Nº 12-001-X au catalogue
ISSN 1712-5685

Techniques d'enquête

Des modes multiples pour les enquêtes à des sources de données multiples pour les estimations

par Constance F. Citro

Date de diffusion : 19 décembre 2014

Statistique Canada Statistics Canada

Canada

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à infostats@statcan.gc.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros sans frais suivants :

- Service de renseignements statistiques 1-800-263-1136
- Service national d'appareils de télécommunications pour les malentendants 1-800-363-7629
- Télécopieur 1-877-287-4369

Programme des services de dépôt

- Service de renseignements 1-800-635-7943
- Télécopieur 1-800-565-7757

Comment accéder à ce produit

Le produit no 12-001-X au catalogue est disponible gratuitement sous format électronique. Pour obtenir un exemplaire, il suffit de visiter notre site Web à www.statcan.gc.ca et de parcourir par « Ressource clé » > « Publications ».

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « À propos de nous » > « Notre organisme » > « Offrir des services aux Canadiens ».

Publication autorisée par le ministre responsable de Statistique Canada

This publication is also available in English.

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, ses entreprises, ses administrations et les autres établissements. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Signes conventionnels

Les signes conventionnels suivants sont employés dans les publications de Statistique Canada :

.	indisponible pour toute période de référence
..	indisponible pour une période de référence précise
...	n'ayant pas lieu de figurer
0	zéro absolu ou valeur arrondie à zéro
0^s	valeur arrondie à 0 (zéro) là où il y a une distinction importante entre le zéro absolu et la valeur arrondie
^{p}	provisoire
^{r}	révisé
x	confidentiel en vertu des dispositions de la *Loi sur la statistique*
^{E}	à utiliser avec prudence
F	trop peu fiable pour être publié
*	valeur significativement différente de l'estimation pour la catégorie de référence (p<0,05)

Techniques d'enquête, décembre 2014
Vol. 40, n° 2, pp. 151-181
Statistique Canada, n° 12-001-X au catalogue

Des modes multiples pour les enquêtes à des sources de données multiples pour les estimations

Constance F. Citro[1]

Résumé

Les utilisateurs et les fournisseurs de statistiques officielles, ainsi que ceux qui en assurent le financement, veulent des estimations « plus vastes, plus approfondies, plus rapides, de meilleure qualité et moins coûteuses » (selon Tim Holt, ancien chef de l'*Office for National Statistics* du Royaume-Uni), attributs auxquels j'ajouterais « plus pertinentes » et « moins fastidieuses ». Depuis la Deuxième Guerre mondiale, nous dépendons dans une large mesure des enquêtes sur échantillon probabiliste - celles-ci étant très bonnes dans les faits - pour atteindre ces objectifs pour les estimations dans de nombreux domaines, y compris le revenu des ménages et le chômage, l'état de santé autodéclaré, l'emploi du temps, les victimes d'actes criminels, l'activité des entreprises, les flux de produits, les dépenses des consommateurs et des entreprises, etc. Par suite des taux de plus en plus faibles de réponse totale et partielle et des preuves d'erreur de déclaration, nous avons réagi de nombreuses façons, y compris en utilisant des modes d'enquête multiples, des méthodes de pondération et d'imputation plus raffinées, l'échantillonnage adaptable, des essais cognitifs des questions d'enquête et d'autres méthodes pour maintenir la qualité des données. Dans le cas des statistiques sur le secteur des entreprises, afin de réduire le fardeau et les coûts, nous avons cessé depuis longtemps de recourir uniquement à des enquêtes pour produire les estimations nécessaires, mais jusqu'à présent, nous ne l'avons pas fait pour les enquêtes auprès des ménages, du moins pas aux États-Unis. Je soutiens que nous pouvons et que nous devons passer du paradigme de production des meilleures estimations possible à partir d'une enquête à la production des meilleures estimations possible pour répondre aux besoins des utilisateurs, à partir de sources de données multiples. Ces sources comprennent les dossiers administratifs et, de plus en plus, des données sur les transactions et des données en ligne. Je me sers de deux exemples - ceux du revenu des ménages et des installations de plomberie - pour illustrer ma thèse. Je propose des moyens d'inculquer une culture de la statistique officielle dont l'objectif est d'aboutir à des statistiques pertinentes, à jour, exactes et peu coûteuses, et qui traite les enquêtes, de même que les autres sources de données, comme des moyens d'atteindre cet objectif.

Mots clés : Enquêtes; dossier administratif; erreur totale; mégadonnées; revenu; logement.

1 Introduction

Tim Holt, ancien chef de l'*Office for National Statistics* du Royaume-Uni et ancien président de la *Royal Statistical Society*, a cerné cinq grands défis pour les statistiques officielles - à savoir qu'elles soient « plus vastes, plus approfondies, plus rapides, de meilleure qualité et moins coûteuses » (Holt 2007) - liste à laquelle j'ajouterais « moins fastidieuses » et « plus pertinentes ». Selon moi, pour relever comme il convient un ou plusieurs de ces défis, sans parler de tous les sept, les organismes statistiques officiels doivent passer du paradigme des enquêtes par échantillon probabiliste en vigueur depuis les 75 dernières années à un paradigme de sources de données multiples. Certains organismes ont procédé à ce changement pour la plupart de leurs programmes statistiques (voir, par exemple, Nelson et West 2014, au sujet de l'usage très étendu de statistiques fondées sur des données de registre au Danemark), et presque tous ont effectué ce changement pour certains de leurs programmes, mais il existe néanmoins des

1. Constance F. Citro, directrice, Committee on National Statistics, U.S. National Academy of Sciences/National Research Council. Courriel : ccitro@nas.edu.

programmes qui ne sont pas encore rendus très loin sur cette voie. Dans le cas des programmes de statistiques sur les ménages des États-Unis, il reste beaucoup à faire.

Cette transition ne devrait pas simplement élever une autre source de données au rang de panacée de la statistique officielle remplaçant l'enquête par échantillon probabiliste. Le recensement de la République allemande de 2011 - le premier réalisé dans ce pays depuis 1983 - nous rappelle justement les dangers d'une telle approche. Les résultats du recensement ont indiqué que les dossiers administratifs sur lesquels l'Allemagne avait fondé les chiffres de population officiels pendant plusieurs décennies surestimaient la population parce que les émigrants nés à l'étranger n'étaient pas enregistrés correctement (voir http://www.nytimes.com/2013/06/01/world/europe/census-shows-new-drop-in-germanys-population.html?_r=0 [November 2014]).

Ma thèse est que les programmes de statistiques officielles doivent prendre pour point de départ l'information dont ont besoin les utilisateurs pour l'élaboration des politiques, l'évaluation des programmes et la compréhension des tendances sociétales, et travailler à rebours des concepts vers les sources de données appropriées. Ces sources peuvent fort probablement inclure des enquêtes à échantillonnage probabiliste, mais aussi un ou plusieurs autres types de données. Ma thèse peut dans un certain sens passer pour un truisme, mais les personnes qui consacrent leur vie à perfectionner un outil particulier pour la collecte des données pourraient trop souvent considérer que cet outil est nécessaire en toute situation, plutôt que d'envisager le moyen le plus rentable d'obtenir les statistiques que souhaitent les décideurs, les chercheurs et d'autres utilisateurs des données.

