

# ETUDE DE LA PRECISION DES ESTIMATEURS DE L'ENQUETE CONTINUE SUISSE SUR LA POPULATION ACTIVE SELON LE CHOIX DU SCHEMA DE ROTATION DES ECHANTILLONS

Jean-Pierre RENFER (\*)

(\*) Office fédéral de la statistique, Service des méthodes statistiques.

## Introduction

L'enquête suisse sur la population active (ESPA) est actuellement une enquête annuelle réalisée chaque 2ème trimestre. Il est prévu de la réviser dans le cadre du projet "Révision ESPA : enquête continue aux normes EUROSTAT" de manière à pouvoir produire des estimateurs trimestriels en plus des estimateurs annuels. L'entrée en vigueur des bilatérales II entre la Suisse et l'UE imposera à l'ESPA de respecter notamment les critères de précision selon les normes EUROSTAT.

Le service des méthodes statistiques (METH) de l'Office fédéral de la statistique (OFS) a été consulté pour mener à bien cette étude préliminaire sur l'estimation de la variance des estimateurs considérés dans l'ESPA révisée afin de satisfaire les critères de précision imposés par EUROSTAT.

Cette étude préliminaire permet d'estimer globalement la taille des échantillons pour satisfaire les critères de précision EUROSTAT. Elle permet également de montrer au travers de quatre schémas de rotation envisagés combien le choix de la variante retenue est important pour les différents estimateurs étudiés. En effet, pour une taille d'échantillon fixée, selon le schéma de rotation choisi, on privilégie tel estimateur plutôt que tel autre du point de vue de la précision. Comme il s'avère que le critère le plus contraignant est celui relatif à la précision de la variation de deux trimestres successifs, une variante avec un chevauchement important entre deux trimestres successifs est certainement à considérer. Cependant, il faut reconnaître que cela se fait au détriment d'autres estimateurs comme l'estimation annuelle par exemple. Un bon compromis peut être réalisé en considérant les variantes qui permettent d'obtenir des résultats satisfaisants pour les différents estimateurs tout en garantissant que les critères EUROSTAT soient respectés pour des tailles d'échantillon raisonnables.

Ce document a pour objectif de déterminer l'ordre de grandeur de la taille des échantillons pour satisfaire les critères de précision. Ainsi le paragraphe 2 est consacré aux critères de précision selon les normes EUROSTAT et au calcul des tailles d'échantillons correspondantes sans tenir compte d'une éventuelle rotation de l'échantillon. Le paragraphe 3 contient les différents estimateurs envisagés et le paragraphe 4 l'estimation de variance de ces estimateurs. Le paragraphe 5 procède à la comparaison des différentes variantes de schémas de rotation envisagés alors que le paragraphe 6 est consacré aux conclusions et à la variante privilégiée.

# 1 Buts de la révision de l'ESPA

L'objectif du projet "Révision ESPA : enquête continue aux normes EUROSTAT" est de réaliser une enquête continue satisfaisant à la réglementation d'EUROSTAT selon l'accord bilatéral entre la Suisse et l'UE dans le domaine de la statistique. Les buts peuvent être décrits ainsi :

- a) Passer à une enquête continue en répartissant de manière uniforme les interviews sur les 52 semaines de l'année (production d'indicateurs trimestriels et annuels).
- b) Respecter les critères de précision portant sur les effectifs annuels moyens de chômeurs par grande région et les variations trimestrielles de sous-groupes de la population active.
- c) Intégrer les modules thématiques développés dans le cadre du programme européen des modules.

Selon le règlement de la Communauté européenne (CE) 577/98 du 9 mars 1998 relatif à l'organisation d'une enquête par sondage sur les forces de travail, il est prévu dans son article premier que l'enquête continue doit fournir des résultats trimestriels et annuels. Les informations recueillies pendant l'enquête concernent généralement la situation au cours de la semaine précédant l'entretien, dite semaine de référence. Pour le cas de l'enquête continue, les semaines de référence doivent être réparties uniformément sur l'ensemble de l'année. A noter que la semaine de référence et la date de l'entretien ne peuvent être distantes de plus de cinq semaines, sauf au cours du troisième trimestre. Enfin, les trimestres et années de référence sont respectivement des ensembles de treize ou cinquante-deux semaines consécutives. Dans son article 2, le règlement stipule que l'enquête est effectuée auprès d'un échantillon de ménages ou d'individus résidant sur le territoire économique au moment de l'enquête et précise les méthodes d'observation selon les unités d'enquête.

## 2 Critères de précision selon les normes EUROSTAT

Les critères de précision selon les normes EUROSTAT ont été précisés dans l'article 3 (point 1 et 2 respectivement) du règlement 577/98 du 9 mars 1998 de la CE évoqué au paragraphe précédent ainsi que dans le règlement 430/2005 du 15 mars 2005 de la CE. Ces trois critères sont repris dans ce paragraphe selon les termes exacts utilisés dans les règlements y relatifs. Pour chacun d'eux, un calcul initial (dans le cas d'un échantillon aléatoire simple, sans effet de rotation et en négligeant le terme de correction pour population finie) de la taille d'échantillon qu'il faut pour satisfaire le critère est présenté.

