cdot \mathbf{D} \cdot \mathbf{Q}$.

– Calculer $f(\mathbf{P}) = \mathbf{Q}^{-1} \cdot f(\mathbf{D}) \cdot \mathbf{Q}$

Les valeurs propres de $\mathbf{P}_{95}$ et $\mathbf{P}_{98}$ étant toutes distinctes, $\mathbf{P}_{95}$ et $\mathbf{P}_{98}$ sont diagonalisables. Elles n'ont qu'une valeur propre réelle, toutes les autres sont complexes, mais cela ne complique pas excessivement le calcul. On trouve donc par la méthode décrite ci-dessus $\ln(\mathbf{P}_{95})$ et $\ln(\mathbf{P}_{98})$. On forme alors la matrice $(1-w_t)\ln(\mathbf{P}_{95}) + w_t \ln(\mathbf{P}_{98})$ que l'on diagonalise à son tour, et on en calcule enfin l'exponentielle, ce qui nous donne $\widetilde{\mathbf{P}}_t$ dans (12).

Toutes ces manipulations pour calculer $\widetilde{\mathbf{P}}_t$ sont relativement aisées à programmer, voir le programme passage5.ssc en Annexe, mais sont quand même plus compliquées que le calcul immédiat de $\widehat{\mathbf{P}}_t$ dans (5).

Approximation linéaire de $\widetilde{\mathbf{P}}_t$

Montrons maintenant que $\widehat{\mathbf{P}}_t$ est l'approximation linéaire de $\widetilde{\mathbf{P}}_t$.

Le rayon spectral d'une matrice $\mathbf{P}$ est égal au plus grand module de ses valeurs propres.

Si la fonction f possède un développement en série dont le rayon de convergence est r, ce développement en série peut s'appliquer à $f(\mathbf{P})$, pour autant que le rayon spectral de $\mathbf{P}$ soit inférieur à r.

Développement à l'ordre 1 : $\begin{array}{ll} \exp(x) \approx 1 + x & r = 1 \\ \ln(x+1) \approx x & r = 1 \end{array}$

Posons $\mathbf{H} = \mathbf{P} - \mathbf{I}$, où $\mathbf{I}$ est la matrice identité. Alors, si le rayon spectral de $\mathbf{H}$ est <1,

$$\begin{aligned} \ln(\mathbf{P}) &= \ln(\mathbf{H} + \mathbf{I}) \approx \mathbf{H} = \mathbf{P} - \mathbf{I} \\ \exp(\mathbf{H}) &= \exp(\mathbf{P} - \mathbf{I}) \approx \mathbf{P}. \end{aligned}$$

$\mathbf{P}_{95} - \mathbf{I}$ et $\mathbf{P}_{98} - \mathbf{I}$ ont des rayons spectraux de 0.24 et de 0.08, respect., donc bien inférieurs à 1. On peut alors écrire les approximations :

$$\begin{aligned} (1-w_t)\ln(\mathbf{P}_{95}) + w_t \ln(\mathbf{P}_{98}) &\approx (1-w_t)(\mathbf{P}_{95} - \mathbf{I}) + w_t(\mathbf{P}_{98} - \mathbf{I}) \\ &= (1-w_t)\mathbf{P}_{95} + w_t\mathbf{P}_{98} - \mathbf{I} \end{aligned}$$

et

$$\begin{aligned} \widetilde{\mathbf{P}}_t &= \exp[(1-w_t)\ln(\mathbf{P}_{95}) + w_t \ln(\mathbf{P}_{98})] \approx \exp[(1-w_t)\mathbf{P}_{95} + w_t\mathbf{P}_{98} - \mathbf{I}] \\ &\approx (1-w_t)\mathbf{P}_{95} + w_t\mathbf{P}_{98} = \widehat{\mathbf{P}}_t \end{aligned}$$

Nous avons ainsi démontré la relation annoncée.

Qualité de l'approximation de $\widetilde{\mathbf{P}}_t$ donnée par $\widehat{\mathbf{P}}_t$

Bien sûr, $\widetilde{\mathbf{P}}_{95} = \widehat{\mathbf{P}}_{95} = \mathbf{P}_{95}$, et de même pour 1998.
Posons donc

$$\begin{aligned} \Delta_{96}(i,j) &= \widetilde{\mathbf{P}}_{96}(i,j) - \widehat{\mathbf{P}}_{96}(i,j) \\ \Delta_{97}(i,j) &= \widetilde{\mathbf{P}}_{97}(i,j) - \widehat{\mathbf{P}}_{97}(i,j) \end{aligned}$$

L'approximation linéaire $\widehat{\mathbf{P}}_t$ de $\widetilde{\mathbf{P}}_t$ est très bonne, puisque

$$\begin{array}{ll} -0.00027 \leq \Delta_{96}(i,j) \leq 0.00617 & \sum_{i,j} \Delta_{96}(i,j) = 0.0166 \\ -0.00022 \leq \Delta_{97}(i,j) \leq 0.00614 & \sum_{i,j} \Delta_{97}(i,j) = 0.0162 \end{array}$$

Elle a de plus l'avantage de permettre le passage aux sous-classes (voir §3) et de garantir l'additivité au niveau NOGA2, ce que ne permet pas l'estimation non-linéaire donnée par $\widetilde{\mathbf{P}}_t$.

Posons encore

$$\widetilde{\mathbf{u}}_t = \widetilde{\mathbf{P}}_t \cdot \widehat{\mathbf{b}}_t \tag{13}$$

qui représente le vecteur des EPT dans le concept entreprise calculé à l'aide de $\widetilde{\mathbf{P}}_t$.

Remarque Le calcul de la matrice de covariance de $\widehat{\mathbf{u}}_t$, $\widetilde{\mathbf{u}}_t$ ou $\widehat{u}_t^{(c)}$ découle aisément de la matrice de covariance $\mathbf{C}_{b_t}$ de $\widehat{\mathbf{b}}_t$. Cette matrice - diagonale - est calculée dans [5, §7]. On a

$$\begin{aligned} \mathbf{C}_{\widehat{u}_t} &= \widehat{\mathbf{P}}_t \cdot \mathbf{C}_{b_t} \cdot \left(\widehat{\mathbf{P}}_t\right)^t \\ \mathbf{C}_{\widetilde{u}_t} &= \widetilde{\mathbf{P}}_t \cdot \mathbf{C}_{b_t} \cdot \left(\widetilde{\mathbf{P}}_t\right)^t \\ \mathbf{C}_{\widehat{u}_t^c} &= \left(w_t \mathbf{P}_{98}^{(c)} + (1-w_t)\,\mathbf{P}_{95}^{(c)}\right) \cdot \mathbf{C}_{b_t} \cdot \left(w_t \mathbf{P}_{98}^{(c)} + (1-w_t)\,\mathbf{P}_{95}^{(c)}\right)^t \end{aligned} \tag{14}$$

