Techniques d'enquête, juin 1978
Vol. 4, nº 1
Statistique Canada, nº 12-001-X au catalogue

Un système de plan d'enquête pour la mesure des flux de marchandises par camion au Pérou[1]

A. Satin et R. Ryan[1]

Résumé

Le présent article décrit un plan d'enquête qui a été élaboré pour mesurer le flux de marchandises transportées par camion au Pérou. L'article considère les traits conceptuels et opérationnels du plan d'enquête, et en décrit les éléments et les techniques d'exécution dans le contexte d'un projet pilote. Enfin, on démontre comment on pourrait utiliser les résultats de ce projet pilote pour élaborer et exécuter une enquête nationale à grande échelle.

Mots-clés : Conception d'enquêtes; flux de commodités; enquête pilote.

1 Introduction

Le présent article documente les éléments d'un système de plan d'enquête proposé qui permettra la collecte de données (sur les produits, l'origine et la destination, c'est-à-dire les P-O-D) sur les mouvements de marchandises par camion au Pérou. Le plan de base a été établi au cours d'une visite de trois semaines au bureau des statistiques sectorielles du ministère des Transports et des Communications à Lima, au Pérou. Il a été décidé que le plan serait testé dans une étude pilote au cours de l'été 1978. En cas de succès, le concept servirait de base potentielle pour la réalisation d'une enquête nationale continue et à grande échelle sur les routes. Les techniques et procédures peuvent également être adaptées à des études sur les P-O-D similaires menées au niveau régional. Les documents de référence utilisés pour le développement du système sont inclus dans les articles [1], [2], [3], [4] et [5].

2 Sources des données et exigences

En ce qui concerne l'élaboration d'un modèle de collecte de renseignements sur les P-O-D, trois sources de données potentielles ont été examinées par les auteurs :

---

1. Ce projet a été commandé par la Commission canadienne des transports dans le cadre d'un programme global d'assistance technique de l'Agence canadienne de développement international dans le domaine des transports.
2. A. Satin, Division du développement des enquêtes auprès des ménages et R. Ryan, Division de la coordination des enquêtes spéciales, Statistique Canada.

(a) le consignateur — le consignataire (l'expéditeur);

(b) la société de transport;

(c) le conducteur.

Selon la source de données (a), il faudrait localiser et obtenir les données de l'expéditeur. De nombreux expéditeurs n'ont pas d'adresse connue, en particulier les agriculteurs de la Sierra, et beaucoup vendent leurs produits (même dans les champs avant la récolte) à un intermédiaire-agent-transporteur; l'intermédiaire transporte et commercialise les marchandises et le producteur initial peut ne pas savoir quelle quantité il devait expédier ni sa destination. Pour ces raisons, la source (a) est considérée comme une source peu probable de renseignements sur les P-O-D fiables.

La société de transport (source [b]) est également considérée comme une source peu probable. Une forte proportion des transporteurs (plus de 50 % et peut-être jusqu'à 85 %) sont des exploitants de véhicules individuels. L'emplacement de bon nombre de ces petits exploitants est inconnu, de sorte qu'ils ne seraient pas disponibles pour recevoir le courrier ou pour avoir des contacts personnels. En outre, on estime que les registres tenus par ces personnes sont médiocres (c'est-à-dire pratiquement inexistants), de sorte que leurs mouvements et leur cargaison au cours de n'importe quelle période, sauf celle qui les précède immédiatement, seraient incertains.

La source de données restante (source [c]), à savoir le conducteur du camion, est considérée pour l'instant comme la meilleure source de données. Il est entendu qu'ils constituent un groupe peu coopératif dans son ensemble, car ses membres considèrent la police et le gouvernement comme une intrusion indésirable dans leurs affaires privées. On les soupçonne de faire des pieds et des mains pour éviter les questions nécessaires à la collecte des données d'une enquête.

L'avantage du conducteur du camion comme source de données est que certaines données sont automatiquement disponibles, à savoir la marchandise, l'origine et la destination et la description du véhicule. Cette approche présente l'avantage supplémentaire d'avoir été utilisée auparavant, de présenter des risques connus et, dans une certaine mesure, d'être connue des conducteurs.

L'une des « solutions » souvent évoquées pour résoudre les problèmes de collecte de données est la « Planica Unica » ou lettre de transport (obligatoire) utilisée pour tous les envois, des copies supplémentaires étant automatiquement fournies aux fins de contrôle et de statistiques. Elle est apparemment utilisée avec succès à l'heure actuelle en Colombie. La planification occasionnelle de son utilisation au Pérou remonte au moins à 1970. Dans ce type de système, l'expéditeur et le transporteur fournissent toutes les données nécessaires, avant le mouvement des marchandises, et des copies des documents sont recueillies à différents points de contrôle au cours d'un voyage.

L'introduction d'un tel système de documents est susceptible d'être un processus complexe, long et coûteux. On estime qu'il a fallu environ quatre ans avant que les statistiques ne commencent à émaner du modèle colombien.

Un tel système présente des avantages évidents, mais compte tenu de la longueur des délais et de l'absence d'accord sur ce point au Pérou dans le passé, il conviendrait mieux, pour le moment, de chercher une autre solution.

Enfin, l'utilisation de modèles peut être considérée comme ayant le potentiel de satisfaire les besoins en matière de données. L'inconvénient fondamental est le manque de statistiques fiables (consommation de produits, production de produits) pour rendre une telle approche réalisable à l'heure actuelle.

En résumé, il a été conclu que, pour l'instant, la seule autre solution viable pour recueillir des statistiques sur les produits, l'origine et la destination (P-O-D) est une enquête routière, dont les précisions sont présentées dans les sections suivantes du présent rapport.

En ce qui concerne les exigences en matière de données, les renseignements les plus fondamentaux requis sur une base régulière et susceptibles d'être fournis par un conducteur de camion sont les suivantes :

(1) Marchandise(s) : Le système de codification proposé est la classification européenne. Un autre système envisagé est la classification

		des marchandises de l'Organisation des Nations Unies.
(2)	Origine(s) destination(s) :	Un nom de lieu, associé à une zone précise du pays, à coder en dernier ressort au niveau provincial (régional).

Seuls les flux majeurs seront mesurés, car la mesure des flux mineurs nécessiterait de placer des stations le long de tous les segments de route, ce qui n'est pas pratique.

(3)	Quantité :	En tonnes métriques, en mètres cubes ou en unités (lorsque d'autres mesures ne sont pas disponibles) à convertir en poids selon le cas.
(4)	Véhicule :	Plaque d'immatriculation, nombre d'essieux, type de carrosserie.