1. Règlement 577/98 du 9 mars 1998, article 3, point 1 :

Pour un groupe de chômeurs représentant 5% de la population d'âge actif, l'erreur-type relative pour l'estimation des moyennes annuelles (ou des estimations de printemps, dans le cas d'une enquête annuelle effectuée au printemps) ne doit pas excéder 8% au niveau de la NUTS II de la sous-population en cause.

Les régions de moins de 300'000 habitants ne sont pas soumises à cette condition.

Selon ce premier critère, par NUTS II (grande région), on utilise la formule :

$$CV(p) = \frac{s}{p} = \frac{1}{p} \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{1-p}{p \cdot n}}$$
$$n = \frac{1-p}{p[CV(p)]^2}$$

Dans notre cas, on trouve par grande région (strate) la taille annuelle d'échantillon minimale suivante :

$$n = \frac{1 - 0.05}{0.05[0.08]^2} \simeq 2969 \text{ annuellement par grande région (NUTS II)}$$

2. Règlement 577/98 du 9 mars 1998, article 3, point 2 :

Dans le cas d'une enquête continue, pour des sous-populations représentant 5% de la population d'âge actif, l'erreur-type relative pour l'estimation des variations entre deux trimestres successifs, au niveau national, ne doit pas excéder 2% de la sous-population en cause.

Pour les Etats membres dont la population est comprise entre un million et vingt millions d'habitants, la condition précédente est allégée de telle sorte que l'erreur-type relative pour l'estimation des variations trimestrielles ne doit pas excéder 3% de la sous-population en cause.

Les Etats membres dont la population est inférieure à un million d'habitants ne sont pas soumis à ces exigences de précision concernant les variations.

Pour calculer la taille d'échantillon trimestrielle minimale pour satisfaire ce second critère, introduisons les notations suivantes : soient  $p_{t_2} - p_{t_1}$  la variation entre deux trimestres successifs,  $\rho$  la corrélation entre les variables  $p_{t_1}$  et  $p_{t_2}$ , et  $q$  le taux de recouvrement entre les deux échantillons au temps  $t_1$  et  $t_2$ . On utilise la formule :

$$\begin{aligned}
\text{Var}(p_{t_2} - p_{t_1}) &= \text{Var}(p_{t_1}) + \text{Var}(p_{t_2}) - 2\text{Cov}(p_{t_1}, p_{t_2}) \\
&= \text{Var}(p_{t_1}) + \text{Var}(p_{t_2}) - 2\rho q \frac{[\text{Var}(p_{t_1}) + \text{Var}(p_{t_2})]}{2} \\
&= [\text{Var}(p_{t_1}) + \text{Var}(p_{t_2})](1 - \rho q) \\
&\approx \frac{2p(1-p)}{n}(1 - \rho q) \\
\Rightarrow n &= \frac{2p(1-p)}{\text{Var}(p_{t_2} - p_{t_1})}(1 - \rho q) \\
\Rightarrow n &= \frac{2p(1-p)}{(p_{t_2} - p_{t_1})^2 [CV(p_{t_2} - p_{t_1})]^2}(1 - \rho q)
\end{aligned}$$

On peut dès lors calculer la taille trimestrielle d'échantillon minimale :

$$n = \frac{2 \cdot 0.05 \cdot 0.95}{(0.05)^2 [0.03]^2} \simeq 42222 \text{ trimestriellement au niveau Suisse}$$

Selon la corrélation entre deux trimestres successifs ( $\rho$ ) et le chevauchement des échantillons ( $q$ ), la taille d'échantillon annuelle peut varier considérablement comme le montre le tableau suivant :

	$q$	0%	25%	50%	75%	100%
$\rho$						
0.5		168889	147778	126667	105556	84444
0.6		168889	143556	118222	92889	67556
0.7		168889	139333	109778	80222	50667
0.8		168889	135111	101333	67556	33778
0.9		168889	130889	92889	54889	16889

Une autre manière de visualiser ce critère est de représenter les courbes de niveau de la fonction  $f(\rho, q) = 168888(1 - \rho q)$ , comme sur la figure 1.