6 Matrices de passage des EPT du concept établissement au concept entreprise entre 1991 et 1995

En 1991, on dispose de la matrice $\mathbf{A}_{91}$, mais pas de $\mathbf{A}_{91}^{(c)}$. L'information existante nous permet donc d'écrire :

$$\begin{aligned} \mathbf{u}_{91} &= \mathbf{A}_{91} \cdot \mathbf{1} \\ \mathbf{b}_{91} &= \mathbf{A}_{91}^t \cdot \mathbf{1} \end{aligned} \tag{15}$$

$$\mathbf{P}_{91} = \mathbf{A}_{91} \cdot \left(\mathbf{A}_{91}^t\right)^{-1} \tag{16}$$

$\mathbf{P}_{91}$ est une matrice de passage de $\mathbf{u}_{91}$ à $\mathbf{b}_{91}$:

$$\mathbf{u}_{91} = \mathbf{P}_{91} \cdot \mathbf{b}_{91} \tag{17}$$

Donc

$$\widehat{\mathbf{u}}_t = \left(w_t \mathbf{P}_{95} + (1 - w_t)\, \mathbf{P}_{91}\right) \cdot \widehat{\mathbf{b}}_t \qquad w_t = (t - 3Q91)/12 \qquad 3Q91 \leq t \leq 3Q95 \tag{18}$$

Passage à une sous-catégorie

Soit c une sous-catégorie d'une classe NOGA2, et $\mathbf{A}_{91}^{(c)}$ le vecteur-ligne inconnu des EPT correspondant, et $\mathbf{P}_{91}^{(c)}$ la ligne c de la matrice de passage inconnue.

$$\begin{aligned} \mathbf{P}_{91}^{(c)} &= \mathbf{A}_{91}^{(c)} \cdot \left(\mathbf{A}_{91}^{t}\right)^{-1} \\ \mathbf{P}_{95}^{(c)} &= \mathbf{A}_{95}^{(c)} \cdot \left(\mathbf{A}_{95}^{t}\right)^{-1} \end{aligned} \tag{19}$$

On veut donc trouver un vecteur $\widehat{\mathbf{A}}_{91}^{(c)}$, tel que

$$\sum_{c \in i} \widehat{\mathbf{A}}_{91}^{(c)}(.) = \mathbf{A}_{91}(i, .) \tag{20}$$

L'idée est d'utiliser la répartition dans les sous-classes donnée par $\mathbf{A}_{95}^{(c)}$, tout en gardant l'additivité des sous-classes en 1991. Il nous faut donc construire une matrice de passage $\mathbf{R}$ entre $\mathbf{A}_{95}$ et $\mathbf{A}_{91}$. Cette matrice est donnée par :

$$\mathbf{R} = \left(\mathbf{A}_{95}\right)^{-1} \cdot \mathbf{A}_{91} \tag{21}$$

Ainsi $\mathbf{A}_{95} \cdot \mathbf{R} = \mathbf{A}_{91}$. On obtient donc les expressions cherchées :

$$\widehat{\mathbf{A}}_{91}^{(c)} = \mathbf{A}_{95}^{(c)} \cdot \mathbf{R} \tag{22}$$

$$\begin{aligned} \widehat{\mathbf{P}}_{91}^{(c)} &= \widehat{\mathbf{A}}_{91}^{(c)} \cdot \left(\mathbf{A}_{91}^{t}\right)^{-1} = \mathbf{A}_{95}^{(c)} \cdot \mathbf{R} \cdot \left(\mathbf{A}_{91}^{t}\right)^{-1} \\ &= \mathbf{A}_{95}^{(c)} \cdot \left(\mathbf{A}_{95}\right)^{-1} \cdot \mathbf{P}_{91} \end{aligned} \tag{23}$$

$$\widehat{u}_t^{(c)} = \left(w_t \mathbf{P}_{95}^{(c)} + (1 - w_t)\, \widehat{\mathbf{P}}_{91}^{(c)}\right) \cdot \widehat{\mathbf{b}}_t \tag{24}$$

La contrainte (20) est respectée par la formule (22). On aura donc aussi

$$\sum_{c \in i} \widehat{u}_t^{(c)} = \widehat{\mathbf{u}}_t\,(i)$$

comme demandé.

Discussion

L'hypothèse qui est à la base de l'estimation (22) de $\mathbf{A}_{91}^{(c)}$ par $\widehat{\mathbf{A}}_{91}^{(c)}$ présuppose une grande stabilité dans la répartition des activités à l'intérieur d'une même classe Noga2 pendant la période 1991-1995. Cette méthode ne permet donc pas de restituer des passages soudains d'une sous-classe à l'autre. Elle n'exclut pas non plus que certains termes de $\widehat{\mathbf{A}}_{91}^{(c)}$ soient négatifs. Cela s'est effectivement produit, mais le résultat est en général de quelques unités et ne dépasse jamais le 2% du total de la classe. Il n'en découle aucune valeur négative au niveau des valeurs agrégées $\widehat{u}_t^{(c)}$.

L'année 1995 a vu le changement de nomenclature de l'ancienne NGAE à la nouvelle NOGA. L'harmonisation des données a été effectuée grâce à une clé de passage [6]. Il est rassurant de constater que les différences entre $\mathbf{A}_{91}$ et $\mathbf{A}_{95}$ sont d'amplitude analogue à celles de la période suivante : entre deux cellules correspondantes, cette amplitude varie entre -43622 et 18035, alors que pour $\mathbf{A}_{95}$ et $\mathbf{A}_{98}$, elle s'étend entre -44024 et 12850.

Dans la pratique, voir aussi l'exemple au § 7, les formules (8) et (18) sont utilisées pour tous les trimestres, et pas seulement pour le 3e. L'estimation de $\widehat{\mathbf{u}}_t$ ainsi obtenue ne tient pas compte d'une éventuelle saisonnalité des matrices de passage. A priori ces variations saisonnières devraient être faibles relativement à la saisonnalité de $\mathbf{b}_t$, qui, elle, est prise en compte.

7 Analyse des résultats sur le calcul des EPT au niveau NOGA5 x forme juridique

Posons, pour les 3e trimestres de la période 1995 à 2000,

$\sum_i \widehat{b}_t(i)$	somme des EPT de la BESTA adaptée pour le 3e trimestre de chaque année sur toutes les classes NOGA2.
$\sum_i \widehat{u}_t(i)$	somme des éléments de $\widehat{\mathbf{u}}_t$ de la formule (5) pour chaque année sur toutes les classes NOGA2.
$\sum_i \sum_{c\in i} \widehat{u}_t^{(c)}$	somme des $\widehat{u}_t^{(c)}$ de la formule (8) sur toutes les sous-classes.
$\sum_i \widetilde{u}_t(i)$	somme des éléments de $\widetilde{\mathbf{u}}_t$ de la formule (13) pour 1996-7 sur toutes les classes NOGA2.
$\Delta_{u.b}$	différence dans le total des EPT attribuable à la méthode d'interpolation linéaire de la matrice de passage.
$\Delta_{u^c.u}$	différence due à l'erreur d'arrondi $\Delta_{\mathbf{A}^c.\mathbf{A}}$.
$\Delta_{u^c.b}$	différence totale entre les EPT selon le concept entreprise au niveau sous-classe et le concept établissement.
$\Delta_{\widetilde{u}.b}$	différence dans le total des EPT attribuable à la méthode d'interpolation non linéaire de la matrice de passage.
$\Delta_{\mathbf{A}^c.\mathbf{A}}$	différence due à l'erreur d'arrondi dans la somme de tous les éléments de $\mathbf{A}_{95}$ ($\mathbf{A}^c_{95}$) et $\mathbf{A}_{98}$ ($\mathbf{A}^c_{98}$).

