

# ***Bilan qualité de l'enquête transport routier de marchandises 2010***

## Fiche descriptive de l'enquête

|                               |  |
|-------------------------------|--|
| <b>« carte d'identité »</b>   |  |
| nom                           | Enquête Transport Routier de Marchandises 2010                           |
| n° visa                       | 2010 H030 EN   |
| année : 2010                  | périodicité annuelle : permanente  |
| service enquêteur responsable | SOeS/OST/BSTV (bureau de la statistique des transports et des véhicules) |
| mode de collecte              | collecte postale et électronique   |

|   |                                |
|---|--------------------------------|
| <b>caractéristiques</b>   |                                |
| Définition du champ :   |                                |
| <p>les véhicules routiers à moteur immatriculés en France métropolitaine (hors DOM) de moins de 15 ans d'âge, des genres suivants :</p> <ul style="list-style-type: none"> <li>- les camions de poids total autorisé en charge (PTAC) supérieur à 3,5 tonnes</li> <li>- les tracteurs routiers de plus de 5 tonnes de poids total roulant autorisé (PTRA).</li> </ul> <p>Sont exclus du champ de l'enquête : les véhicules non immatriculés dans les séries normales (véhicules militaires, des domaines, etc.), les véhicules agricoles, les tracteurs agricoles ou forestiers, les véhicules utilitaires à usages très spéciaux (matériel de travaux publics et d'incendie, bennes à ordures, etc.), les véhicules appartenant à des forains ou à des nomades (cirques, manèges, etc.), à des auto-écoles, les véhicules appartenant à des entreprises de transport de fonds, les véhicules revendus à l'étranger, les véhicules équipés pour un autre usage que le transport de marchandises (camions-ateliers, camions-écoles, etc.) et les véhicules utilisés exclusivement à l'intérieur de chantiers ou d'usines.</p> <p>Les véhicules à moteur étrangers ne sont pas couverts mais une remorque étrangère attelée à un camion français (ou une semi-remorque étrangère couplée à un tracteur français) sera prise en compte dans les évaluations de transport.</p> <p>Sont ainsi exclus du champ les véhicules appartenant à des entreprises dont l'activité principale (<i>en naf Rev. 2</i>) relève des activités financières (Codes APE 6411Z, 6419Z, 6430Z, 6491Z, 6492Z, 6499Z), de l'enseignement de la conduite ( : 8532Z, 8553Z) ainsi que des forains, cirques et spectacle vivant (9001Z, 9002Z, 9321Z)</p> |                                |
| unité statistique <b>véhicule semaine</b> : Le propriétaire (ou l'utilisateur) est interrogé sur l'utilisation du véhicule pendant une semaine de référence   |                                |
| taille de la population   | 527 403                        |
|   | taille de l'échantillon 77 921 |
| temps de réponse médian (estimé)  | entre ½ heure et 1 heure       |
| Caractère obligatoire   | oui                            |

## Les dates-clés

|  | Prévu   | Réalisé                     |
|--|---|-----------------------------|
| réunion de la formation CNIS                   |   | le 11/10 pour d'opportunité |
| réunion du comité du label                     | Septembre 2009  |                             |
| lancement de la collecte                       | Enquête permanente  | Enquête permanente          |
| fin du travail des gestionnaires               | Mai 2011  | 07/05/2011                  |
| base de données finale                         | Mai 2011  | 16/05/2011                  |
| retour d'information aux entreprises           | Pas de retour personnalisé mais plutôt un retour de résultats globaux par l'intermédiaire du verso du questionnaire d'enquête | idem                        |
| mise à disposition du fichier                  | printemps-été 2011  | 21/05/2011                  |
| première publication sur résultats provisoires | 30 juin 2010  | 30 juin 2010                |
| première publication sur résultats définitifs  | 17 juin 2011  | 17 juin 2011                |

## La préparation de l'enquête

|  |  |  |   |
|--|--|--|---|
| existe-t-il un comité d'utilisateurs des résultats de l'enquête?   | avec utilisateurs externes <input type="checkbox"/>                  | sans utilisateur: externe <input type="checkbox"/> | non <input checked="" type="checkbox"/> |
| nombre d'entreprises ayant fait l'objet d'un test de questionnaire | 30   |  |   |
| l'échantillon a-t-il fait l'objet d'une démarche de coordination ? | oui <input checked="" type="checkbox"/> non <input type="checkbox"/> |  |   |

## L'échantillon enquêté

### Type de sondage :

Tirage à probabilités inégales effectué dans la partie du fichier central des automobiles correspondant au champ de l'enquête. Les données du fichier central des automobiles sont issues notamment des cartes grises.