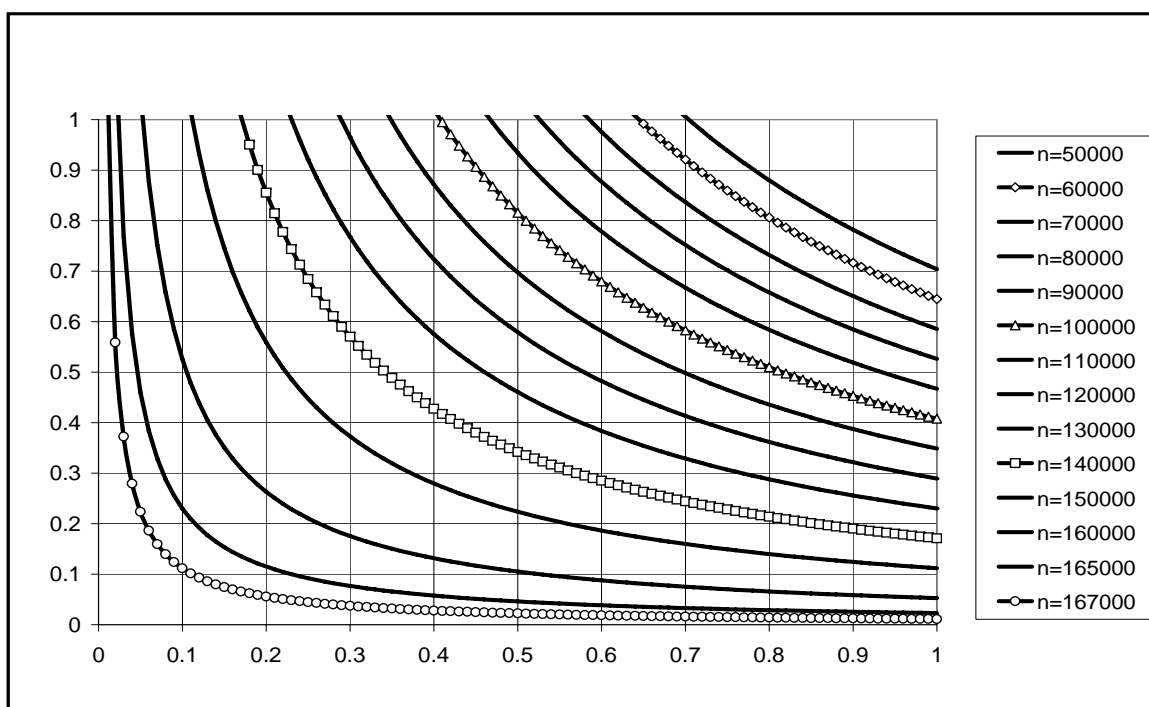


Figure 1 : Courbes de niveau de la fonction  $f(\rho, q) = 168888(1 - \rho q)$ .

3. Règlement 430/2005 du 15 mars 2005, Annexe I, points 1 et 2 :

Pour les variables structurelles (collectées uniquement comme moyennes annuelles sur la base de 52 semaines au moyen d'un sous-échantillon d'observations indépendantes), l'erreur-type relative (en faisant abstraction de l'effet de sondage) de toute estimation annuelle représentant 1% ou plus de la population en âge de travailler ne peut excéder

- a) 9% pour les pays dont la population est comprise entre un et vingt millions d'habitants;
- b) 5% pour les pays dont la population est supérieure à vingt millions d'habitants.

On trouve avec la même formule que pour le premier critère et  $CV(p) = 9\%$  dans le cas de la Suisse :

$$n = \frac{1 - p}{p[CV(p)]^2} = \frac{1 - 0.01}{0.01[0.09]^2} \simeq 12222 \text{ annuellement au niveau Suisse}$$

La figure 2 permet de visualiser ce type de critère pour d'autres valeurs de  $p$ .