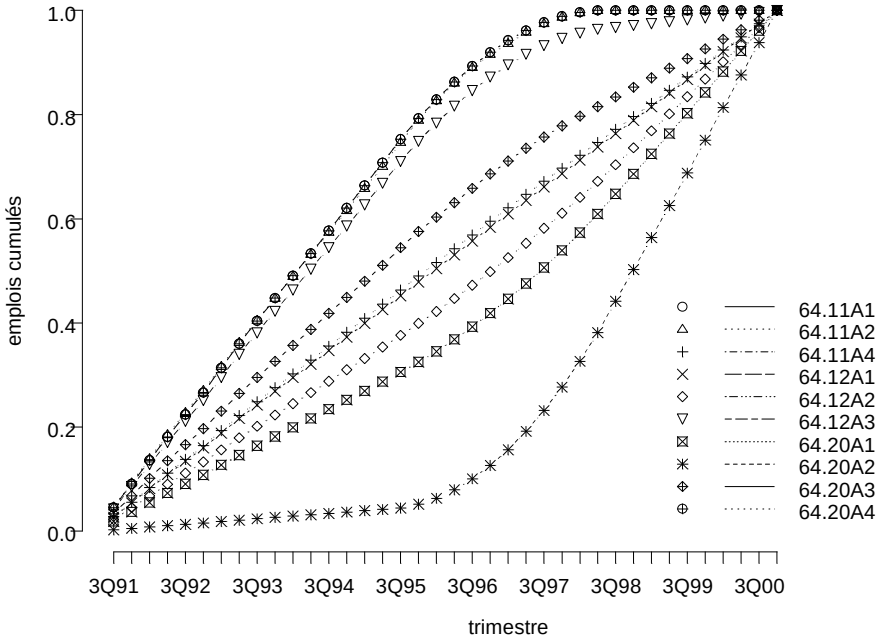
Le tableau ci-dessous présente le résumé des résultats.

Résumé des résultats

t	$\sum_i \widehat{\mathbf{b}}_t$ (a)	$\sum_i \widehat{\mathbf{u}}_t$ (b)	$\sum_i \sum_{c\in i} \widehat{u}_t^{(c)}$ (c)	$\sum_i \widetilde{\mathbf{u}}_t$ (d)	$\Delta_{u.b}$ (b)-(a)	$\Delta_{u^c.b}$ (c)-(a)	$\Delta_{\widetilde{u}.b}$ (d)-(a)	$\Delta_{u^c.u}$ (c)-(b)	$\Delta_{\mathbf{A}^c.\mathbf{A}}$
$3Q95$	3073697	3073697	3073698	-	0	1	-	1	1
$3Q96$	3015883	3014752	3014753	3015277	-1131	-1130	607	1	-
$3Q97$	2981932	2982182	2982183	2982716	251	251	-784	1	-
$3Q98$	2996867	2996867	2996868	-	0	1	-	1	0
$3Q99$	3016101	3016179	3016179	-	78	78	-	0	-
$3Q00$	3066853	3066817	3066817	-	-36	-36	-	0	-

On peut dire que les différences sont acceptables. On remarque que les signes des différences $\Delta_{u.b}$ et $\Delta_{\widetilde{u}.b}$ changent tous deux en 1996 et 1997 : en combinant l'approximation linéaire et non linéaire on pourrait trouver une matrice de passage conservant exactement le nombre total d'emplois en 1996 et 1997. Cependant, seule l'approximation linéaire (5) permet le passage aux sous-catégories.

Exemple



Résultats pour les sous-classes de la NOGA2 64

La figure ci-dessus montre les emplois intégrés sur 38 trimestres des sous-classes de la classe NOGA2 64, Postes et télécommunications[2], c'est-à-dire que la valeur au trimestre t est la somme des emplois du trimestre $3Q91$ au trimestre t. Les séries intégrées peuvent plus facilement être représentées simultanément. Pour des raisons de protection des données, on n'a représenté que les valeurs relatives cumulées, masquant de la sorte les effectifs réels. Le temps se voit sur l'axe horizontal. Une courbe convexe, comme pour la classe 6420A4, par exemple indique une diminution des emplois au cours du temps, ceux-ci ayant pratiquement disparu à partir de 1997. Une répartition linéaire, comme pour la classe 6412A1 montre la stabilité de l'emploi, enfin une pente croissante comme dans le cas de 6420A2 indique une croissance des EPT avec une forte accélération pendant l'année 1996. Il ne faut pas oublier que les ordres de grandeurs sont très différents d'une classe à l'autre : l'emploi est très largement dominant dans la sous-classe 6411A4.

On observe très bien la diminution des activités de la Poste, et le transfert des télécommunications du groupe de formes juridiques du Public (64.20A4) vers des sociétés en capitaux (64.20A2).

[2]64 Poste et télécommunications (NOGA 2002)
64.11A Activités de la Poste
64.12A Activités de courrier autres que celles de la Poste
64.20A Télécommunications sans transmission de programmes de radio et de télévision
64.20B Transmission de programmes de radio et de télévision (n'existe pas encore en 98)
64.20C Fourniture d'accès à Internet (n'existe pas encore en 98).
Le dernier chiffre indique le groupe de formes juridiques, voir Annexe.

Annexes

Nomenclature des groupes de formes juridiques (RFG)

Nomenclature des groupes de formes juridiques (RFG)

RFG	RF Dénomination
RFG 1 :	1 Raison individuelle
	2 Société simple
	3 Société en nom collectif
	4 Société en commandite
RFG 2 :	6 Société anonyme et société en commandite par action
	7 Société à responsabilité limitée
	8 Société coopérative
	11 Filiale de société anonyme étrangère
	12 Filiale de société à responsabilité limitée étrangère ou de coopérative étrangère
	34 Autres formes juridiques
RFG 3 :	9 Association
	10 Fondation
RFG 4 :	20 Administration de la Confédération
	21 Administration cantonale
	22 Administration de district
	23 Administration communale
	30 Etablissement de la Confédération
	31 Etablissement cantonal
	32 Etablissement de district
	33 Etablissement communal