L'unité de base de l'étude est un déplacement de marchandises. Ainsi, pour chaque chargement ou déchargement de marchandises, un ensemble distinct de données sur l'origine et la destination des marchandises serait préparé. Le trajet du véhicule serait défini par le point de premier chargement de la cargaison jusqu'au point de déchargement final du véhicule, pour les marchandises présentes sur le camion au point d'enquête. Le fret recueilli et déchargé avant l'interview ne doit pas être pris en compte, pas plus que le fret qui doit être recueilli et livré plus tard au cours du trajet du véhicule.

3 L'étude pilote

Lors de la conception du plan d'échantillonnage pour l'enquête pilote, deux domaines majeurs ont été pris en compte pour déterminer (1) la couverture géographique ou spatiale de l'enquête pilote, et (2) la couverture temporelle associée à la méthodologie d'échantillonnage. Il a été déterminé que l'enquête pilote et l'extension pour l'avenir comprendraient la mesure du trafic routier de marchandises pour les principaux flux dans le pays. La mesure de tous les flux ne serait pas faisable ou pratique sur le plan financier ou opérationnel.

Pour ce qui est de la couverture spatiale, il était important que les emplacements choisis pour l'enquête pilote et les points d'enquête connexes (stations de contrôle) simulent autant que possible les conditions qui pourraient exister dans le reste du pays. De cette manière, toutes les méthodes associées à l'enquête pilote ont pu être testées de façon réaliste et évaluées en fonction des possibilités d'extension à une enquête nationale. Les critères utilisés pour sélectionner la zone pour l'enquête pilote ont été déterminés comme suit :

(1) un volume élevé de trafic;

(2) une diversité de la gamme de produits;

(3) un test des méthodes de collecte de données dans des conditions climatiques extrêmes;

(4) une zone où des sources de données externes peuvent être utilisées pour la validation.

Les responsables péruviens ont énuméré trois zones à prendre en considération, à savoir : 1) la région d'Ancash au nord; 2) la route Lima-La Rolla dans la zone centrale; 3) la région de Pisco au sud.

Les trois zones ont été discutées à la lumière des critères de sélection et il a été déterminé que la région d'Ancash près de Chimbote répondrait le mieux aux critères; cette région a donc été choisie pour l'enquête pilote qui se déroulera pendant un trimestre complet. Un objet secondaire de l'enquête pilote était de fournir des données pour la région, dans la mesure du possible. Cela dépend de la mesure dans laquelle le plan d'échantillonnage, les procédures sur le terrain, etc. doivent être modifiés et améliorés au cours de l'enquête pilote.

Cinq points de contrôle ont été établis pour mesurer les principaux flux dans la région.

Avant de traiter du plan de base, il convient de mentionner que les contraintes suivantes ont dû être intégrées dans les procédures de plan et d'estimation. Le personnel sur le terrain (brigade), composé de 4 ou 5 personnes dans chaque station de contrôle, pouvait travailler au maximum 12 heures par jour, 20 jours par mois, en ayant une période de repos minimale de 6 heures entre les périodes d'interview.

Afin de ne pas perturber indûment la circulation pendant les périodes d'interview, il n'a pas été possible d'adopter un système d'échantillonnage systématique (notamment en cas de circulation intense). La procédure sur le terrain qu'il est possible de mettre en œuvre peut être décrite brièvement comme l'arrêt du premier camion, l'obtention d'une interview complète, puis l'arrêt du camion suivant qui passe à la station lorsque l'interview est terminée. Pendant une période d'interview sélectionnée, un deuxième membre de la brigade (classificateur) enregistre de manière indépendante les mouvements de tous les camions passant par la station selon le type de camion.

4 Plan d'échantillonnage (temporel)

4.1 Introduction

À chaque station de contrôle, un échantillon indépendant est construit. Le plan peut être décrit comme un échantillon probabiliste stratifié et répété en trois étapes de camions qui passent par une station de contrôle au cours d'une année civile (trimestre dans l'enquête pilote).

4.2 Stratification

La stratification dans un cadre temporel est un processus consistant à classer des unités de temps dans certains ensembles qui s'appellent des strates. Un avantage de l'échantillonnage stratifié est l'augmentation possible de l'efficacité par coût unitaire dans l'estimation des caractéristiques de la population. Dans le contexte d'une future enquête nationale, la stratification offre la flexibilité de redessiner l'échantillon d'une strate précise ou de groupes de strates, sans toucher au plan dans les autres strates.

Il est important que les variables sélectionnées pour la stratification répondent à certaines exigences si la stratification doit aboutir à l'augmentation de l'efficacité du plan. Les exigences de l'enquête sur les P-O-D étaient les suivantes :

(i) Les variables de stratification sélectionnées doivent être fortement corrélées aux mouvements de fret routier. La prise en compte des principales sources de variation des flux de

marchandises devrait aboutir à la création de strates relativement homogènes, ce qui augmente l'efficacité du plan.

(ii) Aux fins de l'enquête nationale, les variables sélectionnées doivent aboutir à des strates qui conservent leur homogénéité pendant toute la durée de l'enquête.

(iii) Le modèle doit être construit de manière à permettre la production de données annuelles (enquête nationale) ainsi que pour chacun des quatre trimestres. Cette contrainte doit également être prise en compte dans la stratification.

Les sources de variation temporelle des flux de marchandises dépendent particulièrement de l'emplacement des stations d'interview (points de contrôle). Aux fins de l'enquête pilote, les sources suivantes de variation temporelle ont été déterminées sur la base des données existantes, d'une visite sur le terrain de plusieurs stations de contrôle et des connaissances des responsables de l'O.S.E. en la matière.

(a)	moment de la journée	- matin, après-midi, soir, nuit
(b)	type de journée	- jour ouvrable, jour non ouvrable
	jour ouvrable	-(1) lundi, vendredi
		(2) mardi, mercredi, jeudi
	jour non ouvrable	-(3) samedi, dimanche, jour férié
(c)	mois	
(d)	trimestres.	

Les variables de stratification retenues étaient le trimestre (4) × les mois (3) × le type de journée (3). Le moment de la journée est incorporé dans le plan par l'échantillonnage des tranches horaires et l'estimation par les quotients.

Au total, il y a 9 strates pour chaque trimestre et donc 36 pour l'année civile.

