

**COMPARAISONS METHODOLOGIQUES DES ENQUETES PAR
SONDAGE EN STATISTIQUE D'ENTREPRISE :
POLOGNE, RUSSIE, MOLDAVIE**

GRUN-REHOMME MICHEL ¹

RÉSUMÉ

L'INSEE coopère depuis le début des années 90 avec de nombreux pays d'Europe Centrale et de l'Est pour mettre en place une statistique d'entreprise conforme aux standards européens. Ces pays ont du passer d'un système statistique de recueil exhaustif de l'information aux enquêtes par sondage.

L'objectif de cette communication est de présenter, à travers trois pays représentatifs de la diversité des situations économiques, politiques et administratives rencontrées sur le terrain, les méthodologies mises en place dans les enquêtes structurales, et plus particulièrement la base de sondage, le plan de sondage et les redressements.

Trois pays vont donc servir d'exemples :

- La Pologne, nouveau membre de l'Union Européenne avec une population de 38,2 millions d'habitants et un revenu par habitants en USD de 5270 en 2003. Son système statistique d'entreprise est calqué sur le système français.
- La Fédération de Russie, grand pays confronté à des particularités économiques et à une grande hétérogénéité régionale. Une population de 143,3 millions d'habitants et un revenu par habitants de 2610.
- La Moldavie, petit pays de l'ex-URSS, rendu célèbre par Tintin, avec une population de 4,2 millions d'habitants et un revenu par habitants de 590.

¹ Grun-Réhomme Michel, Université Paris 2, 92 rue d'Assas 75006 Paris, France.
Courriel : grun@ensae.fr

1. INTRODUCTION

Pratiquement tous les pays où une coopération internationale d'ordre économique et statistique est mise en place effectuent des enquêtes auprès des entreprises. Ces enquêtes permettent d'obtenir de l'information sur l'activité économique de l'entreprise, son fonctionnement, ses financements, son compte de résultat, la répartition du chiffre d'affaires par branches d'activité, les établissements, l'emploi... Ces informations recueillies permettent alors d'élaborer des statistiques sectorielles et éventuellement régionales, et participent à la construction des indicateurs économiques du pays.

Les nombreux pays qui s'ouvrent à l'économie de marché, aux réalités économiques certes différentes, se heurtent toutefois aux mêmes difficultés méthodologiques liées aux enquêtes par sondage. En effet, il s'agit d'obtenir une "bonne" estimation de l'ensemble de la réalité économique des entreprises, à partir d'un échantillon. Ces enquêtes posent toujours des problèmes de qualité des données, de plan de sondage et de redressements.

Les objectifs principaux d'une enquête structurelle permettent de produire des statistiques sectorielles, de fournir des informations à la comptabilité nationale et de corriger le répertoire. Les comparaisons entre ces trois pays seront donc effectuées à travers l'examen de la base de sondage, du plan de sondage et des redressements.

2. BASE DE SONDAGE

Chacun de ces pays dispose d'un répertoire administratif et d'un répertoire statistique. Ce dernier sert de base de sondage.

2.1 Pologne

La coopération de l'INSEE avec le GUS (office statistique polonais) remonte à 1992 et pour ma part depuis 1996. Le champ de l'enquête structurelle auprès des entreprises (personnes physiques et personnes morales) concerne les secteurs habituels des classes B à I et K de la NACE.

Le Répertoire contient environ 43 000 grandes et moyennes entreprises dans le champ de l'enquête et environ 2 millions de petites entreprises. Le répertoire contient encore des entreprises inactives ou liquidées.

Pour les différentes enquêtes structurelles sur les entreprises, la taille de l'entreprise est définie à partir du nombre de personnes occupées (moyenne annuelle en équivalent temps plein) dans l'entreprise. Cette information est enregistrée dans le répertoire à partir du bilan comptable et est utilisée comme critère d'appartenance à l'une ou l'autre des enquêtes.