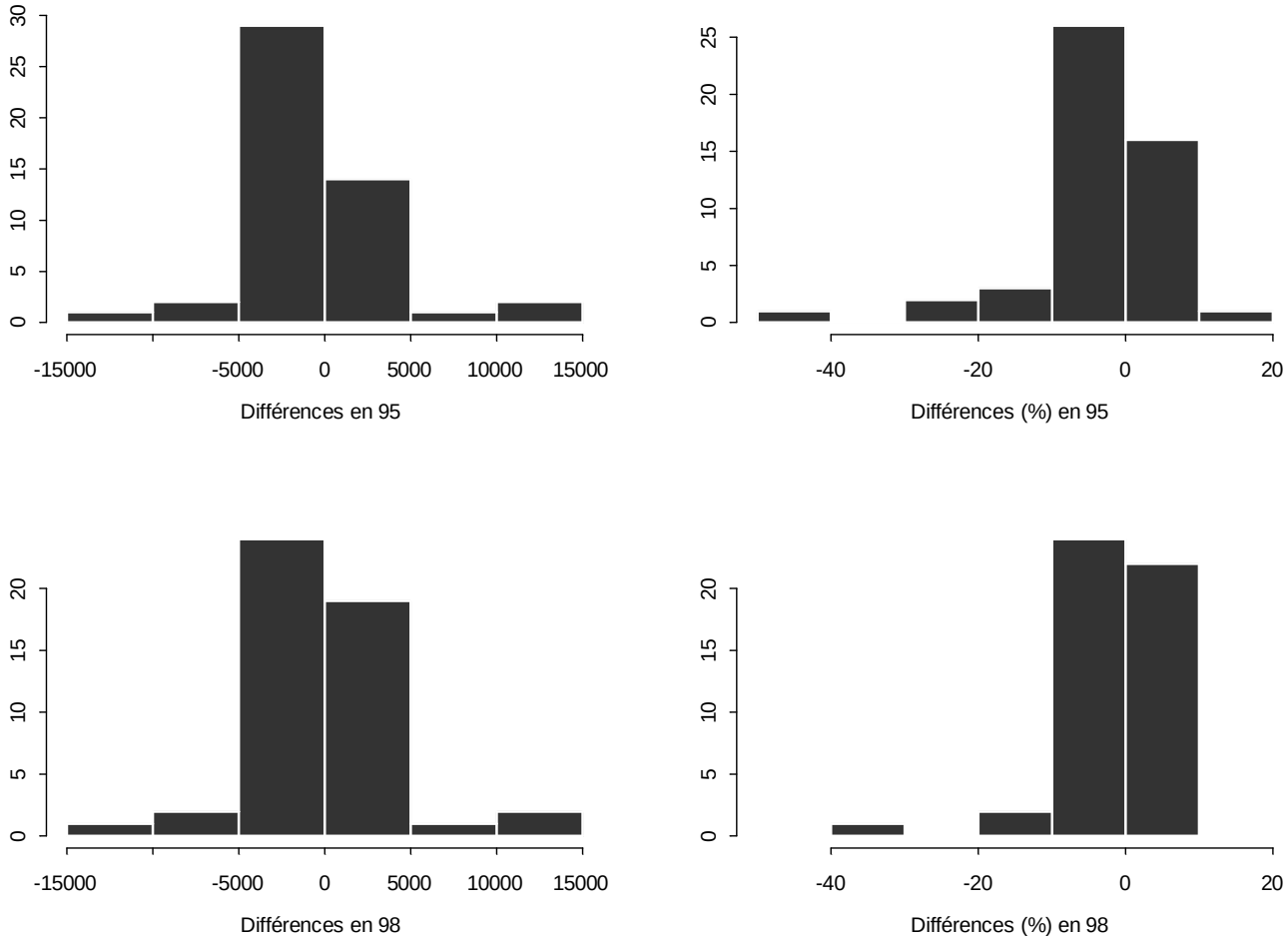
Différences observées en 1995 et 1998 entre les totaux établissements et entreprises

NOGA2	d95	r95	d98	r98	NOGA2	d95	r95	d98	r98
N10-14	-306.0	-5.20	-133.5	-2.68	N50	-1193.2	-1.57	-670.4	-0.89
N15	1331.4	2.29	247.6	0.44	N51	-1368.2	-0.80	-2281.9	-1.32
N16	-54.7	-1.77	52.7	2.06	N52	11514.2	4.18	11001.7	4.25
N17	363.1	1.88	81.2	0.51	N55	-6226.1	-3.19	-6673.6	-3.56
N18	0.5	0.00	-11.8	-0.15	N60	1404.3	1.63	2781.5	3.38
N19	-27.3	-0.67	-12.5	-0.41	N61	86.5	3.43	-89.6	-4.24
N20	-100.4	-0.24	-505.7	-1.39	N62	1194.9	7.24	887.6	9.42
N21	263.5	1.70	32.9	0.22	N63	-5967.7	-14.52	-5118.2	-11.23
N22	-632.2	-1.15	-44.1	-0.09	N64	2116.6	3.13	16.3	0.02
N23	-252.7	-24.76	-228.4	-36.57	N65	-1341.4	-1.16	-278.6	-0.26
N24	1457.4	2.26	512.6	0.87	N66	4020.5	8.04	-144.5	-0.26
N25	-336.6	-1.49	269.5	1.15	N67	-4092.0	-40.33	19.0	0.24
N26	-23.9	-0.11	-2.8	-0.01	N70	-525.3	-3.68	-193.9	-1.26
N27	-303.2	-1.85	-531.8	-3.32	N71	-87.8	-2.81	9.9	0.34
N28	-20.0	-0.02	230.2	0.29	N72	-930.0	-3.16	-847.4	-2.19
N29	2586.8	2.37	2166.7	2.05	N73	-1731.7	-19.61	-1060.1	-10.55
N30	565.2	16.32	77.2	2.52	N74	-3749.0	-1.79	-1459.2	-0.68
N31	2125.3	4.69	2117.2	6.00	N75	-10579.9	-8.61	-11609.0	-9.33
N32	-133.4	-0.79	-62.2	-0.32	N80	11811.4	8.12	10171.1	6.71
N33	-284.4	-0.46	547.1	0.82	N85	8346.2	3.13	8472.0	3.03
N34	-15.3	-0.34	9.3	0.21	N90	-1062.9	-10.06	374.5	3.42
N35	-2374.6	-28.03	-1285.8	-9.06	N91	-509.6	-1.86	-561.1	-2.06
N36-37	173.8	0.55	11.9	0.04	N92	-1623.4	-4.77	-1725.4	-4.55
N40-41	-1010.5	-4.15	-757.7	-3.22	N93	-949.3	-2.58	-948.0	-2.63
N45	-1548.9	-0.47	-2852.5	-1.01					

$d_{95} = u_{95} - b_{95}$ $\quad r_{95} = 100(u_{95} - b_{95})/b_{95}\ \%$ $\quad d_{98} = u_{98} - b_{98}$ $\quad r_{98} = 100(u_{98} - b_{98})/b_{98}\ \%$

Le tableau ci-dessus présente les différences (d_{95} et d_{98}) et les différences relatives (r_{95} et r_{98}) entre les totaux selon le concept établissement et le concept entreprise. On observe des variations absolues entre les concepts supérieures à 10'000 EPT pour les 2 années dans les classes 52, 75 et 80. La plus grande variation relative se situe dans des classes à petits effectifs. Les histogrammes suivants montrent que les différences entre les totaux selon les 2 concepts ont diminué en 1998: les entreprises ont tendance à se concentrer sur leur activité principale.