### Plan de sondage :

La probabilité inégale de tirage est définie de façon à être proportionnelle à trois facteurs dérivés de : l'âge du véhicule, sa charge utile ou la charge utile estimée des semi-remorques attelées, et l'activité principale de l'entreprise propriétaire (APE) voire sa spécialisation éventuelle (transport international, transport de marchandises dangereuses). (Il s'agit de la stratification a priori.).

Le but est de tirer plus de véhicules « jeunes », de véhicules qui ont une grande capacité (charge utile), et de véhicules qui transportent habituellement de gros poids. Ces derniers réalisent plus de tonnes, de tonnes-kilomètres et de véhicules-kilomètres en un an que les autres, et ont donc une contribution plus importante dans le total des tonnes, des tonnes-km et des véhicules-km de tous les véhicules.

Les véhicules tirés sont reconduits par moitié d'une année sur l'autre, la même semaine d'enquête. Un véhicule tiré une semaine donnée ne peut être tiré à nouveau une autre semaine de la même année. Ce recouvrement partiel permet une meilleure estimation des évolutions annuelles. Il se justifie par la stabilité observée de l'utilisation d'un même véhicule à moteur d'une année sur l'autre, malgré la pluridisciplinarité théorique des tracteurs routiers.

Les coefficients de pondération sont ensuite améliorés par un calage sur marges (programme CALMAR) définies par la région et le département d'immatriculation, l'activité du propriétaire du véhicule, la semaine d'enquête, un indicateur de transport international ou d'appartenance au registre des transporteurs internationaux, et enfin, le genre<sup>1</sup> croisé avec l'âge du véhicule et, pour les camions, avec le type de carrosserie.

Chacun des véhicules échantillonnés fait l'objet d'une enquête portant sur son utilisation pendant une semaine de référence. L'unité de sondage est donc le véhicule-semaine.

Pour tenir compte des phénomènes saisonniers, l'enquête est permanente : l'envoi des questionnaires s'étale sur toute l'année.

### Sélection du fichier de tirage (du demi-échantillon renouvelé) d'un trimestre donné

On retient dans ce fichier les véhicules dont les numéros de série du type se terminent parmi les vingt-cinq valeurs attribuées au trimestre (exemple : '00' à '24' pour le premier trimestre)

- et ont la même parité que l'année d'enquête 2 ;

- ou ont l'autre parité mais ont moins d'un an (donc ne figuraient pas dans le fichier de tirage de l'année précédente).

Au ième trimestre n, le fichier de tirage sera ainsi composé des véhicules dont le n° de série du type est de la parité de n entre 25(i-1) et 25i-1 pour les véhicules de plus d'un an, simplement compris entre 25(i-1) et 25i-1 pour les véhicules neufs (mis en circulation depuis le début du trimestre précédent le ième trimestre n-1).

Le fichier est trié par région administrative (22 régions métropolitaines), puis par département, par code APE, par numéro siren, enfin par numéro de série du type.

La taille de l'échantillon trimestriel est notée  $\sqrt{V}$ . Elle peut être réactualisée chaque trimestre.

---

1 genre = tracteur routier vs camion

2 ceci permet d'être sûr de ne pas retirer des véhicules qui avaient été tirés l'année précédente

## Calcul des probabilités inégales

Appelons  $u_k$ ,  $c_k$  et  $v_k$  les trois variables auxiliaires choisies pour le tirage à probabilités inégales (pour chaque véhicule  $k$  de l'échantillon trimestriel, de taille  $v$ ) :

$u_k$  rend compte de l'âge du véhicule.  $u_k = 1$  pour les véhicules de plus de 15 ans, 2 pour les véhicules de 11 à 15 ans, 3 pour les véhicules entre 6 et 10 ans, 6 pour les véhicules de 0 à 5 ans. Cette progression est inspirée de l'échelle des kilométrages annuels mesurée par l'enquête.

