

# Introduction

Les données sur les dépenses en immobilisations fournissent des indications utiles quant à l'état du marché dans l'économie en général et dans ses divers secteurs. Comme ces dépenses représentent une partie importante et relativement variable des dépenses nationales brutes, l'envergure et la teneur des programmes d'investissements donnent une bonne idée de la demande que les producteurs canadiens ont eu à satisfaire au cours de la période en cause. De plus, l'importance relative des immobilisations projetées, au total et dans chaque secteur, permet de savoir ce que prévoient les dirigeants d'entreprise en ce qui concerne la demande future par rapport à la capacité de production actuelle.

On trouve dans la section « Sources et méthodes » les renseignements nécessaires pour tirer le maximum des tableaux statistiques. La section « Qualité des données, concepts et méthodologie — Concepts » présente les définitions de base employées pour la collecte et la publication des données, les unités d'enquête visées et la classification des secteurs d'activité et des zones géographiques. Cette dernière section aborde en outre la question de la comparabilité de la série des dépenses en immobilisations avec les autres sources de données.

À la section « Qualité des données, concepts et méthodologie — Sources », on peut prendre connaissance des différents types de questionnaires employés pour recenser les données, des sources des données non recensées et des modes de recueil de données. La section « Qualité des données, concepts et méthodologie — Assurance de la qualité » présente les mesures prises pour assurer la qualité des données pendant et après la collecte.

La section « Qualité des données, concepts et méthodologie — Méthodologie » expose les actions entreprises et les sources utilisées pour définir la base de sondage et la méthode employée pour en tirer un échantillon stratifié. On trouvera dans la même section les procédés d'imputation et d'estimation concernant les non répondants compris dans l'échantillon et l'inflation concernant la partie non recensée de la base de sondage. Les deux dernières sections, « Qualité des données, concepts et méthodologie — Utilisateurs et utilisations » et « Qualité des données, concepts et méthodologie — Chronologie de la série des dépenses en immobilisations », fournissent respectivement des renseignements sur l'utilité des données et sur les données historiques qui existent en matière de dépenses en immobilisations.

## Concepts

### Définitions

#### Dépenses en immobilisations

Les dépenses en immobilisations sont constituées par le coût d'acquisition, de construction et d'installation d'usines, de matériel et d'outillage durables, que ce soit ou non à des fins de remplacement ou de location. Les dépenses immobilisées servant par exemple à payer des études de faisabilité, des honoraires d'architecte, d'avocat ou d'ingénieur et des frais d'installation, la valeur des biens immobilisés implantés par une entreprise, soit par voie de sous-traitance, soit par son propre personnel, ainsi que les frais d'intérêt capitalisés découlant des prêts qui financent les immobilisations sont aussi comprises dans les dépenses en immobilisations. Les dépenses brutes ont été indiquées sans déduction de la valeur de rebut ou d'échange des biens remplacés et comprennent les subventions reçues le cas échéant.

Les dépenses en immobilisations des ministères ne comprennent pas les subventions accordées aux entités externes (municipalités, organismes, établissements, entreprises) ni les articles budgétaires liés à un organisme ministériel ou à une société d'État de propriétaires, qui font l'objet d'une enquête distincte. Les dépenses en immobilisations ministérielles fédérales concernent les frais acquittés par le ministère, peu importe le ministère ayant octroyé le contrat. Les dépenses ministérielles provinciales comprennent les dépenses en immobilisations consacrées à des bâtiments, au matériel et à l'outillage devant être utilisées au Canada, financées par un fonds renouvelable, par un prêt rattaché à un fonds renouvelable, par un autre genre de prêt, par le Trésor ou par un compte spécial.

Le but est d'inclure le coût des usines, du matériel et de l'outillage neufs dont la durée d'utilisation doit dépasser un an. C'est pourquoi on demande aux répondants d'inscrire comme immobilisations tous les achats

imputés aux comptes d'immobilisations. Cette méthode passe sous silence certains types d'équipements dont l'achat est imputé aux comptes courants.

## Constructions permanentes

Une dépense de construction correspond à une activité humaine se traduisant par l'érection, l'assemblage ou l'achèvement d'un bâtiment ou d'un autre genre d'ouvrage autoportant et immobile dont les fondations, l'assise ou l'emplacement sont permanents. Le prix d'achat du terrain est exclu des dépenses de construction, qui comprennent toutefois les frais de raccordement aux services publics et d'aménagement des lieux. Les constructions comprennent en outre les modifications, les rénovations et les ajouts majeurs comportant des changements de structure ou prolongeant la durée d'utilisation prévue d'un bien. La structure peut se trouver au-dessus ou au-dessous de la surface terrestre et peut servir à faire passer ou à loger des biens ou des personnes. Une structure, non considérée comme du matériel, sous forme de bâtiment ou d'« autre structure », peut être définie comme étant le résultat d'une activité de construction ayant pour but de protéger, de soutenir, de retenir ou de faire passer quelque chose. Toute construction peut être classée comme bâtiment ou comme ouvrage de génie civil.

Un **bâtiment** est une construction permanente qui comporte des murs et un toit protégeant des biens ou des personnes d'une réalité sociale ou physique. Il peut aussi s'agir d'un abri mobile ou temporaire destiné à rester à un endroit donné pour une période prolongée, ou encore d'une annexe ou des dépendances nécessaires pour leurs fonctions de contenance, de soutien, d'accès ou de protection, ainsi que du matériel et de l'outillage faisant partie intégrante de la structure et servant par exemple à la plomberie, aux installations électriques, à la climatisation ou aux ascenseurs. Par exemple, un hangar pour avions, une usine, un hôpital, un hôtel, un immeuble à bureaux, une gare ferroviaire, une école et un centre commercial sont des bâtiments.

Un **ouvrage de génie civil** sert directement ou indirectement au transport de personnes, d'équipement, de matériel, de gaz ou de signaux électriques. Il peut aussi s'agir de structures autoportantes logeant ces mêmes réalités, que ce soit ou non lors du transport comme tel. Les structures autoportantes érigées pour la transmission de signaux électriques peuvent aussi comprendre les structures conçues pour fournir un éclairage statique ou des signaux lumineux périodiques à partir d'un endroit immobile. De plus, les frais engagés pour apporter des modifications importantes à un terrain en vue d'un usage spécialisé sont des frais d'ouvrage de génie civil. Sont des ouvrages de génie civil entre autres les ponts, les routes, les aqueducs, les systèmes d'égouts, les barrages, les installations d'éclairage de rue, les chemins de fer et les pipelines.

Ces définitions présentent une conception générale de ce qu'est une construction permanente; cependant, plusieurs industries sont caractérisées par des conditions particulières qui nécessitent des considérations spéciales. Ainsi, outre les définitions qui précèdent, les entreprises minières engagent des dépenses d'exploration, de mise en valeur et de location des sites miniers, des frais d'activités sur le terrain et des frais généraux qui rentrent dans les dépenses en constructions permanentes. Les dépenses consacrées par les entreprises gazières et pétrolières au forage d'exploration et d'exploitation, aux installations de production, à la récupération assistée et aux usines de traitement du gaz naturel rentrent aussi dans cette catégorie. Pour les entreprises de services publics, les dépenses en constructions permanentes comprennent les frais liés à la transformation, aux postes d'aiguillage et aux usines de production ainsi que les frais généraux imputables aux usines.

Bien que le **logement** ne rentre pas dans les dépenses en immobilisations selon les définitions qui précèdent, il a été inclus en l'occurrence car il occupe une place importante dans les dépenses en construction et présente des fluctuations similaires à celles qui caractérisent les dépenses en immobilisations des entreprises, des établissements et des administrations publiques.

## Matériel et outillage immobilisés

Le matériel et l'outillage sont constitués par toute combinaison de pièces interreliées qui sont mues physiquement ou électromagnétiquement et emploient la pression, la chaleur, la mécanique, l'électricité ou une autre forme d'énergie pour accomplir un travail ou, si elles sont statiques, pour faire partie du milieu de travail de personnes.

Les dépenses en immobilisations rentrant dans la catégorie du matériel et de l'outillage représentent le total des frais capitalisés consacrés à des appareils tels qu'automobiles, chaudières, compresseurs, engins de

terrassement ou de manutention, génératrices, moteurs, fournitures de bureau et de magasin, matériel professionnel et scientifique, pompes, outils et transformateurs.

De plus, les dépenses en matériel et en outillage comprennent le coût de tout autre matériel ou outillage qui ne sont pas déjà déclarés comme faisant partie d'un bâtiment ou d'un ouvrage de génie civil, le coût des travaux d'exploration ou de mise en valeur (installations non productrices), le coût des biens pouvant être considérés comme du matériel de transformation ou d'exploitation minière et le coût des autres biens d'équipement que l'entreprise utilise elle-même ou loue à un tiers. Les dépenses immobilisées liées à l'outillage, les paiements proportionnels effectués avant l'achèvement des travaux ainsi que les résidus de compte et les retenues de garantie comptabilisés dans l'année en cause sont aussi incluses. Les dépenses brutes sont indiquées sans déduction des produits de la vente des actifs immobilisés ni de la valeur de rebut ou d'échange des biens remplacés.