4.3 Taille de l'échantillon

Compte tenu des contraintes financières et du manque de données historiques fiables, il a été décidé de retenir un maximum de 20 jours par mois pour les trois mois du trimestre. Les données résultant de

l'enquête pilote serviront à déterminer la taille de l'échantillon nécessaire pour l'enquête nationale. La taille de l'échantillon pour l'enquête nationale sera basée sur (1) les exigences liées à la fiabilité des données;(2) le plan d'échantillonnage et les procédures d'estimation utilisées;(3) les systèmes mis en œuvre pour les classifications géographiques et des produits;(4) les contraintes financières.

4.4 Attribution d'un échantillon

Les exemples de stratégies de répartition suivants ont été traités en détail :

(i) Répartition Neyman;

(ii) Répartition proportionnelle à X (proportionnelle au volume de trafic);

(iii) Répartition proportionnelle à la taille (nombre d'unités de temps dans les strates).

L'approche de répartition Neyman a été éliminée en raison du fait que (a) il n'existe pas de base de renseignements historiques fiables (c'est-à-dire aucune statistique de variance), (b) la répartition pourrait varier considérablement en fonction du produit mesuré et (c) les schémas de flux pourraient changer au fil du temps, ce qui aurait des répercussions négatives sur l'efficacité du plan.

La répartition proportionnelle à X pourrait être envisagée lorsque des renseignements fiables sur les flux de trafic seront disponibles. Les renseignements sur les flux de trafic n'ont pas été jugés suffisamment fiables pour que cette approche soit utilisée dans le cadre de l'enquête pilote. De plus, bien que cette approche puisse améliorer la fiabilité des estimations de nombreux produits, la mesure des mouvements de produits qui ont une corrélation négative avec le flux de trafic total serait touchée de manière négative. En outre, il serait préférable de disposer des flux de trafic des camions plutôt que du trafic total (camions et autres véhicules) si cette information était disponible. Là encore, si le schéma de flux était instable pendant la durée de l'enquête, le plan devrait être mis à jour ou éventuellement être modifié pour tenir compte de ces changements.

La stratégie choisie pour l'enquête pilote était proportionnelle à la taille (c'est-à-dire proportionnelle au nombre de jours dans chaque strate). Cette approche est prudente et s'appuie le moins possible sur des renseignements historiques fiables. Lorsque l'enquête pilote sera terminée, une décision sera prise quant à la possibilité de choisir un compromis entre cette répartition et la répartition proportionnelle à X pour l'enquête nationale, compte tenu des considérations exposées ci-dessus.

4.5 Étapes de l'échantillonnage et méthodes de sélection des échantillons

Après avoir déterminé la taille des échantillons (jours) et la stratégie de répartition, on procède à la première étape de sélection des échantillons. Dans chaque strate, un échantillon indépendant de jours (20 par mois) est sélectionné systématiquement. Pour faciliter l'estimation de la variance, compte tenu du plan d'échantillonnage complexe, l'échantillon de jours est systématiquement sélectionné (sans remise) dans chacun des deux rééchantillonnages pour chaque strate.

La deuxième étape de la sélection consiste à choisir au hasard une période de six heures (matin, après-midi, soir ou nuit) au cours de chacun des jours sélectionnés. La sélection de périodes de six heures dans l'enquête pilote était telle qu'elle permettait de satisfaire à la contrainte de terrain exigeant une période de repos minimale de six heures entre les périodes d'interview.

Après avoir sélectionné les jours et les tranches horaires de chaque jour, la dernière étape de la sélection consiste à sélectionner les camions à l'intérieur des tranches horaires sélectionnées.

4.6 Le problème de la direction

Pour traiter le problème de la direction, il a été décidé de mener des interviews dans chaque direction pendant trois heures au cours de la tranche horaire de six heures. (Dans le cadre de l'élaboration d'une enquête nationale, il est possible de s'accommoder de flux légers en faisant des interviews simultanément dans les deux directions sans perturber indûment les flux de trafic ou l'efficacité statistique).

4.7 Classification des camions

Pendant la période d'interview de six heures, un deuxième membre de la brigade de terrain enregistre le trafic de camions par type. Cela est nécessaire pour l'estimation (extension de l'échantillon à la population). Pour tenir compte de la méthodologie de l'interview, ces renseignements de classification doivent être enregistrés dans des tranches horaires suffisamment courtes pour assurer la précision statistique. Pour l'enquête pilote, la classification doit être effectuée pour chaque période d'une heure au lieu des périodes entières de six heures. Comme elle porte sur des tranches horaires suffisamment courtes, la méthode d'échantillonnage se rapproche de l'échantillonnage systématique.

L'heure de l'interview est enregistrée par l'intervieweur sur chaque questionnaire, aux fins d'estimation. Il convient aussi de mentionner que si les estimations pour un flux majeur ne sont pas suffisamment fiables, deux options sont disponibles pour ajuster le plan : (1) ajuster la répartition de l'échantillon; (2) augmenter la taille de l'échantillon (c'est-à-dire une deuxième brigade à la même station de contrôle).

5 Les formulaires

5.1 Conception des formulaires

Les deux principaux formulaires sont : (1) le document de contrôle des échantillons et (2) le calendrier des interviews.

(1) Le document de contrôle de l'échantillon (DCE) contient toute l'information nécessaire pour pondérer l'échantillon à la population. Un DCE distinct est rempli, lequel correspond à chaque période d'interview de six heures. Tous les renseignements de contrôle nécessaires sont insérés avant le début des opérations sur le terrain. Ces renseignements de contrôle concernent le lieu, la date, la tranche horaire, la strate, le rééchantillonnage, la direction et le sous-poids (voir pondération). Le DCE est ensuite transféré au membre de la brigade responsable de la classification du camion. Pour chaque heure, il inscrit le nombre total de camions par type qui passent par la station de contrôle.

(2) Le formulaire d'interview doit être rempli par le membre de la brigade responsable de l'interview avec les « cammioneros ». Les renseignements de contrôle décrits ci-dessus sont transférés sur les questionnaires en lots avant le début des opérations sur le terrain. Le questionnaire permet actuellement les envois de marchandises multiples. Le nombre total de marchandises est enregistré, et jusqu'à trois marchandises sont énumérées avec l'origine « majeure » et la destination « majeure ». L'heure de l'interview est inscrite sur le questionnaire aux fins d'estimation.

Il convient également de mentionner les points suivants :

(i) Il reste encore à mettre au point de façon définitive les types de camions et un système de classification des marchandises. Il a été indiqué que plus la classification des produits est détaillée, plus la fiabilité des estimations des produits sera faible, en général. En outre, la complexité des opérations sur le terrain, du codage, du contrôle de la qualité et des procédures de traitement des données va s'accroître.

