

Analyser les disparités spatiales sur un indicateur conjoncturel avec le recensement : l'exemple du taux de chômage.

Stéphane JUGNOT

DARES, Département du marché du travail

Introduction

Le recensement de population fournit des informations essentielles pour l'analyse des marchés locaux de l'emploi, qu'il s'agisse des évolutions sur le moyen terme ou d'analyses des disparités spatiales. C'est en effet la seule source qui fournit des données cohérentes en matière d'emploi, de population active et de chômage, données qui permettent notamment de calculer des taux de chômage (« au sens du recensement ») à un niveau local fin.

Parmi ces utilisations, les analyses de l'évolution du bouclage « emploi – chômage - population active » au niveau des zones d'emploi pour distinguer les effets des facteurs démographiques, de l'évolution du tissu économique (*cf.* [3] et [6]). Autre exemple, la mobilisation des différences de taux de chômage, parmi d'autres indicateurs, pour l'analyse des disparités entre quartiers dans le cadre de la politique de la ville, ou entre territoires composés d'agréats de communes, dans le cadre du zonage des fonds structurels européens.

La relance de la décentralisation et plus particulièrement la plus grande place laissée aux acteurs locaux en matière de politiques d'emploi, comme le développement du souci de l'évaluation, pourraient renforcer les attentes.

Comme l'a souligné un groupe de travail *ad hoc* du Cnis [1], la mise en place du recensement rénové de la population et sa collecte étalée sur cinq ans, présente une difficulté par rapport à l'analyse des disparités spatiales puisque les territoires ne seront plus recensés tous à la même date. Cette difficulté concerne en particulier les communes de moins de 10 000 habitants, recensées exhaustivement, une fois tous les cinq ans, pour un cinquième d'entre elles chaque année mais elle concerne aussi les agrégats de communes qui comportent des communes de moins de 10 000 habitants. Pour les variables très cycliques, en particulier les indicateurs conjoncturels du marché du travail, les écarts observés entre deux territoires recensés à des dates différentes résultent alors à la fois de disparités spatiales, supposées structurelles, et des différences dans la situation conjoncturelle du marché du travail au moment de la collecte.

La première partie illustre cette difficulté dans le cas du taux de chômage. La deuxième partie montre que celle-ci peut être en partie surmontée si l'on accepte de conserver l'hypothèse implicite que les utilisateurs du recensement traditionnel faisaient en utilisant ses résultats plusieurs années après la collecte, hypothèse selon laquelle les disparités spatiales reflètent des inégalités structurelles, donc stables dans le temps. Cette hypothèse n'est cependant pas forcément pertinente à un échelon très fin. La troisième partie aborde succinctement ce point à partir d'un exercice de simulation frustré.

1. Illustration de la difficulté

Après quelques données de cadrage, cette partie illustre le problème que peut poser l'utilisation de données collectées à des dates différentes pour l'analyse des disparités spatiales dans le cas du taux de chômage, à partir d'un cas d'école, forcément irréaliste, mais qui permet de préciser ce qui doit être pour les utilisateurs, un point d'attention particulier par rapport aux habitudes prises avec le recensement traditionnel.

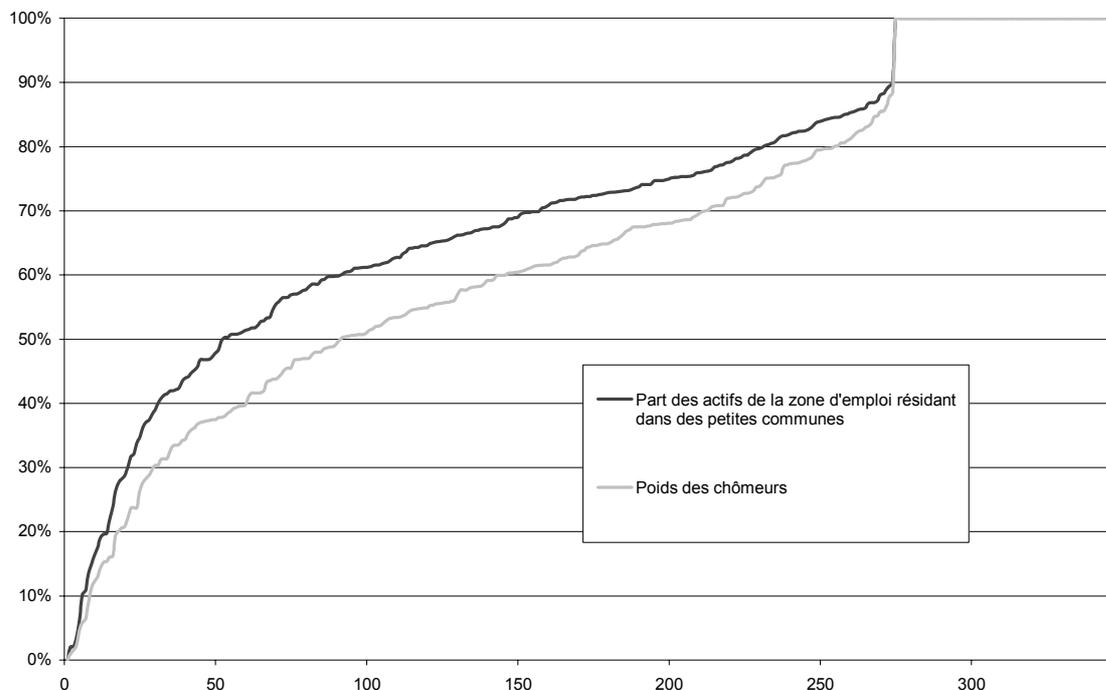
Quelques données de cadrage

Dans ses grandes lignes, la nouvelle méthode de collecte du recensement consiste à recenser exhaustivement une fois tous les cinq ans les « petites communes », de moins de 10 000 habitants. Ces communes sont réparties en cinq groupes de façon à recenser une commune sur cinq chaque année, donc toutes sur cinq ans. Les « grandes communes » ne sont plus recensées de façon exhaustive. Chaque année, un échantillon d'adresses représentant près de 8% des logements de la commune est enquêté. Sur cinq ans, 40% des logements sont donc enquêtés (pour une présentation plus développée, on peut se reporter à [2], [4] et [5]).