## **Contrats de location**

Conformément aux recommandations de l'Institut canadien des comptables agréés, les contrats de location sont divisés en contrats de location-exploitation et contrats de location-acquisition. Les immobilisations acquises pour être employées par l'acheteur ou pour être louées à autrui, soit à des fins d'exploitation, soit à des fins d'acquisition, sont classées comme des dépenses en immobilisations nouvelles. L'Institut canadien des comptables agréés recommande que les biens acquis par contrat de location-acquisition soient déclarés par le preneur. Toutefois, pour les besoins de l'enquête, ces biens sont déclarés par le bailleur.

## **Immobilisations usagées**

Les immobilisations usagées sont définies comme étant des bâtiments, des constructions, du matériel ou de l'outillage ayant déjà servi à un autre organisme. Les dépenses liées à des immobilisations canadiennes usagées sont exclues puisqu'elles représentent un transfert de biens à l'intérieur du Canada et n'ont aucun effet sur le total national. Par contre, les dépenses liées à des biens importés au Canada augmentent le stock national et sont donc incluses dans la série des dépenses en immobilisations.

## **Travaux en cours**

La série des dépenses en immobilisations comprend les dépenses consacrées aux travaux en cours, soit les coûts accumulés des projets d'immobilisations non achevés mais destinés à être immobilisés lorsqu'ils seront terminés.

## **Dépenses en entretien et réparations**

Les dépenses en entretien et réparations portant sur des ouvrages, du matériel et de l'outillage sont indiquées séparément. **Ces dépenses ne sont pas considérées comme des immobilisations.**

Les dépenses en entretien et réparations sont constituées par la partie des frais courants qui sont portés en diminution des revenus de l'année où ils ont été engagés et qui sont consentis afin de maintenir le stock des immobilisations ou leur capacité productive (prévention) tout au long de la durée d'utilisation prévue. L'entretien et les réparations permettent de maintenir la productivité des immobilisations jusqu'à la fin de leur période de service sans temps d'arrêt indus. Ils ont aussi pour but de remettre une partie quelconque des immobilisations en bon état de marche après une défaillance ou une réduction de l'efficacité, ou de prolonger sensiblement la durée d'utilisation ou la productivité des biens. Ces sorties de fonds donnent une meilleure idée de la demande portant sur la main-d'oeuvre et le matériel.

## **Réparation des constructions**

Les dépenses d'entretien et de réparations liées aux constructions comprennent les dépenses qui ne prolongent pas la durée d'utilisation prévue de la structure et qui n'en augmentent pas la capacité. Les dépenses d'entretien des bâtiments et des autres structures peuvent comprendre le soin courant des biens comme les services de conciergerie, l'enlèvement de la neige ou l'épandage de sel ou de sable par le personnel de l'entreprise ou par d'autres personnes engagées par elle.

## **Réparation du matériel et de l'outillage**

Les dépenses d'entretien et de réparations liées au matériel et à l'outillage comprennent les dépenses qui ne prolongent pas la durée d'utilisation prévue du bien et n'en augmentent pas la capacité. Elles peuvent comprendre les vidanges d'huile et le graissage des véhicules et outillage.

## **Amortissement cumulé**

L'amortissement cumulé est la somme des annuités imputées à la dépréciation du bien depuis son achat.

## **Utilisation de la capacité**

On calcule l'utilisation de la capacité en divisant le niveau de production effectif d'une entreprise (en dollars ou en unités), par son niveau de production normal.

## **Impartition ou exécution autonome**

L'impartition est l'action de confier des travaux à un entrepreneur en construction. L'exécution autonome consiste à faire effectuer les travaux par le personnel de l'entreprise.

## **Aliénation, vente et réduction des immobilisations**

Il s'agit ici de la valeur comptable brute des immobilisations qui ont été aliénées, vendues, mises hors service, détruites ou sorties de quelque autre façon (y compris les réductions) ou qui ont été échangées à l'occasion de l'achat de nouvelles immobilisations. Le coût en capital cumulé représente le total des dépenses en immobilisations engagées pour un bien au moment de la construction ou de l'achat et après.

## **Durée d'utilisation prévue**

La durée d'utilisation prévue d'un bien n'a pas de rapport avec la durée d'utilisation déclarée aux fins de l'impôt sur le revenu. Dans le cas d'une mine, la durée d'utilisation prévue est la période pendant laquelle la mine est censée être productive. Il s'agit ici des dépenses amorties (ou passées en charges dans certain cas) liées à l'exploration et à la mise en valeur du site minier. La durée prévue est basée sur le nombre d'années pour lequel l'entreprise s'engage dans la production au départ (exemple : méthode de l'amortissement proportionnel au rendement) en supposant qu'aucune diminution (ou augmentation) du prix des minéraux ne vienne réduire (ou prolonger) cette durée. La durée d'exploitation ou de productivité peut ne pas correspondre à la durée déclarée aux fins de l'impôt sur le revenu ni à la teneur mesurée des gisements.

## **Durée d'utilisation restante prévue des immobilisations**

La durée d'utilisation restante prévue des immobilisations représente le nombre d'années qui restent dans la durée d'utilisation d'un bien usagé au moment de l'acquisition.

## **Valeur comptable brute**

La valeur comptable brute est le prix d'achat du bien à l'origine.

## **Classification**

Dans l'enquête sur les dépenses en immobilisations, l'établissement sert d'unité statistique principale dans la mesure des dépenses en immobilisations et en réparations. Par définition, l'établissement est la plus petite entité opérationnelle produisant un ensemble de biens et de services aussi homogène que possible, dont les registres indiquent la valeur de la production ainsi que le coût des matériaux utilisés et le coût et la qualité des ressources humaines employées pour cette production et pour laquelle des registres ou des ventilations estimées permettent d'obtenir toutes les variables de production nécessaires pour calculer la valeur ajoutée.

Le terme **établissement** fait référence à une capacité de production organisée comportant un certain degré de spécialisation. Pour tenir compte de la production diversifiée, on utilise le **Système de classification des industries de l'Amérique du Nord** (SCIAN, n° 12-501-X au catalogue), qui établit une distinction entre les activités principales, secondaires et auxiliaires et qui, en définitive, regroupe les établissements selon leur activité principale. Dans cette version, les établissements sont regroupés en industries, en grands groupes et en divisions selon la production de biens ou de services homogènes ou selon leur participation à une activité économique similaire. Ce mode de regroupement s'applique à tous les établissements privés ou publics ainsi qu'aux entreprises gouvernementales. Toutes les autres activités gouvernementales sont classées comme fédérales, provinciales ou municipales dans la division des services gouvernementaux. Les notions et définitions employées dans la série des dépenses en immobilisations sont celles qui sont décrites dans les **Concepts and Definitions of Capital Stock and Capital Formation Series F No. 3** de 1953 des Nations-Unies.

Puisqu'un établissement peut être fixé dans plusieurs provinces, la **Classification géographique type** (CGT, n° 12-571-X au catalogue) a été employée dans l'enquête sur les dépenses en immobilisations. La CGT a été conçue pour diviser le Canada en régions en fonction des provinces, des divisions et subdivisions de recensement et des régions métropolitaines de recensement. Pour l'enquête sur les dépenses en immobilisations, on a retenu la classification géographique au niveau provincial comme base pour l'échantillonnage stratifié des établissements. Pour descendre au niveau des divisions et des subdivisions de recensement, il aurait fallu accroître l'échantillon pour de nombreuses industries.

## Comparabilité

Bien que la série des dépenses en immobilisations soit conforme aux normes préconisées par Statistique Canada en matière de classification géographique et de classification des industries, on observe des différences entre les valeurs des dépenses en immobilisations qui y sont présentées et celles que l'on trouve dans d'autres sources.

Les nouveaux investissements enquêtés par la Division de l'investissement et du stock de capital (DISC) de Statistique Canada comprennent toutes les dépenses d'investissement des entreprises privées et des organismes gouvernementaux qui acquièrent des biens corporels durables. Toutefois, les totaux ne correspondent pas exactement aux données publiées pour la formation brute de capital fixe dans la comptabilité nationale des revenus et dépenses à cause des rajustements effectués pour le système de comptabilité nationale. Ces rajustements comprennent des déductions pour les constructions servant à la défense, les ventes nettes de véhicules automobiles usagés et le matériel mis au rebut ou à récupérer et des ajouts pour les coûts de transfert des terrains et des bâtiments existants.

Les totaux des dépenses en immobilisations publiés par la Division de l'organisation et des finances de l'industrie (DOFI) ne correspondent pas non plus exactement aux résultats présentés ici du fait que la DOFI se concentre sur les données à l'échelle de l'entreprise dans le secteur privé. De plus, contrairement à la série des dépenses en immobilisations, la DOFI compte le prix d'achat des terrains et des bâtiments usagés.