L'enquête est adaptée à la taille de l'entreprise. On distingue trois catégories de questionnaire selon la taille des entreprises : les grandes entreprises de plus de 100 personnes, les moyennes (de 10 à 6) et les petites (moins de 6). Un tiers des entreprises se trouvent dans le secteur industriel et un tiers dans le secteur du commerce. Au fil des années, on a noté une augmentation des entreprises du commerce et une diminution dans l'industrie.

3,3% des entreprises, celles employant plus de 500 salariés, réalisent 42% du chiffre d'affaires total du champ de l'enquête. A contrario, les petites entreprises ne représentent que 23% du chiffre d'affaires. Le secteur du commerce représente environ la moitié des petites entreprises du répertoire.

2.2 Russie

La coopération entre l'Insee et le Goskomstat de Moscou, sur les statistiques d'entreprise, a débuté en 1995.. A partir du répertoire administratif qui contient les nouvelles entreprises, mais donne avec du retard

les informations sur les fermetures, est géré un répertoire statistique. Chaque région dispose de son registre statistique.

Le nombre actuel de petites entreprises est de l'ordre de 830 000. L'essor des petites entreprises a été très rapide entre 1991 et 1994, surtout dans le commerce, la restauration et les services aux particuliers, mais depuis la tendance semble stable.

La définition des petites entreprises est déterminée de façon réglementaire. La dernière réglementation date de juin 1995. Les petites entreprises sont définies à partir de différents critères : taille, secteur d'activité et d'un critère d'appartenance au secteur privé. Ce dernier stipule qu'aucun des quatre principaux détenteurs du capital social de l'entreprise ne doit appartenir pour plus du quart de son capital au secteur public. Plus précisément, l'enquête concerne les petites entreprises de moins de 100 personnes dans l'industrie et la construction, moins de 60 dans le commerce de gros et moins de 30 dans le commerce de détail et d'autres secteurs.

Par ailleurs, les entreprises individuelles sans salarié ne sont pas considérées en Russie comme des entreprises, mais des activités indépendantes, et donc n'appartiennent pas au champ de l'enquête.

Pour la taille, il faut distinguer les emplois principaux et secondaires, la nature du contrat ainsi que les effectifs qui proviennent de créations provisoires de collectivités de travail. En effet, un contrat peut être signé entre une entreprise et une personne pour un travail bien déterminé. Les exonérations fiscales sont différentes selon l'effectif de cette collectivité de travail. On constate un écart de 20% entre les enquêtes et le répertoire sur la variable de taille. On peut estimer en moyenne à environ 25 % le nombre d'entreprises du répertoire qui ne sont plus en activité.

2.3 Moldavie

La coopération entre l'Insee et le Département Statistique de Moldavie sur les statistiques d'entreprise a commencé en avril 1996. Le champ de l'enquête structurelle auprès des entreprises (personnes morales et depuis 2004 les personnes physiques) concerne tous les secteurs d'activité (y compris l'agriculture, les activités financières et l'administration, à l'exception des services domestiques et des activités extra-territoriales). Ce champ est limité aux entreprises qui ont communiqué leur bilan comptable aux services fiscaux (déclaration obligatoire).

La taille de l'entreprise est définie à partir du nombre de salariés dans l'entreprise. Cette information est enregistrée dans le répertoire à partir du bilan comptable. On distingue trois catégories d'entreprises : les grandes (effectif salarié supérieur à 75), les petites (de 74 à 10) et les micro entreprises (effectif salarié inférieur à 10). Cette définition a été modifiée en 2003. Le questionnaire est simplifié pour les micro entreprises et 90% de ces micro entreprises sont mono actives.