Au recensement de 1999, près de 900 communes, sur moins de 37 000, dépassaient les 10 000 habitants en France métropolitaine. Les « petites » communes rassemblaient la moitié des 26 millions d'actifs et 40% des 3 millions de chômeurs recensés.

Les zones d'emploi ont été dessinées à l'origine pour disposer d'un zonage d'études adapté à l'analyse des marchés locaux du travail, en s'appuyant sur les navettes domicile - travail pour se rapprocher de la notion de « bassin d'emploi ». Si l'on s'intéresse à ce découpage, près d'une zone d'emploi sur cinq n'est composée que de « petites communes ». Dans plus de neuf cas sur dix, plus de 90% des communes qui composent la zone d'emploi sont des « petites communes » et elles rassemblent plus de 40% des actifs de la zone.

Graphique 1 - L'importance des « petites communes » dans les zones d'emploi

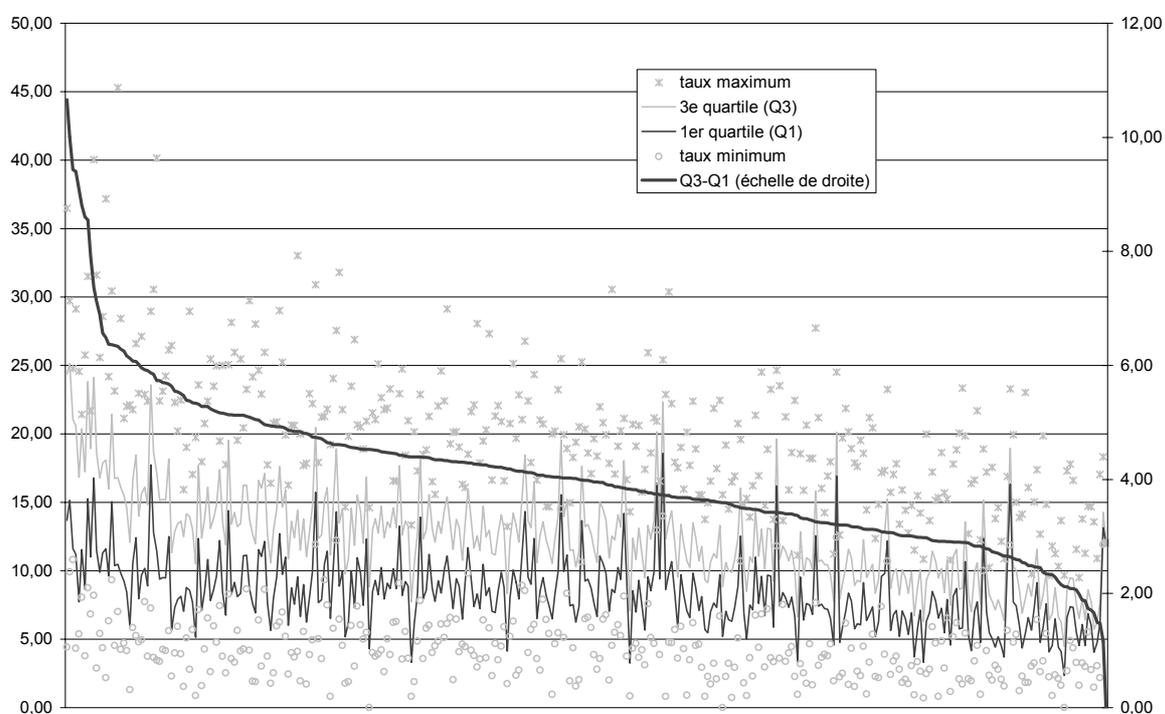


Source : Insee-RP99.

Les aires urbaines sont également un zonage pertinent pour étudier les marchés du travail locaux, étant elles aussi construites à partir des navettes domicile - travail. Dans les trente premières aires urbaines de métropole, au moins huit communes sur dix sont des « petites communes » (sauf pour Toulon, où leur part est de sept sur dix). Dans vingt-deux aires urbaines, les « petites communes » rassemblent entre le tiers et six dixièmes des actifs. C'est notamment le cas pour l'aire urbaine de Marseille. Dans les aires urbaines de Lyon et de Paris, les « petites communes » rassemblent un peu moins de 20% des actifs.

Les disparités territoriales sont relativement marquées. Dans la moitié des zones d'emploi, l'écart interquartile de taux de chômage est supérieur à 4 points (dit autrement, l'écart de taux de chômage entre les 25% de communes de la zone d'emploi ayant le taux de chômage le plus faible et les 25% des communes ayant le taux de chômage le plus élevé, est supérieur à 4 points). Dans un peu moins de 20% des zones d'emploi, l'écart interquartile est inférieur à 3 points, ce qui correspond à une amplitude de variation conjoncturelle du taux de chômage qui s'est déjà observée sur une période de cinq ans aux niveaux régionaux et départementaux.