Les chiffres de la présente publication de la DISC diffèrent sous plusieurs rapports des dépenses connexes d'activités d'amont publiées par le Secteur de la politique énergétique de Ressources naturelles Canada (RNCAN) et par la Division des comptes des industries de Statistique Canada. D'abord, la comparabilité des statistiques de l'exploration et de la mise en valeur dans les secteurs du pétrole et du gaz naturel se trouve limitée par le fait que la Division des comptes des industries de Statistique Canada tient compte des terrains achetés à des fins de construction, ainsi que des acquisitions et locations de terres. Dans le secteur non traditionnel, la Division des comptes des industries inclut également les acquisitions de logements. De plus, le Secteur de la politique énergétique de Ressources naturelles Canada et la Division des comptes des industries incluent les dépenses liées aux activités géologiques et géophysiques. Or, ces dépenses ne sont pas considérées comme faisant partie de la « formation de capital » au sens de la comptabilité nationale et ne sont pas comptées dans la présente publication. Ensuite, RNCAN et la Division des comptes des industries recueillent les données sur les « autres dépenses en immobilisations » au niveau national, tandis que la DISC le fait à l'échelle provinciale. Enfin, la Division des comptes des industries recueille ses données pour l'année civile, dans la mesure du possible, contrairement à RNCAN et à la DISC. Cette dernière différence ne devrait cependant pas causer des écarts considérables.

L'enquête sur les dépenses en immobilisations respecte autant que possible les recommandations de l'Institut canadien des comptables agréés (ICCA); cependant, les données fournies par les établissements comportent souvent des coûts passés en charge pour des articles qui auraient dû être immobilisés. Les biens loués à bail sont déclarés par le bailleur dans l'enquête sur les dépenses en immobilisations, alors que l'ICCA recommande que les biens acquis par location-acquisition soient comptabilisés par le preneur.

## Sources

### Données enquêtées

La majorité des secteurs couverts par les séries sur les investissements sont enquêtés. Tous les établissements sélectionnés durant les trois périodes d'enquête (voir « Périodes d'enquête ») reçoivent soit le questionnaire ordinaire (long ou abrégé), soit un questionnaire spécialisé (long ou abrégé) soit le questionnaire de nouveau projet. Le genre de questionnaire reçu par un établissement dépend de son secteur d'activité, du niveau de dépenses anticipé, de l'enquête en cause et du fait que l'établissement fait partie ou non des nouveaux projets (hors de la base de sondage ou cas aberrant).

Le questionnaire ordinaire abrégé est le plus utilisé pour les trois périodes d'enquête. Ce questionnaire vise à recueillir des renseignements de base portant sur les constructions, le matériel et l'outillage immobilisés, les réparations apportées aux constructions ainsi qu'au matériel et à l'outillage, la valeur comptable brute, l'utilisation de la capacité dans les secteurs manufacturier et minier, les motifs des changements de dépenses, les travaux en cours et les contrats de location à bail. Soulignons que l'établissement est tenu de revoir ses dépenses en réparations durant l'enquête sur les données réelles seulement. L'établissement reçoit l'un des autres types de questionnaire si on l'on croit qu'il consacre des sommes importantes aux immobilisations, s'il est engagé dans un secteur spécialisé ou s'il fait partie des nouveaux projets.

Le questionnaire ordinaire long ne sert que pour l'enquête sur les données réelles et vise les établissements ayant déjà déclaré des dépenses importantes. Il vise à rassembler des renseignements plus approfondis que le questionnaire abrégé et touche le détail des dépenses par type d'actif, la valeur brute des actifs, les motifs d'aliénation, les frais d'intérêt capitalisés, le nombre de robots et les contrats de location selon le genre de bien (voir l'enquête 2803).

Les questionnaires spécialisés concernent le secteur minier et celui du pétrole et du gaz naturel. Les questionnaires de projets nouveaux sont destinés aux nouveaux établissements qui ne font pas encore partie de la base de sondage parce qu'ils ne produisent pas ou sont classés comme des cas aberrants.

Outre le sondage des établissements, la série des dépenses en immobilisations repose aussi sur des modes de cueillette particuliers. En effet, certains répondants opérant au Canada sont dans l'impossibilité de fournir une ventilation provinciale de leurs dépenses pour les périodes en cause. On utilise alors un rapport consolidé. La répartition entre les provinces est calculée à partir de ce rapport selon les caractéristiques pertinentes. Il peut aussi arriver que le nombre de sites administrés par l'établissement soit trop élevé pour un sondage classique. Dans ce cas, pour simplifier la déclaration, on fait appel, pour la collecte des données, à une entité appelée établissement provincial. Cependant, les sites représentés par l'établissement provincial doivent tous faire partie de la même industrie.

On demande aux répondants de déclarer les dépenses pour l'exercice financier de douze mois dont le dernier jour se situe entre le 1<sup>er</sup> avril de l'année demandée et le 31 mars de l'année suivante.

### Données non enquêtées

Bien que la série des dépenses en immobilisations fournisse une estimation des dépenses attribuables à chaque division de la SCIAN, toutes ne sont pas enquêtées. Lorsqu'elles ne le sont pas, l'estimation des dépenses est fondée sur les indices de production, de consommation et de coûts associés au secteur en cause.

La valeur des dépenses en immobilisations du secteur de la **pêche** (division B), pour toutes les périodes d'enquête, est fondée sur la modélisation statistique des données obtenues du ministère des Pêches et Océans du Canada et de la Division de l'organisation et des finances de l'industrie de Statistique Canada. Toutefois, les groupes industriels 032 et 033, soit les services relatifs à la pêche et le piégeage respectivement, ne sont

pas couverts par ces sources et ne font pas l'objet d'estimations dans la série des dépenses en immobilisations.

### **Modifications relatives aux données non enquêtées :**

#### Agriculture

Les données relatives aux industries des cultures agricoles et de l'élevage présentées dans les estimations des *Investissements privés et publics au Canada* sont fournies par la Division de l'agriculture de Statistique Canada. En effet, c'est la Division de l'agriculture qui recueille les données sur les dépenses en immobilisations pour ces deux industries, à l'aide de l'Enquête financière sur les fermes (EFF).

La Division de l'agriculture a effectué des modifications à la base de sondage pour la période de référence de 2007 qui ont influé les estimations des dépenses en immobilisations.

Les comparaisons par rapport aux estimations des dépenses en immobilisations précédant cette période devraient être faites avec prudence, considérant le bris des séries chronologiques.

#### Construction

Le secteur de la construction est non enquêté pour ce qui est de l'Enquête sur les investissements privés et publics. Les estimations sont basées sur les tendances observées des activités dans le secteur de la construction par rapport à l'ensemble des activités économiques. Nous avons remarqué que la valeur des dépenses en immobilisations par province obtenue par la modélisation statistique des tendances pouvait être biaisée. Pour résoudre ce problème, à compter de la période de référence de 2007, nous utiliserons les données fiscales afin de dériver une répartition provinciale de l'investissement qui s'appliquera à l'estimation pour cette industrie obtenue par modélisation statistique.

Finalement, les chiffres sur les investissements dans le domaine du **logement** sont produits par la Section des indicateurs courants de l'investissement et sont fondés sur les mises en chantier projetées, les coûts de construction et la valeur des rénovations dans chaque province. L'**infrastructure résidentielle** mise en place par les promoteurs est estimée et la valeur des infrastructures qui seront remises aux municipalités une fois achevées est comprise dans la série des dépenses en immobilisations sous la rubrique des administrations locales.

### **Mode de collecte des données**

Plusieurs divisions de Statistique Canada participent à la collecte des données comprises dans les estimations finales des dépenses en immobilisations produites par la Division de l'investissement et du stock de capital. La Division de l'agriculture recueille des données sur les dépenses réelles et réelles préliminaires en immobilisations à l'aide de l'Enquête financière sur les fermes et de l'Enquête sur les cultures. La Division des institutions publiques accélère le processus en fournissant les données de son Enquête sur les dépenses en immobilisations des administrations locales, tandis que la Division des comptes des industries apporte des données sur les petites entreprises en se fondant sur son Enquête sur les dépenses nettes de l'industrie pétrolière et gazière. Les estimations portant sur le logement sont produites par la Section des indicateurs courants de l'investissement (Division de l'investissement et du stock de capital).

De plus, la série des dépenses en immobilisations comprend des données recueillies par des organismes et ministères extérieurs à Statistique Canada. Ainsi, les données rassemblées par chaque coordonnateur statistique provincial ou territorial dans le domaine de l'enseignement (écoles provinciales et territoriales), de la santé et de l'administration publique sont incorporées dans les séries des dépenses en immobilisations. Les chiffres portant sur l'industrie minière sont recueillis à l'échelle provinciale par les ministères provinciaux de l'énergie, des mines et des richesses naturelles de Terre-Neuve et Labrador, de la Nouvelle-Écosse, du Québec et du Manitoba.

## Périodes d'enquête

Les deux périodes d'enquête sont organisées de telle sorte qu'on recueille trois ensembles de données annuelles ayant trait aux perspectives, aux données réelles préliminaires et aux données réelles relativement aux dépenses en immobilisations et en réparations pour toutes les branches de l'économie (voir tableau explicatif 1).

Tableau explicatif 1

Collecte des données pour la série des dépenses en immobilisations

Données	Période de collecte	Date de publication
Perspectives (A <sup>1</sup> )	novembre (A-1) à février (A) <sup>1 1</sup>	février
Réelles préliminaires (A <sup>1-1</sup> )	novembre (A-1) à février (A) <sup>1 1</sup>	février
Réelles (A <sup>1-2</sup> )	mars (A-1) à octobre (A-1) <sup>1 1</sup>	février

1.

A = année civile courante.