Je ne doute pas un seul instant que Joe Waksberg, que j'ai eu l'honneur de connaître alors qu'il participait à un panel d'experts sur la méthodologie du recensement décennal du *Committee on National Statistics* (CNSTAT) au milieu des années 1980, approuverait mon sujet. Non seulement Joe était un être doué d'une bienveillance et d'un charme hors du commun, mais il possédait aussi une aptitude de premier plan à résoudre les problèmes et à innover. Joe insistait sur le fait qu'il importe d'examiner non seulement ce que l'on vous demande, mais aussi ce à quoi, selon vous, pense l'analyste (Morganstein et Marker 2000). Joe s'écartait invariablement des sentiers battus afin de cerner des sources de données et des modèles répondant aux besoins sous-jacents d'information au lieu de partir d'un concept a priori quant aux outils qu'il convenait d'utiliser.

Dans l'exposé qui suit, à la section 2, je passe brièvement en revue l'ascension et les avantages de l'échantillonnage probabiliste pour la statistique officielle aux États-Unis et, à la section 3, les menaces croissantes qui pèsent sur la pertinence, l'exactitude, l'actualité et la rentabilité des estimations fondées sur des enquêtes, ainsi que leur acceptation par le public. À la section 4 et à la section 5, j'examine les points forts et les points faibles des dossiers administratifs et d'autres sources de données non probabilistes qui pourraient être utiles, seules ou regroupées, pour la production de statistiques officielles. À la section 6, je décris de belles occasions pour les États-Unis de transformer les programmes d'enquêtes-ménages courants afin d'utiliser de multiples sources de données pour fournir de l'information d'une plus grande valeur. Je conclus à la section 7 en énumérant les obstacles à l'évolution vers le paradigme des sources de données multiples et propose des moyens de les aplanir.

Je me concentre sur ce que je connais le mieux, à savoir la statistique officielle aux États-Unis et les programmes de statistiques sur les ménages en particulier. D'autres programmes statistiques et d'autres organismes découvriront des analogies avec leurs propres travaux. Je critique le paradigme des enquêtes dans une perspective d'amélioration des statistiques officielles, tout en continuant d'apprécier grandement

la valeur des enquêtes à échantillonnage probabiliste, seules ou combinées à d'autres sources de données, et d'éprouver une profonde admiration pour le travail important des organismes statistiques voués à servir le bien public (voir National Research Council 2013c).

2 L'ascension de l'échantillonnage probabiliste en statistique officielle aux États-Unis

Il n'est pas exagéré de dire que les enquêtes par échantillonnage probabiliste à grande échelle ont été la réponse du 20e siècle au besoin de statistiques officielles plus vastes, plus approfondies, plus rapides, de meilleure qualité, moins coûteuses, plus pertinentes et moins fastidieuses. Ces enquêtes fournissaient des renseignements d'une précision connue contrairement aux enquêtes non probabilistes; en outre, ils fournissaient des renseignements détaillés, plus rapidement et à moindre coût, que les recensements. Duncan et Shelton (1978) et Harris-Kojetin (2012) passent en revue l'ascension de l'échantillonnage probabiliste dans les statistiques officielles aux États-Unis.

Il n'était pas évident durant les années 1930, époque à laquelle ont été élaborées la théorie et la pratique de l'échantillonnage probabiliste moderne aux États-Unis, que les enquêtes probabilistes seraient acceptées de manière aussi générale. L'arrivée de Jerzy Neyman au milieu des années 1930 a donné un élan aux travaux de W. Edwards Deming, Calvin Dedrick, Morris Hansen et leurs collègues au *Census Bureau* qui cherchaient à élaborer la théorie nécessaire pour l'échantillonnage de populations finies. Des enquêtes-échantillons à petite échelle, réalisées au cours des années 1930 par des universités et des organismes fédéraux sur des sujets comme les achats des consommateurs, le chômage, le logement urbain et la santé, ont fourni des preuves de concept et des conseils pratiques.

Les statisticiens innovateurs de l'administration fédérale avaient encore à surmonter les obstacles bureaucratiques jusqu'à la Maison-Blanche avant de pouvoir faire entrer l'échantillonnage dans la statistique fédérale officielle. Donc, les « vétérans » du *Census Bureau* étaient sceptiques quant à la possibilité d'utiliser des méthodes d'enquête pour obtenir des renseignements sur le chômage, tandis que les politiciens avaient des avis partagés quant à l'idée d'accepter des estimations (Anderson 1988). En 1937, une percée importante a eu lieu quand un échantillonnage de 2 % des ménages inclus dans les tournées postales non commerciales conçu par Dedrick, Hansen et d'autres a donné une estimation nettement plus élevée - et plus crédible - du nombre de chômeurs qu'un recensement « complet » de toutes les adresses résidentielles mené sur une base volontaire. S'inspirant de cet effort, de 1940 à 1942, la *Works Progress Administration* a produit le *Monthly Report on the Labor Force* fondé sur un échantillon, qui était le précurseur de la *Current Population Survey* (CPS). La CPS demeure aujourd'hui la source des estimations mensuelles officielles du chômage aux États-Unis produites par le *Census Bureau* et publiées par le *Bureau of Labor Statistics* (BLS).

Une autre percée a eu lieu quand le *Census Bureau*, qui essayait depuis des décennies de répondre aux demandes de questions supplémentaires dans le recensement décennal sans transformer le questionnaire en un cauchemar pour les répondants et les intervieweurs, a posé six questions à un échantillon de 5 % de la population au recensement de 1940. En raison du succès de l'échantillonnage, il a été décidé d'administrer les deux cinquièmes des questions du recensement de 1950 à un échantillon, et la même décision a été prise pour les recensements suivants. Le tableau 2.1 énumère certaines enquêtes-ménages,

certaines enquêtes-entreprises et certaines enquêtes par panel en cours aux États-Unis, ainsi que la date de leur lancement. La variété des sujets abordés et la longévité de ces enquêtes attestent de la dominance et de la valeur accordée au paradigme des enquêtes en statistique officielle aux États-Unis.

Tableau 2.1
Quelques enquêtes probabilistes menées par les organismes statistiques aux États-Unis, selon l'année de lancement

Décennie et année/type d'enquête	Enquête auprès des ménages transversale répétée	Enquête auprès des établissements commerciaux transversale répétée	Enquête par panel de particuliers
1940	1940 - Current Population Survey (CPS)	1946 - Monthly Wholesale Trade Survey	
	1947 - CPS Annual Social and Economic Supplement (CPS/ASEC)		
1950	1950 - Consumer Expenditure Survey (CE)	1953 - Advance Monthly Retail Sales Survey	
	1955 - National Survey of Fishing, Hunting, and Wildlife-Associated Recreation	1953 - Business R&D and Innovation Survey (BRDIS)	
	1957 - National Health Interview Survey (NHIS)	1959 - Building Permits Survey	
1960	1960 - Decennial Census Long-Form Sample (devenu l'American Community Survey en 2005)	1965 - National Hospital Care Survey	1966-1990 - National Longitudinal Survey of Older Men
1970	1972 - National Crime Victimization Survey (NCVS)	1975 - Farm Costs and Returns Survey and Cropping Practices and Chemical Use Surveys (combiné à l'Agricultural Resource Management Survey en 1996)	1972-1986 - National Longitudinal Survey of High School Class of 72
	1973 - American Housing Survey (AHS);		Depuis 1973 - Survey of Doctorate Recipients (SDR)
	1973 - National Survey of College Graduates (NSCG)		
	1979 - Residential Energy Consumption Survey (RECS)	1979 - Commercial Buildings Energy Consumption Survey (CBECS)	Depuis 1979 - National Longitudinal Survey of Youth (NLSY79)
1980	1983 - Survey of Consumer Finances (SCF)	1985 - Manufacturing Energy Consumption Survey (MECS)	Depuis 1984 - Survey of Income and Program Participation (SIPP)
1990	1991 - Medicare Current Beneficiary Survey (MCBS)	1996 - Agricultural Resource Management Survey (ARMS)	Depuis 1997 - National Longitudinal Survey of Youth (NLSY97)
2000	2005 - American Community Survey (ACS)		2001-2008 - Early Childhood Longitudinal Study (Birth Cohort)

Notes : Le nom courant de l'enquête est utilisé; la périodicité de l'interview pour les enquêtes transversales et par panel répétées varie; certaines enquêtes transversales répétées ont une composante de panel (groupes de renouvellement); la durée des enquêtes par panel (nombre d'années durant lesquelles les répondants sont dans l'échantillon) varie.