Graphique 2 – Les disparités de taux de chômage communaux par zone d'emploi (communes ayant au moins 100 actifs)



Source : Insee-RP99.

Un cas d'école

La « réalité » inobservée

Dans cet exemple, on considère donc cinq communes de moins de 10 000 habitants appartenant à un groupe de rotation différent, nommées A, B, C, D et E. On considère également une commune, F, de plus de 10 000 habitants, recensée par sondage tous les ans. Pour cette commune, on met de côté le problème de l'imprécision de l'observation liée au sondage. La non réponse est également laissée de côté (il existait déjà dans le cas du recensement traditionnel). Tous ces aspects ne changent *a priori* pas le problème spécifique qui nous intéresse. Ils ne font qu'ajouter du flou.

On suppose que les six communes ont le même taux de chômage. Pour que les variations annuelles soient réalistes, on retient pour l'exemple le taux de chômage observé au niveau national les cinq dernières années. Pour simplifier, pour la grande commune, on suppose par ailleurs que les cinq groupes de rotation ont la même composition socio-démographique et connaissent de ce fait le même taux de chômage que la grande commune dans son ensemble.

On considère par ailleurs deux petites communes, l'une "structurellement" moins touchée par le chômage que la moyenne (commune G), l'autre davantage (commune H). Dans l'exemple, l'écart à la moyenne est supposé constant dans le temps. Cette stabilité totale n'est évidemment pas réaliste. Si les disparités territoriales n'évoluent que lentement, cette simplification n'est pas gênante pour le propos. De même, on pourrait complexifier l'exemple en raisonnant à partir des données individuelles, donc du nombre de chômeurs et du nombre d'actifs recensés. Pour faciliter la lecture, l'exemple est présenté directement sur le taux de chômage.

Le tableau ci-dessous présente l'évolution « réelle » du taux de chômage dans ces huit communes au cours des 5 années du cycle de collecte.

Tableau 1 – évolution « réelle »

	An1	An2	An3	An4	An5
Taux de chômage des petites communes A, B, C, D, E, et de la grande commune F, (taux de chômage dans la moyenne nationale)	12,2%	12,0%	11,2%	10,3%	8,9%
Taux de chômage de la petite commune G, moins touchée par le chômage que la moyenne nationale (écart de - 1 point)	11,2%	11,0%	10,2%	9,3%	7,9%
Taux de chômage de la petite commune H, plus touchée par le chômage que la moyenne nationale (écart de +1 point)	13,2%	13,0%	12,2%	11,3%	9,9%

Les informations collectées

Le tableau ci-dessous présente le taux de chômage résultant des informations collectées. La première année, alors que la conjoncture est mauvaise, les petites communes A et G et le premier groupe de rotation de la grande commune F sont recensés. Sur les données de collecte, le taux de chômage observé est respectivement de 12,2%, 11,2% et 12,2%. La quatrième année du cycle, année plus favorable au niveau de la conjoncture du marché du travail, les petites communes D et H et le quatrième groupe de rotation de la grande commune F sont recensés. Sur les données de collecte, le taux de chômage observé est respectivement de 10,3%, 11,3% et 10,3%.

Tableau 2 – information collectée

	An1	An2	An3	An4	An5
Petite commune A - groupe 1	12,2%	//	//	//	//
Petite commune B - groupe 2	//	12,0%	//	//	//
Petite commune C - groupe 3	//	//	11,2%	//	//
Petite commune D - groupe 4	//	//	//	10,3%	//
Petite commune E - groupe 5	//	//	//	//	8,9%
Grande commune F - groupe 1	12,2%	//	//	//	//
Grande commune F - groupe 2	//	12,0%	//	//	//
Grande commune F - groupe 3	//	//	11,2%	//	//
Grande commune F - groupe 4	//	//	//	10,3%	//
Grande commune F - groupe 5	//	//	//	//	8,9%
Petite commune G - groupe 1	11,2%	//	//	//	//
Petite commune H - groupe 4	//	//	//	11,3%	//

// : pas de collecte.

Une utilisation « empilée » des données collectées biaise l'analyse

Pour les petites communes, recensées exhaustivement une seule fois pendant les cinq années du cycle, les évolutions de population peuvent être interpolées ou extrapolées entre deux recensements à partir de sources externes au recensement, comme la taxe d'habitation, ce qui modifie alors le poids des individus interrogés pour « corriger » la population totale.

Mobiliser des sources administratives pour extrapoler ou interpoler d'autres informations, comme l'emploi et le chômage, la pyramide des âges, etc., ne semble pas faisable facilement au moins à court terme. Par exemple, aucune source ne couvre la totalité du champ de l'emploi et du chômage ; les sources « emplois » posent souvent le problème du passage du lieu de travail au lieu de domicile et la fiabilité de la localisation géographique est d'autant plus problématique que l'on cherche une localisation fine.

De ce fait, pour les petites communes, les informations structurelles disponibles dans les données de synthèse relative au cycle de collecte de 5 ans, qui seront ici datées à l'année médiane « An3 », seront à peu de chose près celles observées lors de l'année de collecte, l'année « An1 » pour la commune A, « An2 » pour la commune B, etc.

Pour les grandes communes, la donnée de synthèse sera faite à partir des données collectées sur les 5 groupes de rotation. Une extrapolation - interpolation est aussi possible pour tenir compte des évolutions de la population intervenues dans les groupes de rotation par rapport à l'année de leur collecte mais, comme pour les petites communes, on peut supposer que cela ne devrait pas modifier de façon significative les informations structurelles collectées initialement. De ce fait, le taux de chômage de synthèse sera une sorte de moyenne pondérée des taux de chômage observés successivement dans chacun des groupes de rotation.