$c_k$  est la charge utile du véhicule.  $c_k = 0,5 \times \text{PTAC} / 10$  tonnes pour les camions,  $(\text{PTRA} - 6) \times 0,88 / 10$  tonnes pour les tracteurs routiers.

$v_k$  marque l'intérêt spécifique pour une activité ou une spécialité.  $v_k = 1$  par défaut, 1,5 pour les véhicules dont les APE relève du secteur des transports ou dont les entreprises sont inscrites au registre des transporteurs, 2 pour l'APE déménagement (4942Z) ou dont les entreprises appartiennent aux ensembles " transport international " ou " matières dangereuses ".

La probabilité de tirage,  $\pi_k$ , est définie comme suit :

$\pi_k = \alpha * c_k * u_k * v_k$  pour chaque véhicule  $k$ ,

$$\alpha = \frac{v}{\sum_k c_k u_k v_k}$$

avec :

la somme portant sur tous les véhicules  $k$  du fichier de tirage

$\alpha$  est une constante de normalisation ; elle est définie simplement de façon à ce

$$\sum_k \pi_k = v$$

que  $\sum_k \pi_k = v$ , cette condition devant toujours être réalisée, quel que soit le mode de tirage.

- Quand un véhicule a une probabilité de tirage ( $\pi_k$  ainsi calculée) supérieure ou égale à 1, il est versé dans l'échantillon avec  $\pi_k$  tronqué à 1, et il est retiré du fichier de tirage.

- On recalcule alors une taille cible de l'échantillon :  $v' = v - \mu$ , où  $\mu$  est le nombre de véhicules qui ont été versés dans l'échantillon avec probabilité 1, et on recalcule en conséquence  $\alpha$  et  $\pi_k$ .

Plusieurs itérations peuvent avoir lieu jusqu'à ce que plus aucun véhicule restant dans le fichier de tirage n'ait une probabilité de tirage supérieure ou égale à 1.

### Taux de sondage

La proportion de véhicules tirés étant de  $77\,921 / 527\,403 \approx 14,8\%$  et chaque véhicule étant sondé une semaine sur 52 (l'année 2010 comportait 52 semaines), le taux de sondage global est égal à  $14,8\% \times 1/52 \approx 2,8\text{‰}$ .

### Prise en compte des non réponses

Les poids de tirage affectés aux véhicules de l'échantillon sont corrigés des non-réponses (ou, plus largement des questionnaires inexploitable (relatifs aux véhicules introuvables (NPAI, reventes imprécises, ...), refus de réponse, non-réponses, véhicules loués à de nombreux locataires dans la semaine d'enquête, ...).

Cette opération de re-pondération est répétée chaque semaine pour l'estimation des résultats intermédiaires utilisés en conjoncture, mêlant échantillon reconduit et tiré.

La règle consiste à diviser pour chaque semaine, chaque poids de tirage de questionnaire exploitable par la somme des poids de tirage des questionnaires exploitables et à le multiplier, par la somme des poids de tirage de tous les questionnaires. On annule le poids des questionnaires non exploitables.

Pour prendre en compte l'information sur les véhicules utilisés mais non encore saisis, la pondération des questionnaires utilisés et saisis, est divisée par la somme des poids de tirage des questionnaires utilisés et saisis et multipliée par la somme des poids de tirage des questionnaires utilisés.

---

3 car il reste  $v - \mu$  places à remplir dans l'échantillon.

|   |  |
|---|--|
| variables objectif pour la diffusion (ou pour les études)               |  |
| ▶ nombre : 6  |  |
| ▶ liste :   | tonnes transportées, tonnes-kilomètres réalisées, véhicules-kilomètres parcourus, kilométrage moyen réalisé par les véhicules, parc en service, parc utilisé   |
| sous-populations de diffusion   |  |
| ▶ nombre : 15   |  |
| ▶ liste :   |  |
| 1) variables relatives au véhicule :                                    | genre du véhicule, charge utile, PTAC, âge, carrosserie, consommation  |
| 2) variables relatives à l'entreprise :                                 | regroupement d'APE de l'utilisateur du véhicule  |
| 3) variables relatives à la nature des parcours ou de la marchandise :  | nature du transport (national/international) compte d'autrui /compte propre, lieux de chargement et de déchargement, nature de la marchandise, l'appartenance de la marchandise à une catégorie, marchandise dangereuse, type de parcours (un seul chargement, circuit, à vide, etc.) conditionnement de la marchandise, classe de distance parcourue par la marchandise |
| part de la population de la strate exhaustive dans la population totale | Sans objet   |
| distribution des taux de sondage par strate de sondage : 1er quartile : | 0,15   |
| taux médian :   | 0,26   |
| 3e quartile :   | 0,39   |