Source : Liste compilée par l'auteur.

3 Des défauts dans la cuirasse : menaces croissantes pesant sur le paradigme des enquêtes

Les enquêtes par échantillonnage probabiliste sont indispensables aux organismes statistiques officiels et autres pour de nombreux types de mesures : par exemple, pour suivre des phénomènes tels que l'approbation du public à l'égard du président des États-Unis ou les sentiments exprimés de bien-être. En outre, les enquêtes probabilistes visant principalement à produire des mesures de concept, comme le revenu du ménage, qui pourraient être obtenues à partir d'autres sources, offrent deux grands avantages : 1) elles permettent d'obtenir des données sur une grande variété de covariables pouvant être utilisées dans l'analyse de la ou des variables principales d'intérêt et 2) elles sont sous le contrôle de leur concepteur. Pourtant, les menaces qui pèsent sur le paradigme des enquêtes par échantillonnage probabiliste font boule de neige d'une façon qui ne présage rien de bon pour l'avenir. Manski (2014) va jusqu'à accuser les organismes statistiques d'enfouir sous le tapis les principaux problèmes liés à leurs données et de sous-estimer nettement l'incertitude présente dans leurs estimations. Il considère la non-réponse aux enquêtes comme un exemple d'« incertitude permanente ».

3.1 Caractérisation de la qualité des enquêtes

Une classification des erreurs et des autres problèmes qui peuvent compromettre la qualité des estimations d'enquête est essentielle à la compréhension et à l'amélioration des statistiques officielles. Brackstone (1999) a écrit un article majeur concernant le développement des cadres de qualité des données. Plus récemment, Biemer, Trewin, Bergdahl et Lilli (2014) ont passé en revue la littérature sur les cadres systématiques de la qualité, en soulignant, en particulier, les six dimensions proposées par Eurostat (2000), à savoir la pertinence, l'exactitude, l'actualité et la ponctualité, l'accessibilité et la clarté, la comparabilité (temporelle et géographique) et la cohérence (normes cohérentes). Iwig, Berning, Marck et Prell (2013) ont examiné les cadres de la qualité établis par Eurostat, l'*Australian Bureau of Statistics*, l'*Office for National Statistics* du Royaume-Uni, Statistique Canada et d'autres organismes, et élaboré des questions fondées sur six dimensions de la qualité de leur cru - la pertinence, l'accessibilité, la cohérence, l'intelligibilité, l'exactitude et l'environnement institutionnel - destinées à être utilisées par les organismes statistiques américains pour évaluer l'utilité des dossiers administratifs. Daas, Ossen, Tennekes et Nordholt (2012) ont construit un cadre d'évaluation de l'utilisation de dossiers administratifs pour produire des données de recensement pour les Pays-Bas.

Biemer et coll. (2014) sont allés plus loin et ont utilisé le cadre d'Eurostat (en combinant la comparabilité et la cohérence en une seule dimension) comme fondement pour concevoir, tester et mettre en œuvre un système de cotes numériques pour évaluer et améliorer continuellement la qualité des produits de données de Statistics Sweden. Pour que l'évaluation soit complète, elle devrait aussi porter sur les dimensions de la qualité en regard du coût et du fardeau de réponse. Utilement en ce qui concerne mes objectifs, Biemer et coll. ont décomposé la dimension d'« exactitude », conçue comme étant l'erreur totale d'enquête (ou l'erreur totale de produit pour les programmes statistiques non fondés sur des enquêtes, comme les comptes nationaux), en une erreur d'échantillonnage et sept types d'erreurs non dues à l'échantillonnage, à savoir 1) l'erreur de base de sondage, y compris le sous-dénombrement et le surdénombrement, ainsi que les variables auxiliaires manquantes ou erronées dans la base de sondage; 2) l'erreur due à la non-réponse (totale et partielle); 3) l'erreur de mesure (surdéclaration, sous-

déclaration, autre); 4) l'erreur de traitement des données; 5) l'erreur de modélisation/estimation, telle que celle découlant de l'ajustement de modèles pour l'imputation ou de l'ajustement des valeurs des données afin qu'elles concordent avec les valeurs de référence; 6) l'erreur de révision (la différence entre les estimations publiées provisoires et définitives); et 7) l'erreur de spécification (la différence entre la variable réelle non observable et la variable indicatrice observée). Pour les enquêtes permanentes, j'ajouterais l'*erreur de concept dépassé*, qui est apparentée à l'erreur de spécification mais différente de celle-ci. Par exemple, le concept de revenu monétaire ordinaire du *Census Bureau* pour le calcul des estimations officielles du revenu des ménages et de la pauvreté d'après *l'Annual Social et Economic Supplement* (ASEC) de la CPS est devenu progressivement dépassé en raison de l'évolution des programmes d'imposition et de transferts des États-Unis (voir, par exemple, Czajka et Denmead 2012; National Research Council 1995).

3.2 Quatre sources d'erreur dans les statistiques américaines sur les ménages

3.2.1 Déficiences des bases de sondage

Obtenir une base de sondage complète et exacte pour les enquêtes peut être aussi difficile qu'obtenir des réponses auprès des unités sélectionnées dans l'échantillon à partir de la base de sondage et, dans de nombreux cas, la difficulté a persisté, voire même augmenté, au fil du temps. Joe Waksberg serait d'accord sur le problème des déficiences des bases de sondage : non seulement il a élaboré, en collaboration avec Warren Mitofsky, la méthode de composition aléatoire (CA) pour créer des bases de sondage et des échantillons pour réaliser des enquêtes téléphoniques résidentielles de haute qualité durant les années 1970 (voir Waksberg 1978; Tourangeau 2004), mais il a aussi assisté aux premiers signes de déclin de la popularité de la méthode en raison de phénomènes tels que l'existence de ménages ne possédant qu'un téléphone mobile.

L'une des bases de sondage utilisées fréquemment pour réaliser les enquêtes-ménages aux États-Unis est le Fichier maître des adresses (FMA) du *Census Bureau* élaboré pour le recensement décennal. Lors des quelques derniers recensements, la couverture nette des adresses résidentielles dans le FMA n'a cessé de s'améliorer, particulièrement pour les logements occupés (Mule et Konicki 2012). Le problème que continuent de poser les enquêtes-ménages est celui du sous-dénombrement des membres individuels dans les logements échantillonnés. Les ratios de couverture (c.-à-d. les estimations avant ajustement des ratios sur les chiffres de population de contrôle) dans le cas de la CPS de mars 2013, par exemple, ne sont que de 85 % pour l'ensemble de la population, et il existe des écarts prononcés entre les hommes et les femmes, les jeunes et les personnes âgées, ainsi que les blancs et les groupes minoritaires, les ratios de couverture étant aussi faibles que 61 % pour les hommes et les femmes de race noire âgés de 20 à 24 ans (voir http://www.census.gov/prod/techdoc/cps/cpsmar13.pdf [November 2014]). Aucune étude systématique de la série chronologique de ratios de couverture pour les enquêtes-ménages américaines n'a été réalisée, mais il existe des preuves que les ratios se sont dégradés.