Dans le répertoire statistique, on distingue les entreprises actives qui ont présenté un bilan comptable (environ 16 000 entreprises dont 12 540 de moins de 20 salariés en 2003), les entreprises en sommeil qui ont présenté un bilan comptable (4 200 entreprises) et les nouvelles entreprises sans bilan comptable (1 700 entreprises). A ces personnes morales qui appartiennent au champ de l'enquête, il faut ajouter dans le répertoire environ 10 000 entreprises sans bilan comptable et supposées inactives. Les personnes physiques sont de l'ordre de 18 000, dont 10 000 environ sont supposées actives.

3. PLAN DE SONDAGE

Dans chacun de ces pays le plan de sondage est stratifié.

3.1 Pologne

L'enquête est exhaustive pour les grandes et les moyennes entreprises.

Le plan de sondage pour les petites entreprises est stratifié et la constitution des strates tient compte de 3 facteurs : les résultats des enquêtes précédentes, les normes européennes et les exigences particulières internes ou externes au GUS. Les strates croisent le secteur d'activité (76 modalités) et la région. Un

facteur de taille peut intervenir dans certains secteurs. Le plan de sondage contient aussi une strate pour les entreprises atypiques et des strates pour les entreprises nouvelles.

Le taux de sondage global pour les petites entreprises est de l'ordre de 5%, soit environ 100 000 entreprises.

L'allocation dans les strates suit la méthode de Neyman. Plus précisément : Au début, on choisit une taille d'échantillon inférieure à la taille définitive. On effectue une allocation de Neyman en utilisant comme variable le chiffre d'affaires. On obtient ainsi un premier poids dans chaque strate. A l'issue de cette optimisation, des aménagements sont effectués : poids de sondage décroissant avec la tranche d'effectif, recherche d'un nombre minimum d'unités dans les strates, augmentation du taux de sondage dans les strates où le coefficient de variation reste important, prise en compte de l'importance économique de la strate et examen des taux de non réponse dans l'enquête précédente. On obtient ainsi un bon compromis global-local. Le logiciel SAS est utilisé dans cette démarche. Le tirage de l'échantillon est alors effectué.

3.2 Russie

Les grandes entreprises sont enquêtées exhaustivement

Pour chacune des 89 régions, une stratification est effectuée à partir de trois variables : le secteur d'activité (37 modalités), la forme de propriété (privée ou mixte) et la taille de l'entreprise, basée sur les effectifs salariés. Dans chaque région, les tranches de taille varient selon le secteur d'activité. Cette stratification est pertinente dans presque toutes les régions, sauf les zones off shore. Comme en Pologne, l'allocation est optimale avec des aménagements.

Le Goskomstat a élaboré son propre logiciel pour les enquêtes. Il permet de construire le plan de sondage, d'effectuer un certain nombre de contrôles de cohérence interne et de contrôles agrégés, de traiter les non réponses (différentes méthodes sont proposées), d'analyser les résultats et de calculer la précision des estimations. Ce logiciel est exploité sous DOS ou Windows et se trouve à la disposition des régions. Le logiciel SPSS est également utilisé.

Afin d'éviter une surcharge de travail pour les petites entreprises enquêtées par sondage à l'aide des différents questionnaires (production, investissements, effectifs), une stratégie de coordination des enquêtes (algorithme de Jales) et de rotation des échantillons est mise en place.

3.3 Moldavie

L'enquête est exhaustive pour toutes les entreprises de plus de 20 salariés, ainsi que dans les strates où le nombre d'entreprises est inférieur à 20 et dans la strate des petites entreprises et micro entreprises atypiques. Sinon le sondage est stratifié selon la taille (5 modalités), le secteur d'activité et la région (2 modalités : Chisinau et le reste du pays). Le taux de sondage global sur les personnes morales est de 40%. L'allocation est optimale dans les strates avec des aménagements.

4. REDRESSEMENTS

Les contrôles sont plus ou moins importants selon les pays. Les non réponses totales sont traitées par repondération avec une post-stratification pour la Pologne et la Russie. Les entreprises atypiques sont recherchées.