## Indicateurs liés à la production

|   |  |
|---|--|
| y a-t-il des relances ?   | Oui  |
| si oui, de quel type et combien ?   | 31 565 rappels<br>18 678 mises en demeure<br>1 872 constats de non réponse                           |
| Nombre d'unités de l'échantillon initial  | 77 921   |
| Nombre de retours   | 60 842   |
| Nombre de « hors champ » (y compris unités disparues) dans les retours  | 10 909   |
| Nombre d'entrées dans le champ  | Sans objet   |
| Nombre de questionnaires utilisables  | 49 933   |
| Taux de non-réponses selon les regroupements de strates de départ   | $22,3 \% = 1 - [49\,933 / (77\,921 - 10\,909 - 2\,744^*)]$<br>* le chiffre de 2 744 désigne les Npai |
| Pourcentage d'unités pour lesquelles la valeur « finale » est différente de la valeur brute, pour chaque variable-cible et selon les regroupements de strates | Il n'y a pas de compteur dans l'application de saisie-contrôle                                       |

Pourcentage de non réponses-partielles, pour chaque variable et selon les regroupements de strates : sondage non stratifié

| Variable   | % non réponse partielle |
|--|-------------------------|
| Consommation moyenne de carburant aux 100 km   | 64 %                    |
| Taux de remplissage (plein-pas plein)  | 24 %                    |
| Classe de marchandise dangereuse   | 22 % *                  |
| Lieu de chargement/déchargement  | < 1 %                   |
| Lieu d'embarquement et de débarquement du transport intermodal   | < 1 %                   |
| Indication de remorque ou semi-remorque  | < 1 %                   |
| Motif de non utilisation du véhicule la semaine d'enquête  | < 1 %                   |
| <i>* surestimation possible car il est parfois impossible de distinguer la non réponse du non concerné</i> |                         |

Part de l'agrégat, pour les variables-cibles, venant de l'imputation et du redressement des non-réponses totales selon les regroupements de strates

Il n'y a pas d'imputation des non réponses totales.

La part de l'agrégat pour chacune des variables-cibles venant du redressement des non réponses totales est de 24,1% pour les tonnes-kilomètres, 23,5% pour les tonnes et de 23,6% pour les véhicules-kilomètres

Liste des variables (pas forcément variables-cibles) pour lesquelles le pourcentage global de non-réponse partielle est le plus élevé :

Consommation moyenne de carburant aux 100 km

Taux de remplissage (plein-pas plein)

Classe de marchandise dangereuse

## Indicateurs de précision

Coefficient de variation pour les estimations de chaque variable cible, et pour les domaines de diffusion

| Année                | 2010 |
|----------------------|------|
| Tonnes               | 0,77 |
| Tonnes-kilomètres    | 0,48 |
| Véhicules-kilomètres | 0,47 |

## La diffusion

|   |  |  |                              |
|---|--|--|------------------------------|
| l'enquête a-t-elle fait l'objet d'une procédure d'archivage ?                           |  | oui <input checked="" type="checkbox"/>  | non <input type="checkbox"/> |
| existe-t-il une documentation pour la diffusion ?                                       |  |  |                              |
| aucune documentation <input type="checkbox"/>   |  | une documentation simple <input checked="" type="checkbox"/>   |                              |
| un dictionnaire de données complet <input type="checkbox"/>                             |  | une documentation à usage externe <input checked="" type="checkbox"/>  |                              |
| nombre de publications, par type  |  |  |                              |
| ➔ type « Infos Rapides »  |  | 4 trimestrielles et<br>1 annuelle  |                              |
| ➔ type « 4 Pages » (Insee-Première, Le 4 pages des sta<br>industrielles, ou équivalent) |  | 1 sur le TRM européen  |                              |
| ➔ type CD-Rom, « Insee-Résultats », Minitel, web  |  | Trimestriellement et annuellement<br>mise sur Internet de résultats<br>détaillés   |                              |
| ➔ articles, études  |  | Selon l'actualité du moment, des<br>analyses sont réalisées soit à la<br>demande notamment pour<br>évaluer des politiques affectant le<br>transport routier de<br>marchandises, soit diffusées<br>largement (ex. logistique,<br>émissions polluantes...)<br>Enfin, les résultats sont repris<br>chaque année dans le rapport de<br>la Commission des comptes des<br>Transports |                              |