Il est certes utile de corriger les erreurs de couverture pour tenir compte de l'âge, du sexe, de la race et du groupe ethnique, mais les ajustements des ratios effectués à l'heure actuelle pour les enquêtes-ménages ne fournissent indubitablement pas de correction pour d'autres écarts de couverture conséquents. (Les chiffres de contrôle pour l'ajustement des ratios, dans le cadre de l'un des usages les moins controversés et

les plus anciens des dossiers administratifs dans les enquêtes-ménages des États-Unis, sont tirés des estimations démographiques produites d'après les données du recensement précédent et mises à jour au moyen de dossiers administratifs et de données d'enquête.) Donc, tout ce que l'on sait au sujet du sous-dénombrement au recensement décennal des États-Unis indique que, si l'on maintient constantes la race et l'origine ethnique, les populations désavantagées sur le plan socioéconomique sont moins bien dénombrées que les autres (voir, par exemple, National Research Council 2004, annexe D). Il est peu probable que de meilleurs résultats soient obtenus dans le cas des enquêtes-ménages - par exemple, Czajka, Jacobson et Cody (2004) constatent que la *Survey of Income and Program Participation* (SIPP) sous-représente considérablement les familles à revenu élevé comparativement à la *Survey of Consumer Finances* (SCF), qui comprend un échantillon de ménages à revenu élevé tiré d'une liste basée sur les dossiers fiscaux. En tenant compte des différences de couverture socioéconomique, Shapiro et Kostanich (1988) estiment au moyen de simulations que les estimations de la pauvreté présentent un important biais à la baisse pour les hommes noirs dans la CPS/ASEC. Par ailleurs, comparativement à l'échantillon ayant reçu le questionnaire complet du Recensement de 2000, Heckman et LaFontaine (2010) constatent que le sous-dénombrement au supplément d'octobre sur les études de la CPS de 2000 contribue peu à la sous-estimation des taux d'achèvement des études secondaires; d'autres facteurs sont plus importants.

3.2.2 Tendance à la baisse de la réponse totale

Un groupe d'étude du National Research Council des États-Unis (2013b) vient d'achever un examen complet des causes et conséquences de la non-réponse totale aux enquêtes-ménages, qui confirme le phénomène bien connu voulant que le public soit de moins en moins disponible et disposé à répondre aux enquêtes, même celles menées par les organismes statistiques officiels jugés fiables. Aux États-Unis, déjà durant les années 1980, il existait des preuves que les taux de réponse ont été à la baisse depuis pratiquement le début de l'usage répandu des enquêtes par échantillonnage probabiliste (voir, par exemple, Steeh 1981; Bradburn 1992). De Leeuw et De Heer (2002) ont estimé un taux séculaire de diminution de la participation aux enquêtes de 3 points de pourcentage par année en examinant les enquêtes permanentes menées dans 16 pays occidentaux du milieu des années 1980 à la fin des années 1990. Le taux de participation mesure la réponse des cas échantillonnés admissibles effectivement contactés; les taux de réponse (il existe plusieurs variantes acceptées) possèdent des dénominateurs plus généraux, comprenant les cas admissibles qui n'ont pas été rejoints (National Research Council 2013c, p. 9-12). Le National Research Council (2013b, tableaux 1 et 2, p. 104) fournit les taux de réponse initiaux ou à la présélection pour une gamme d'enquêtes américaines officielles pour 1990-1991 (alors que les taux de réponse avaient déjà diminué considérablement pour de nombreuses enquêtes) et pour 2007-2009 montrant clairement que le problème ne disparaît pas.

On a longtemps supposé que des taux de réponse plus faibles, même avec repondération pour tenir compte de la non-réponse, entraînent inévitablement un biais dans les estimations d'enquête. Selon des travaux de recherche récents (voir, par exemple, Groves et Peytcheva 2008), la relation entre la non-réponse et le biais est complexe. Lorsqu'on prend des mesures extraordinaires pour accroître le taux de réponse, il est possible qu'on augmente aussi le biais, par inadvertance, si l'on obtient une réponse plus importante auprès de certains groupes seulement et non d'autres (voir, par exemple, Fricker et Tourangeau 2010). Toutefois, il serait imprudent de la part des organismes officiels de statistique de supposer que l'accroissement de la non-réponse n'a que peu d'effet, voire aucun, sur l'exactitude des estimations,

particulièrement si la non-réponse totale est couplée à la non-réponse partielle. Par exemple, on estime que les non-répondants aux enquêtes sur la santé sont en moins bonne santé, en moyenne, que les répondants, et que les non-répondants aux enquêtes sur le bénévolat sont moins susceptibles de faire du bénévolat que les répondants (National Research Council 2013b, p. 44-45). De surcroît, les études des effets de la non-réponse sur les associations bivariées ou multivariées ou sur la variance sont peu nombreuses, sauf en ce qui concerne le fait évident - et non sans importance - que la non-réponse totale réduit la taille effective de l'échantillon.

3.2.3 Réponse partielle souvent faible et à la baisse

Ni les enquêtes ni les recensements ne peuvent s'attendre à obtenir que les répondants fournissent une réponse à chacune des questions. Dans le cas du recensement des États-Unis, la vérification de certaines questions pour s'assurer de la cohérence est une pratique de longue date, mais jusqu'au milieu du 20e siècle, aucun ajustement n'était effectué pour la non-réponse partielle - les tableaux contenaient des lignes intitulées « pas de réponse » ou un énoncé similaire. Le premier recours à l'imputation a eu lieu en 1940 quand Deming a élaboré une méthode « cold deck » pour imputer l'âge en sélectionnant aléatoirement une valeur d'âge dans un ensemble approprié de cartes sélectionnées en fonction des autres renseignements connus au sujet de la personne dont l'âge manquait. À partir de 1960, grâce à l'émergence des ordinateurs à haute vitesse, des méthodes d'imputation « hot deck » ont été utilisées pour imputer les valeurs manquantes pour de nombreuses questions du recensement (Citro 2012). La méthode hot deck consiste à utiliser la valeur la plus récente enregistrée dans une matrice pour la personne ou le ménage traité précédemment et, par conséquent, ne requiert pas l'hypothèse que les données manquent entièrement au hasard (MCAR pour *missing completely at random*), bien qu'il soit nécessaire de supposer que les données manquent au hasard (MAR pour *missing at random*) dans les catégories définies par les variables dans la matrice hot deck. Des méthodes d'imputation fondées sur un modèle ne nécessitant pas d'aussi fortes hypothèses que celles de type MAR ou MCAR ont été élaborées (voir National Research Council 2010b), mais leur usage n'est pas très répandu dans les enquêtes-ménages aux États-Unis. Font exception la *Survey of Consumer Finance*s (SCF) (Kennickell 2011) et la *Consumer Expenditure (CE) Interview Survey* (Passero 2009).