A noter la classe 67 (services auxiliaires des activités financières et des assurances) dont l'effectif en 95 est de 40% inférieur dans le concept entreprise relativement au concept établissement, alors qu'il s'équilibre en 98.



Histogrammes des différences du tableau précédent

Une étude des matrices $\mathbf{A}_{95}$ et $\mathbf{A}_{98}$ (voir tableau ci-dessous) éclaire l'évolution des emplois de la classe 67. On voit qu'en 1995, ces services auxiliaires se situaient presque pour moitié dans des assurances (43% dans la classe 66) et à 56% dans leur propre classe (colonne N67 de $\mathbf{A}_{95}$), et qu'ils sont devenus à 98% des entreprises indépendantes en 1998 (colonne N67 de $\mathbf{A}_{98}$). Cela explique l'évolution des marges: la classe 67 perd 2000 emplois environ dans le concept établissement (comparer $\mathbf{b}_{95}$ et $\mathbf{b}_{98}$) et en gagne autant dans le concept entreprise (comparer $\mathbf{u}_{95}$ et $\mathbf{u}_{98}$).

Cases des matrices $\mathbf{A}_{95}$ et $\mathbf{A}_{98}$ et marges correspondant aux classes 66 et 67

	A95		u95		A98		u98
NOGA2	N66	N67		NOGA2	N66	N67	
N66	49226.8	4300.1	54024.6	N66	55556.4	90.7	55749.1
N67	383.0	5660.1	6053.1	N67	0.0	7849.5	8062.0
b95	50004.1	10145.1		b98	55893.6	8043.0	