## Moyens relatifs à l'opération

Taille de l'équipe (en équivalent temps plein)

- équipe « statistique » : 1
- gestionnaires : 10

Moyens informatiques (nombre de jours de développement, maintenance dans le cas d'une opération régulière) : 6 mois de maintenance



### Annexe :

Description détaillée de la rénovation du tirage de l'enquête TRM et la mise en place du calage sur marge sont décrits dans le document joint, qui est un article de la lettre du SSE n° 49 du 2ème trimestre 2002, rédigé par Alain GALLAIS.

## **Tirage à probabilités inégales et calage : les nouveaux principes de l'enquête « Transports routiers de marchandises »**

**Depuis le deuxième semestre 2001, l'enquête sur les transports routiers de marchandises, dite « TRM », a adopté un nouveau plan de sondage assorti d'une procédure de calage sur marges. Le Service économique et statistique du ministère de l'Équipement, des Transports, du Logement, du Tourisme et de la Mer a rénové la méthodologie de cette enquête quinquagénaire pour améliorer la précision des résultats en évolution ou au niveau régional. Le nouveau mode de tirage et le calage se basent sur l'exploitation du fichier central des automobiles.**

L'enquête française sur les transports routiers de marchandises, dite « TRM », fête ses cinquante ans en 2002. Chaque année, le Service économique et statistique (SES) du ministère de l'Équipement, des Transports, du Logement, du Tourisme et de la Mer enquête environ 80 000 véhicules à ce titre. L'utilisateur du véhicule est interrogé sur l'activité de son véhicule pendant une semaine déterminée. L'enquête TRM a beaucoup évolué depuis sa naissance. La rénovation la plus importante est certainement la plus récente, engagée sous le titre générique de « TRM 2000 » (encadré). Elle répond aux exigences d'un règlement européen de 1998 qui encadre très strictement le dispositif et en fait la pièce maîtresse de l'observation statistique européenne des transports, tant en structure qu'en conjoncture.

Pour la deuxième phase de la rénovation « TRM 2000 », le SES a voulu améliorer la précision de l'enquête TRM, sans augmenter la charge statistique des entreprises. Pour cela, il a adopté un plan de

### **Trois vagues de rénovation « TRM 2000 » pour satisfaire au règlement européen**

En France, la première enquête par sondage sur le transport routier a été réalisée en 1952. Depuis lors, l'enquête sur les transports routiers de marchandises (TRM) a connu de grands changements, de l'extension de la période d'enquête (au début une semaine dans l'année, maintenant cinquante-deux) au genre de véhicule enquêté (autrefois le véhicule porteur : camion, remorque ou semi-remorque ; à présent le véhicule à moteur et son attelage éventuel).

L'objectif est de mesurer les transports routiers de marchandises, sur le territoire national ou à l'étranger, pour le compte d'autrui et pour le compte propre, et de suivre les prix du transport pour le compte d'autrui. L'utilisateur du véhicule est interrogé sur l'activité de son véhicule pendant une semaine déterminée.

En 1998, le règlement européen n° 1172/98 a édicté les variables à collecter, leur définition, leur format et les critères de qualité. La première vague de rénovation « TRM 2000 » portait sur les variables supplémentaires à introduire et la gestion de l'enquête, depuis l'interface avec le fichier central des automobiles jusqu'à l'envoi à Eurostat. La deuxième vague de rénovation, développée dans cet article, concernait toutes les techniques de stratification a priori (tirage à probabilités inégales, reconduction d'échantillon) et a posteriori (calages sur marges). La troisième vague de rénovation, encore en cours de réalisation, vise à utiliser les techniques EDI (échange de données informatisé) ou web-EFI (échange de formulaires informatisés).

Eurostat mesure lui-même les critères de qualité de l'enquête TRM, car les Etats-membres sont tenus de fournir les données individuelles collectées.

sondage plus sophistiqué, associé à un calage sur marges.