## Suivi des questionnaires non retournés

Les mesures de suivi sont le plus souvent motivées par un faible taux de réponse dans un secteur ou dans une province ou un territoire en particulier. Pour commencer, on envoie un rappel général par la poste à toutes les unités n'ayant pas encore répondu. Si la situation ne se règle pas, on téléphone aux établissements faisant partie des secteurs les plus lacunaires pour les inciter à retourner le questionnaire rempli. Le fait d'intervenir activement auprès des établissements enquêtés qui ne répondent pas accroît le taux de réponse et permet ainsi d'obtenir des résultats plus précis pour l'estimation de la partie non enquêtée de la base de sondage (voir « Qualité des données, concepts et méthodologie — Méthodologie »).

## Vérification

Lorsque les questionnaires remplis sont reçus, le processus d'assurance de la qualité se poursuit par la vérification statistique des données, qui consiste à scruter les données pour détecter les incohérences internes et les anomalies par rapport aux autres enquêtes ou aux autres années.

La vérification des additions sert à repérer les chiffres indiqués en unités plutôt qu'en milliers, les pourcentages dont la somme n'arrive pas à 100 % et les totaux inexacts. La vérification des différences consiste à évaluer la vraisemblance des dépenses déclarées en les comparant avec les dépenses indiquées lors des enquêtes antérieures de la même année et d'une autre année. Pour l'enquête sur les données réelles auprès des répondants qui reçoivent le questionnaire long, la vérification du détail des biens permet de découvrir les établissements qui présentent des renseignements sur les biens d'une manière qui ne concorde pas avec les déclarations des questionnaires antérieurs ou avec les réalités communes dans l'industrie. Ces vérifications permettent d'identifier les données pour confirmation selon des seuils fixés après évaluation de la



portion couverte par l'enquête par rapport à l'industrie et selon l'emplacement géographique. De plus, des données sur les grands projets et sur les projets nouveaux sont recueillies dans les journaux, dans les revues spécialisées et dans les rapports d'industrie. Ces renseignements sont confrontés aux données déclarées et toute incohérence est identifiée pour confirmation.

Lorsque les données d'un établissement ont été identifiées à l'étape de la vérification, on consulte les autres données du questionnaire pour éclaircir le cas. Par exemple, une section du questionnaire demande au répondant d'indiquer les motifs des changements dans ses dépenses en immobilisations. Toutefois, si l'incohérence ne peut être expliquée à partir du questionnaire ni par d'autres renseignements portant sur l'industrie, on appelle directement l'établissement pour demander une confirmation. Les données sont alors mises à jour, soit par une correction, soit par une explication.

D'autres vérifications des microdonnées peuvent être effectuées pour les valeurs comptables brutes ou l'utilisation de la capacité. On fait une vérification des valeurs comptables brutes lorsque la valeur déclarée des immobilisations d'un établissement n'égale pas la valeur comptable brute déclarée antérieurement plus les investissements courants en immobilisations nouvelles déduction faite des aliénations. Dans un cas de ce genre, on appelle l'établissement pour qu'il confirme ou corrige les données. La vérification de l'utilisation de la capacité consiste à repérer tous les établissements manufacturiers et miniers qui fonctionnent à une capacité inférieure aux perspectives. Si les déclarations antérieures présentent des différences considérables par rapport aux données du questionnaire examiné, on appelle l'établissement pour qu'il confirme ou corrige les chiffres.

## Évaluation des macrodonnées

L'exercice d'estimation (voir « Estimation ») permet d'obtenir un ensemble complet de données pour les parties enquêtées et non enquêtées de l'univers (base de sondage), donc il est possible de commencer l'analyse des tendances des divers secteurs. On commence par une évaluation des changements annuels (ou en pourcentage) dans chaque industrie pour repérer les provinces et territoires comportant des industries ou des sous-industries qui présentent une activité inhabituelle. Ce genre d'analyse met aussi en relief les industries qui ont le plus de poids dans les totaux canadiens.

La macroanalyse se poursuit par l'ajout des facteurs pouvant avoir un effet sur les dépenses dans une province ou dans une industrie en particulier. Ces renseignements supplémentaires peuvent prendre la forme d'indices économiques comme le PIB, la productivité, l'utilisation de la capacité, les bénéfices ou l'innovation technologique. D'autres facteurs peuvent influencer sur les dépenses, par exemple les politiques gouvernementales (politique fiscale, politique monétaire, subventions) ou les données propres à une industrie comme le nombre de mètres forés, l'import-export et les permis de construction. Sans établir un lien de cause à effet, cette analyse vise à mettre en rapport les informations ayant directement et indirectement trait à l'industrie avec les tendances récentes en matière de dépenses d'immobilisations. À la suite de cet exercice, les industries qui présentent une activité exceptionnelle sont soumises à des évaluations de microdonnées plus poussées en vue d'expliquer les différences importantes entre les années.

## Méthodologie

### Introduction

L'enquête sur les dépenses en immobilisations (EDI) produit des données sur les investissements qui se font au Canada et dans tous les types d'industries canadiennes. Ces données sont recueillies deux fois par année, à deux moments bien précis. Ceci permet d'avoir un suivi sur les perspectives et les réalisations en termes d'investissements et ce, sur une base annuelle. Un seul et même échantillon sert à recueillir des données pour trois années financières distinctes. Un premier questionnaire est envoyé aux unités de l'échantillon en mars de l'année financière A. Ce questionnaire sert à recueillir les données réelles de l'année financière A-1 qui vient de se terminer. Un deuxième questionnaire est ensuite envoyé à ces mêmes unités au mois d'octobre suivant,

toujours dans l'année financière A. Ce questionnaire sert à recueillir les données réelles provisoires de l'année financière A qui s'achève bientôt et les données sur les perspectives pour l'année financière A+1. Notons que cet échantillon est tiré au mois de novembre de l'année financière A-1.

De même qu'un échantillon recueille des données pour trois années financières différentes, une année financière est couverte par trois échantillons différents. Un premier échantillon produit des données sur les perspectives pour une année financière A. Un an plus tard, un second échantillon produit des données réelles préliminaires pour cette même année financière A. Enfin un an plus tard, un troisième échantillon produit les données réelles pour cette même année financière A.

À chaque année, la Division des investissements et stock de capitaux (DISC) publie, au mois de février de l'année A, les résultats de l'enquête sur les données réelles (ER) de l'année financière A-2, de l'enquête sur les données réelles préliminaires (ERP) de l'année financière A-1 et de l'enquête sur les perspectives (EP) de l'année financière A.

Lors des enquêtes EP et ERP, les variables d'intérêt sont les capitaux immobilisés pour de la nouvelle construction (CC) ainsi que les capitaux immobilisés pour de la nouvelle machinerie et du nouvel équipement (CM). Lors de l'enquête ER, on ajoute les capitaux immobilisés pour la réparation reliée à la construction (RC) et les capitaux immobilisés pour la réparation reliée à la machinerie et à l'équipement (RM). De plus, lors de l'enquête ER, on produit des estimations plus détaillées pour les nouveaux capitaux. En effet, des estimations par type d'actifs sont également disponibles dans la publication **Dépenses en immobilisations par type d'actif**, no 61-223-X au catalogue.

## Méthodologie par secteur industriel

Comme dans toute enquête qui couvre plusieurs secteurs industriels, la méthodologie de l'EDI diffère d'un secteur à un autre et devient par conséquent très fastidieuse à expliquer à l'intérieur d'une même section. Voici comment se scinde la méthodologie pour les différents secteurs industriels selon le Système de classification industrielle de l'Amérique du Nord (SCIAN) :

Secteur 11, sous-secteurs 111 et 112 (Industrie de la production animale et agricole) :

- L'enquête est menée par la Division de l'agriculture (DA) qui ajoute les questions d'investissements à certaines de leurs enquêtes auprès des fermiers. Les données sont traitées par la DA et les estimations sont rapatriées dans la publication biannuelle. Consultez la section « Données non enquêtées » sous la rubrique Qualité des données, concepts et méthodologie — Sources pour plus de détails.

Secteur 11, sous-secteur 114 (Industrie de la pêche, de la chasse et du piégeage) et secteur 23 (Industrie de la construction):

- Il n'y a pas d'enquête. Les données publiées sont basées sur des indicateurs économiques. Consultez la section « Données non enquêtées » sous la rubrique Qualité des données, concepts et méthodologie — Sources pour plus de détails.

Secteur 91, sous-secteur 913 (Administrations locales):

- L'enquête est menée par la Division des institutions publiques (DIP) qui en profite pour demander la distribution des dépenses en investissements par fonction qui servira pour leur propre publication "Finance du secteur public". Les données sont cependant traitées par la DISC et sont habituellement dans le même format que la plupart des données recueillies par la DISC. Pour plus de détails quant à la méthodologie d'échantillonnage, consultez Pandher (1995). Notons que dans le cas du Québec, un arrangement spécial permet d'obtenir les valeurs d'investissements pour la province.

Secteur 21, sous-secteurs 211 (Pétrole brut et gaz naturel) et 212 (Mines) et secteur 91, sous-secteurs 911, 912 et 914 (Gouvernements fédéraux, provinciaux et territoriaux et autochtones) :

- Un échantillon suivant une méthodologie complètement basée sur un modèle a été préservé. Le traitement est le même que pour le reste de l'échantillon à quelques exceptions près. Pour plus de détails, consultez Lacroix (1991).