Quelle que soit la méthode, l'imputation a l'avantage de créer un enregistrement de données complet pour chaque répondant, ce qui facilite l'analyse multivariée et réduit la probabilité que les chercheurs utilisent différentes méthodes de traitement des données manquantes donnant des résultats différents. Cependant, l'imputation peut introduire un biais dans les estimations, et la mesure dans laquelle des données manquent accentuera vraisemblablement l'importance de tout biais. Par conséquent, il est troublant de constater que la non-réponse a augmenté pour des questions importantes des enquêtes-ménages, comme celles sur le revenu, les actifs, les impôts et les dépenses de consommation, qui obligent les répondants à fournir des montants en dollars - par exemple, Czajka (2009, tableau A-8) compare les taux d'imputation d'une question pour le revenu total et pour plusieurs sources de revenus dans le cas de la CPS/ASEC et la SIPP pour 1993, 1997 et 2002 - un bon tiers des données sur le revenu sont imputées à l'heure actuelle dans le cas de la CPS/ASEC, en hausse par rapport à environ le quart en 1993 - et la situation n'est pas meilleure pour la SIPP. Clairement, étant donné des taux d'imputation aussi élevés, il est impératif de procéder à une évaluation minutieuse des effets des méthodes d'imputation. Hoyakem, Bollinger et Ziliak (2014), par exemple, estiment que la méthode d'imputation hot deck pour les revenus

dans la CPS/ASEC a systématiquement entraîné une sous-estimation de un point de pourcentage, en moyenne, de la pauvreté en se basant sur l'évaluation des revenus manquants dans les enregistrements des revenus de la CPS/ASEC et de la sécurité sociale.

3.2.4 L'erreur de mesure pose problème et n'est pas bien étudiée

Même en cas de déclaration complète, ou, plus fréquemment, d'ajustements pour tenir compte de la non-réponse totale et partielle, les estimations d'après les données d'enquête contiendront encore une erreur découlant des déclarations inexactes faites par les répondants qui devinent la réponse, évitent délibérément de donner une réponse correcte ou ne comprennent pas l'intention de la question. Même si les organismes statistiques reconnaissent l'existence de l'erreur de mesure, la portée de celle-ci est habituellement moins bien étudiée que celle de l'erreur d'échantillonnage ou des données manquantes. De nombreuses études de l'erreur de mesure comparent les estimations agrégées provenant d'une enquête à des estimations similaires provenant d'une autre enquête ou à un ensemble approprié de dossiers administratifs, ajustés autant qu'il est possible pour qu'ils soient comparables. Il est impossible de dégager de ces études le rôle joué par l'erreur de mesure comparativement à d'autres facteurs, mais les résultats indiquent l'ordre de grandeur des problèmes. Les auteurs de certaines études arrivent à apparier des enregistrements individuels et par conséquent à examiner les composantes de l'erreur de mesure.

Il est connu qu'une erreur de mesure importante affecte les estimations socioéconomiques clés produites d'après les enquêtes-ménages américaines. Donc, une foule d'études ont donné des preuves, enquête après enquête, d'une sous-estimation nette du revenu des ménages américains et, constatation encore plus troublante, d'une diminution de la complétude des déclarations, même après imputation et pondération. Ainsi, Fixler et Johnson (2012, tableau 2) ont estimé qu'entre 1999 et 2010, les estimations moyennes et médianes calculées d'après la CPS/ASEC sont devenues progressivement inférieures aux estimations des *National Income and Product Accounts* (NIPA) en raison de facteurs tels que 1) la sous-représentation des ménages à revenu très élevé dans l'échantillon de la CPS/ASEC, 2) la non-déclaration ou la sous-déclaration par les ménages à revenu élevé qui sont inclus dans l'échantillon et 3) la non-déclaration ou la sous-déclaration par les ménages à revenu moyen ou faible. Les études portant sur les sources individuelles de revenu révèlent une erreur encore pire. Par exemple, Meyer et Goerge (2011) constatent, en appariant les enregistrements du *Supplemental Nutrition Assistance Program* (SNAP) obtenus dans deux États, que près de 35 % et 50 %, respectivement, de véritables bénéficiaires ne déclarent pas avoir reçu des prestations dans le cadre de l'*American Community Survey* (ACS) ou de la CPS/ASEC. De même, Meyer, Mok et Sullivan (2009) fournissent des preuves d'écarts importants et souvent croissants entre les estimations d'enquête et les estimations fondées sur les dossiers administratifs correctement ajustés des bénéficiaires du revenu et des montants totaux pour de nombreuses sources.

La richesse est, comme on le sait, difficile à mesurer dans les enquêtes-ménages, et de nombreux organismes n'essaient pas de le faire. Czajka (2009, p. 143-145) résume les travaux de recherche sur la qualité des estimations de la richesse d'après la SIPP en les comparant aux estimations d'après la SCF et la *Panel Study of Income Dynamics* (PSID). En simplifiant considérablement les résultats, historiquement, la SIPP s'est avérée assez efficace pour mesurer les éléments de passif, comme la dette hypothécaire, et la valeur d'éléments d'actif possédés tels que les logements, les véhicules et les obligations d'épargne. Par contre, la SIPP n'a pas fourni de bonnes mesures de la valeur des actifs détenus principalement par les ménages à revenu élevé, comme les actions, les fonds communs de placement, ainsi que les comptes IRA

et KEOGH, tandis que la PSID a donné d'un peu meilleurs résultats. Sur une base nette, la SIPP sous-estime considérablement la valeur nette.

Une étude menée par le National Research Council (2013a) sur la *CE Interview and Diary Surveys* du BLS sur les dépenses de consommation comportant une interview et la tenue d'un journal a révélé des différences de qualité de la déclaration de divers types de dépenses comparativement aux estimations des dépenses de consommation personnelles (PCE pour *personal consumption expenditure*) ajustées de manière appropriée provenant des NIPA. Bee, Meyer et Sullivan (2012, tableau 2) ont également constaté une diminution de la déclaration de certaines dépenses - par exemple, la déclaration des dépenses en essence dans l'estimation des dépenses de consommation des ménages est passée de plus de 100 % de l'estimation des PCE comparables en 1986 à un peu moins de 80 % en 2010, tandis que la déclaration des dépenses en meubles et accessoires d'ameublement est passée de 77 % à 44 % au cours d'une période comparable.

4 Que peut-on faire?

Les spécialistes de la recherche sur les enquêtes ne sont pas restés inactifs face aux menaces multiples et croissantes qui pèsent sur le paradigme des enquêtes. Durant au moins les 15 dernières années, ils ont cherché activement des moyens de réduire ou de compenser l'erreur de couverture, la non-réponse totale et partielle, l'erreur de mesure et, plus récemment, le fardeau de réponse. Les stratégies adoptées comprenaient 1) consacrer plus d'argent à l'achèvement des cas (mais les contraintes budgétaires limitent la viabilité de cette stratégie), 2) utiliser les paradonnées et l'information auxiliaire afin de déterminer et de corriger plus efficacement le biais dû à la non-réponse totale, 3) employer des ajustements plus perfectionnés pour tenir compte des données manquantes qui ne reposent pas sur l'hypothèse qu'elles sont de type MAR, 4) utiliser des méthodes reposant sur un plan de collecte adaptatif afin d'optimiser le coût et la qualité des réponses, 5) utiliser de multiples bases de sondage pour réduire l'erreur de couverture (p. ex. listes de numéros de téléphone mobile et de numéros de téléphone fixe pour les enquêtes téléphoniques), 6) utiliser de multiples modes de collecte pour que la réponse soit plus efficace par rapport au coût, comme dans l'ACS, qui a ajouté récemment une option de réponse en ligne aux options d'envoi par la poste, d'ITAO et d'IPAO, 7) réduire le fardeau de réponse en optimisant les nombres d'appels et de visites de suivi, et 8) décrire les besoins de données d'enquête. Aux États-Unis, on fait souvent appel aux utilisateurs des données pour plaider la cause devant le Congrès et d'autres parties prenantes. Par exemple, l'*Association of Public Data Users*, le *Council of Professional Associations on Federal Statistics* et la *Population Association of America* mobilisent fréquemment les utilisateurs des données au nom des programmes des organismes statistiques.