4.1 Pologne

Ces résultats concernent les années 1998-2000. La coopération avec l'INSEE sur l'enquête structurelle est maintenant terminée.

Le taux de réponse sur les grandes et moyennes entreprises est de l'ordre de 87% (de l'ordre de 76% pour les moyennes entreprises et de 98,5% pour les grandes).

On obtient un taux de réponse pour les petites entreprises de l'ordre de 55%, avec 27% de non réponse, dont 7% de refus et 18% d'entreprises hors champ. Compte tenu des hors champ, le taux de non réponse est de 33% dans l'échantillon. Ces résultats sont variables selon les secteurs et les régions. Pour les "zéro"

salariés, les résultats ne sont pas de bonne qualité. Ce mauvais score général peut s'expliquer aussi ici par un vieillissement du répertoire, la fermeture d'offices statistiques régionaux liés à la réforme du découpage administratif et la lourdeur du questionnaire.

Les entreprises atypiques sont détectées à l'aide des graphiques Boxplot de SAS/Insight en s'appuyant sur une connaissance sectorielle.

Des calculs de précision sont effectués sur le chiffre d'affaires avec le coefficient de variation.

4.2 Russie

A partir d'un codage, au niveau régional, sur les raisons de la non réponse, il est possible d'effectuer une distinction entre les entreprises hors champ et les entreprises non répondantes du champ de l'enquête. Ce codage permet de regrouper les situations en 3 groupes : les entreprises hors champ, les entreprises actives non répondantes et les non répondantes sur lesquelles la région ne possède pas d'information.

Le logiciel permet de choisir entre différentes méthodes classiques de redressement, mais ce choix est effectué au niveau fédéral pour un enquête donnée et imposé aux régions. Le redressement sera alors automatique. Au niveau régional, il est seulement possible d'intervenir sur le codage des raisons de la non réponse et sur le choix ou non d'appliquer le redressement automatique.

La détection des entreprises atypiques utilise les quartiles. Les quartiles sont estimés à partir de la fonction de répartition du chiffre d'affaires.

Les entreprises atypiques se trouvent à l'extérieur d'un intervalle de la forme $[Q_2, Q_2 + c_u d_u]$, avec $d_u = Q_3 - Q_2$ ou $d_u = \text{Max}(Q_3 - Q_2, a Q_2)$, a est un paramètre, par exemple égal à 0,05. D'autre part, c_u est à déterminer de façon empirique. La variable retenue est le chiffre d'affaires.

4.3 Moldavie

Le taux de non réponse est faible (champ de l'enquête réduit aux personnes morales avec bilan comptable) et les changements d'activité principale de l'ordre de 10%. Les non réponses totales sont traitées par repondération sans post stratification. Le coefficient de variation est calculé sur le chiffre d'affaires dans chacune des strates.

5. CONCLUSIONS

En Pologne, il resterait à effectuer quelques contrôles agrégés supplémentaires pour repérer au mieux les entreprises atypiques, afin de les traiter en liaison avec les départements sectoriels concernés et de continuer à suivre la démographie des entreprises.

Le Goskomstat de Moscou se heurte à deux difficultés principales : la diversité et l'hétérogénéité des provinces et d'autre part, aux rouages administratifs. Les questions qui vont se poser à moyen terme sont les suivantes : Etude des estimations sur petits domaines (estimations régionales sur quelques indicateurs), prise en compte des entreprises individuelles dans l'enquête structurelle et mesure de la part du secteur informel.

Les faibles effectifs du Département de Statistique de Moldavie et le manque de moyens financiers sont les deux principales difficultés rencontrées pour construire plus vite un véritable système statistique d'entreprise.

Si une méthodologie statistique rigoureuse, ordonnée dans le temps, est nécessaire pour obtenir une bonne qualité des données d'enquêtes, il ne faut pas oublier les réalités économiques des pays et les contraintes que doivent respecter les statisticiens d'enquête.