Le tableau ci-dessous présente ce que pourraient alors être les données de synthèse diffusées au niveau communal.

Tableau 3 – Utilisation des données empilées simples

	Données diffusées « An 3 »	Conclusion de l'utilisateur naïf sur les disparités spatiales
Petite commune A	12,2%	1,3%
Petite commune B	12,0%	1,1%
Petite commune C	11,2%	0,3%
Petite commune D	10,3%	-0,6%
Petite commune E	8,9%	-1,1%
Grande commune F	10,9%*	Réf.
Petite commune G	11,2%	0,3%
Petite commune H	11,3%	0,4%

* donnée calculée en supposant que chaque groupe de rotation de la grande commune a la même importance numérique, ce qui ne sera pas forcément le cas si l'échantillon n'est pas tout à fait équilibré ou si la commune connaît des évolutions démographiques importantes au cours du cycle quinquennal de la collecte. Le taux de chômage synthétique sera alors tiré par la situation conjoncturelle prévalant l'année de collecte du groupe pesant le plus dans l'estimation.

Le taux de chômage diffusé est celui prévalant l'année de collecte pour les petites communes. C'est une sorte de taux de chômage moyen sur 5 ans pour la grande commune. Le taux de chômage étant cyclique, les données ne sont pas directement comparables entre elles. L'utilisation « naïve » des données sans tenir compte de la date de collecte peut alors conduire alors à des conclusions erronées sur les disparités territoriales.

2. Comparer des écarts à la moyenne

Biais de conjoncture ou biais de déformation des disparités

Soit $u(c,t)$, le taux de chômage dans la « petite commune » c à la date t . S'intéresser aux disparités entre la commune c , recensée en t , et la commune c' , recensée en t' , à la date t_0 , c'est regarder :

$$\begin{aligned}\Delta u(c,c'/t_0) &= u(c,t_0) - u(c',t_0) \\ &= [u(c,t) - u(c',t')] + [u(c,t_0) - u(c,t)] - [u(c',t_0) - u(c',t')] \\ &= [\hat{u}(c,t) - \hat{u}(c',t')] + [\varepsilon(c,t) - \varepsilon(c',t')] + [u(c,t_0) - u(c,t)] - [u(c',t_0) - u(c',t')]\end{aligned}$$

\Rightarrow écart « naïf » observé = écart réel + biais de conjoncture + solde d'erreurs de mesure (A)

Pour un indicateur peu conjoncturel et stable, le biais de conjoncture sera faible. Il sera également faible dans le cas très particulier où le point d'observation est à un point de retournement conjoncturel par rapport auquel les dates de collecte des deux communes les placeraient dans des phases symétriques du cycle.

Dans le cas général, les biais peuvent ne pas être négligeables dans la mesure où l'ampleur des disparités mesurées peut être du même ordre de grandeur que les variations conjoncturelles (ou tendancielle) de l'indicateur. Dans ce cas, il faut alors éviter d'utiliser l'écart « naïf » présenté dans la section précédente.

Si z désigne une zone géographique de référence et $E[c,z,t]$, l'écart entre le niveau de l'indicateur pour la commune c et « la moyenne » de la zone de référence z à la date t , on peut aussi écrire :

$$\begin{aligned}\Delta u(c,c'/t_0) &= [u(c,t) - u(z,t)] - [u(c',t') - u(z,t')] + [E(c,z,t_0) - E(c,z,t)] - [E(c',z,t_0) - E(c',z,t')] \\ &= [\hat{u}(c,t) - \hat{u}(z,t)] - [\hat{u}(c',t') - \hat{u}(z,t')] + [\varepsilon(c,t) - \varepsilon(c',t')] + [\varepsilon(z,t) - \varepsilon(z,t')] \\ &\quad + [E(c,z,t_0) - E(c,z,t)] - [E(c',z,t_0) - E(c',z,t')]\end{aligned}$$

\Rightarrow différence des écarts à « la moyenne » observée
= écart réel + biais lié à la déformation des disparités
+ solde d'erreurs de mesure sur les indicateurs communaux
+ solde d'erreur de mesure sur l'indicateurs de référence (B)

Le solde d'erreurs de mesure sur les indicateurs communaux est le même que dans l'approche précédente. Si le solde d'erreur de mesure sur l'indicateur de référence est faible et si la déformation des disparités territoriales est faible, alors la différence des écarts à « la moyenne » observée permet de mesurer sans trop d'erreur les disparités territoriales réelles.

Retour sur l'exemple

Dans le cas d'école présenté plus haut, pour les « petites communes », on peut mesurer l'écart entre le taux de chômage observé dans la commune à la date de la collecte et, par exemple, le taux de chômage national à cette même date. Pour les « grandes communes », on pourrait calculer de même les écarts sur les résultats de la collecte annuelle, donc groupe de rotation par groupe de rotation, et en déduire un écart moyen sur l'ensemble du cycle de collecte. On aurait alors le tableau suivant :

Tableau 4 – Utilisation des données empilées avec indicateurs de référence

	Données diffusées « An 3 »		Conclusion de l'utilisateur sur les disparités spatiales
	Données communales	Indicateur à l'échelon de référence	
Petite commune A	12,2%	12,2%	0%
Petite commune B	12,0%	12,0%	0%
Petite commune C	11,2%	11,2%	0%
Petite commune D	10,3%	10,3%	0%
Petite commune E	8,9%	8,9%	0%
Grande commune F	10,9%*	10,9%*	Réf.
Petite commune G	11,2%	12,2%	-1,0%
Petite commune H	11,3%	10,3%	1,0%

* donnée calculée en supposant que chaque groupe de rotation de la grande commune a la même importance numérique.