Selon moi, quoique louables et nécessaires, ces étapes ne suffisent pas à restaurer le paradigme fondé sur l'enquête par échantillonnage probabiliste pour la production de statistiques officielles sur les ménages et d'autres types de répondants. Je propose plutôt que les organismes statistiques commencent systématiquement par cerner les besoins des décideurs et du public et qu'ils travaillent à rebours afin de déterminer quelles sont les sources de données appropriées pour répondre aux besoins de la façon la plus rentable et la moins lourde possible. Ce paradigme des sources de données multiples devrait s'appliquer à

tous les programmes statistiques qui sont habituellement fondés sur des enquêtes, des dossiers administratifs ou d'autres sources.

Certains programmes statistiques importants, comme les NIPA et l'Indice des prix à la consommation (voir Horrigan 2013) aux États-Unis et dans d'autres pays, utilisent des sources de données multiples depuis des décennies. L'une des raisons est que ces programmes s'appuient sur un cadre conceptuel généralement accepté qui détermine les éléments requis pour constituer un ensemble satisfaisant d'estimations. Il n'est pas acceptable d'omettre une ou plusieurs composantes du revenu des NIPA simplement parce que les données ne peuvent pas être obtenues à partir d'une source unique. En outre, comme les estimations clés des NIPA sont révisées périodiquement afin d'ajouter des données, d'améliorer la méthodologie et de peaufiner les concepts, il existe un biais positif intégré en faveur de la recherche de sources de données nouvelles et améliorées pour combler les lacunes et accroître l'exactitude. Les recensements économiques menés aux États-Unis s'appuient aussi sur des sources multiples, plus précisément les données de l'impôt sur le revenu pour les entreprises individuelles et les très petits employeurs, ainsi que des données d'enquête pour les plus grandes entreprises. En revanche, les programmes de statistiques sur les ménages des États-Unis ont adhéré plus étroitement au paradigme des enquêtes par échantillonnage probabiliste. De surcroît, comme les intervalles sont habituellement longs entre les révisions des concepts et du plan des enquêtes-ménages, les enquêtes perdent trop souvent du terrain en ce qui concerne leur capacité à servir les décideurs et le public, alors que l'utilisation de sources de données additionnelles permettrait d'importantes améliorations.

5 Quelles sources de données utiliser pour soutenir les enquêtes?

Pendant des décennies après l'introduction de l'échantillonnage probabiliste en statistique officielle, la seule autre source de données était les dossiers administratifs - provenant de divers paliers de gouvernement, selon la structure gouvernementale du pays (fédéral, État et local aux États-Unis), et de diverses entités non gouvernementales (p. ex. dossiers de paye des employeurs ou dossiers d'admission des hôpitaux). Un certain nombre d'organismes statistiques nationaux dans le monde ont commencé à intégrer des dossiers administratifs dans leurs programmes - cette intégration allant de leur utilisation accessoire au transfert, sans distinction aucune, des enquêtes et des recensements à un paradigme axé sur les dossiers administratifs.

Grâce aux innovations technologiques des années 1970 et des années 1980, certaines sources de données supplémentaires, comme les enregistrements des dépenses aux caisses (rendus possibles par le développement des codes à barres et des scanneurs), et les images aériennes et par satellite pour catégoriser l'utilisation des terres, sont devenues disponibles, du moins potentiellement, pour la production de statistiques officielles. Cependant, l'univers des sources de données demeurait relativement limité. À partir des années 1990, l'avènement d'Internet et de la technologie de l'informatique à haute-vitesse a donné le jour à un extraordinaire éventail de nouvelles sources de données, dont les données envoyées par les caméras de circulation, la localisation des téléphones mobiles, les termes de recherche utilisés sur le Web et les affichages sur les sites des médias sociaux. Le défi pour les organismes statistiques consiste à classer et à évaluer toutes ces sources de données d'une manière qui les aide à en déterminer l'utilité.

5.1 Le concept des « mégadonnées » est-il utile?

Bon nombre de nouvelles catégories de données devenues disponibles au cours des quelque 15 dernières années sont souvent de très grande taille, ce qui a donné naissance au terme de « mégadonnées ». Je soutiens que ce terme à la mode n'aide que fort peu, voire nullement, les organismes statistiques à déterminer quelles sont les combinaisons de données convenant pour leurs programmes. En sciences informatiques, « les mégadonnées sont des fonds d'information à grand volume, grande vélocité et/ou grande variété qui nécessitent de nouvelles formes de traitement pour permettre la prise de meilleures décisions, la découverte d'idées et l'optimisation des processus » [*Traduction*] (Laney 2001). Ces propriétés ne sont pas inhérentes à un type particulier de données ou à une plateforme particulière, telle qu'Internet. Ce qui peut être considéré comme des « mégadonnées » est plutôt une cible en évolution à mesure que l'informatique à haute vitesse et les techniques d'analyse des données progressent. Dans l'environnement informatique actuel, les données de recensement, d'enquête et de dossiers administratifs peuvent rarement être qualifiées de « mégadonnées », même si elles auraient pu l'être à une époque antérieure. Aujourd'hui, les gens ont tendance à considérer comme étant des « mégadonnées » les flux de données provenant de caméras, de détecteurs et d'interactions en grande partie libres avec Internet, comme les messages sur les médias sociaux. À l'avenir, bon nombre de ces types de données pourraient ne plus rentrer dans cette catégorie. De plus, en ce qui concerne Internet, celui-ci génère non seulement une grande quantité de « mégadonnées » contemporaines, mais il facilite aussi l'accès à des données de volume plus habituel - par exemple, accès aux sondages d'opinion ou aux registres fonciers locaux.

À mon avis, les organismes statistiques souhaiteront le plus souvent, et devraient, être des « adeptes suivant de près les leaders » plutôt que des leaders de l'utilisation des mégadonnées. Il me paraît plus approprié que le milieu universitaire et le secteur privé soient les premiers à s'attaquer à l'utilisation de données aussi volumineuses et d'une telle vélocité et variété qu'elles nécessitent de grands pas en avant dans l'élaboration de nouvelles formes de traitement et d'analyse. Les organismes statistiques devraient se tenir au courant des avancées dans le domaine des mégadonnées qui pourraient être prometteuses pour leurs programmes et ils seraient bien avisés d'appuyer la recherche dans ces domaines pour s'assurer que les applications pertinentes pour leurs programmes voient le jour. Toutefois, je pense que les ressources des organismes statistiques devraient être consacrées principalement à l'utilisation de sources de données qui offrent des avantages dont l'utilité est plus immédiate.