Programmes S-Plus

```
##############
# passage1.ssc
##############

# Création des piliers et des matrices de transformation
# pour le passage du concept établissement au niveau NOGA2
# au concept entreprise au niveau des sous-classes.

# Auteur: Monique Graf, 18.07.01

# fichiers input :

# pour 1998:
import.data(DataFrame="A98", FileName="H:\\Daten\\Besta\\passage\\A98.xls",
  FileType = "EXCEL")     # EPT en 1998

import.data(DataFrame="IN1", FileName="H:\\Daten\\Besta\\passage\\A98c.xls",
  FileType = "EXCEL")     # fichier contenant l'identificateur de la sous-classe c en 1ère colonne
                          # et les EPT A98c dans les colonnes suivantes
# idem 1995:
import.data(DataFrame="A95", FileName="H:\\Daten\\Besta\\passage\\A95.xls",
  FileType = "EXCEL")     # EPT en 1995

import.data(DataFrame="IN2", FileName="H:\\Daten\\Besta\\passage\\A95c.xls",
  FileType = "EXCEL")     # fichier contenant l'identificateur de la sous-classe c en 1ère colonne
                          # et les EPT A95c dans les colonnes suivantes
#_____________________________________________________________________________

id98 <- IN1[,1]            # identificateurs des NOGA5 x forme juridique présentes en 98
d1 <- dim(IN1)[1]
d2 <- dim(IN1)[2]
d2m1 <- d2-1
# A98c

id95 <- IN2[,1]             # identificateurs des NOGA5 x forme juridique présentes en 95
d3 <- dim(IN2)[1]
d4 <- dim(IN2)[2]
if (d2 != d4) print("erreur! pas le meme nombre de classes NOGA2 en colonnes")
id98 <- as.character(id98)
id95 <- as.character(id95)
range(diff(order(id95)))               # doit valoir 1 1
range(diff(order(id98)))               # doit valoir 1 1

id <- union(id98,id95)                 # ensemble des sous-classes présentes en 95 ou en 98
ncla <-length(id)
orid <-order(id)

ind95 <-pmatch(id,id95)[orid]
ind98 <-pmatch(id,id98)[orid]
# ind95;ind98
B98c <- IN1[ind98 , ]                 # insère des valeurs manquantes dans A98c pour
                                       # les sous-classes non représentées en 1998
B95c <- IN2[ind95 , ]                 # idem 95
#_____________________________________________________
# recoder interactivement les NA en 0 dans B95c et B98c
#_____________________________________________________


A95c <- as.matrix(B95c[,2:d2])
A98c <- as.matrix(B98c[,2:d2])
```

```
u98c <- A98c %*% rep(1,d2m1)                                    # EPT par sous-classes en 1998
u95c <- A95c %*% rep(1,d2m1)                                    # EPT par sous-classes en 1995
rm(IN1)
rm(IN2)

# Stockage des piliers
Ut <- data.frame(id[orid], u95c, u98c)
names(Ut) <- c("classe", "u95c", "u98c")

# Passage
A98 <- as.matrix(A98)
A95 <- as.matrix(A95)

valp98 <- svd(A98)
invA98t <- valp98$u%*%diag(1/valp98$d)%*%t(valp98$v)            # inverse de la transposée de A98
valp95 <- svd(A95)
invA95t <- valp95$u%*%diag(1/valp95$d)%*%t(valp95$v)            # inverse de la transposée de A95

calculut <-function(wt, bt){
       # fonction calculant les EPT dans les sous-classes selon le concept emtreprise
       # à partir de la BESTA adaptée au trimestre t: vecteur bt
       # wt=(t-3Q95)/12

       # Input : les paramètres d'appel +
       #       A98c; A95c; invA98t; invA95t
       #       id ; id98; id95
       #       d1 ; d3

       # Output : ut, vecteur des EPT


       v98t <- A98c %*% invA98t %*% bt
       v95t <- A95c %*% invA95t %*% bt

       ut <- wt*v98t + (1-wt)*v95t
       return(ut)
}

# fin de passage1.ssc
```

```
##############
# passage2.ssc
##############

# Calcul des EPT ventilés dans les sous-classes
# dans le concept entreprise pour les années entre les RE de 95 et 98.

# Auteur: Monique Graf, 13.08.01

# Input: les matrices calculées dans passage1.ssc
#        invA95t, invA98t, A95c, A98c;
#        le data frame Ut contenant les EPT en 95 et 98 pour les sous-classes
#        selon le concept entreprises (données du RE), calculé dans passage1.ssc
#        Ce data frame sera complété des EPT pour les années intermédiaires.

import.data(DataFrame="Btall", FileName="H:\\Daten\\Besta\\passage\\b9500.xls",
  FileType = "EXCEL")                                  # BESTA adaptée aux trimestres 3Q95 à 4Q00
Bt <- Btall[,c(1,5,9,13,17,21)]                        # BESTA adaptée, extrait des 3e trimestres
Bt[,1]
# année 1996
wt96 <- 4/12                                           # valeur de (3Q96-3Q95)/12
bt96 <- Bt[,2]
u96c <- calculut(wt96, bt96)
Ut <- data.frame(Ut,u96c)                              # années intermédiaires

# année 1997
wt97 <- 8/12                                           # valeur de (3Q97-3Q95)/12
bt97 <- Bt[,3]
u97c <- calculut(wt97, bt97)
Ut <- data.frame(Ut,u97c)                              # années intermédiaires

# année 1999
wt99 <-1
bt99 <- Bt[,5]
u99c <- calculut(wt99, bt99)
Ut <- data.frame(Ut,u99c)                              # années ultérieures au 2e pilier

# année 2000
wt00 <-1
bt00 <- Bt[,6]
u00c <- calculut(wt00, bt00)
Ut <- data.frame(Ut,u00c)                              # années ultérieures au 2e pilier

# Stockage des résultats
Ut <- data.frame(Ut[,c(1:2,4:5,3,6:7)])                # années remises dans l'ordre

#
# Création du fichier EXCEL
export.data(DataSet = "Ut", FileName = "H:\\Daten\\Besta\\passage\\U9500.xls", FileType = "EXCEL")


# fin de passage2.ssc
```

```
##############
# Passage3.ssc
##############
# Vérifications des résultats de passage1 et passage2

# vérification de la fusion des matrices par sous-classes
N5.fj <- data.frame(id95[ind95],id98[ind98],id[orid])
N5.fj


P98 <- A98 %*% invA98t               # définition des matrices de passage
P95 <- A95 %*% invA95t
prod(svd(P98)$d)                     # déterminant de P98 (vaut bien 1)
sum(P95)                             #|
sum(P98)                             #| ces sommes valent environ 49
sum(diag(P95))                       #| (nb. de cl. Noga2)
sum(diag(P98))                       #|

range(t(A98)%*%rep(1,49)-Bt[,4])     #|
range(A98c%*%rep(1,49)-Ut[,5])       #| les sommes des lignes et des colonnes
range(t(A95)%*%rep(1,49)-Bt[,1])     #| reproduisent bien Bt et Ut en 95 et 98
range(A95c%*%rep(1,49)-Ut[,2])       #|

sum(A98c)-sum(A98)                   #| les sommes des EPT des sous-classes
sum(A95c)-sum(A95)                   #| = les sommes des EPT des classes

# H95 est presque nulle, c'est-à-dire P98 est presque l'identité:
H95 <- P95 - diag(rep(1,49))
range(H95)
sum(diag(H95))
p95.s <- sort(P95-diag(diag(P95)))
l1 <- length(p95.s[p95.s<0]) + 1
l2 <- l1+48
p95.s[l1:l2]
length(p95.s)-length(p95.s[-l1:-l2])
sort(p95.s[-l1:-l2])
hist(p95.s[-l1:-l2],breaks=c(-0.2,seq(-0.05,0.05,0.005),seq(0.1 , 0.7 , 0.1)))
summary(diag(P95))                  # distribution des poids sur la diagonale
summary(p95.s[-l1:-l2])             # distribution des poids hors diagonale
quantile(p95.s[-l1:-l2], c(0.005,0.05,0.95,0.995))
range(P95)

# H98 est presque nulle, c'est-à-dire P98 est presque l'identité:
H98 <- P98 - diag(rep(1,49))
range(H98)
sum(diag(H98))
p98.s <- sort(P98-diag(diag(P98)))
l1 <- length(p98.s[p98.s<0]) + 1
l2 <- l1+48
p98.s[l1:l2]
length(p98.s)-length(p98.s[-l1:-l2])
sort(p98.s[-l1:-l2])
hist(p98.s[-l1:-l2],breaks=c(-0.2,seq(-0.05,0.05,0.005),seq(0.1 , 0.7 , 0.1)))
summary(diag(P98))                  # distribution des poids sur la diagonale
summary(p98.s[-l1:-l2])             # distribution des poids hors diagonale
quantile(p98.s[-l1:-l2], c(0.005,0.05,0.95,0.995))

wt <-c(0,4/12,8/12,1,1,1)
sumu <-rep(NA,6)
for (i in 1:6){
   w <- wt[i]
   sumu[i] <- sum((w*P98+(1-w)*P95)%*%Bt[,i])   # somme des EPT ds le concept entrepr.
}

# tableau résumé des résultats
```

```
preci <- data.frame(
c(sum(Bt[,1]),sum(Bt[,2]),sum(Bt[,3]),sum(Bt[,4]),sum(Bt[,5]),sum(Bt[,6])),sumu,
c(sum(Ut[,2]),sum(Ut[,3]),sum(Ut[,4]),sum(Ut[,5]),sum(Ut[,6]),sum(Ut[,7])) )
names(preci) <-c("sum.b","sum.u","sum.uc")
row.names(preci) <- c(1995:2000)

preci$difu.b <- preci$sum.u-preci$sum.b
preci$difuc.b <- preci$sum.uc-preci$sum.b
preci$difuc.u <- preci$sum.uc-preci$sum.u
preci$difA <-c(sum(A95c)-sum(A95),NA,NA,sum(A98c)-sum(A98),NA,NA)
preci <- preci[,-8]

round(preci)                                    # impression

# définition des noms des 49 classes considérées
attr(Noms98)
Noga2 <- as.vector(Noms98)
Noga2 <- as.vector(Noga2$V1[c(3:25,28:53)])
Noga2
Noga2[c(1,23,24)] <-c("N10-14","N36-37","N40-41")

# PP98 est P98 avec les noms des classes Noga2
PP98 <- data.frame(P98)
names(PP98) <- Noga2
row.names(PP98) <- Noga2
round(PP98,3)
rm(PP98)

# taux d'occupation
Acti <- data.frame(diag(A98)/(A98%*%rep(1,49)))
row.names(Acti) <- Noga2
names(Acti) <- "Rapport"
order(Acti$Rapport)
s.Acti <- data.frame(sort(Acti$Rapport))
row.names(s.Acti) <- Noga2[order(Acti$Rapport)]
names(s.Acti) <- "Rapport"
Acti;s.Acti

# fin de passage3
```

```
##############
# Passage4.ssc
##############

# Analyse de la décomposition en valeurs singulières de A95 et A98.

# Les résultats ne sont pas montrés dans le rapport, cette décomposition étant
# essentiellement effectuée pour inverser A95 et A98.
# L'analyse montre qu'il n'y a aucun problème numérique.


##############
# passage5.ssc
##############

# Comparaison des approximations linéaires et non linéaires des matrices de passage.

ln.P <- function(P){
   # calcul du logarithme d'une matrice à val.pr. complexes
   # à val. propres distinctes

   va <- eigen(P)
   Q <- va$vectors
   val.P <- va$values

   # P a 1 val.propre réelle et 48 val. pr. complexes

   # calcul de l'inverse de Q
   val.Q <- svd(Q)
   Qm1 <- val.Q$v %*% diag(1/val.Q$d) %*% Conj(t(val.Q$u)) # inverse de Q

   # calcul du logarithme de P
   ln.P <- Q%*%diag(log(va$values))%*%Qm1
   # ln.P est bien réelle:
   del <-max(Mod(ln.P-Re(ln.P)))
   print(c("max(Mod(ln.P-Re(ln.P))) =",del))
   if (del < 1.0e-010) ln.P <- Re(ln.P)


   # Qualité de l'approximation linéaire de ln.P par P-I:
   print(range(ln.P-P+diag(rep(1,49)))
   return(ln.P)
}

ln.P98 <- ln.P(P98)
ln.P95 <- ln.P(P95)

wt <- c(0,4/12,8/12,1)
for (i in 1:4){
   w <- wt[i]
   print(summary((1-w)*(ln.P95-P95) + w*(ln.P98-P98) + diag(rep(1,49))) )
}
exp.P <- function(P){
   # calcul de l'exponentielle d'une matrice à val.pr. complexes
   # à val. propres distinctes

   va <- eigen(P)
   Q <- va$vectors
   val.P <- va$values

   # P a 1 val.propre réelle et 48 val. pr. complexes

   # calcul de l'inverse de Q
   val.Q <- svd(Q)
   Qm1 <- val.Q$v %*% diag(1/val.Q$d) %*% Conj(t(val.Q$u)) # inverse de Q
```

```
    # calcul de l'exponentielle de P
    exp.P <- Q%*%diag(exp(va$values))%*%Qm1
    # exp.P est bien réelle:
    del <- max(Mod(exp.P-Re(exp.P)))
    print(c("max(Mod(exp.P-Re(exp.P))) =",del))
    if (del < 1.0e-010) exp.P <- Re(exp.P)

    print(range(exp.P) )

    return(exp.P)
}

    w <- wt[2]
    P <- (1-w)*ln.P95 + w*ln.P98
    exp.P96 <- exp.P(P)
    range(exp.P96-((1-w)*P95+w*P98))
    sum(exp.P96-((1-w)*P95 + w*P98))

    w <- wt[3]
    P <- (1-w)*ln.P95 + w*ln.P98
    exp.P97 <- exp.P(P)
    range(exp.P97-((1-w)*P95 + w*P98))
    sum(exp.P97-((1-w)*P95 + w*P98))
    rm(P)
# L'approximation est très bonne.

# Complément à preci (voir passage3.ssc)
preci2 <- data.frame(c(sum(exp.P96%*%Bt[,2]),sum(exp.P97%*%Bt[,3])),c(preci[2,1],preci[3,1]),
                     c(preci[2,1] - sum(exp.P96%*%Bt[,2]),preci[3,1]  -sum(exp.P97%*%Bt[,3])))
names(preci2) <- c("sum.utild","sum.b","difb.utild")
row.names(preci2) <- c("3Q96","3Q97")
round(preci2)

# Rayon spectral
Mod(eigen(P95-diag(rep(1,49)))$values[1])
Mod(eigen(P98-diag(rep(1,49)))$values[1])


# Autres analyses
# dessins des valeurs propres de P98
plot(0,0,xlim=c(0,1), ylim=c(-0.2,0.2), ylab="", xlab="",sub="Valeurs propres de P98 dans le plan
   complexe")
arrows(rep(0,49),rep(0,49),Re(eigen(P98)$values), Im(eigen(P98)$values),size=.1, open=F)
lines(cos(seq(-pi/20,pi/20,pi/80)),sin(seq(-pi/20,pi/20,pi/80)),lty=1)

par(mfrow=c(1,2))
plot(Re(eigen(P98)$values), Im(eigen(P98)$values),xlim=c(0.8,1.2), ylim=c(-0.2,0.2), ylab="",
   xlab="",sub="Valeurs propres de P98 dans le plan complexe")
lines(cos(seq(-pi/20,pi/20,pi/80)),sin(seq(-pi/20,pi/20,pi/80)),lty=1)
plot(0,0,xlim=c(0,1), ylim=c(-0.2,0.2),plot=F, ylab="", xlab="",sub="Valeurs propres de P98 dans le
   plan complexe")
arrows(rep(0,49),rep(0,49),Re(eigen(P98)$values), Im(eigen(P98)$values),size=.1, open=F)
lines(cos(seq(-pi/20,pi/20,pi/80)),sin(seq(-pi/20,pi/20,pi/80)),lty=1)

# fin de passage5.ssc
```

```
##############
# Passage6.ssc
##############

# Illustration pour les sous-classes de la Noga2 64
# création de la fig. 1.

# Input: fichier Excel des EPT dans le concept entreprises "EPT 1991-2000secteurs II et III.xls"
#        transmis par E. Capezzali (EWL).

# Les lignes correspondant à la classe Noga 64 sont collées dans le data editor
# (fichier DS3)
D64<-DS3
dim(D64)

# Définition des entêtes de colonnes (les trimestres)
lab9 <- paste(c(3,4,rep(1:4,9)),"Q", sep="")
lab9 <- paste(lab9,c(rep(91,2),rep(92,4),rep(93,4),rep(94,4),rep(95,4),rep(96,4),
             rep(97,4),rep(98,4),rep(99,4),rep("00",4)),sep="")

names(D64) <-c("sclasse",lab9)
row.names(D64) <- D64$sclasse

# transposition et suppression de 2 classes vides
tD64 <-t(D64[c(1:2,4:7,9:12),2:39])

# Calcul des emplois cumulés et graphique
# 1ère courbe
ctD64 <- data.frame(cumsum(tD64[,1]/sum(tD64[,1])))
plot(1:38, ctD64, axes=F,type="b",xlim=c(1,45),ylim=c(0,1), xlab="trimestre", ylab="emplois
cumulés")

axis(1, at=1:38, labels=names(D64)[2:39])
axis(2, at=c(0.0,0.2,0.4,0.6,0.8,1.0),labels=T)
legend(30,0.47,legend=row.names(D64)[c(1:2,4:7,9:12)], lty=1:10, marks=1:10,bty="n")

# Courbes suivantes
for (i in 2:10){
       ctD64 <- data.frame(cumsum(tD64[,i]/sum(tD64[,i])))
       lines(1:38,ctD64,lty=i,type="b",pch=i)
}

# fin de passage6.ssc
```