## Coefficient de variation d'une proportion $p$ selon la taille $n$ de l'échantillon

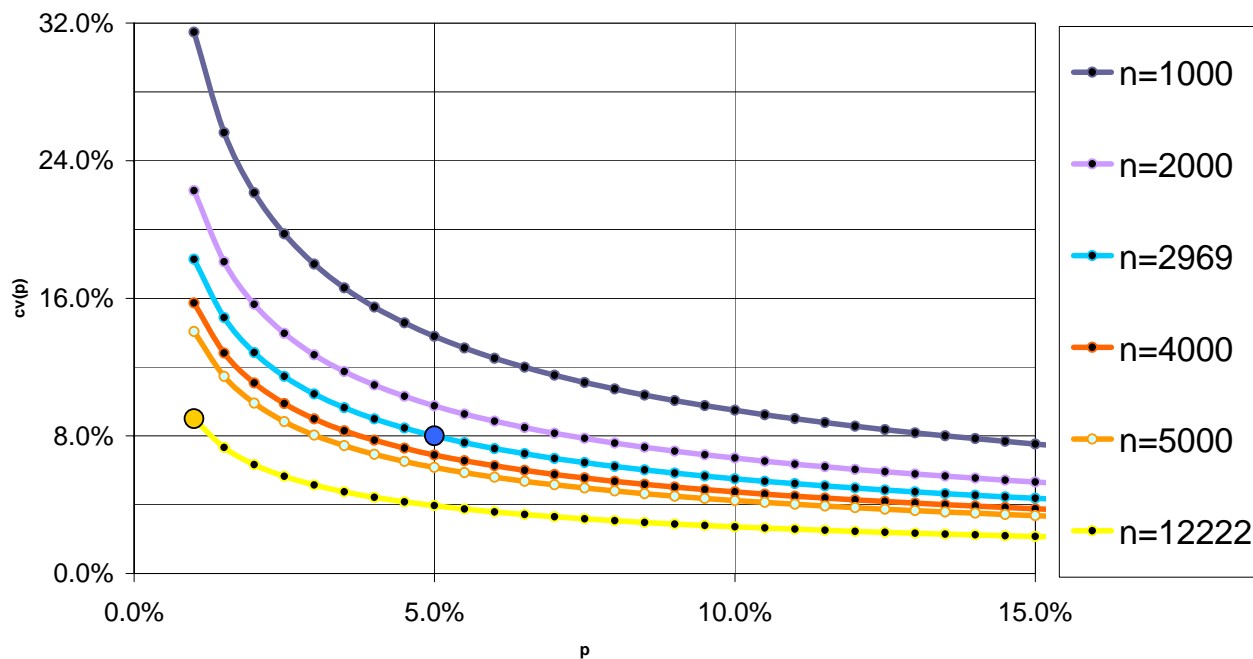


Figure 2 :  $CV(p)$  en fonction de  $p$ .

### 3 Estimateurs considérés

Avant d'expliciter les estimateurs importants pour cette étude, introduisons les notations suivantes :

$Y$	: Variable d'intérêt du total
$V$	: Variance
$C$	: Covariance
$q$	: Taux de recouvrement
$\rho$	: Corrélacion
$\Delta$	: Différence de totaux

En outre, trois types d'indices temporels seront utilisés :  $m$  désigne le mois,  $t$  le trimestre et  $a$  l'année. A noter qu'un estimateur trimestriel se calcule à partir de trois estimateurs mensuels issus de trois échantillons mensuels **indépendants**.

Il y a huit estimateurs considérés dans ce travail qui sont décrits ci-dessous :

1.	$\hat{Y}_m$
2.	$\hat{Y}_t = \frac{4}{13}(\hat{Y}_m + \hat{Y}_{m+1}) + \frac{5}{13}(\hat{Y}_{m+2})$
3.	$\hat{Y}_a = \frac{\hat{Y}_m + \hat{Y}_{m+1} + \dots + \hat{Y}_{m+11}}{12}$
4.	${}_1\hat{\Delta}_m = \hat{Y}_m - \hat{Y}_{m-1}$
5.	${}_{12}\hat{\Delta}_m = \hat{Y}_m - \hat{Y}_{m-12}$
6.	${}_1\hat{\Delta}_t = \hat{Y}_t - \hat{Y}_{t-1}$
7.	${}_4\hat{\Delta}_t = \hat{Y}_t - \hat{Y}_{t-4}$
8.	${}_1\hat{\Delta}_a = \hat{Y}_a - \hat{Y}_{a-1}$

## 4 Estimation de la variance des estimateurs

Le calcul de l'estimation de variance (sous l'hypothèse que la variance empirique de la variable  $Y$  est constante dans le temps) des huit estimateurs considérés figure ci-dessous en fonction de  $\hat{V}(\hat{Y}_m)$  qui peut être calculé à partir du mois concerné :

$$1. \hat{V}(\hat{Y}_m) = \hat{V}(\hat{Y}_m);$$

$$\begin{aligned}
 2. \hat{V}(\hat{Y}_t) &= \hat{V} \left[ \frac{4}{13}(\hat{Y}_m + \hat{Y}_{m+1}) + \frac{5}{13}(\hat{Y}_{m+2}) \right] = \frac{16}{169} \hat{V}(\hat{Y}_m + \hat{Y}_{m+1}) + \frac{25}{169} \hat{V}(\hat{Y}_{m+2}) \\
 &\approx \left( \frac{16}{169} \cdot 2 + \frac{25}{169} \right) \hat{V}(\hat{Y}_m) \\
 &= \frac{57}{169} \hat{V}(\hat{Y}_m) \\
 &\approx \frac{1}{3} \hat{V}(\hat{Y}_m);
 \end{aligned}$$