Comme les marchandises transportées par les différents types de camions varient en fonction du type et du poids, la définition des principaux types de camions aux fins de l'estimation permettra d'améliorer l'efficacité statistique du plan.

(ii) Il est reconnu qu'il existe une certaine perte d'information, en particulier en ce qui concerne les produits agricoles, pour lesquels il est probable que certains camions prendront la même marchandise à différents points de l'itinéraire et la déposeront à plusieurs destinations le long de l'itinéraire. Pour réduire la complexité de la collecte des données sur le terrain, du codage, des contrôles de qualité et du traitement des données, une seule origine et une seule destination seront nécessaires.

(iii) Un système normalisé d'enregistrement du poids ainsi qu'un code de la classification géographique type doivent être élaborés et clairement utilisables sur le terrain par le personnel chargé du traitement des données.

5.2 Mise en lots, codage par couleur

Il est suggéré d'attribuer aux formulaires un code de couleur correspondant aux stations d'interview afin d'accroître le contrôle des documents d'enquête.

Les questionnaires doivent être attachés et joints au DCE correspondant à chaque période de six heures sélectionnée dans l'échantillon. Une fois remplis, les formulaires doivent être attachés de la même manière, puis transmis.

5.3 Opérations de bureau

Les formulaires doivent être vérifiés pour s'assurer qu'ils sont complets et exacts dès leur retour du terrain. Des instructions doivent être préparées pour que le personnel puisse vérifier les données et résoudre les divergences (voir section 10).

Après vérification et correction des données, la taille effective de l'échantillon (c'est-à-dire les questionnaires utilisables) est déterminée en fonction de l'heure de l'interview qui a été inscrite sur les questionnaires. Ce numéro est inscrit sur le DCE à côté du total correspondant au type de camion qui a été enregistré par l'agent de classification des camions sur le terrain.

Le personnel de bureau vérifie ensuite que le DCE est complet, effectue les regroupements nécessaires (lorsque l'échantillon effectif est égal à zéro) et, enfin, ajuste le sous-poids en fonction de huit catégories prédéterminées dans chaque trimestre défini par la tranche horaire (4) et la direction (2).

Après que l'on a ajusté le sous-poids, le poids final (voir pondération) est déterminé par le personnel de bureau et inséré sur les questionnaires remplis et vérifiés. Les questionnaires sont alors prêts pour la saisie des données.

6 Formation du personnel sur le terrain

Il a été reconnu que le succès du système statistique repose fondamentalement sur la fiabilité de la collecte des données. L'importance des programmes de formation centralisés ou régionaux (des

séances de formation, des manuels de l'intervieweur, des manuels de procédures) a été abordée dans le contexte de l'exercice de collecte des données. Des dispositions ont été prises dans le lancement du programme de formation pour la mise en œuvre de tous ces aspects.

7 Estimation

7.1 Introduction

Dans un échantillon probabiliste, le plan d'échantillonnage lui-même détermine les poids qui peuvent être utilisés pour produire des estimations sans biais. Chaque enregistrement peut être pondéré par l'inverse de la probabilité de sélectionner l'unité à laquelle l'enregistrement fait référence.

Le fichier à créer aux fins de totalisation contient un enregistrement par produit pour chaque camion sélectionné dans l'échantillon. Au lieu de reproduire physiquement les enregistrements de l'échantillon, un coefficient de pondération global est saisi dans chaque enregistrement. Par exemple, si l'on veut connaître le poids total du cuivre expédié par camion entre une origine et une destination données, il faut trier les enregistrements relatifs aux camions de l'échantillon qui transportent du cuivre entre l'origine et la destination et additionner le produit des poids de l'échantillon par les pondérations de cuivre inscrites sur ces enregistrements.

Puisque des renseignements objectifs concernant les tranches horaires et les directions sont disponibles pour l'univers, la fiabilité des estimations peut être améliorée en utilisant ces renseignements auxiliaires. L'estimation par le quotient est l'une des techniques les plus répandues pour utiliser des renseignements pertinents.

Lors de la comparaison des estimations dérivées de l'enquête avec celles obtenues à l'extérieur, les estimations provenant de sources extérieures sont divisées par les estimations de l'échantillon pour chaque classification et les poids des enregistrements dans chaque classification sont ajustés en multipliant les poids par ce facteur. Après l'ajustement des poids, les agrégats estimés seront désormais conformes à l'estimation de la source indépendante pour chaque

classification. L'estimation par le quotient est assez simple par rapport à d'autres méthodes d'utilisation de renseignements externes et se traduit en même temps par une efficacité accrue. Le choix de l'information externe est cependant très crucial pour la procédure, car il ne conduit à une plus grande efficacité que si cette information est fortement corrélée avec les caractéristiques d'intérêt de l'enquête.

7.2 Pondération

Dans l'étude sur le fret, le poids final lié à chaque enregistrement est calculé en cinq étapes et est le produit de cinq facteurs. Il s'agit du poids du jour, du poids de la tranche horaire, du poids de la direction, du facteur de correction pour la non-réponse et du poids du camion.

(i) Poids du jour

Ce poids correspond au rapport d'échantillonnage inverse des jours sélectionnés pour chaque rééchantillonnage dans chaque strate. Comme chaque rééchantillonnage correspond à un 1/2 échantillon, le poids est divisé par 2.

(ii) Poids de la tranche horaire

Puisque l'une des quatre tranches horaires correspondant à 0:01 à 6:00, 6:01 à 12:00, 12:01 à 18:00 ou 18:01 à 24:00 est sélectionnée pour chaque jour échantillonné, ce poids est de quatre dans tous les cas.

(iii) Poids de la direction

Étant donné que la classification et les interviews dans l'enquête pilote se font dans une seule direction pour chacune des trois périodes consécutives d'une heure, ce poids est de deux dans tous les cas. Pour la classification et l'interview dans les deux directions, le poids est de 1.

Le produit des trois facteurs ci-dessus sera appelé le sous-poids. Ce sous-poids est saisi dans le DCE avant le début des opérations sur le terrain.