Groves (2011) a tenté de passer à une classification plus pertinente pour les organismes statistiques que celle comprenant les « mégadonnées », d'une part, et toutes les autres données, d'autre part, en faisant la distinction entre ce qu'il appelle les « données conçues » qui sont « produites pour découvrir ce qui n'est pas mesuré » et les « données organiques » qui sont « produites secondairement aux processus, pour enregistrer le processus ». Keller, Koonin et Shipp (2012) énumèrent des exemples de sources de données sous les deux en-têtes de Groves. Leur liste de données conçues comprend les données administratives (p. ex. dossiers fiscaux), les enquêtes fédérales, les recensements de la population et les « autres données recueillies pour répondre à des questions stratégiques particulières ». Leur liste de données organiques comprend les données de localisation (« données externes » de téléphones mobiles, de transpondeurs pour postes de péage, de caméras de surveillance), les préférences politiques (dossiers d'enregistrement des électeurs, votes aux élections primaires, contributions aux partis politiques), les renseignements commerciaux (transactions sur carte de crédit, ventes de propriété, recherches en ligne, identification de radiofréquences), les renseignements sur la santé (dossiers médicaux électroniques, admissions à l'hôpital,

appareils pour surveiller les signes vitaux, ventes des pharmacies), et autres données organiques (imagerie optique, infrarouge et spectrale, mesures météorologiques, mesures sismiques et acoustiques, rayonnements ionisants biologiques et chimiques). Sans omettre, sous chaque catégorie, des données telles que les messages affichés sur Facebook ou Twitter, bien qu'ils puissent se retrouver sous la rubrique plus générale des « recherches en ligne ».

La question est de savoir si la classification en deux catégories de Keller et coll. (2012) est plus utile que celle de « mégadonnées » pour les besoins des organismes statistiques. Par exemple, classer les dossiers d'inscription des électeurs ou les dossiers de santé électroniques comme des données organiques plutôt que comme des données administratives conçues semble ne pas tenir compte des façons dont elles diffèrent de sources telles que les recherches en ligne et des façons dont elles sont similaires aux dossiers administratifs de l'administration fédérale et des États. En outre, même les données organiques sont « conçues », si ce n'est que de manière minimale, en ce sens que le fournisseur a spécifié certains paramètres, tels que les 140 caractères pour un message sur Twitter ou un angle de vision particulier pour une caméra de circulation. Néanmoins, la distinction entre données conçues et données organiques met en relief une dimension utile, qui est le degré auquel les organismes statistiques ont déjà accès à une source de données, contrôlent les changements apportés à une source de données et sont capables de comprendre facilement les propriétés d'une source de données.

5.2 Dimensions des sources de données : illustrations pour quatre grandes catégories

Établir une nomenclature et des critères d'évaluation satisfaisants qui peuvent aider les organismes statistiques à évaluer l'utilité éventuelle de diverses sources de données pour leurs programmes, dans le but de comprendre aussi bien les propriétés d'erreur des sources de données de rechange qu'ils ne comprennent l'erreur totale dans le cas des enquêtes, demandera un effort considérable de la part des organismes statistiques du monde entier (Iwig et coll. 2013 et Daas et coll. 2012, sont des exemples de tels efforts). Je ne prétends pas pouvoir m'approcher de ce but dans le présent article. Mon objectif est plus modeste - à savoir donner certaines illustrations afin que ceux et celles qui sont des inconditionnels du paradigme des enquêtes par échantillonnage probabiliste (ou du paradigme des dossiers administratifs) puissent voir que la tâche de comprendre d'autres sources de données est à la fois faisable et souhaitable. Je fournis des illustrations pour quatre sources de données variant du classique à l'avant-garde :

(1) Enquêtes et recensements, ou un ensemble de données tirées des réponses de particuliers qui sont interrogés sur un ou plusieurs sujets selon le plan établi par l'enquêteur (organisme statistique, autre organisme gouvernemental ou organisme universitaire ou privé d'enquête) conformément aux principes de la recherche par enquête dans le but de produire des données généralisables pour une population définie.

(2) Dossiers administratifs ou un ensemble de données obtenues au moyen de formulaires conçus par un organisme administratif conformément à une loi, un règlement ou une politique pour exploiter un programme, comme le versement de prestations à des bénéficiaires admissibles ou pour le versement de salaires. Les dossiers administratifs sont habituellement permanents et peuvent être gérés par des organismes gouvernementaux ou des organisations non gouvernementales.

(3) Dossiers de transactions commerciales, ou un ensemble de données obtenues par saisie électronique d'achats (p. ex. épicerie, biens immobiliers) effectués par un acheteur, mais sous une forme déterminée par un vendeur (p. ex. renseignements sur les produits et prix sous forme de codes à barres enregistrés par les scanneurs des caisses, enregistrements de renseignements sur les produits et les prix provenant des ventes en ligne, comme par l'intermédiaire d'Amazon).

(4) Interactions des particuliers avec le Web en utilisant des outils fournis commercialement, comme un navigateur Web ou un site de média social. Cette catégorie englobe un éventail vaste et en constante évolution de sources de données possibles pour lesquelles il n'existe aucune classification simple. L'une des caractéristiques déterminantes est que les personnes qui fournissent l'information, comme un message sur Twitter, agissent de manière autonome : elles ne doivent pas répondre à un questionnaire ou fournir des renseignements administratifs, mais choisissent plutôt de lancer une interaction.

Je commence par classer chaque source en fonction de deux dimensions, qui sont liées au cadre décrit dans Biemer et coll. (2014). J'attribue le classement en supposant qu'un organisme statistique n'a pas encore pris de mesure proactive afin de l'améliorer (p. ex. en intégrant du personnel dans un organisme administratif afin qu'il se familiarise en profondeur avec les dossiers de cet organisme). Les deux dimensions sont les suivantes :

(1) Degré d'accessibilité de l'organisme statistique national à la source et de contrôle qu'il exerce sur la source : élevé (l'organisme statistique conçoit la source de données et contrôle les changements qui y sont apportés); moyen (l'organisme statistique est autorisé à utiliser la source de données et influe sur les changements qui y sont apportés); faible (l'organisme statistique doit s'arranger pour obtenir la source de données conformément aux conditions établies par le fournisseur et n'a que peu d'influence, voire aucune, sur les changements qui y sont apportés). Une gradation peut être ajoutée à chacune de ces catégories selon, par exemple, la force de l'autorité dont dispose l'organisme pour acquérir un ensemble de dossiers administratifs.

(2) Degré possible de détermination et de mesure des composantes de l'erreur : élevé, comme dans le cas des enquêtes et des recensements conçus par l'organisme; moyen, comme dans le cas des dossiers administratifs des secteurs public et privé; et faible, comme dans le cas des flux de données provenant de choix autonomes de particuliers.

Je détermine ensuite des aspects de la qualité des données pour chaque source, à l'instar de Biemer et coll. (2014). J'indique aussi les variations pour la plupart des dimensions selon le fournisseur, comme un organisme statistique national, une autre unité gouvernementale nationale, un autre palier de gouvernement, une institution universitaire ou une entité commerciale. Toute cette information est regroupée dans le tableau 5.1 au mieux de mes connaissances.

Une source idéale pour un organisme statistique, toutes choses étant égales par ailleurs, est une source qui est fournie, conçue et contrôlée par l'organisme, et pour laquelle les erreurs peuvent être identifiées et mesurées et sont généralement maîtrisées, comme dans le cas d'une enquête à échantillonnage probabiliste de haute qualité, mise sur pied par l'organisme. À l'autre extrême se trouve une source de données qui est contrôlée par une ou plusieurs entreprises privées (p. ex. données de scanneur) ou, peut-être, des centaines

ou des milliers d'administrations publiques locales (p. ex. caméras de circulation), pour laquelle les données résultent de choix autonomes ou de mouvements non contrôlés, et pour laquelle il est difficile de conceptualiser, sans parler de mesurer, les erreurs dans la source de données. Pourtant, étant donné qu'un organisme statistique est chargé de fournir aux décideurs et aux membres du public des statistiques pertinentes, à jour et exactes dont le coût et le fardeau de réponse sont réduits au minimum, il pourrait fort bien exister des sources de données autres que les enquêtes qui justifient l'effort de les rendre utilisables à des fins statistiques. Je soutiens que les menaces qui pèsent sur le paradigme des enquêtes passées en revue plus haut rendent impérative la prise en considération d'autres sources de données, car il n'est plus possible de démontrer que les enquêtes représentent en tout temps et en toutes circonstances un meilleur choix que d'autres sources - elles n'obtiennent pas systématiquement une cote « élevée » sur les dimensions prises en compte dans le tableau 5.1.