### **Le besoin d'améliorer la précision des évolutions annuelles et des résultats régionaux**

Jusqu'à la rénovation, l'enquête TRM reposait sur un tirage d'échantillon stratifié assez simple. Chaque trimestre, les échantillons de véhicules étaient tirés dans un quart différent du fichier central des automobiles, puis répartis sur treize semaines. Des échantillons mensuels complétaient le dispositif, pour prendre en compte les véhicules neufs. Les strates de tirage croisaient deux ou trois critères pour chaque genre de véhicules : la classe d'âge et la classe d'activité pour les tracteurs routiers ; la classe

d'âge, la classe de carrosserie et la classe de charge utile pour les camions. Le fichier était préalablement trié par région et département afin de favoriser une répartition géographique équilibrée. Les véhicules non répondants étaient considérés comme ayant le même profil que les véhicules répondants de la strate de tirage. Le poids de chaque véhicule s'obtenait donc simplement par le rapport entre la population de la strate pour le mois et le nombre de véhicules répondants correspondants. Les tirages d'échantillons étaient effectués par l'Association auxiliaire de l'automobile, maître d'œuvre du SES pour le fichier central des automobiles.

La précision des résultats nationaux annuels, était considérée comme satisfaisante pour les analyses structurelles, et très correcte en comparaison de la moyenne européenne. En revanche, certains utilisateurs critiquaient la précision des résultats en évolution ou au niveau régional. En effet, aucune coordination des échantillons n'était assurée entre deux années successives. A l'échelon régional, les régions étaient impliquées dans le tri préalable, mais ne participaient ni à la définition des strates, ni au traitement des non réponses. Or, la création des observatoires régionaux des transports et la mise en place d'un accès Intranet à la base de données SITRAM (Système d'information sur les transports de marchandises) ont accru l'exigence d'une bonne précision des résultats régionaux.

#### **Le fichier central des automobiles**

Le fichier central des automobiles est constitué à partir des informations des certificats d'immatriculation (cartes grises) des véhicules, transmises par le ministère de l'intérieur. Ce fichier a été créé en 1950. Il était géré initialement par l'INSEE. Depuis 1973, le Service économique et statistique (SES) en assure la maîtrise d'ouvrage.

#### **Un nouveau mode de tirage, basé sur l'exploitation directe du fichier central des automobiles**

Dans la deuxième phase de la rénovation « TRM 2000 », le Service économique et statistique (SES) a cherché à améliorer la précision des résultats sans augmenter la charge statistique des entreprises. Le SES a sollicité l'Insee, en la personne de Christian Hesse à la direction des statistiques d'entreprises, pour l'aider à mieux prendre en compte les dimensions régionale et temporelle dans le plan de sondage de l'enquête TRM.

Le nouveau plan de sondage est entré en vigueur au deuxième semestre 2001. Il marque l'aboutissement d'un projet ambitieux, qui a nécessité au préalable que le service statistique sache exploiter en interne une copie du fichier central des automobiles.

Dans le nouveau plan de sondage, la moitié de l'échantillon de l'année n-1 est reconduite en n, pour chaque semaine d'observation. Ce recouvrement partiel permet une meilleure estimation des évolutions annuelles. Il se justifie par la stabilité observée de l'utilisation d'un même véhicule à moteur d'une année sur l'autre, malgré la pluridisciplinarité théorique des tracteurs routiers.

Cette méthode améliore fortement la précision des évolutions en glissement annuel, sans que le gain en soit encore exactement mesuré. En revanche, la gestion des questionnaires est alourdie. Il faut distinguer les questionnaires reconduits des questionnaires nouvellement tirés et des questionnaires routés, en s'épargnant d'effectuer des relances pour des véhicules détruits, vendus, hors champ ou introuvables tels qu'observés en n-1.

Dans la nouvelle base de production de l'enquête TRM, les différents poids (de tirage, après correction des non réponses, après calage) sont affectés au véhicule enquêté pour sa semaine d'observation. Tous les véhicules du demi-échantillon n-1 sont reconduits avec leur poids de tirage n-1.

A l'occasion de la refonte du plan de sondage, le champ de l'enquête a été élargi aux véhicules de plus de trois tonnes et demie de poids total autorisé en charge (trois tonnes de charge utile précédemment).