Références

[1] La Statistique de l'emploi : Bases Méthodologiques. *OFS* 1997, 31pp

[2] Revision der Beschäftigungstatistik Grobkonzept : 1999/2000. *OFS* 1999, 56pp

[3] Révision de la Statistique de l'emploi, Conception des détails. *OFS* 2000, 53 pp

[4] Graf, M. (2001). Désaisonnalisation : Aspects méthodologiques et application à la statistique de l'emploi.*Rapport de Méthodes 338-0003, OFS*

[5] Salamin, P.A. (1997). Evaluation de la Statistique de l'emploi.

[6] Salamin, P.A. (1998). Etablissement d'une clef de passage pondérée entre l'ancienne (NGAE 85) et la nouvelle nomenclature (NOGA 95) générale des activités économiques.

Methodenberichte des Dienstes Statistische Methoden des BFS
Rapports de méthodes du Service de méthodes statistiques de l'OFS
Methodological reports of the Statistical Methods Unit of SFSO

Hulliger, B., Peters, R. (1994). Enquête sur la structure des salaires en Suisse: stratégie d'échantillonnage pour le secteur privé.

Peters, R., Hulliger, B. (1994). La technique de pondération des données: application à l'enquête suisse sur la santé.

Schwendener, P. (1995). Verbrauchserhebung 1990 - Vertrauensintervalle.

Hulliger, B. (1995). Konjunkturelle Mietpreiserhebung: Stichprobenplan und Schätzverfahren.

Peters, R. (1995). Enquête de structure sur les loyers: établissement d'un plan d'échantillonnage stratifié.

Salamin, P.A. (1995). Estimation des flux pour le module II des comptes globaux du marché de travail.