(iv) Facteur de correction pour la non-réponse

Étant donné que certains DCE seront perdus (inutilisables, etc.) ou que, dans certains cas, il n'a pas été possible d'effectuer l'interview, les renseignements correspondant à la

période d'échantillonnage seront perdus, à moins que les poids pour les autres périodes d'échantillonnage ne soient ajustés pour compenser cette perte. Les sous-poids pour les DCE terminés sont ajustés au sein de huit classes définies par la tranche horaire (4) et la direction (2), en partant du principe que les caractéristiques de volume et de marchandises des camions pour lesquels un interview a été réalisé avec succès représentent les caractéristiques de volume et de marchandises des camions qui auraient dû être visés par une interview, mais qui ne l'ont pas été au sein des classes ci-dessus. Cependant, si cette hypothèse n'est pas vraie, les estimations seront biaisées et le biais sera important avec un taux de non-réponse élevé. L'ampleur exacte du biais introduit par l'ajustement pour la non-réponse est impossible à calculer. Par conséquent, un tel ajustement doit être considéré comme un dernier recours, et tous les efforts doivent être déployés pour le réduire sur le terrain.

Les sous-poids multipliés par le facteur de correction donnent le sous-poids corrigé.

(v) Poids du camion

Après vérification et correction des questionnaires, l'échantillon effectif est enregistré sur le DCE en fonction du type de camion, à côté du nombre total de camions. Le rapport fournit la fraction de sondage pour les camions, et l'inverse fournit le poids du camion.

Le poids final correspondant à un camion pour lequel une interview a été réalisée constitue le produit des cinq facteurs ci-dessus. Ce poids est calculé manuellement et est lié à chaque marchandise d'un camion échantillonné.

L'estimation du poids total d'une marchandise donnée déplacée entre une origine et une destination données au cours d'une tranche horaire donnée est obtenue en triant d'abord tous les enregistrements se rapportant à cette marchandise au cours de la tranche horaire (trimestre ou année) ayant les codes d'origine et de destination correspondants, puis en agrégeant le produit du poids final et du poids de la marchandise.

Un ajustement du quotient a été envisagé en utilisant comme source auxiliaire les renseignements fournis par les compteurs de trafic. L'idée a été abandonnée pour trois raisons, à savoir :

(1) les compteurs de trafic enregistrent le nombre d'essieux et non de véhicules;

(2) les procédures d'incorporation de l'information étaient complexes sur le plan opérationnel et donc assez coûteuses;

(3) une validation de la fiabilité physique des compteurs serait nécessaire avant toute tentative sérieuse d'utilisation des renseignements.

7.3 Description de l'estimation

La procédure d'estimation décrite ci-dessus peut être énoncée de manière algébrique, et la notation suivante peut être utilisée à cette fin.

h - strate (h = 1, 2, ... 36)

i — rééchantillonnage (i = 1, 2)

j - jour (j = 1, 2, ... 20)

k - tranche horaire (k = 1, 2, 3, 4)

ℓ — direction (ℓ = 1, 2)

m — heure (m = 1, 2, ... 6)

r - type de camion (r = 1, 2, 3, 4)

o - camion visé par une interview

t - trimestre (t = 1, 2, 3, 4)

Le poids final d'un camion visé par une interview, représenté par $W_{hijk\ell mro}$, peut être exprimé comme le produit de cinq facteurs de pondération comme suit :

$$W_{hijk\ell mro} = W_{hij}\, W_{hijk}\, W_{hijk\ell}\, W_{k\ell t}\, W_{hijk\ell mr}$$

Chacun des facteurs de pondération des composants peut être exprimé comme suit :

(i) $W_{hij} = N_h / 2n_{hi}$;

N_h : nombre de jours dans la strate h;

n_{hi} : nombre de jours sélectionnés dans la strate h dans le rééchantillonnage i;

(ii) $W_{hijk} \equiv 4$ (une des quatre tranches horaires est sélectionnée);

(iii) $W_{hijk} \equiv 2$ (sélection d'une direction comme dans l'enquête pilote; flux importants dans l'enquête nationale);

$\equiv 1$ (deux directions pour les flux légers);

(iv) $W_{k\ell t} = N_{k\ell t} / \hat{N}_{k\ell t}$;

$N_{k\ell t}$: nombre de tranches horaires de type k correspondant à la direction ℓ dans le t^e trimestre de l'univers

$\hat{N}_{k\ell t} =$;

p. ex. $\hat{N}_{k\ell t}$ est la somme des sous-poids des DCE utilisables se rapportant $\sum_{h \in t} \sum_{i} \sum_{j} W_{hij} W_{hijk} W_{hijk\ell}$ au type de tranche horaire k^e correspondant à la direction ℓ dans le t^e trimestre;

(v) $W_{hijk\ell mr} = N_{hijk\ell mr} / n_{hijk\ell mr}$;

$N_{hijk\ell mr} =$ nombre total de camions de type r passant devant la station de contrôle au cours de la m^e heure correspondant à la direction ℓ dans la k^e tranche horaire type du j^e jour sélectionné appartenant au rééchantillonnage i dans la h^e strate, comme l'a enregistré l'agent de classification des camions.

$n_{hijk\ell mr} =$ nombre correspondant de camions visés par une interview. Le poids final est répété pour chaque marchandise d'un camion échantillonné.

7.4 Regroupement des heures et des types de camions en cas d'observations nulles

Soit $T_{k\ell mr}$ désignant le trafic total de camions de type r dans la m^e heure dans la ℓ^e directive dans la tranche horaire de type k et $t_{k\ell mr}$, le nombre correspondant d'interviews.

S'il arrive que, pour certains r et m, $T_{k\ell mr} = 0$, il n'y a aucun problème; le poids du camion est défini comme étant égal à 0 et ne figure tout simplement pas dans l'estimation.

Si, toutefois, $T_{k\ell mr} \neq 0$ et $t_{k\ell mr} = 0$, la procédure ci-dessus ne s'applique pas, car une partie du trafic serait omise dans toutes les estimations.

Il est donc nécessaire d'élaborer une stratégie de regroupement des heures ou des types de camions pour tenir compte des observations nulles.

La procédure à suivre dans l'enquête sur les P-O-D consiste à regrouper un type de camion donné sur des heures adjacentes. Si, par exemple, $t_{k\ell\ 11} = 0$ et $t_{k\ell\ 21} = 0$, les deux premières heures sont regroupées pour le type de camion 1 pour obtenir la quantité suivante nécessaire au calcul des poids des camions : $T_{k\ell m^*1}/t_{k\ell m^*\ 1}$ où m* se rapporte à un segment temporel de deux heures regroupées.

Si le regroupement de toutes les heures dans une direction ne donne toujours aucune observation, le regroupement est alors effectué entre les types de camions adjacents. Enfin, si aucune observation n'est conservée après un tel regroupement, le nombre d'interviews dans une direction au cours d'une tranche horaire donnée est nul et la période correspondante est traitée comme une non-réponse. Cette composante de la non-réponse est traitée par un ajustement des sous-poids discutés dans la section sur l'estimation.