En pratique, raisonner en écart à la moyenne est donc préférable, sous les deux conditions suivantes :

1) Les points de référence doivent être fiables, c'est-à-dire le moins susceptible d'erreur de mesure. Comme ces points de référence résultent, chaque année, des données de la collecte annuelle et non des données de synthèse sur cinq ans, la zone géographique de référence doit être choisie avec soin ; d'une part, au regard de l'imprécision liée à l'échantillonnage des logements dans les grandes villes et à l'échantillonnage des « petites communes » dans les territoires supra communaux (aires urbaines, zones d'emploi, départements, régions, etc.) ; d'autre part, au regard du plan de sondage du recensement, qui n'a évidemment pas été équilibré dans le but d'obtenir des indicateurs fiables avec la collecte annuelle à n'importe quel échelon géographique. Cela pourrait conduire à privilégier l'utilisation des indicateurs nationaux comme indicateurs de référence pour les calculs des écarts à « la moyenne » .

2) Les disparités territoriales pour l'indicateur retenu doivent être de nature structurelle, autrement dit suffisamment stable dans le temps. De fait, c'était l'hypothèse implicite qui était faite quand les informations du recensement classique étaient utilisées pour des analyses structurelles plusieurs années après le recensement. De ce point de vue, cette approche ne change donc pas le cadre d'hypothèses sous-jacentes à l'utilisation des données pour l'utilisateur.

3. Simuler n'est pas trancher

Cette section présente une tentative de simulation visant à illustrer ce qui précède par un exemple plus réaliste que le cas d'école présenté. Peu concluante par rapport au problème posé, elle conduit à nuancer la force de l'hypothèse de stabilité des disparités territoriales.

La méthode

Le principe consiste à reconstruire des évolutions annuelles fictives sur 5 ans de l'emploi et du chômage, donc de la population active et du taux de chômage, en partant d'une situation observée au recensement de 1999. La simulation a été réalisée sur les communes de la région Rhône-Alpes, qui présente l'intérêt de disposer de bassins d'emploi connaissant des situations contrastées sur le marché du travail.

Les données du chômage et de l'emploi du recensement de 1999 sont supposées représenter la situation au 31 décembre 1998.

L'évolution du chômage est extrapolée commune par commune pour les années 1997 et 1999 à 2001, en fonction de l'évolution du nombre de demandeurs d'emploi inscrits à l'ANPE (situation au 31 décembre, inscription en catégorie 1, 2 ou 3, sans activité réduite).

Pour l'emploi, aucune source administrative n'est disponible. Les sources existantes (DADS, UNEDIC, fichier des agents de l'Etat, enquêtes COLTER, etc.) couvrent des champs partiels. La localisation géographique du lieu de travail n'y est pas toujours très bonne et il faut ensuite passer de l'emploi au lieu de travail à l'emploi au lieu de domicile. Dans ces conditions, il est difficile d'estimer une évolution de l'emploi au niveau communal.

Par simplicité, la variation de l'emploi au lieu de domicile est donc déduite de la variation du chômage simulée, commune par commune¹. On peut ajouter un effet d'appel pour prendre en fait le compte que lors que la situation du marché du travail s'améliore, des inactifs peuvent décider d'entrer ou de revenir sur le marché du travail.

Déduire la variation de l'emploi de celle du chômage revient en particulier à ne pas tenir compte des mobilités résidentielles. La baisse du chômage est imputée à la hausse de l'emploi local et non à des départs vers d'autres zones. Réciproquement, l'arrivée de nouveaux actifs en emploi n'est pas pris en compte.

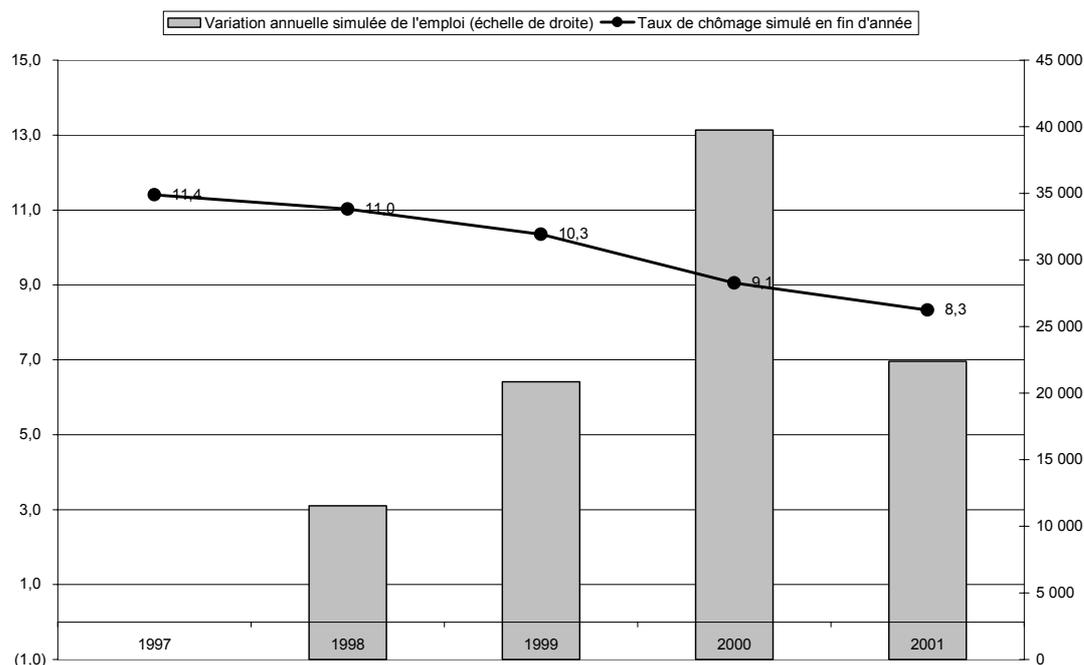
Les 2800 « petites communes » ont par ailleurs été réparties aléatoirement entre cinq groupes de collecte.