Secteur 21, industrie canadienne 213119 (Services reliés aux mines), secteur 55, industrie canadienne 551114 (Bureaux-chefs) et secteur 81 sous-secteur 814 (Ménages privés) :

- Il n'y a pas d'enquête ni aucune estimation pour ces secteurs.

Autres secteurs industriels:

- On utilise la méthodologie décrite dans cette section, en particulier une méthode d'estimation assistée d'un modèle.

Les sections suivantes présentent la méthodologie utilisée pour l'échantillonnage, la calendrialisation, l'imputation et l'estimation des autres secteurs. L'information concernant la méthodologie des secteurs industriels autres que ceux décrits au dernier point, est disponible dans les documents de référence cités.

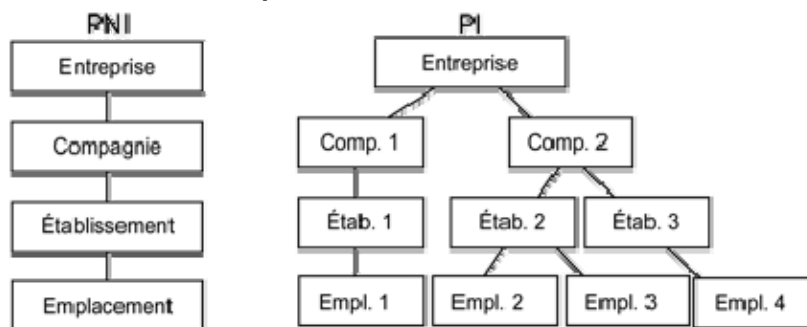
## Base de sondage

La base de sondage est principalement formée du Registre des entreprises (RE) développé à Statistique Canada. La Division du Registre des entreprises (DRE) est chargée d'en faire l'entretien et la mise à jour. Le registre est utilisé par un grand nombre d'enquêtes qui ne manquent pas de lui retourner de la rétroaction pour s'assurer que les plus récents changements dans le monde des entreprises soient incorporés au RE dans les plus brefs délais.

On retrouve sur le RE les unités nécessaires à l'établissement de notre base de sondage finale. La hiérarchie s'y lit comme suit : Entreprise - Compagnie - Établissement - Emplacement. Une entreprise peut avoir plusieurs compagnies qui peuvent chacune avoir plusieurs établissements qui peuvent à leur tour avoir plusieurs emplacements. Cette structure dite « statistique » est en fait une modélisation de la structure opérationnelle décrite par l'entreprise elle-même. Selon l'information disponible pour chaque niveau de la structure opérationnelle, on définit le niveau statistique correspondant. Par exemple, pour être considéré comme un établissement, on doit pouvoir fournir au RE les salaires et taux de rémunération, le revenu et les intrants principaux dans le processus d'exploitation.

Dans le cas des unités formant la portion non intégrée (PNI) du RE, la structure statistique est linéaire : une entreprise est reliée à une seule compagnie, à un seul établissement et à un seul emplacement. Dans la portion intégrée (PI), la structure peut être linéaire mais est généralement plus complexe. Le schéma [1](#) illustre les deux structures.

**Figure 1**  
**Structures statistiques**



Dans le cadre de l'EDI, l'unité d'échantillonnage choisie est l'établissement, celle-ci correspondant le mieux au besoin de collecte et de divulgation des données d'investissements. Pour plus de détails concernant le RE, consultez Cuthill (1996).

Lorsque l'on tire l'échantillon au mois de novembre, on prend une nouvelle « image » du RE. Avec l'avènement de l'enquête unifiée sur les entreprises, le RE a amélioré sa couverture et cette « image » s'en trouve d'autant plus complète et à jour. L'enquête sur les dépenses en immobilisations fait partie de l'enquête unifiée et se sert de cette nouvelle image pour procéder à l'échantillonnage.

Puisque les questionnaires sont envoyés au mois de mars et au mois d'octobre suivants, et étant donné la nature dynamique des entreprises, il est certain que de nouveaux projets se mettront en branle après la sélection originale de l'échantillon. Pour être certain que d'importants investissements ne soient pas « oubliés », on ajoute des unités à l'échantillon, même après le premier envoi postal, lorsque le projet est jugé assez important. Ces « nouveaux projets » comme nous les appelons, sont trouvés grâce à la lecture de journaux, de rapports de compagnie ou encore grâce aux listes de permis de construction. Ils sont échantillonnés avec certitude et nous permettent d'éviter de sous-estimer grossièrement la valeur des investissements des industries où on les retrouve.

Il est à noter que certaines unités qu'on tient à mettre dans l'échantillon, tels que les nouveaux projets, ont une information incomplète. Le revenu, qui est connu pour toutes les unités de la base, peut être inconnu pour ces unités. Comme le revenu est utilisé au cours de différents processus (imputation, estimation, etc.), ces unités sont regroupées afin de les traiter séparément lors du traitement des données.

## **Regroupement**

Avant de procéder à l'échantillonnage, toutes les unités dans le secteur privé qui ne faisaient pas partie des industries minières et manufacturières ont été regroupées selon la méthode suivante. Tous les établissements opérant dans la même province, dans le même secteur industriel codé à six chiffres et sous la même entreprise ont été regroupés en un seul super-établissement. Le revenu du super-établissement est la somme de tous les revenus des établissements qui le composent et le reste de l'information est tiré de la tête du regroupement, soit le bureau-chef si c'est possible ou, sinon, l'établissement avec le plus grand revenu. Pour le secteur public, toutes les unités font partie de l'échantillon.

Une fois le nouvel univers construit avec les nouveaux super-établissements, toutes les unités qui ont un revenu inférieur à une certaine limite sont éliminées de la base à moins qu'elles ne correspondent à des bureaux-chefs ou à des laboratoires, ces unités étant choisies avec certitude. Cette procédure est mise en place pour éviter de « perdre » ces unités qui ne génèrent pratiquement aucun revenu, mais qui pourraient être l'objet d'investissements substantiels.

La limite inférieure déterminant les unités non-observées est construite en fonction de la province et du secteur industriel. Celle-ci varie de 100 000 \$ à 3 185 000 \$ selon la taille des unités qui composent l'industrie et la province. En gros, la limite est calculée de sorte à exclure au maximum 10 % du revenu total du secteur industriel et provincial de l'échantillonnage. Ceci permet d'alléger le fardeau des répondants de petite taille et répond ainsi aux lignes directrices du bureau. La portion non couverte est estimée en utilisant des données administratives lorsque disponibles (se référer à la section « Estimation » pour plus de détails).

Lorsque tous les regroupements ont été effectués et que les petites unités ont été éliminées, la population observée est prête à être stratifiée.

## **Échantillonnage**

L'échantillonnage se divise selon les trois parties traditionnelles : la stratification, la répartition et la sélection. Celles-ci sont décrites dans le texte qui suit.

### **Stratification**

On a tout d'abord stratifié selon le lieu géographique, la classification industrielle et également en fonction du pays d'origine du contrôle afin de tenir compte de nouveaux besoins. La division géographique se fait selon les 13 provinces et territoires, sans autre raffinement (aucune stratification infra-provinciale). Douze pays d'origine de contrôle ont été considérés dans la stratification cette année : Canada, États-Unis, Allemagne, Japon, France, Grande-Bretagne, Suède, Italie, Pays-Bas, Chine, Hong Kong et Australie. Les autres pays ont été regroupés. Pour ce qui est de la stratification industrielle, le SCIAN de 1997 est utilisé selon le niveau

requis pour les estimations. Si par exemple, pour une certaine industrie, le niveau le plus désagrégé publié correspond au SCIAN à 3 chiffres, ce sera le niveau de stratification. Notons que pour le reste de la section, le SCIAN à 6 chiffres sera abrégé par SCIAN-6, le SCIAN à 5 chiffres par SCIAN-5, etc...

Le tableau explicatif [1](#) indique, par industrie, quels sont les niveaux de publication possibles les plus désagrégés pour les estimations provinciales et canadiennes.

Tableau explicatif 1

**Niveaux de publication les plus désagrégés**

Secteur industriel	Code de SCIAN	Niveau de publication SCIAN
Agriculture, foresterie, pêche et chasse	11	3
Extraction minière, de pétrole et gaz	21	3 à 6
Services publics	22	4
Fabrication (SCIAN -3 316 et 323)	31-33	3
Commerce de gros	41	3
Commerce de détail	44-45	3
Transport et entreposage	48-49	3
Information et culture	51	3
Finance et assurances	52	3
Services immobiliers, de location et de location à bail	53	4
Services professionnels, scientifiques et techniques	54	4
Gestion de sociétés et d'entreprises	55	2

Secteur industriel	Code de SCIAN	Niveau de publication SCIAN
Services administratifs, de soutien, de gestion des déchets et d'assainissement	56	3
Services d'enseignement	61	4
Soins de santé et assistance sociale	62	3
Arts, spectacles et loisirs	71	3
Hébergement et services de restauration	72	3
Autres services	81	3
Administrations publiques	91	3

Les niveaux de publication provinciaux correspondent aux secteurs industriels à l'exception de l'industrie de la fabrication où on publie au niveau SCIAN-3 pour quatre provinces : Québec, Ontario, Alberta et Colombie-Britannique.