$$\begin{aligned}
3. \widehat{V}(\widehat{Y}_a) &= \widehat{V} \left[ \frac{\widehat{Y}_m + \widehat{Y}_{m+1} + \cdots + \widehat{Y}_{m+11}}{12} \right] \\
&= \frac{1}{144} \widehat{V}(\widehat{Y}_m + \widehat{Y}_{m+1} + \cdots + \widehat{Y}_{m+11}) \\
&= \frac{1}{144} \sum_{i=1}^{12} \sum_{j=1}^{12} \widehat{C}(\widehat{Y}_i, \widehat{Y}_j) \\
&= \frac{1}{144} \left[ \widehat{C}(\widehat{Y}_1, \widehat{Y}_1) + \cdots + \widehat{C}(\widehat{Y}_{12}, \widehat{Y}_{12}) \right] \\
&\quad + \frac{2}{144} \left[ \widehat{C}(\widehat{Y}_1, \widehat{Y}_4) + \widehat{C}(\widehat{Y}_1, \widehat{Y}_5) + \widehat{C}(\widehat{Y}_1, \widehat{Y}_6) + \cdots + \widehat{C}(\widehat{Y}_9, \widehat{Y}_{10}) + \widehat{C}(\widehat{Y}_9, \widehat{Y}_{11}) + \widehat{C}(\widehat{Y}_9, \widehat{Y}_{12}) \right] \\
&\quad + \frac{2}{144} \left[ \widehat{C}(\widehat{Y}_1, \widehat{Y}_7) + \widehat{C}(\widehat{Y}_1, \widehat{Y}_8) + \widehat{C}(\widehat{Y}_1, \widehat{Y}_9) + \cdots + \widehat{C}(\widehat{Y}_6, \widehat{Y}_{10}) + \widehat{C}(\widehat{Y}_6, \widehat{Y}_{11}) + \widehat{C}(\widehat{Y}_6, \widehat{Y}_{12}) \right] \\
&\quad + \frac{2}{144} \left[ \widehat{C}(\widehat{Y}_1, \widehat{Y}_{10}) + \widehat{C}(\widehat{Y}_1, \widehat{Y}_{11}) + \widehat{C}(\widehat{Y}_1, \widehat{Y}_{12}) + \cdots + \widehat{C}(\widehat{Y}_3, \widehat{Y}_{10}) + \widehat{C}(\widehat{Y}_3, \widehat{Y}_{11}) + \widehat{C}(\widehat{Y}_3, \widehat{Y}_{12}) \right] \\
&\approx \frac{1}{144} \left\{ 12\widehat{V}(\widehat{Y}_m) \right\} + \frac{2}{144} \left[ 9\widehat{C}(\widehat{Y}_m, \widehat{Y}_{m+3}) + 6\widehat{C}(\widehat{Y}_m, \widehat{Y}_{m+6}) + 3\widehat{C}(\widehat{Y}_m, \widehat{Y}_{m+9}) \right] \\
&= \frac{1}{12} \widehat{V}(\widehat{Y}_m) + \frac{1}{72} \{ 9(q\widehat{\rho})_{m,m+3} + 6(q\widehat{\rho})_{m,m+6} + 3(q\widehat{\rho})_{m,m+9} \} \widehat{V}(\widehat{Y}_m);
\end{aligned}$$

$$4. \widehat{V}({}_1\widehat{\Delta}_m) = \widehat{V}(\widehat{Y}_m - \widehat{Y}_{m-1}) \stackrel{!}{=} \widehat{V}(\widehat{Y}_m) + \widehat{V}(\widehat{Y}_{m-1}) \approx 2\widehat{V}(\widehat{Y}_m);$$

$$5. \widehat{V}({}_{12}\widehat{\Delta}_m) = \widehat{V}(\widehat{Y}_m - \widehat{Y}_{m-12}) \approx 2\widehat{V}(\widehat{Y}_m) - 2(q\widehat{\rho})_{m,m-12}\widehat{V}(\widehat{Y}_m) = 2(1 - (q\widehat{\rho})_{m,m-12})\widehat{V}(\widehat{Y}_m);$$

$$6. \widehat{V}({}_1\widehat{\Delta}_t) = \widehat{V}(\widehat{Y}_t - \widehat{Y}_{t-1}) \approx 2\widehat{V}(\widehat{Y}_t) - 2(q\widehat{\rho})_{t,t-1}\widehat{V}(\widehat{Y}_t) \approx \frac{2}{3}(1 - (q\widehat{\rho})_{m,m-3})\widehat{V}(\widehat{Y}_m);$$

$$7. \widehat{V}({}_4\widehat{\Delta}_t) = \widehat{V}(\widehat{Y}_t - \widehat{Y}_{t-4}) \approx 2\widehat{V}(\widehat{Y}_t) - 2(q\widehat{\rho})_{t,t-4}\widehat{V}(\widehat{Y}_t) \approx \frac{2}{3}(1 - (q\widehat{\rho})_{m,m-12})\widehat{V}(\widehat{Y}_m);$$