La simulation

Au niveau régional, sur les cinq années de collecte, le taux de chômage simulé baisse, avec une accélération en 2000. Il passe de 11,4% à 8,3%, soit une baisse de 3,1 points.

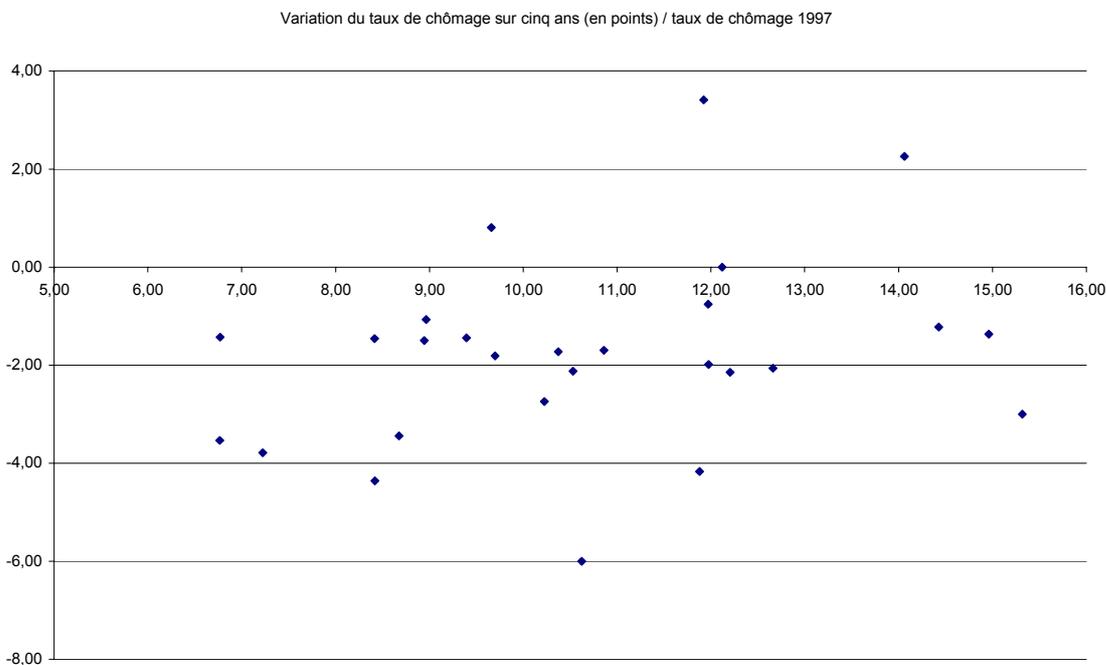
¹ Une variante a en fait été réalisée en utilisant les évolutions annuelles de l'emploi salarié du secteur privé de l'Unedic, disponible à la commune. Outre la réserve déjà formulée sur la qualité de la localisation géographique, il faut également des hypothèses pour passer du lieu de travail au lieu de domicile (à partir des données du recensement, l'emploi au lieu de domicile a été défini comme fonction de l'emploi au lieu de travail de la commune, de la zone d'emploi et de la région). Ils faut également faire des hypothèses sur les autres types d'emploi, salarié non marchand, agricole et emploi non salarié (de stabilité ou de prolongation de tendance). Cette simulation est plus complexe, pas forcément plus réaliste et conduit *in fine* aux mêmes conclusions que la simulation présentée.

Graphique 3 – Evolution simulée du taux de chômage et de l'emploi dans la région



La situation des 27 zones d'emploi est contrastée. Le taux de chômage initial, lors de la première année du cycle de collecte s'étend de 6,8% à 15,3%. La variation du taux de chômage sur 5 ans va de -6,0 points à +3,4 points. Elle est comprise entre 0 et -3 points pour la plupart des zones.

Graphique 4 – Situation simulée des 27 zones d'emploi



L'écart interquartile du taux de chômage des « petites communes » au recensement de 1999 (situation simulée datée 1998) est inférieur à 3 dans huit zones d'emploi et supérieur à 4 dans sept zones d'emploi.