8 Considérations spéciales

8.1 Camions vides

La valeur de code correspondant à « vide » est insérée dans le questionnaire. La marchandise transportée est « sans marchandise », et aucune distinction n'est faite aux fins de la pondération. Les camions vides sont ensuite traités par estimation de domaine.

8.2 Dédoublement

Étant donné que plus d'un point de contrôle peut être posté le long d'un itinéraire donné entre une origine et une destination définies, le problème du dédoublement se pose. Le problème du dédoublement vient du fait qu'un camion a une probabilité non nulle d'être arrêté à chaque station le long de l'itinéraire. Le problème ne peut pas être résolu en échantillonnant chaque station à des moments différents ou en rejetant les camions en double dans l'échantillon. Les quatre méthodes suivantes sont proposées pour résoudre ce problème.

1) Sélectionner à l'avance la station de contrôle qui fournira l'estimation de P-O-D. Cette option est facile à appliquer. Le principal inconvénient réside dans le fait qu'une station est choisie arbitrairement et que les renseignements provenant des autres stations de contrôle sont perdus.

2) Sélectionner la station de contrôle pour laquelle l'estimation de P-O-D est basée sur la plus grande taille d'échantillon. L'inconvénient, là encore, est que les renseignements provenant d'autres stations de contrôle sont perdus et qu'il faut déterminer le nombre d'enregistrements pour chaque estimation afin de sélectionner la station.

3) Créer une moyenne simple des estimations des stations de contrôle le long d'un itinéraire entre l'origine et la destination. L'avantage de cette approche est que tous les renseignements sont utilisés. Les inconvénients sont que la contribution de chaque station à l'estimation est la même (c'est-à-dire un poids égal) quelle que soit la taille de l'échantillon et qu'une programmation est nécessaire pour construire ces estimations.

4) Créer une moyenne pondérée des estimations des stations de contrôle le long d'un itinéraire entre l'origine et la destination sur la base de la taille relative de l'échantillon. Cette méthode est la plus solide sur le plan statistique. L'avantage, là encore, est que tous les renseignements sont utilisés de manière à minimiser la variance des estimations de P-O-D. L'inconvénient vient du fait que le nombre d'enregistrements doit être déterminé pour toutes les estimations et qu'une programmation supplémentaire est nécessaire pour construire ces estimations.

8.3 Autres itinéraires d'une origine à une destination

Il existe quelques autres routes (pas beaucoup) entre une origine majeure et une destination majeure au Pérou. Un tel réseau routier a été inclus dans l'essai pilote. Les estimations de P-O-D obtenues par d'autres routes doivent être agrégées. Après agrégation, la situation se réduit à celle correspondant à une seule route.

9 Estimation de la variance

Sur la base des données de l'échantillon, il est possible de calculer la variabilité d'échantillonnage des estimations de P-O-D.

La méthodologie adoptée pour l'estimation des variances dans le test pilote est la pseudorépétition. Pour les besoins de l'estimation de la variance, les deux rééchantillonnages sélectionnés dans chaque strate sont censés avoir été sélectionnés indépendamment. En effet, les unités de chacun des deux rééchantillonnages ont été sélectionnées sans remise. En outre, l'ajustement pour la non-réponse est effectué pour les deux rééchantillonnages ensemble plutôt qu'indépendamment pour chaque rééchantillonnage. Cet ajustement est effectué au niveau trimestriel plutôt qu'au niveau de la strate. Les rééchantillonnages sont donc quelque peu corrélés, mais cela a été considéré comme négligeable dans l'enquête pilote. Une formulation plus précise pour la construction des estimations de la variance nécessite une programmation plus complexe. Un tel système peut être mis au point pour l'enquête nationale, comme il est expliqué dans une section ultérieure de ce rapport. L'avantage de la présente formulation est qu'elle devrait être relativement facile à programmer et devrait fournir des approximations utiles en ce qui concerne la fiabilité des données.

Soit $\hat{X}_t$ représentant une estimation de P-O-D donnée pour le trimestre t, l'estimation de la variance désignée par $\hat{V}(\hat{X}_t)$ peut être déterminée comme suit :

$$\hat{V}(\hat{X}_t) = \sum_{\substack{h=1 \\ h \varepsilon t}}^{9} (\hat{X}_{h1} - \hat{X}_{h2})^2$$

où

X_{h1} = estimation du demi-échantillon correspondant au rééchantillonnage 1 dans la strate h,

X_{h2} = estimation du demi-échantillon correspondant au rééchantillonnage 2 dans la strate h,

et $\sum_{\substack{h=1 \\ h \varepsilon t}}^{9}$ désigne la sommation de toutes les strates du trimestre t.

L'estimation de la variance de l'estimation annuelle correspondante de P-O-D $\hat{X}_A$ désignée par $\hat{V}(\hat{X}_A)$ peut être déterminée comme suit :

$$\hat{V}(\hat{X}_A) = \sum_{t=1}^{4} \hat{V}(\hat{X}_t)$$

où $\sum_{t=1}^{4}$ désigne la sommation sur les quatre trimestres.

10 Contrôle et imputation

10.1 Introduction

Le contrôle et l'imputation ont pour but de déterminer et de corriger les entrées ou les codes non valides, de concilier les données contradictoires et, dans la mesure du possible, de combler les valeurs manquantes dans les champs d'information des enregistrements comportant une non-réponse partielle. Cette procédure est distincte des ajustements de pondération pour la non-réponse complète qui ne prennent pas la forme de changements dans les enregistrements « individuels » mais peuvent changer le poids relatif d'un enregistrement entier dans les estimations.

Le contrôle et l'imputation doivent être effectués selon un ensemble précis d'étapes et de règles de décision qui reposent principalement sur des connaissances externes et des règles logiques.

Toutes les opérations de contrôle et d'imputation à effectuer sur les réponses à l'enquête pilote doivent être réalisées au niveau de l'enregistrement individuel, ce qui permet d'éviter la nécessité d'une approche d'imputation de fréquence « hot-deck ».

Les sections suivantes illustrent l'approche générale du contrôle et de l'imputation, car les spécifications exactes dépendront du contenu définitif du questionnaire.

10.2 Erreurs de codage et de transcription

Toutes les modifications seront effectuées manuellement. Des codes non valides peuvent s'afficher dans des champs d'identification uniques. La correction doit être effectuée au moyen d'autres renseignements d'identification disponibles. Par ailleurs, plusieurs champs peuvent

contenir individuellement des codes autorisés, mais, pris ensemble, ils indiquent un cas inexistant. Par exemple, il ne suffit pas de s'assurer que le « lieu » est l'un des codes autorisés; il doit correspondre à une combinaison de jours et de tranches horaires sélectionnée. Ce type de conflit est facile à préciser au moyen d'une liste de combinaisons de codes valides.