$$\begin{aligned}
8. \widehat{V}({}_1\widehat{\Delta}_a) &= \widehat{V}(\widehat{Y}_a - \widehat{Y}_{a-1}) \\
&= \widehat{V} \left( \left[ \frac{\widehat{Y}_m + \widehat{Y}_{m+1} + \cdots + \widehat{Y}_{m+11}}{12} \right] - \left[ \frac{\widehat{Y}_{m-1} + \widehat{Y}_{m-2} + \cdots + \widehat{Y}_{m-12}}{12} \right] \right) \\
&= \frac{1}{144} \widehat{V}(\widehat{Y}_m + \cdots + \widehat{Y}_{m+11}) + \frac{1}{144} \widehat{V}(\widehat{Y}_{m-1} + \cdots + \widehat{Y}_{m-12}) \\
&\quad - \frac{1}{144} \cdot 2\widehat{C}(\widehat{Y}_m + \cdots + \widehat{Y}_{m+11}, \widehat{Y}_{m-1} + \cdots + \widehat{Y}_{m-12}) \\
&\approx 2 \cdot \left[ \frac{1}{12} \widehat{V}(\widehat{Y}_m) + \frac{1}{72} \{9(q\widehat{\rho})_{m,m-3} + 6(q\widehat{\rho})_{m,m-6} + 3(q\widehat{\rho})_{m,m-9}\} \widehat{V}(\widehat{Y}_m) \right] \\
&\quad - \frac{2}{144} \widehat{C}(\widehat{Y}_m + \cdots + \widehat{Y}_{m+11}, \widehat{Y}_{m-1} + \cdots + \widehat{Y}_{m-12}) \\
&\approx \left[ \frac{1}{6} \widehat{V}(\widehat{Y}_m) + \frac{1}{72} \{18(q\widehat{\rho})_{m,m-3} + 12(q\widehat{\rho})_{m,m-6} + 6(q\widehat{\rho})_{m,m-9}\} \widehat{V}(\widehat{Y}_m) \right] \\
&\quad - \frac{1}{72} [3 \cdot \{(q\widehat{\rho})_{m,m-3} + 2(q\widehat{\rho})_{m,m-6} + 3(q\widehat{\rho})_{m,m-9} + 4(q\widehat{\rho})_{m,m-12}\}] \\
&\quad - \frac{1}{72} [3 \cdot \{3(q\widehat{\rho})_{m,m-15} + 2(q\widehat{\rho})_{m,m-18} + (q\widehat{\rho})_{m,m-21}\}] \\
&= \frac{1}{6} \widehat{V}(\widehat{Y}_m) + \frac{1}{72} \{15(q\widehat{\rho})_{m,m-3} + 6(q\widehat{\rho})_{m,m-6} - 3(q\widehat{\rho})_{m,m-9}\} \widehat{V}(\widehat{Y}_m) \\
&\quad + \frac{1}{72} \{-12(q\widehat{\rho})_{m,m-12} - 9(q\widehat{\rho})_{m,m-15} - 6(q\widehat{\rho})_{m,m-18} - 3(q\widehat{\rho})_{m,m-21}\} \widehat{V}(\widehat{Y}_m).
\end{aligned}$$

## 5 Comparaison des schémas de rotation envisagés

Ce paragraphe est consacré à l'étude de quatre variantes de schémas de rotation pour les variables conjoncturelles (relevées chaque trimestre).

### 1. Schéma de rotation 3-3-3-3 :

Sous-échantillon	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4	...
1	X				
2	X	X			
3	X	X	X		
4	X	X	X	X	
5	X	X	X	X	...
6		X	X	X	...
7			X	X	...
8				X	...
⋮					...

Du point de vue du chevauchement des échantillons, on a un chevauchement de 80% à trois mois, de 60% à six mois de 40% à neuf mois et de 20% à douze mois.

### 2. Schéma de rotation 3-9 :

Sous-échantillon	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4	...
1	X				
2		X			
3			X		
4	X			X	
5	X	X			...
6		X	X		
7			X	X	
8				X	...
⋮					...

On a : chevauchement de 33% à trois mois, 0% à six mois, 33% à neuf mois et 33% à douze mois.

### 3. Schéma de rotation 3-9-3 :

Sous-échantillon	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4	...
1	X				
2	X	X			
3		X	X		
4			X	X	
5	X			X	...
6	X	X			...
7		X	X		
8			X	X	
⋮				X	...

On a : chevauchement de 50% à trois mois, 0% à six mois, 25% à neuf mois, 50% à douze mois et 25% à quinze mois.

4. Schéma de rotation 12-12-12-12 :

Sous-échantillon	Trimestre 1	Trimestre 2	Trimestre 3	Trimestre 4	...
1	X				
2		X			
3			X		
4				X	
5	X				...
6		X			
7			X		
8				X	
⋮					...

On a : chevauchement de 0% à trois mois, 0% à six mois, 0% à neuf mois et 50% à douze mois.

En faisant l'hypothèse que la corrélation entre deux trimestres successifs est de  $\rho = 60\%$  et que cette corrélation diminue de 3% à chaque trimestre, on peut comparer les trois variantes du point de vue de leur variance dans le tableau suivant :

Estimateur	Sans effet rotation	Avec effet de rotation			
	$\widehat{V}(\widehat{Y}_m)$	Variante 3-3-3-3	Variante 3-9	Variante 3-9-3	Variante 12-12-12-12
$\widehat{Y}_m$	1.00	1.00	1.00	1.00	1.00
$\widehat{Y}_t$	0.33	0.33	0.33	0.33	0.33
$\widehat{Y}_a$	0.08	0.18	0.12	0.13	0.08
${}_1\widehat{\Delta}_m$	2.00	2.00	2.00	2.00	2.00
${}_{12}\widehat{\Delta}_m$	2.00	1.79	1.66	1.49	1.49
${}_1\widehat{\Delta}_t$	0.67	0.35	0.53	0.47	0.67
${}_4\widehat{\Delta}_t$	0.67	0.60	0.55	0.50	0.50
${}_1\widehat{\Delta}_a$	0.17	0.27	0.17	0.17	0.12