L'analyse des disparités

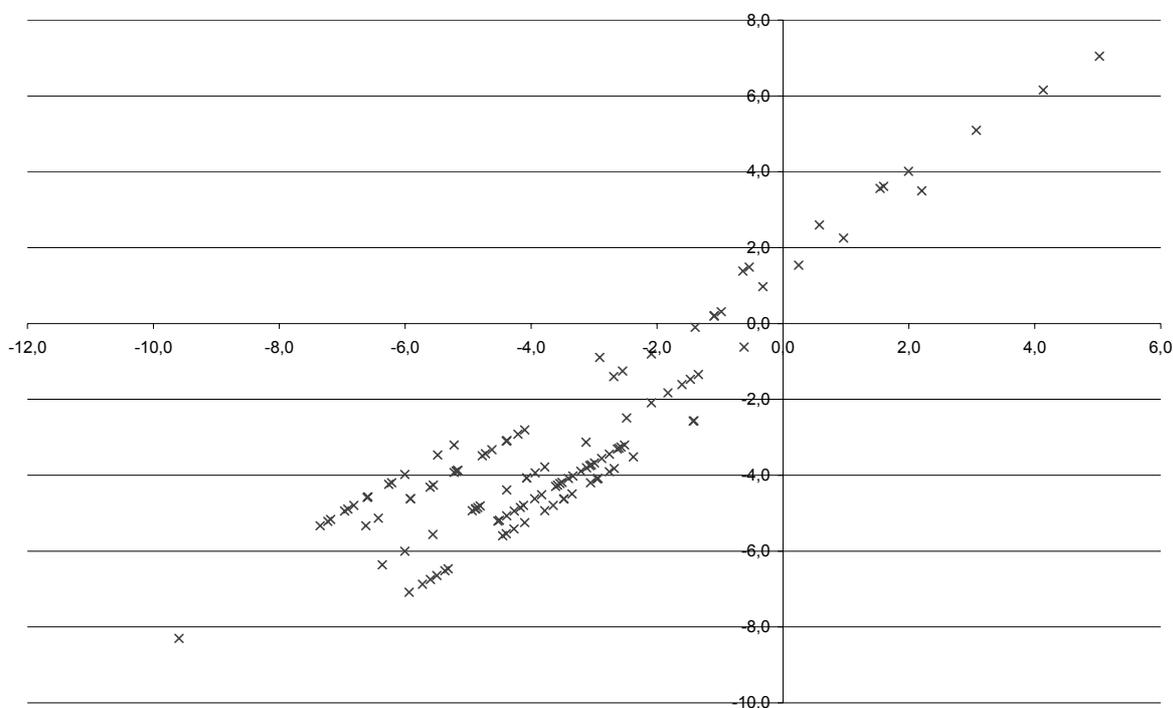
On cherche à comparer les deux approches présentées dans la deuxième partie : l'approche « naïve » qui s'appuie sur les données collectées sans tenir compte de l'année de collecte et une approche, qu'on qualifiera par la suite de « synthétique », qui s'appuie sur les écarts « à la moyenne » observés au moment de la collecte.

Afin de ne pas introduire trop de bruit, on suppose qu'il n'y a pas d'erreur de mesure sur l'estimation du calcul de référence, ici le taux de chômage régional simulé, et on ne s'intéresse qu'aux « petites communes », là aussi, sans erreur de mesure.

Analyser les disparités spatiales revient à classer les communes en fonction de l'écart de leur taux de chômage à « la moyenne ».

Le graphique ci-dessous présente un nuage de points des communes qui les positionne en fonction de la valeur de l'écart calculé selon les deux méthodes, « naïve » et « synthétique ». Pour plus de lisibilité, on ne s'intéresse ici qu'aux communes d'une seule zone d'emploi, la plus importante en nombre de « petites communes » (Lyon). De plus, ne sont retenues dans le graphique que les communes ayant un taux de chômage « réel » compris dans l'intervalle interquartile des valeurs prises dans la zone d'emploi, soit près de 123 communes.

Graphique 5 – Répartition des communes selon les deux indicateurs d'écart à la moyenne (en point de taux de chômage)

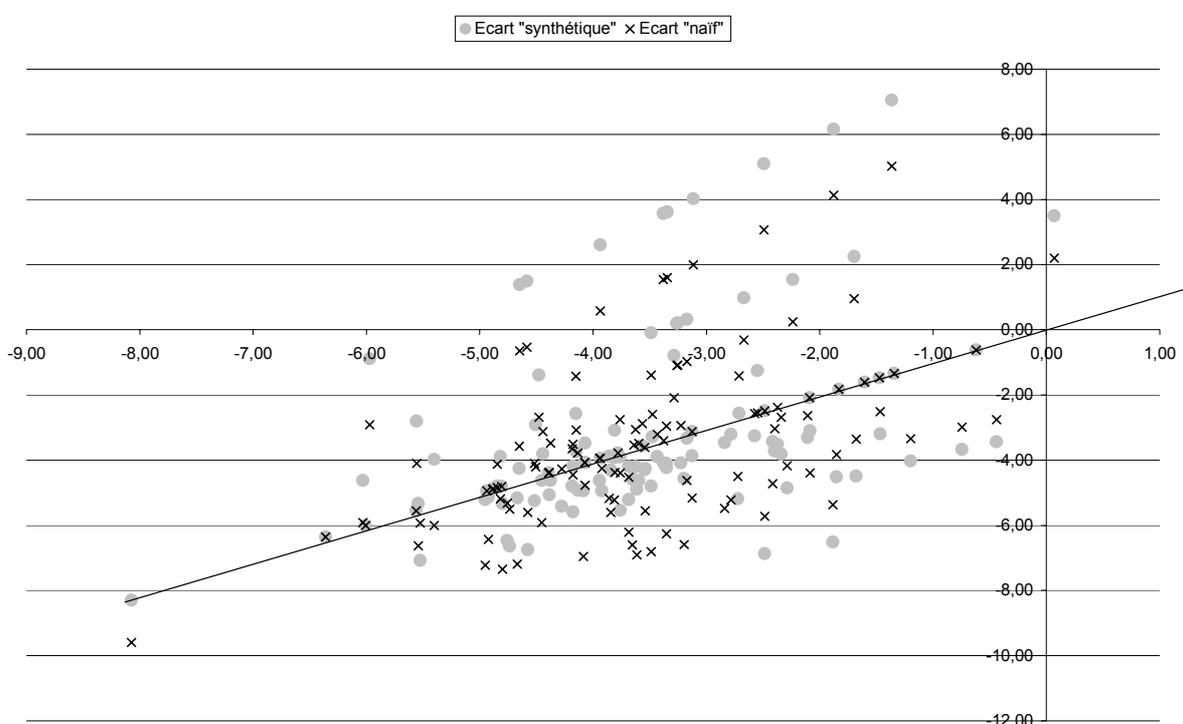


Dans un certain nombre de cas, les disparités sont d'une telle ampleur que les conclusions que l'on peut tirer en terme de disparités ne sont guère sensibles au choix de la méthode. En revanche, ce n'est pas le cas quand les écarts sont du même ordre de grandeur que l'amplitude des variations conjoncturelles du taux de chômage pour la zone de référence.