10.3 Spécifications relatives au contrôle et à l'imputation

Tous les contrôles doivent être précisés avant que le contrôle manuel ne puisse être effectué. La première étape consiste à déterminer toutes les réponses admissibles. Cela est traité au moyen d'une liste codée de stations de contrôle, d'origines, de destinations, de marchandises, de classification de poids, etc. Dans le cas du poids réel de la marchandise, tout nombre entier non négatif, par exemple, peut être acceptable, mais une limite doit généralement être fixée. Les rejets peuvent essentiellement prendre l'une des trois formes suivantes : un champ unique contenant un code non valide qui doit être corrigé; un champ unique contenant un blanc non valide, ou encore, deux champs ou plus en conflit, c'est-à-dire ayant chacun un code légal mais formant une combinaison invalide. Dans le cas des variables qui figurent dans le DCE et le questionnaire, on exigerait normalement un accord sur les renseignements contenus dans ces enregistrements. En cas de désaccord, le problème est de décider quel enregistrement contient les bons codes.

De plus, il y a quelques contrôles logiques. Considérons, par exemple, la situation où le nombre total de produits enregistrés est inférieur au nombre de produits répertoriés.

Enfin, on peut imposer des contrôles du « caractère raisonnable ». Un camion à deux essieux peut-il transporter 10 marchandises, dont 3 sont inscrites sur la liste et dont le poids total est de 1 000 tonnes? La connaissance du sujet déterminera le nombre de contrôles de ce type qu'il sera utile d'imposer. Comme le contrôle doit se faire manuellement, la complexité de la structure de contrôle doit être réduite au minimum.

Après avoir déterminé les enregistrements qui ne respectent pas une ou plusieurs règles de contrôle, le problème est de déterminer la façon de les corriger. La première tâche consiste à séparer les questionnaires en

trois groupes : (a) les enregistrements qui respectent toutes les règles de contrôle, (b) les enregistrements qui nécessitent quelques corrections et (c) les retours inutilisables. Les deux premiers groupes fourniront les données pour un traitement ultérieur. Il est nécessaire d'établir des règles concernant la manière dont les rejets au contrôle sont corrigés pour le groupe (b).

La correction d'un champ unique doit être résolue en consultant le DCE pour les mêmes renseignements ou de manière déterministe sur la base d'un ensemble défini d'actions d'imputation. La situation la plus difficile à résoudre est celle qui suppose le contrôle de plusieurs champs. Dans certains cas, on peut simplement modifier le nombre minimum de champs requis pour obtenir une combinaison valide. Dans d'autres cas, un champ ou une relation entre des champs peuvent être déterminés comme étant d'une importance primordiale. En tout état de cause, toutes les règles d'imputation doivent être élaborées de manière à pouvoir être appliquées manuellement de façon simple et cohérente.

11 Traitement informatique et totalisation

À partir de l'ensemble de questionnaires propres et pondérés qui en découle, un fichier propre peut être créé sous réserve des erreurs de saisie des données qui peuvent être minimisées au moyen d'un contrôle clé, etc. La construction de toute estimation souhaitée est alors relativement simple : l'estimation pour toute origine-destination se trouve simplement en additionnant le produit des poids finaux et les poids des produits pour la période à l'étude. La totalisation revient en fait à déterminer les estimations à produire et à organiser un format de sortie approprié.

12 Considérations relatives au plan d'une enquête nationale continue sur les produits, l'origine et la destination

Vous trouverez ci-dessous des recommandations concernant l'analyse des données de l'enquête pilote et l'évaluation des opérations de l'enquête qui serviront de base à l'élaboration d'une enquête nationale ou régionale sur les P-O-D.

12.1 Concepts et définitions

L'adéquation des concepts et des définitions utilisés dans l'enquête pilote pour obtenir l'origine, la destination, les envois d'une seule et de plusieurs marchandises doit être évaluée du point de vue de leur efficacité statistique et opérationnelle.

Plus précisément, il convient d'évaluer la manière dont le problème associé à plusieurs origines ou destinations correspondant à la même marchandise est traité dans l'enquête pilote. En outre, les décisions doivent être prises sur la base de l'enquête pilote quant au nombre maximal de produits qui peuvent être obtenus de manière fiable au cours d'une interview routière.

12.2 Systèmes de classification

Les systèmes de classification des marchandises, des camions, des poids et mesures et de la géographie adoptés dans le cadre du projet pilote doivent être évalués en fonction de leur capacité à satisfaire les besoins en données et de leur efficacité opérationnelle et statistique. Les systèmes révisés peuvent ensuite être mis en œuvre aux échelons régional et national.

12.3 Opérations sur le terrain

Les méthodes et procédures utilisées pour le recrutement et la formation des brigades de terrain dans le cadre de l'enquête pilote doivent être évaluées et éventuellement révisées. Plus précisément, les procédures de terrain doivent être revues en ce qui concerne leur capacité à gérer des flux de trafic importants et à effectuer une enquête dans des conditions climatiques défavorables. Il devrait être possible de déterminer combien d'interviews peuvent être traitées avec succès par une brigade en fonction du volume de trafic. Ces renseignements peuvent être utilisés pour suivre le rendement des brigades dans l'enquête nationale. Des programmes systématiques d'observation sur le terrain devraient également faire partie intégrante d'un exercice national de collecte de données.

12.4 Plan d'échantillonnage

(a) Stratification

Sur la base d'une analyse des estimations des principales marchandises, l'efficacité des variables de stratification sélectionnées pour l'enquête pilote peut être évaluée. Les variables qui s'avèrent importantes peuvent être incorporées dans le plan d'échantillonnage révisé.

(b) Taille de l'échantillon et répartition de l'échantillon

L'analyse des estimations de la variance pour les estimations des produits sélectionnés devrait guider les décisions concernant les ajustements de la taille de l'échantillon ou de la répartition. Par exemple, les flux importants peuvent être gérés par deux brigades au lieu d'une.

La stratégie de répartition, qui devrait être considérée dans le contexte d'une enquête nationale, est un compromis entre une répartition proportionnelle à la taille des strates et une répartition proportionnelle au volume de trafic estimé des strates.

(c) Sélection des tranches horaires

Si de grandes différences de flux de trafic sont constatées entre les quatre tranches horaires de la journée, les tranches horaires des jours sélectionnés peuvent être choisies avec une probabilité proportionnelle aux flux de trafic estimés. Une telle procédure de sélection ne doit être envisagée que si les renseignements sur les flux de trafic sont jugés suffisamment précis et raisonnablement stables dans le temps.