En revanche, il a fallu renoncer à satisfaire une demande récurrente des entreprises enquêtées, grandes et moyennes : procéder à un tirage à deux niveaux, les entreprises puis les véhicules, afin de n'interroger chaque entreprise que sur une semaine donnée, pour plusieurs véhicules. La perte de précision aurait été trop importante en raison des spécialisations sur les origines-destinations et les produits transportés de beaucoup d'entreprises, y compris les grandes et les moyennes.

#### **Un tirage combinant l'âge du véhicule, sa charge utile et l'activité de l'entreprise**

Le tirage à probabilités inégales intervient sur la moitié renouvelée de l'échantillon. Celui-ci est tiré dans le quart affecté au trimestre du fichier central des automobiles, soit sur un huitième du fichier total (annexe). La probabilité inégale de tirage est proportionnelle au produit de trois facteurs dérivés de :

l'âge du véhicule, sa charge utile ou la charge utile estimée de la semi-remorque attelée, l'activité NAF de l'entreprise propriétaire voire sa spécialisation éventuelle (transport international, transport de marchandises dangereuses).

Certaines de ces variables sont utilisées avec plus de modalités que dans l'ancien mode de tirage ; d'autres sont continues, comme la charge utile, qui est prise en compte sans transformation. Pour les tracteurs routiers, il est calculé une pseudo-charge utile. Ceci permet de constituer un seul fichier de tirage regroupant les tracteurs routiers et les camions. Ce fichier est préalablement trié par régions, départements et entreprises.

L'échantillon nouvellement tiré se répartit sur les treize semaines du trimestre en complément de l'échantillon reconduit, de façon à équilibrer le nombre de questionnaires par semaine d'enquête.

#### **Variables explicatives, variables de diffusion**

Les principales variables observées sont le nombre de tonnes transportées, la nature de la marchandise transportée, le conditionnement, les prix du transport, le kilométrage du véhicule en charge et à vide, les trajets parcourus (type et origine-destination). Les chargements et les parcours sont mesurés selon trois unités : les tonnes-kilomètres, les tonnes et les véhicules-kilomètres.

La diffusion standard met en relief la nature du transport (compte d'autrui, compte propre) ; les classes de distance ; les marchandises transportées ; les origines/destinations par régions (voire départements) ; enfin, quelques activités NAF.

Les principales variables explicatives prises en compte dans le tirage ou le calage sur marges sont l'activité NAF de l'entreprise et sa spécialisation (transport international, transport de marchandises dangereuses), le type du véhicule, son âge, son département d'immatriculation. Deux variables ne concernent que les camions : la classe de charge utile et la carrosserie.

#### **Un calage sur quatre marges simultanées**

Dans l'enquête TRM rénovée, l'échantillonnage se complète d'une procédure de calage sur marges. En utilisant des informations auxiliaires issue du fichier central des automobiles, cette méthode permet entre autres de rectifier le décalage entre la période de référence de l'enquête et la période de référence des populations de l'échantillon.

En effet, la date de référence du quartier du fichier central des automobiles servant au tirage trimestriel retarde d'au moins un trimestre par rapport à la période de référence de l'enquête. Ainsi, la moitié de l'échantillon de l'enquête de la semaine n°27 du mois de juillet a été tirée dans une population datée du 1er avril. Le SES dispose d'une copie trimestrielle intégrale du fichier central des automobiles, mais le service peut également interroger ce fichier à distance pour récupérer les résultats mensuels agrégés. La population de l'échantillon trimestriel est calée sur des marges constituées par ces résultats agrégés, pour le mois correspondant à la semaine d'enquête. Les deux sous-populations, donnant l'une la moitié reconduite de l'échantillon, l'autre la partie renouvelée, ne sont pas traitées de façon distincte, bien qu'elles soient décalées d'un an. Les véhicules reconduits et tirés sont donc calés ensemble. Le calage est effectué sur quatre marges simultanément.

La première marge de calage porte sur les régions, camions et tracteurs routiers mêlés.

La deuxième marge de calage porte sur l'activité du propriétaire en six modalités (602L, 602M, 602N, 602P, autres dans [60-63], hors transport), camions et tracteurs routiers mêlés.