Pour juger de la performance comparée des deux méthodes, il convient de comparer les écarts à « la moyenne » estimés par l'une et l'autre méthode à l'écart « réel » qui a été simulé. C'est l'année du milieu du cycle de collecte qui sert ici de référence, soit 1999.

Dans le graphique ci-dessous, l'écart « réel » du taux de chômage de la commune par rapport à la moyenne (valeur régionale) figure en abscisse. L'écart estimé par la méthode « naïve » (nuage de points avec les croix) et l'écart estimé par la méthode « synthétique » (nuage de points avec les ronds) figurent en ordonnée. Par souci de lisibilité, le champ est là encore restreint aux 123 communes de la zone d'emploi de Lyon dont le taux de chômage est compris dans l'intervalle interquartile.

Graphique 6 – Ecart à la valeur « réel » selon la méthode (en point de taux de chômage)



A la lecture du graphique, il est difficile de conclure à la supériorité d'une méthode sur l'autre. Comme les erreurs de mesure sont supposées nulles, l'erreur d'estimation faite avec la méthode « naïve » correspond à ce que nous avons appelé le « biais de conjoncture » dans la relation (A), tandis que l'erreur d'estimation faite avec la méthode « synthétique » correspond au « biais lié à la déformation des disparités » de la relation (B). Dans la base simulée, les deux biais seraient donc du même ordre. Ce que le tableau ci-dessous présentant la distribution des erreurs selon les deux méthodes pour l'ensemble des zones d'emploi conduirait également à soutenir (on ne peut cependant conclure sur ce point puisqu'il s'agit ici de données simulées qui peuvent mal rendre compte de la stabilité réelle des disparités spatiales).

Tableau 5 – Distribution des erreurs d'estimation selon les méthodes (en point de taux de chômage)

Année de collecte	indicateur	1er quartile	médiane	3e quartile	maximum
1	"naïf"	0,56	1,36	2,84	16,18
	"synthétique"	0,81	1,66	2,99	15,04
2	"naïf"	0,52	1,19	2,12	11,11
	"synthétique"	0,64	1,05	2,41	10,43
3	"naïf"	0,00	0,00	0,00	0,00
	"synthétique"	0,00	0,00	0,00	0,00
4	"naïf"	0,64	1,45	3,50	10,29
	"synthétique"	0,99	1,87	2,70	11,59
5	"naïf"	2,32	3,50	5,33	17,50
	"synthétique"	1,08	2,41	6,80	19,52
Ensemble	"naïf"	0,16	1,29	2,81	17,50
	"synthétique"	1,08	1,06	2,95	19,52

Conclusion

Le nouveau mode de collecte du recensement pose un problème théorique pour l'étude des disparités spatiales pour des indicateurs conjoncturels tels que le taux de chômage. En pratique, ce problème est en partie atténué dans la mesure où, dans de nombreux cas, l'ampleur des disparités est plus forte que l'amplitude de variation de l'indicateur sur un cycle de cinq ans.

Par ailleurs, on peut toujours raisonner en s'appuyant sur les écarts à « la moyenne » par rapport à une zone de référence à la date de collecte, ce qui revient à faire la même hypothèse de stabilité des disparités que les utilisateurs faisaient, explicitement ou non, quand ils utilisaient le recensement traditionnel plusieurs années après la collecte. Il y a alors sans doute intérêt dans ce cas à privilégier l'indicateur national comme référence afin d'atténuer les erreurs de mesure.

Cette hypothèse de stabilité des disparités ne va pas de soi. Si cette hypothèse n'est pas vérifiée, il est possible que les erreurs faites soient du même ordre de grandeur que celles que l'on ferait si on utilisait les informations du recensement rénové de façon « naïve » sans tenir compte de la date de collecte. Elle présente toutefois l'intérêt d'être posée et d'être usuellement utilisée.

Bibliographie

- [1] Rapport du groupe de travail sur la dynamique des territoires, *Actes de la concertation sur le recensement rénové de la population*, CNIS, 2000.
- [2] Dumais Jean, Durr Jean-Michel, « la rénovation du recensement français », recueil du Symposium de Statistique Canada « la qualité des données d'un organisme statistique : une perspective méthodologique », 2001.
- [3] Stéphane Méloux, Jean Laganier, « Zones d'emploi de Provence-Alpes-Côte-d'Azur : des marchés du travail en croissance et de plus en plus ouverts », *Sud Insee, l'essentiel*, n°49, INSEE PACA, mars 2002
- [4] Bertrand Philippe, Christian Barbara, Chauvet Guillaume, Grosbras Jean-Marie, « Les plans de sondage du nouveau recensement », contribution aux journées de méthodologie statistique 2002.
- [5] Bertrand Philippe, Christian Barbara, Chauvet Guillaume, Grosbras Jean-Marie, « Données produites par le recensement rénové de la population », contribution aux journées de méthodologie statistique 2002.
- [6] Gitton François, « Trois années d'embellie sur le marché du travail franc-comtois », *L'essentiel*, n°70, INSEE Franche-Comté, mai 2004.