## Répartition

Une fois la stratification initiale mise en place, on calcule le coefficient de variation (CV) (voir « Estimation » pour plus d'information sur les CV) à viser en utilisant la variable revenu de façon à atteindre le CV fixé pour le niveau de publication le plus désagrégé, dans notre cas, par province et différents niveaux industriels SCIAN tel que définis précédemment. Un exemple aide à mieux comprendre la situation.

Supposons qu'on veuille publier des estimations pour le secteur industriel 72 (Hébergement et services de restauration) pour lequel on publie au niveau SCIAN-3 pour le Canada et au niveau de l'industrie complète par province / territoire. On construit alors le tableau 2, où le nombre de provinces a été simplifié à 3 et le nombre de sous-secteurs (SS) SCIAN-3 pour l'industrie au complet est 2 (721 et 722).

Tableau explicatif 2  
**Croisements de publication pour le secteur 72**

	Province 1	Province 2	Province 3	CV
SS721	...	...	...	15%
SS722	...	...	...	15%
CV	15%	15%	15%	...

La stratification initiale correspond à chacune des cellules du tableau 2 et les marginales correspondent aux estimations qu'on désire publier. Si on désire, par exemple, publier des estimations avec un CV cible de 15 %, on doit d'abord calculer le CV à viser pour chacune des cellules de telle sorte que les CV des marginales soient respectés.

Avant de pouvoir calculer le CV nécessaire au niveau des cellules pour atteindre le CV fixé au niveau des marginales, on doit d'abord ajuster ces CV marginaux. En effet, on ne peut obtenir des CV de 15 % dans les deux directions, car lorsque l'on fixe la variance dans une direction pour obtenir le CV visé, on fixe automatiquement la variance (donc le CV) pour l'autre direction et on « subit » le CV résultant. Sachant que les CV des deux directions ne peuvent être simultanément égaux au CV visé (à moins d'un hasard), nous avons choisi de minimiser la distance des CV des marginales au CV cible. On obtient donc, dans une direction, un CV résultant supérieur au CV cible et dans l'autre, un CV inférieur à ce même CV. Ceci est fait en minimisant la distance entre les CV résultants et le CV cible sous la contrainte d'avoir des variances égales dans les deux directions. D'une façon mathématique:

**Figure 2**

$$\text{Minimiser } (CV^C - CV^A)^2 + (CV^C - CV^B)^2$$

sous la contrainte  $V^A = V^B$

où  $CV^A$  et  $CV^B$  représentent les CV atteignables dans les deux directions,  $CV^C$  représente le CV cible et  $V^A$  et  $V^B$  représentent les variances dans les deux directions.

Appelons le CV résultant le nouveau CV cible. Dans l'exemple précédent, on pourrait se retrouver avec de nouveaux CV cibles comme dans le tableau 3.

Tableau explicatif 3  
**Nouveau CV cibles (les plus près du CV visé)**

	Province 1	Province 2	Province 3	CV
SS721	...	...	...	11%
SS722	...	...	...	11%
CV	18%	18%	18%	...

Pour atteindre le nouveau CV cible, on doit calculer ce que devraient être les CV visés pour chacune des strates initiales en utilisant l'algorithme itératif du quotient tel que décrit dans Latouche (1988).

En utilisant à nouveau les lettres A et B pour désigner les deux directions (A la direction géographique et B la direction industrielle par exemple), on recalcule les CV des cellules jusqu'à ce que la combinaison des CV sur une même ligne ou une même colonne soit assez près du CV cible de la marginale correspondante.

**Figure 3**

$$CV_r^B(\hat{Y}_{ij}) = CV_{(r-1)}^A(\hat{Y}_{ij}) * \frac{CV(\hat{Y}_j)\hat{Y}_j}{\sqrt{\sum_j (CV_{(r-1)}^A(\hat{Y}_{ij}))^2 \hat{Y}_{ij}^2}}$$

$$CV_r^A(\hat{Y}_{ij}) = CV_{(r-1)}^B(\hat{Y}_{ij}) * \frac{CV(\hat{Y}_i)\hat{Y}_i}{\sqrt{\sum_j (CV_{(r-1)}^B(\hat{Y}_{ij}))^2 \hat{Y}_{ij}^2}}$$

où :

- r désigne l'itération courante,
- r-1 désigne l'itération précédente,
- i. désigne la marginale dans la direction A,
- .j désigne la marginale dans la direction B,
- ij désigne un croisement des directions A et B et
- Y correspond au total de la variable revenu pour un groupement donné.

L'algorithme s'arrête lorsque le critère de convergence (0,1 %) est rencontré ou après un maximum de 10 itérations. Notons ici que l'algorithme converge très rapidement et on atteint presque à coup sûr le CV visé au niveau des marginales. Le tableau 4 illustre le résultat du procédé itératif.



Tableau explicatif 4  
**CV des cellules après itérations**

	Province 1	Province 2	Province 3	CV
SS721	20%	23%	24%	11%
SS722	17%	20%	21%	11%
CV	18%	18%	18%	...

Maintenant que le CV est fixé pour chacune des strates initiales (celles-ci correspondent aux cellules du tableau précédent), on peut les stratifier en deux strates de taille : grande taille où l'échantillonnage se fait avec certitude et petite taille où l'échantillonnage se fait selon une probabilité de sélection permettant d'atteindre le nouveau CV cible. La méthode préconisée pour séparer les cellules en deux est celle d'Hidiroglou (1986) qui a le mérite de minimiser la taille d'échantillon tout en atteignant le CV cible. La technique est simple : on part de l'équation qui donne le CV de la strate initiale

**Figure 4**

$$CV(\hat{Y})^2 = \frac{(N-t)*(N-n(t))}{(n(t)-t)} \frac{S^2_{(N-t)}}{\hat{Y}^2}$$

où N désigne la taille de la population,  
n(t) désigne le nombre total d'unités à échantillonner,  
t désigne le nombre total d'unités dans la strate à tirage complet,  
 $S^2_{(n-t)}$  désigne la variance dans la strate à tirage partiel et  
Y correspond au total de la variable revenu pour la strate.

On peut la réécrire de façon à isoler n(t), le nombre total d'unités à échantillonner en fonction de t, le nombre d'unités échantillonnées avec certitude :

**Figure 5**

$$n(t) = t + \frac{(N-t)^2 S^2_{(N-t)}}{CV^2 \hat{Y}^2 + (N-t) S^2_{(N-t)}}$$

Il s'agit alors de bien comprendre la fonction pour trouver son point minimum. Celui-ci peut être atteint selon un processus itératif qui calcule, après avoir convergé, les deux paramètres suivants : la borne qui sépare la strate initiale en deux strates finales ainsi que la taille d'échantillon de chacune des strates. On aura t unités dans la strate à tirage complet et n(t) - t unités à tirer dans la strate à tirage partiel. On aura ainsi tiré le nombre minimal d'unités pour atteindre le CV cible fixé.

Il est fort probable qu'on n'obtienne pas exactement le CV cible au niveau des cellules. Le CV atteint est habituellement près, mais peut être jusqu'à 2 % au-dessus du CV cible pour certaines cellules. Ceci a pour effet de changer légèrement les CV visés au niveau des marginales. Le tableau 5 reprend les résultats du tableau 4 après l'application de l'algorithme d'Hidiroglou.

Tableau explicatif 5

**CV final des cellules après itérations**

	Province 1	Province 2	Province 3	CV
SS721	20,1 %	22,8 %	24%	10,8 %
SS722	17,2 %	21,5 %	20,4 %	11,7 %
CV	18,1 %	18,9 %	17,8 %	...

Lorsque cette étape est complétée, on peut alors procéder à la sélection proprement dite de l'échantillon.

**Sélection**

Pour les strates à tirage partiel, la sélection se fait selon un processus aléatoire simple sous les contraintes d'un chevauchement minimum avec l'échantillon de l'enquête unifiée auprès des entreprises (EUE) (Pour plus de détails sur cette enquête, se référer à Simard et al. (2001)), d'une fraction de sondage minimale de 1 % et d'un minimum de 3 unités échantillonnées par strate. Dans les strates à tirage complet, toutes les unités sont échantillonnées avec certitude.

**Vérification des données**

Une fois l'échantillon tiré, un questionnaire est envoyé par la poste invitant le répondant à le retourner dûment rempli. Les unités n'ayant pas répondu font l'objet d'un suivi postal et téléphonique pour s'assurer d'obtenir leurs investissements. Un effort spécial est fait pour les unités faisant partie de strates à tirage complet.

Une fois les données saisies, certaines vérifications sont faites pour chaque établissement. Entre autres, plusieurs règles de cohérence sont en place permettant de s'assurer que si certains champs sont codés, tous ceux qui y sont reliés sont également codés. On peut par exemple, vérifier que la somme des parties égale le tout, que certaines cellules sont bien remplies, etc...

Certaines vérifications ont trait directement aux données d'investissement. Par exemple, si des données historiques sont disponibles, certaines règles de tolérance sont appliquées.

Lorsque aucune donnée historique n'est disponible, tous les répondants présentant des investissements de 10 000 000 \$ et plus font l'objet de vérifications approfondies. Notons que ces règles sont sujettes à changements.

Finalement, un grand nombre de règles de vérification de nature qualitative (et non quantitative) sont aussi en place. Pour plus de détails sur les règles de vérification, consulter Corneau (1995).