De manière à pouvoir comparer les différentes variantes et leur impact sur la précision des estimateurs considérés, nous présentons dans le tableau ci-dessous les intervalles de confiance à 95% selon différentes tailles d'échantillon trimestriel. Pour savoir si les critères de précision selon le règlement d'EUROSTAT sont satisfaits pour une proportion de 5%, il faut considérer le principe suivant : pour l'estimateur  $\widehat{Y}_a$ , au niveau NUTS II, l'intervalle de confiance devrait être de  $\pm 0.8\%$  (8% de 5% équivaut à un écart-type de 0.4%). L'intervalle de confiance à 95% a approximativement une demi-longueur de  $2 \times 0.4\% = 0.8\%$ . En admettant que dans la grande région la plus petite (en l'occurrence le Tessin), le plan d'échantillonnage soit tel qu'il y ait une densification qui ramène la taille au Tessin à celle que l'on aurait de manière proportionnelle en Suisse centrale (de l'ordre de 8%), il est raisonnable de considérer un facteur 12 au maximum entre la taille au niveau Suisse et celle dans la plus petite des NUTS II. Ainsi le critère 1 EURO-STAT est satisfait quand l'intervalle de confiance pour  $\widehat{Y}_a$  est plus petit que  $\pm 0.8\% / \sqrt{12}$ , soit  $\pm 0.231\%$ .

Le deuxième critère de précision s'applique à la précision de  ${}_1\widehat{\Delta}_t$  dont l'intervalle de confiance doit être inférieur à  $\pm 0.300\%$  en terme absolu.

Variante	Est.	Taille de l'échantillon trimestriel				
		$n = 20000$	$n = 25000$	$n = 30000$	$n = 35000$	$n = 40000$
		en %	en %	en %	en %	en %
3-3-3-3	$\widehat{Y}_a$	$\pm 0.227$	$\pm 0.203$	$\pm 0.185$	$\pm 0.172$	$\pm 0.161$
3-3-3-3	${}_1\widehat{\Delta}_t$	$\pm 0.314$	$\pm 0.281$	$\pm 0.257$	$\pm 0.238$	$\pm 0.222$
3-3-3-3	${}_4\widehat{\Delta}_t$	$\pm 0.413$	$\pm 0.369$	$\pm 0.337$	$\pm 0.312$	$\pm 0.292$
3-3-3-3	${}_1\widehat{\Delta}_a$	$\pm 0.277$	$\pm 0.248$	$\pm 0.226$	$\pm 0.209$	$\pm 0.196$
3-9	$\widehat{Y}_a$	$\pm 0.182$	$\pm 0.163$	$\pm 0.148$	$\pm 0.137$	$\pm 0.128$
3-9	${}_1\widehat{\Delta}_t$	$\pm 0.390$	$\pm 0.349$	$\pm 0.318$	$\pm 0.295$	$\pm 0.276$
3-9	${}_4\widehat{\Delta}_t$	$\pm 0.397$	$\pm 0.355$	$\pm 0.324$	$\pm 0.300$	$\pm 0.281$
3-9	${}_1\widehat{\Delta}_a$	$\pm 0.222$	$\pm 0.198$	$\pm 0.181$	$\pm 0.167$	$\pm 0.157$
3-9-3	$\widehat{Y}_a$	$\pm 0.190$	$\pm 0.170$	$\pm 0.155$	$\pm 0.144$	$\pm 0.134$
3-9-3	${}_1\widehat{\Delta}_t$	$\pm 0.365$	$\pm 0.326$	$\pm 0.298$	$\pm 0.276$	$\pm 0.258$
3-9-3	${}_4\widehat{\Delta}_t$	$\pm 0.376$	$\pm 0.336$	$\pm 0.307$	$\pm 0.284$	$\pm 0.266$
3-9-3	${}_1\widehat{\Delta}_a$	$\pm 0.227$	$\pm 0.203$	$\pm 0.185$	$\pm 0.172$	$\pm 0.160$
12-12-12-12	$\widehat{Y}_a$	$\pm 0.154$	$\pm 0.138$	$\pm 0.126$	$\pm 0.116$	$\pm 0.109$
12-12-12-12	${}_1\widehat{\Delta}_t$	$\pm 0.436$	$\pm 0.390$	$\pm 0.356$	$\pm 0.330$	$\pm 0.308$
12-12-12-12	${}_4\widehat{\Delta}_t$	$\pm 0.376$	$\pm 0.336$	$\pm 0.307$	$\pm 0.284$	$\pm 0.266$
12-12-12-12	${}_1\widehat{\Delta}_a$	$\pm 0.188$	$\pm 0.168$	$\pm 0.153$	$\pm 0.142$	$\pm 0.133$