Le poids du camion (quatrième facteur du poids total) est très important si un camion est sélectionné pendant une tranche horaire où le trafic est intense. La sélection des tranches horaires par échantillonnage avec probabilité proportionnelle à la taille entraînerait des poids plus faibles des tranches horaires (deuxième facteur dans le poids global) correspondant à ces périodes de fort trafic. Les poids globaux des camions seraient alors plus équilibrés, ce qui donnerait des estimations plus stables.

La sélection des tranches horaires a été effectuée dans le cadre de l'enquête pilote pour satisfaire aux contraintes d'exploitation sur le

terrain. Il est suggéré que la sélection des tranches horaires soit effectuée sur une base aléatoire pour chaque jour sélectionné, puis ajustée pour satisfaire aux contraintes du terrain.

12.5 Estimation

Il est recommandé que l'ajustement des sous-poids pour la non-réponse soit effectué au niveau de la strate dans chaque rééchantillonnage plutôt qu'au niveau trimestriel sur les deux rééchantillonnages. Cette procédure, bien sûr, ajoute à la complexité de l'opération et pourrait être envisagée lorsque l'expérience aura été acquise et que la mise au point de procédures automatisées deviendra possible et pratique.

L'hypothèse qui sous-tend l'ajustement pour la non-réponse est que les renseignements obtenus pour certaines tranches horaires représentent d'autres tranches horaires pour lesquelles les renseignements ne sont pas disponibles. L'hypothèse est plus valable si la correction est effectuée au niveau de la strate.

Aux fins de l'estimation de la variance décrite ci-dessous, les rééchantillonnages construits à cette fin seront corrélés, à moins que l'ajustement pour la non-réponse ne soit effectué indépendamment pour chaque rééchantillonnage. Les estimations de la variance seraient autrement soumises à un biais, dont l'ampleur pourrait être élevée si chacun des rééchantillonnages présente des taux de réponse très différents.

12.6 Estimation de la variance

Pour tenir compte de l'ajustement du poids du quotient par type de tranche horaire et par direction afin d'améliorer la précision statistique des estimations de P-O-D, des estimations de variance plus précises que celles utilisées dans l'enquête pilote peuvent être construites sur la base de la formulation suivante.

Soit $\hat{X}_t$ et $\hat{X}_A$ se rapportant aux estimations trimestrielles et annuelles de P-O-D du poids du produit X et $\hat{V}(\hat{X}_t)$ et $\hat{V}(\hat{X}_A)$, à leurs estimations de variance, on calcule ce qui suit pour chaque strate.

$$\Delta_h = \hat{X}_{h1} - \hat{X}_{h2} - \sum_{k=1}^{4} \sum_{l=1}^{2} \frac{\hat{X}_{tk\ell}}{T_{tk\ell}} (\hat{T}_{h1k\ell} - \hat{T}_{h2k\ell})$$

où $\hat{X}_{hi}$: se rapporte à l'estimation par demi-échantillon du poids total de la marchandise X dérivée du i^e rééchantillonnage dans la strate h (i = 1, 2);

$\hat{X}_{tk\ell}$: se rapporte au poids total du produit X au cours du trimestre t correspondant à la k^e tranche horaire type ℓ^e dans la direction;

$T_{tk\ell}$: désigne le nombre total de tranches horaires de type k dans la direction ℓ^e au cours du trimestre t;

$\hat{T}_{hik\ell}$: désigne l'estimation par demi-échantillon du nombre de périodes de type k dans la direction ℓ^e obtenue à partir du i^e rééchantillonnage dans la strate h (i = 1, 2).

L'estimation de la variance correspondant à $\hat{X}_t$ est calculée comme suit :

$$\hat{V}(\hat{X}_t) = \sum_{h=1}^{9} \Delta_h^2$$

L'estimation de la variance correspondante pour l'estimation annuelle $\hat{X}_A$ est calculée comme suit :

$$\hat{V}(\hat{X}_A) = \sum_{t=1}^{4} \hat{V}(\hat{X}_t).$$

12.7 Traitement des données

Les problèmes éprouvés dans le cycle de traitement manuel des données : codage, nettoyage des formulaires, pondération, etc. doivent être évalués en fonction du coût, de la rapidité et de la précision. Certaines procédures (p. ex. une partie du contrôle et de l'imputation ou de la pondération) peuvent être traitées de manière plus efficace et efficiente par le développement et la mise en œuvre de progiciels automatisés.

12.8 Analyse des délais et des coûts

Le temps et le coût de toutes les phases de l'enquête pilote doivent être enregistrés et évalués ultérieurement pour estimer ces paramètres dans l'enquête nationale proposée. Ces données serviront à alimenter toutes les phases de l'élaboration et de la mise en œuvre de l'enquête nationale.

Remerciements

Les travaux de recherche présentés dans ce document ont été menés sous la direction générale de M. K.W. Studnicki-Gizbert, directeur exécutif de la Direction de la recherche de la Commission canadienne des transports (CCT). Ce projet fait partie d'un programme global d'assistance technique dans le domaine des transports qui a été commandé par l'Agence canadienne de développement international au Pérou.

Les auteurs tiennent à reconnaître les suggestions utiles de MM. M. Nargundkar, H. Gough et Mlle J. Forgie de Statistique Canada, ainsi que le soutien et l'orientation de Mlle M. Fleming (CCT) et de M. R. Platek de Statistique Canada. Les auteurs souhaitent également remercier MM. A. Gemmell et D. Napier, les représentants de la CCT au Pérou, pour la recherche de fond et le soutien qu'on leur a apporté pendant leur séjour à Lima.

Bibliographie

[1] Gough, J.H., et Ghangurde, P.D. (1976), "An Alternative Method of Surveying International Travellers At Frontier Points - Rapport de méthodologie", Statistique Canada.

[2] Platek, R. et Singh, M.P. (1976), "Methodology of Canadian Labour Force Survey", Household Surveys Development Division, Statistique Canada.

[3] "For-Hire Trucking Survey (1976)", Transportation and Communications Division, Statistique Canada.

[4] Koop, J.C. (1960), "On Theoretical Questions Underlying the Technique of Replicated or Interpenetrating Samples", Institute of Statistics, North Carolina State College.

[5] Gemmell, A., et Napier, D. (1977), "Background Document On Inter-Urban Cargo Statistics (Road)", Research Branch, Canadian Transportation Commission.