**Calendrialisation**

Une fois la collecte effectuée et les données vérifiées, on procède à la calendrialisation des données. Ce processus permet d'obtenir des données pour la période de janvier à décembre de l'année de référence dans

les cas où le répondant fournit ses données sur une autre base. En effet, pour diminuer le fardeau de réponse du répondant, on accepte qu'il fournisse ses données sur la base de sa période fiscale. Pour une année donnée, sa période fiscale doit se terminer entre le 1er janvier de l'année visée et le 31 mars suivant l'année visée.

Pour éviter de produire des estimations se rapportant à plusieurs périodes fiscales différentes, on a recours à la calendrialisation. Le principe directeur est assez simple : tout d'abord "briser" les données annuelles en données mensuelles, extrapoler si nécessaire puis sommer les mois qui composent l'année d'intérêt pour ainsi obtenir les données calendrialisées du répondant.

La méthode développée par Cholette (1984) est utilisée pour "briser" les données de façon mensuelle et faire l'interpolation. La méthode s'apparente à la technique d'annualisation de données ("Benchmarking"). En gros, on peut résumer l'algorithme comme suit:

On cherche à minimiser la fonction

**Figure 6**

$$O(x) = \sum_{m=2}^T ((x_m - x_{m-1}) - (z_m - z_{m-1}))^2$$

de telle sorte que la somme des valeurs mensuelles ( $x_m$ ) sur la période fiscale soit égale à la valeur rapportée par le répondant.

La série des  $z_m$  correspond à l'information auxiliaire connue pour le répondant comme possiblement son cycle ou sa tendance. Pour l'enquête, on n'utilise pas cette option et la série est tout simplement mise égale à une constante ce qui correspond tout simplement à minimiser le changement d'un mois à l'autre (tout en s'assurant de respecter le total fiscal).

Le nombre de mois disponibles ( $T$ ) sur lequel on calcule la fonction à minimiser dépend de l'information historique que l'on détient sur le répondant. Cependant, étant donné qu'habituellement un répondant reçoit au moins deux questionnaires correspondant à deux années civiles distinctes,  $T$  devrait être au minimum égal à 24. Les périodes qui ne sont pas couvertes par les données fiscales (au début ou à la fin de la série) sont extrapolées par la dernière (ou la première) valeur mensuelle calculée. Le reste du traitement décrit peut donc être appliqué sur les données calendrialisées ou sur les données civiles des répondants.

## Détection de données aberrantes

Une fois les données transformées sur une base calendrier, on procède à la détection de données aberrantes. La détection peut se faire à quatre niveaux. On commence la détection au niveau le plus désagrégé. S'il n'y a pas au moins 25 unités à ce niveau, on passe au niveau suivant. Dans la définition de ces niveaux, jusqu'à trois variables peuvent être impliquées : le niveau industriel, la taille et la région géographique.

En ce qui a trait à la taille, il y a trois catégories : la strate à tirage complet avec revenu connu, la strate à tirage complet avec revenu inconnu et la strate à tirage partiel.

Quant à la région géographique, l'unité se retrouve soit dans les grosses provinces (Qc, Ont., Alb. et C.-B.), dans les provinces moyennes (N.-É., N.-B., Man. et Sask.) ou dans les petites provinces (I.-P.-É., Yn, T.N.-O., Nt et T.-N.-L.).

Les quatre niveaux de détection sont :

- Niveau 1 : SCIAN-3 \* Taille \* Qc, Ont., Alb., C.-B., petites et moyennes provinces (séparées)
- Niveau 2 : SCIAN-3 \* Taille \* grosses provinces et petites et moyennes provinces (ensemble)
- Niveau 3 : SCIAN-3 \* Taille \* Canada
- Niveau 4 : Secteur \* Canada

Lorsque dans une industrie, on publie au niveau du secteur, on commence la détection au niveau le plus agrégé, soit le niveau 4.

De plus, la détection des valeurs aberrantes se fait avant et après imputation. La détection suite à l'imputation se fait avec les données imputées et permet ainsi de détecter les données imputées aberrantes.

La méthode d'Hidiroglou-Berthelot (1986) est utilisée pour les détecter. L'établissement "i" est considéré aberrant si une des deux relations est vérifiée:

$$Y_i < M - C * DQ_1$$

$$Y_i > M + C * DQ_3$$

où

$$DQ_1 = \text{Max}(M - Q_1, |A * M|),$$

$$DQ_3 = \text{Max}(Q_3 - M, |A * M|),$$

M est la médiane (c'est à dire le point où exactement 50 % des établissements se trouvent de part et d'autres),

Q<sub>1</sub> est le premier quartile (25 % des établissements sont plus petits et 75 % sont plus grands),

Q<sub>3</sub> est le troisième quartile (75 % des établissements sont plus petits et 25 % sont plus grands),

A et C prennent respectivement les valeurs : 0,5 et 20.

La détection des données aberrantes est effectuée pour les quatre (ratios) suivants. CC calendrialisé sur le revenu, CM calendrialisé sur le revenu, CC sur le revenu, CM sur le revenu. Si un établissement est trouvé aberrant pour un de ces ratios, alors il est automatiquement considéré aberrant pour les deux variables d'investissements, CC et CM, calendrialisées et fiscales. Dans le cas de l'ER, la même procédure est répétée pour les variables RC et RM en plus des variables CC et CM.

## Imputation

Les enregistrements trouvés aberrants ne sont pas imputés puisque les règles de cohérence ont déjà été appliquées et que l'on considère valide l'investissement déclaré par le répondant. On ne fait qu'exclure ces enregistrements du calcul de moyenne lors de l'imputation des non répondants. De plus, si certains des établissements trouvés aberrants font partie de strates à tirage partiel, alors ils sont promus dans la strate à tirage complet avec revenu connu et la probabilité de sélection des unités résiduelles est recalculée.

Pour les enregistrements à imputer, trois méthodes d'imputation sont utilisées pour procéder à l'évaluation des données manquantes. Il n'existe pas d'imputation partielle : les deux variables d'intérêts, CC et CM (on ajoute RC et RM dans le cas de l'ER) sont disponibles ou manquantes pour chacun des établissements. Les trois méthodes permettent donc d'imputer parallèlement toutes les variables. La première méthode est tout simplement la substitution par valeur historique. Pour les enquêtes suivantes, on utilise la valeur historique, en autant que celle-ci couvre la même année de référence :

$$Y_{its} = Y_{it(s-1)}$$

où t correspond à l'année de référence, s correspond à l'enquête courante, s-1 à l'enquête la plus récente pour laquelle des données ont été rapportées et y est la variable d'intérêt.

Dans le cas de l'enquête sur les perspectives (EP), comme il s'agit de la première enquête pour une année de référence et donc qu'aucune donnée historique n'est disponible pour cette même année, on utilisera l'information historique de l'année précédente :

$$Y_{its} = Y_{i(t-1)(s-1)}$$

Où t-1 correspond à l'année de référence précédente.

Notons que cette dernière imputation est aussi utilisée pour les variables RC et RM puisque celles-ci sont requises seulement pour l'enquête sur les données réelles et donc, aucune valeur historique n'est disponible pour la même année de référence.

Dans les deux cas, l'imputation est faite (là où elle est possible) avant de procéder à la calendrialisation. Ainsi les données imputées historiquement à partir de données pouvant provenir d'une période autre que l'année calendrier sont également calendrialisées.

La deuxième méthode est utilisée lorsqu'aucune donnée historique n'est disponible pour une unité. Dans ce cas, on impute en utilisant la méthode du quotient des valeurs courantes :

**Figure 7**

$$y_{it} = \frac{\bar{y}_t}{\bar{x}_t} x_{it}$$

où x correspond au revenu.

La troisième méthode est utilisée pour les unités sans valeur historique et dont le revenu n'est pas connu. Dans ce cas, on utilise l'imputation par la moyenne des valeurs courantes :

**Figure 8**

$$y_{it} = \bar{y}_t$$

Un facteur important lors du calcul de la valeur imputée est le niveau auquel se fait l'imputation. En effet, l'imputation a lieu si le groupe d'imputation comprend au moins 10 établissements dont le questionnaire est complet et que ceux-ci représentent au moins 25 % des unités du groupe.

## Groupes d'imputation

Le groupe d'imputation initial correspond à la strate utilisée pour l'échantillonnage une fois mise à jour avec les nouvelles données recueillies. Si l'une ou l'autre des contraintes précédentes (10 unités, 25 % des unités) n'est pas satisfaite, on passe à un groupe d'imputation plus agrégé se situant dans le même groupe industriel et dans le même groupe de taille mais où toutes les provinces sont combinées. Comme dans la détection des valeurs aberrantes, les tailles possibles sont strate à tirage complet avec revenu connu, strate à tirage complet avec revenu inconnu et strate à tirage partiel.

Si encore une fois les contraintes ne sont pas satisfaites, on regroupe les industries. On combine par exemple tous les SCIAN-6 d'un même SCIAN-5. On demeure au niveau canadien et à l'intérieur du même groupe de taille. Le niveau le plus agrégé qu'on peut atteindre correspond aux groupements de tous les SCIAN-3 d'un même secteur, au niveau canadien, pour un groupe de taille où au dernier niveau les tailles strate à tirage complet « avec revenu connu » et « inconnu » sont regroupées. Deux exemples permettront de mieux comprendre.