La troisième marge porte sur les genres de véhicules (tracteurs routiers vs camions), les carrosseries pour les camions, et les tranches d'âge en quatre modalités, dont la tranche « moins de trois ans » qui correspond à un usage intensif des véhicules. En effet, les véhicules sont utilisés de façon relativement similaire et constante durant leurs trois premières années ; les véhicules de moins de trois ans peuvent ainsi être assimilés à des véhicules neufs. En revanche, l'usage d'un véhicule décroît très fortement après trois ans d'âge. A ce stade, les véhicules sont souvent déclassés et revendus. Les contrats de crédit-bail ou de location financière, notamment, tiennent souvent compte de ce seuil de trois ans.

Enfin, la quatrième marge porte sur les semaines d'enquête, populations des camions et tracteurs routiers mêlées, afin de prendre en compte la forte saisonnalité qu'induisent les semaines de vacances ou les jours fériés.

Le calage est réalisé à l'aide de la macro CALMAR de l'Insee, ici adaptée à SAS sous Unix. La méthode « raking ratio » a été choisie pour répondre à une contrainte spécifique : il fallait que les poids restent tous positifs à l'issue du processus de calage afin d'assurer la cohérence des résultats, y compris pour les croisements fins par origine-destination. La méthode « logit » aboutissait à des

résultats agrégés très proches du raking ratio ; elle n'a pas été retenue car elle aurait imposé de régler les bornes inférieure et supérieure des rapports entre les pondérations ex ante et ex post calage. Or, cette opération de calage sur marges est entièrement automatisée. Elle a lieu tous les week-ends, préalablement à la diffusion hebdomadaire des données de conjoncture.

La procédure converge au bout de quatre ou cinq itérations, dès qu'un nombre suffisant de questionnaires est rentré pour le mois d'enquête, ce qui est généralement accompli un mois après. Cette méthode a permis de supprimer les tirages d'échantillons mensuels auparavant réalisés.

#### ALAIN GALLAIS

Alain Gallais est responsable du bureau de la statistique des transports du Service économique et statistique (SES), au sein de la direction des affaires économiques et internationales du ministère de l'Équipement, des Transports, du Logement, du Tourisme et de la Mer.

**Non-retours, non-réponses :**

Pour les enquêtes réalisées par courrier (ce qui est le cas de la plupart des enquêtes auprès des entreprises), il existe des « non-retours » qui peuvent comprendre différents types d'unités : des non-répondantes (c'est-à-dire des unités dans le champ de l'enquête et qui refusent de répondre), mais également des « hors champ » (au sens large, cette notion recouvrant les unités cessées). Parmi les retours, peuvent également se trouver des « hors champ ».

Dans certaines enquêtes existent également des « entrées dans le champ » : c'est le cas des EAE, avec les transferts d'unités entre services enquêteurs (une entreprise pouvant avoir changé d'activité entre celle connue dans la base de sondage et celle constatée à l'enquête).

Parmi les questionnaires reçus, certains s'avèrent, in fine, « non utilisables » car jugés d'une qualité insuffisante pour être exploités.

Le **nombre de questionnaires utilisables** vaut donc :

*Nombre de questionnaires utilisables = « retours » + « entrées de champ » - « hors champ » - « non utilisables »*

Pour calculer le **taux de réponse** (totale) pour une strate, on rapporte le nombre d'unités utilisables en provenance de la strate (donc, on ne tient pas compte des « entrées de champ ») au nombre d'unités éligibles (calculé comme le nombre d'unités de départ de l'échantillon moins le nombre d'unités constatées comme « hors champ ») :

*Pour une strate, taux de réponse = ( « retours » - « hors champ » - « non utilisables ») / (nombre d'unités de l'échantillon - « hors champ »)*

Le **taux de non-réponse partielle** peut être calculé sur les unités « utilisables » ou sur l'ensemble des questionnaires éligibles. Il est plus facile de produire le calcul en se limitant aux unités « utilisables », mais quel que soit le choix opéré, il est demandé d'indiquer sur quelle population ce taux est calculé (en effet, un taux de non-réponse partielle calculé sur les seules unités « utilisables » donnera une image trop optimiste de la qualité de la variable étudiée).

Le **coefficient de variation** relatif à l'estimation d'une statistique est égal à l'écart-type de l'estimateur de cette statistique rapporté à la valeur estimée. On explicitera les différentes composantes de l'erreur (échantillonnage, non-réponse totale, non-réponse partielle) prises en compte dans le calcul de la précision de l'estimateur.