On constate que le premier critère est toujours satisfait (bien que très proche de la limite dans la variante 3-3-3-3 avec  $n = 20000$  observations). Par contre le deuxième critère qui s'applique à la variation entre deux trimestre successifs est plus restrictif : s'il est déjà satisfait avec une taille de 25000 dans la variante 3-3-3-3, il faut en revanche 35000 observations pour le satisfaire avec la variante 3-9 et 30000 avec la variante 3-9-3, alors que ce critère n'est pas satisfait même avec 40000 observations avec la variante 12-12-12-12. Etant donné que le chevauchement sera inférieur à celui fait dans nos hypothèses (en raison de la non-réponse au cours du temps), que la corrélation se détériore à un rythme plus important que présenté ici et que l'effet de plan n'est peut-être pas négligeable, il est dans un premier temps raisonnable de prendre comme taille d'échantillon trimestriel environ 40000 observations, au moins en ce qui concerne les variantes 3-9 et 3-9-3. La première variante 3-3-3-3 est incontestablement celle qui privilégie le plus la précision de la variation entre deux trimestres consécutifs mais au détriment de tous les autres estimateurs et pourrait se satisfaire d'une taille trimestrielle d'environ 30000 observations, alors que la dernière variante 12-12-12-12 privilégie l'estimation annuelle ainsi que la variation annuelle au détriment de la variation trimestrielle devrait compter environ 50000 observations trimestriellement.

## 6 Conclusions et variante privilégiée

Cette étude préliminaire sur les estimateurs de l'ESPA continue et leur précision afin de satisfaire les critères d'EUROSTAT montre que le choix de la variante de schéma de rotation est cruciale. En effet, selon le schéma de rotation choisi, on privilégie tel estimateur plutôt que tel autre du point de vue de la précision (pour une taille d'échantillon fixée). Comme le critère le plus contraignant est celui relatif à la précision de la variation de deux trimestres successifs, il semble à priori logique de se focaliser sur une variante avec un chevauchement important entre deux trimestres successifs, comme la variante 3-3-3-3. Il ne faut cependant pas perdre de vue que cela se fait au détriment d'autres estimateurs comme l'estimation annuelle par exemple. A l'inverse, un schéma de rotation ne présentant aucun chevauchement à trois mois comme la variante 12-12-12-12 produit d'excellents estimateurs annuels ou encore de variation entre deux années successives en négligeant la précision de la variation trimestrielle. Cette dernière variante requiert ainsi une taille d'échantillon relativement grande pour garantir la précision requise par EUROSTAT du point de vue de la variation trimestrielle. Un bon compromis semble se dessiner pour les variantes 3-9 et 3-9-3 respectivement pour lesquelles la précision de l'un des estimateurs n'est pas drastiquement privilégiée par rapport à un autre.

**La préférence est donnée à la variante 3-9-3** puisque parmi les deux variantes qui offrent un bon compromis, cette dernière fournit une estimation un peu plus précise de la variation trimestrielle, l'estimation la plus contraignante selon les normes EUROSTAT.

## Bibliographie

Bosredon, J. et Février, P. (2000). *Estimations dans l'enquête emploi en continu*. Actes des journées de méthodologie statistique, 4 et 5 décembre 2000, Méthodologie statistique, n°100, pp. 393-413, Méthodologie statistique, Paris, INSEE.

Caron, N. et Ravalet, P. (2000). *Estimation dans les enquêtes répétées : application à l'enquête emploi en continu*. Actes des journées de méthodologie statistique, 4 et 5 décembre 2000, Méthodologie statistique, n°100, pp. 327-392, Méthodologie statistique, Paris, INSEE.

Christine, M. (2000). *La construction de l'échantillon de la future enquête emploi en continu à partir du recensement de 1999*. Actes des journées de méthodologie statistique, 4 et 5 décembre 2000, n°100, pp. 175-229, Méthodologie statistique, Paris INSEE.

Kytir, J. und Stadler, B. (2004). *Die kontinuierliche Arbeitskräfteerhebung im Rahmen des neuen Mikrozensus. Vom "alten" zum "neuen" Mikrozensus*. In : Statistische Nachrichten 6/2004, S. 511-518.

Labour Force Survey User Guide (2003). *Volume 1 : Background & Methodology*.

Mirza, H. and Hörngren, J. (2002). The sampling- and the estimation procedure in the Swedish Labour Force Survey. R&D Report 2002:4, Statistics Sweden.

RÈGLEMENT (CE) N° 577/98 DU CONSEIL DU 9 mars 1998 relatif à l'organisation d'une enquête par sondage sur les forces de travail dans la Communauté.

RÈGLEMENT (CE) N° 430/2005 DE LA COMMISSION DU 15 mars 2005 concernant la mise en oeuvre du règlement (CE) n° 577/98 du conseil relatif à l'organisation d'une enquête par sondage sur les forces de travail dans la Communauté en ce qui concerne la codification à utiliser pour la transmission des données à compter de 2006 et l'utilisation d'un sous-échantillon pour la collecte de données de variables structurelles.

*The European Union Labour Force Survey : main characteristics of the national surveys*. Edition 2005.

Tillé, Y., Qualité, L. (2009). *Estimation de la précision d'évolutions dans l'enquête sur la valeur ajoutée*. Rapport de méthodes 338-00XX, OFS, à paraître.

De Vitiis, C., Di Consiglio, L. Falorsi, S. (2005). *Studio del disegno campionario per la nuova rilevazione continua sulle Forze di lavoro*. ISTAT, Servizio Progettazione e Supporto Metodologico.