Hulliger, B. (1996). Gütertransportstatistik 1993: Schätzverfahren mit Kompensation der Antwortausfälle.

Hulliger, B., Peters, R. (1996). Enquête sur le comportement de la population suisse en matière de transport en 1994: plan d'échantillonnage et pondération.

Peters, R. (1996). Schéma de pondération des indices PAUL.

Peters, R., Hulliger, B. (1996). Schätzverfahren für die Lohnstruktur-Erhebung 1994 / Procédure d'estimation pour l'enquête de 1994 sur la structure des salaires.

Hulliger, B. (1996). Haushalterhebung Familie 1994: Stichprobenplan, Stichprobenziehung und Reservestichproben.

Comment, T., Hulliger, B., Ries, A. (1996). Gewichtungsverfahren für die Schweizerische Arbeitskräfteerhebung (1991-1995).

Peters, R. (1996). Pondération des données de l'enquête sur la famille en Suisse.

Peters, R. (1997). Enquête 1996 sur la structure des salaires en Suisse: établissement du plan d'échantillonnage.

Peters, R. (1997). Etablissement du plan d'échantillonnage pour l'enquête 1996 sur la recherche et le développement dans l'économie privée en Suisse.

Salamin, P.A. (1997). Evaluation de la Statistique de l'emploi.

Bender, A., Hulliger, B. (1997). Enquête suisse sur la population active: rapport de pondération pour 1996.

Peters, R. (1998). Extrapolation des données de l'enquête de structure sur les loyers.

Salamin, P.A. (1998). Etablissement d'une clef de passage pondérée entre l'ancienne (NGAE 85) et la nouvelle nomenclature (NOGA 95) générale des activités économiques.

Hulliger, B., Kassab, M. (1998). Evaluation of Estimation Methods for the Survey on Environment Protection Expenditures of Swiss Communes.

de Rossi, F.-X. (1998). Méthodes statistiques pour le compte routier suisse.

Hulliger, B., Eichenberger, P. (2000). Stichprobenregister für Haushalterhebungen: Umstellung auf Telefonnummern ohne Namen und Adressen, Abläufe für Erstellung und Stichprobenziehung. Bestellnummer: 338-0000

Renaud, A. (2001). Statistique suisse des bénéficiaires de l'aide sociale. Plan d'échantillonnage des communes. Numéro de commande: 338-0001

Hüsler, J., Müller, S. (2001). Schlussbericht Betriebszählung 1995 (BZ 95), Mehrfach imputierte Umsatzzahlen. Bestellnummer: 338-0002

Graf, M. (2001). Désaisonnalisation. Aspects méthodologiques et application à la statistique de l'emploi. Numéro de commande: 338-0003

Potterat, J., Hulliger, B. (2001). Schätzung der Sägereiproduktion mit der Sägerei-Erhebung PAUL. Bestellnummer: 338-0004

Peters, R., Renfer, J.-P. et Hulliger, B., (2001). Statistique de la valeur ajoutée 1997-1998. Procédure d'extrapolation des données. Numéro de commande: 338-0005

Salamin, P.A. (2001). La technique de la double enquête pour la statistique du transport routier de marchandise. Numéro de commande: 338-0006

Graf, M. (2002). Passage du concept établissement au concept entreprise. Numéro de commande: 338-0007

Publikationsprogramm BFS

Das Bundesamt für Statistik (BFS) hat – als zentrale Statistikstelle des Bundes – die Aufgabe, statistische Informationen breiten Benutzerkreisen zur Verfügung zu stellen.

Die Verbreitung der statistischen Information geschieht gegliedert nach Fachbereichen (vgl. Umschlagseite 2) und mit verschiedenen Mitteln:

Programme des publications de l'OFS

En sa qualité de service central de statistique de la Confédération, l'Office fédéral de la statistique (OFS) a pour tâche de rendre les informations statistiques accessibles à un large public.

L'information statistique est diffusée par domaine (cf. verso de la première page de couverture); elle emprunte diverses voies:

Diffusionsmittel	*Kontakt* *N° à composer*	*Moyen de diffusion*
Individuelle Auskünfte	032 713 60 11	Service de renseignements individuels
Das BFS im Internet	http://www.statistik.admin.ch	L'OFS sur Internet
Medienmitteilungen zur raschen Information der Öffentlichkeit über die neusten Ergebnisse	032 713 60 11	Communiqués de presse: information rapide concernant les résultats les plus récents
Publikationen zur vertieften Information (zum Teil auch als Diskette)	032 713 60 60	Publications: information approfondie (certaines sont disponibles sur disquette)
Online-Datenbank	032 713 60 86	Banque de données (accessible en ligne)

Nähere Angaben zu den verschiedenen Diffusionsmitteln liefert das alle 2 Jahre nachgeführte **Publikationsverzeichnis.** Es ist auf dem Internet abzurufen oder kann gratis über Tel. 032 713 60 60 bezogen werden.

La **Liste des publications**, qui est mise à jour tous les deux ans, donne davantage de détails sur les divers moyens de diffusion. Elle est proposée sur Internet et peut être obtenue gratuitement en téléphonant au n° 032 713 60 60.

Methodenberichte des Dienstes Statistische Methoden
Rapports de méthodes du Service de méthodes statistiques
Methodological reports of the Statistical Methods Unit

Die Methodenberichte beschreiben die mathematischen und statistischen Methoden, die den Resultaten und Analysen der öffentlichen Statistik zu Grunde liegen. Sie enthalten ausserdem die Evaluation und Entwicklung von neuen Methoden im Hinblick auf eine zukünftige Anwendung. Diese Publikationen sollen einerseits die verwendeten Methoden dokumentieren, um Transparenz und Wissenschaftlichkeit sicher zu stellen, und sie sollen andererseits die Zusammenarbeit mit Hochschule und Wissenschaft fördern.

Zur Illustration der beschriebenen mathematischen Konzepte, werden im Bericht numerische Resultate aufgeführt. Diese sind allerdings nicht als offizielle Resultate der betreffenden Erhebungen zu verstehen. Ebenfalls können die tatsächlich angewendeten Methoden leicht von den hier beschriebenen abweichen.

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.

La statistique de l'emploi (BESTA) est basée sur un échantillon d'établissements. Elle est utilisée pour calibrer d'autres statistiques fédérales, comme la statistique de la valeur ajoutée qui, elle, est établie à partir d'un échantillon d'entreprises. De plus, le cadre d'échantillonnage n'est pas exactement le même dans les deux cas.

Le présent rapport décrit une méthode d'harmonisation ou de passage entre le concept établissement et le concept entreprise. Le recensement des entreprises fournit une matrice donnant le nombre d'emplois par combinaison des classes d'activité des entreprises et des établissements. C'est à partir de cette matrice qu'une matrice de passage d'un concept à l'autre est déduite, et que l'univers de sondage est adapté. La statistique de l'emploi fournit l'évolution conjoncturelle.

N° de commande:
338-0007

Commandes:
032 713 60 60 Fax: 032 713 60 61

Prix:
Gratuit

ISBN 3-303-00232-0