Si un établissement de l'industrie minière canadienne 212114 en Ontario faisant partie du groupe à tirage partiel est à imputer, on a la séquence suivante:

212114 - Ontario - strate à tirage partiel

212114 - Canada - strate à tirage partiel

21211 - Canada - strate à tirage partiel

2121 - Canada - strate à tirage partiel

212 - Canada - strate à tirage partiel

21 - Secteur de l'extraction minière, de pétrole et gaz - Canada - strate à tirage partiel

Si un établissement du secteur 55 (Gestion de sociétés et d'entreprises) au Québec faisant partie du groupe à tirage complet avec revenu inconnu est à imputer, on a la séquence suivante :

Secteur 55-Québec - strate à tirage complet (revenu inconnu)

Secteur 55-Canada - strate à tirage complet (revenu inconnu)

Secteur 55-Canada - strate à tirage complet (revenu connu et inconnu)

Notons aussi qu'un enregistrement imputé à un niveau désagrégé peut servir dans le calcul des moyennes lors de l'imputation d'un autre enregistrement à un niveau plus agrégé. Par exemple, si on réussit à imputer tous les enregistrements de l'Alberta au premier niveau d'imputation et qu'on doit passer au niveau suivant pour les enregistrements du Nouveau-Brunswick, ceux-ci seront imputés au niveau canadien et les enregistrements imputés de l'Alberta seront utilisés dans le calcul des moyennes au niveau canadien.

Une fois les valeurs manquantes des établissements imputées, on peut passer à l'étape d'estimation.

## **Estimation**

L'estimation se fait par la méthode du quotient avec le revenu comme variable de contrôle. Cette méthode assure que le poids final multiplié par le revenu de chaque unité de l'échantillon, respectera le total connu de la variable revenu pour la population complète du groupe. Les groupes utilisés correspondent cette fois-ci au plus bas niveau industriel publié, à l'intérieur d'un même groupe de taille, au niveau canadien. La différence avec la strate originale est le groupement au niveau canadien. L'exemple suivant permet de mieux comprendre.

Pour un établissement dont la strate correspond au SCIAN-3 323 du secteur de la fabrication dans la Nouvelle-Écosse pour le groupe de strate à tirage partiel, on utilise le groupe d'estimation

323 - Canada - strate à tirage partiel

Lors de l'enquête, il est possible qu'un établissement soit reclassifié dans une nouvelle industrie ou dans une nouvelle province. On utilise cette nouvelle classification pour définir les domaines de publication et c'est cette classification qui déterminera où les investissements seront présents dans les tableaux finaux. L'exemple suivant permet de mieux comprendre.

Si un établissement, échantillonné au Québec selon le SCIAN-3 411, est retrouvé en Ontario selon le SCIAN-3 444, il aura les caractéristiques suivantes :

strate : 411 - Québec

groupe pour calcul de données aberrantes : 444 - Ontario

groupe d'imputation initial : 444 - Ontario

groupe d'estimation : 411 - Canada

domaine de publication : 444 - Ontario

Figure 9

Voici la formule de l'estimateur par le quotient utilisé

$$\hat{Y}_d = \sum_h \sum_{i \in s_h} w_i y_i(d)$$

où pour chaque unité  $i$  d'un groupe  $g$ ,

$$w_i = D_i \times G_i, D_i = \frac{Nh}{nh}, G_i = \frac{\sum_{j \in P_g} x_j}{\sum_{j \in P_g} p_j} \text{ et } y_i(d) = \begin{cases} y_i & \text{si } i \in d \\ 0 & \text{si non} \end{cases}$$

où :

$x$  est la variable auxiliaire (le revenu),

$h$  désigne la strate,

$g$  désigne le groupe d'estimation,

$d$  désigne le domaine de publication,

$n$  désigne la taille échantillonnale,

$N$  désigne la taille de la population,

$s$  désigne l'échantillon,

$P$  désigne la population,

$w$  désigne le poids final,

$D$  désigne le poids de sondage,

$G$  désigne le poids de contrôle ("G-weight")

$y$  est la variable d'intérêt (l'investissement) et

$p$  désigne la probabilité de sélection.

Notons que le calcul du poids-G est de telle sorte que le poids final  $w_i$  est borné inférieurement à 1. Ainsi on évite que pour des raisons mathématiques les valeurs de certains répondants valent moins que les valeurs rapportées une fois pondérées.

## Estimation de la variance et calcul de CV

La variance est estimée à l'aide de la formule de linéarisation de Taylor dans le cas de l'estimateur par le quotient. Celle-ci est disponible dans Estevao (1991). En suivant la même notation que précédemment :

**Figure 10**

$$\hat{V}(\hat{Y}(d)) = \sum_h \frac{N_h - n_h}{n_h - 1} \frac{n_h}{N_h} \sum_{i \in S_h} (u_{hi} - \bar{u}_h)^2$$

$$\text{où } u_{hi} = \frac{N_h}{n_h} G_i \left( y_i(d) - x_i * \frac{\sum_{i \in S_g} y_i / p_i}{\sum_{i \in S_g} x_i / p_i} \right)$$

$$\text{et } \bar{u}_h = \frac{\sum_{i \in S_h} u_{hi}}{n_h}$$

Le coefficient de variation (CV) est calculé à l'aide du quotient:

$$CV(\hat{Y}(d)) = \frac{\sqrt{\hat{V}(\hat{Y}(d))}}{\hat{Y}(d)}$$

### Ajustement de l'estimation pour la portion non observée

Lorsque disponibles, on se sert de données administratives pour la portion non observée par l'enquête.

Pour l'enquête sur les données réelles, les données administratives des trois années précédentes sont utilisées afin de construire un modèle qui sert à dériver les dépenses en immobilisations.

Pour les enquêtes sur les perspectives et les données réelles provisoires, aucune donnée administrative ne couvre les périodes de référence de ces enquêtes. La portion non enquêtée est alors estimée en utilisant la tendance de la portion enquêtée entre les données réelles et les perspectives et données provisoires qu'on applique à l'estimation de la portion non observée calculée pour l'enquête sur les données réelles.

L'estimation de la portion non observée contribue en moyenne à 2 % de l'estimation totale.

### Indicateur de qualité

Lors de la publication des estimations, une échelle permet de distinguer entre les différentes qualités de précision. Celle-ci combine l'effet dû à l'échantillonnage (puisque l'on n'a pas effectué de recensement) et le taux d'imputation (chaque imputation (autre qu'historique) ajoute à l'incertitude des résultats). L'échelle est reproduite au tableau 6.

Tableau explicatif 6

#### Interprétation de la cote de qualité

	Taux d'imputation



	Taux d'imputation			
CV	0,00 à 0,10	0,10 à 0,33	0,33 à 0,60	0,60 et plus
0,00 à 0,05	A	B	C	F
0,05 à 0,10	B	C	D	F
0,10 à 0,15	C	D	E	F
0,15 à 0,25	D	E	F	F
0,25 à 0,50	E	F	F	F
0,50 et plus	F	F	F	F

**Note(s) :**

<sup>A</sup>Excellent; <sup>B</sup>Très bon; <sup>C</sup>Bon; <sup>D</sup>Acceptable; <sup>E</sup>À utiliser avec prudence; F Trop peu fiable pour être publié.

**À cause de considérations techniques, la côte de qualité ne sera pas disponible pour cette publication.**

### Confidentialité

Certaines règles de confidentialité sont évidemment utilisées pour supprimer toute information qui pourrait mener à la divulgation des données fournies par un répondant. Ces règles permettent à Statistique Canada de respecter son mandat de non-divulgation d'information fournie par les répondants. Les règles elles-mêmes sont confidentielles et ne sont pas disponibles pour consultation.

### Erreur d'échantillonnage et non due à l'échantillonnage

La différence entre l'estimation produite à partir de données échantillonnées et de données recensées est appelée erreur d'échantillonnage. Cette différence varie plus ou moins selon la taille de l'échantillon, la variabilité des dépenses, le plan de sondage et la méthode d'estimation. En général, un échantillon plus grand produit une erreur d'échantillonnage plus petite. Si la population est très hétérogène, une taille d'échantillon plus grande est requise pour produire une estimation fiable. L'erreur d'échantillonnage est mesurée par une quantité appelée écart-type. Cette quantité mesure la variabilité anticipée de l'estimation produite si on fait un échantillonnage répété des dépenses. La vraie valeur de l'écart-type est inconnue mais peut être estimée à partir de l'échantillon.

Une deuxième mesure de précision est le coefficient de variation (CV). Ce coefficient est simplement l'écart-type exprimé en pourcentage de la valeur de l'estimation. Il donne donc une mesure de précision relative et comparable entre différentes industries ou provinces. Notons qu'un plus petit CV indique une plus grande

fiabilité de l'estimation. (Voir la section « Qualité des données, concepts et méthodologie — Mesures de la qualité »).

En plus de l'erreur d'échantillonnage, il existe des erreurs non dues à l'échantillonnage sur lesquelles on tente de conserver un contrôle des plus stricts. Néanmoins, celles-ci existent toujours et ne sont pas comptabilisées lorsque l'on calcule le coefficient de variation et ne sont pas incluses dans cette mesure de précision. Certaines mesures telles que les taux de réponses, de couverture et d'imputation peuvent être utilisées comme indicateurs du niveau potentiel des erreurs non liées à l'échantillonnage.