Je soutiens aussi que les dossiers administratifs gouvernementaux, qui, comme l'indique le tableau 5.1, possèdent plus souvent les propriétés souhaitables pour la production de statistiques officielles que d'autres sources de données non issues d'enquêtes, devraient être considérés par les organismes statistiques comme une option toute désignée pour une intégration aussi étendue que possible dans leurs programmes d'enquêtes s'ils ne l'ont pas déjà fait. Les dossiers administratifs sont créés conformément à des règles concernant la population admissible, les personnes qui doivent fournir quel type d'information, les mesures qui doivent être prises par l'organisme administratif pertinent en se basant sur l'information (p. ex. remboursement d'impôt, versement de prestations), et ainsi de suite. Cela devrait permettre à un organisme statistique, moyennant l'effort requis, de se familiariser avec les structures d'erreur des dossiers administratifs comme ils le sont avec l'erreur totale d'enquête. Couper (2013) offre une discussion utile quelque peu semblable à la mienne. Il découvre des failles dans la capacité des sources de données organiques à être aussi utiles qu'on l'affirme, sans parler des affirmations quant à leur capacité de remplacer les enquêtes par échantillonnage probabiliste, mais il avertit les chercheurs d'enquête que s'ils ignorent les sources de données organiques, ils le font à leurs risques et périls. Ironiquement, sa conclusion qu'il faut utiliser certaines sources organiques est renforcée par l'erreur qu'il commet en classant les dossiers administratifs comme étant des données organiques. Leur classification correcte est celle de données conçues, même si elles ne le sont pas par un organisme statistique.

Tableau 5.1
Classement (ÉLEVÉ, MOYEN, FAIBLE, TRÈS FAIBLE ou VARIABLE) de quatre sources de données sur les dimensions d'utilisation dans les statistiques officielles

Dimension/ Source de données	Recensement/enquête probabiliste (p. ex. CPS/ASEC, ACS, NHIS - voir tableau 2.1)	Dossiers administratifs (p. ex. impôt sur le revenu, sécurité sociale, chômage, paye)	Enregistrements de transactions commerciales (p. ex. données de scanneurs et de cartes de crédit)	Interactions des particuliers avec Internet (p. ex. Twitter; volumes de termes de recherche dans Google)
Degré de contrôle/ d'accessibilité de la source par l'organisme statistique	ÉLEVÉ (enquête menée pour l'organisme statistique); MOYEN à FAIBLE (enquête menée pour un organisme privé).	ÉLEVÉ à MOYEN (dossiers d'un organisme national); MOYEN à FAIBLE (dossiers d'État ou dossiers locaux); MOYEN à FAIBLE (dossiers commerciaux).	MOYEN à FAIBLE	TRÈS FAIBLE

Dimension/ Source de données	Recensement/enquête probabiliste (p. ex. CPS/ASEC, ACS, NHIS - voir tableau 2.1)	Dossiers administratifs (p. ex. impôt sur le revenu, sécurité sociale, chômage, paye)	Enregistrements de transactions commerciales (p. ex. données de scanneurs et de cartes de crédit)	Interactions des particuliers avec Internet (p. ex. Twitter; volumes de termes de recherche dans Google)
Degré de capacité de l'organisme statistique à déterminer/ évaluer les propriétés/ erreurs	ÉLEVÉ (enquête menée pour l'organisme statistique); VARIABLE (enquête menée pour un organisme privé, dépend de la documentation et de la transparence).	ÉLEVÉ à MOYEN (dossiers d'un organisme national); MOYEN à FAIBLE (dossiers d'État ou dossiers locaux); MOYEN à FAIBLE (dossiers commerciaux).	MOYEN (dans la mesure où les enregistrements sont conformes aux normes reconnues (p. ex. pour les codes à barres et les renseignements sur les prix).	TRÈS FAIBLE
Attributs de la qualité des données (Biemer et coll. 2014)				
Pertinence pour les décideurs et les membres du public - Concepts et mesures	ÉLEVÉE pour une enquête menée pour l'organisme statistique, en supposant qu'elle est bien conçue et que les concepts et les mesures sont à jour; VARIABLE pour des enquêtes menées pour des organismes privés.	VARIABLE d'un système de dossiers à l'autre et à l'intérieur des systèmes de dossiers (p. ex. les dossiers de versement de prestations peuvent être très pertinents, tandis que les renseignements sur la composition de la famille peuvent s'appuyer sur un concept différent).	VARIABLE	VARIABLE, mais TRÈS FAIBLE dans l'état actuel des moyens d'acquérir, évaluer et analyser ces types de données.
Pertinence - Covariables utiles	ÉLEVÉE pour la plupart des enquêtes.	VARIABLE, mais rarement aussi élevée que pour la plupart des enquêtes.	VARIABLE, mais rarement aussi élevée que pour la plupart des enquêtes.	VARIABLE, mais habituellement FAIBLE.
Fréquence de collecte des données	D'hebdomadaire à toutes les deux ou trois années (toutes les décennies pour le recensement de la population des États-Unis); quelques enquêtes privées, comme les sondages électoraux, peuvent être exécutées à chaque jour.	En général, les dossiers sont mis à jour fréquemment (p. ex. quotidiennement) et continuellement.	En général, les enregistrements sont mis à jour fréquemment (p. ex. au moment de la transaction ou quotidiennement) et continuellement.	Les interactions sont saisies instantanément.
Actualité des données diffusées	VARIABLE, dépend de l'effort de l'organisme statistique ou de l'organisme privé, mais un certain décalage par rapport à la période de référence de la réponse est inévitable.	VARIABLE, mais un certain décalage par rapport à la date de référence à laquelle les dossiers ont été acquis par l'organisme statistique est probable.	VARIABLE, mais vraisemblablement de longs délais pour l'acquisition de données exclusives par l'organisme statistique.	VARIABLE, mais vraisemblablement de longs délais (quoique le *Billion Prices Project* du MIT ait établi des modalités d'accès très rapide aux prix sur Internet; voir bpp.mit.edu).

Dimension/ Source de données	Recensement/enquête probabiliste (p. ex. CPS/ASEC, ACS, NHIS - voir tableau 2.1)	Dossiers administratifs (p. ex. impôt sur le revenu, sécurité sociale, chômage, paye)	Enregistrements de transactions commerciales (p. ex. données de scanneurs et de cartes de crédit)	Interactions des particuliers avec Internet (p. ex. Twitter; volumes de termes de recherche dans Google)
Comparabilité et cohérence	ÉLEVÉES dans le temps et dans l'espace (géographie) au sein d'une enquête (sauf en cas de changement délibéré ou de changement sociétal affectant les mesures qui n'est pas pris en compte); VARIABLES, selon les enquêtes.	ÉLEVÉES au sein du système de dossiers (changements apportés aux dossiers gouvernementaux généralement annoncés par un changement juridique/ réglementaire/ de politique; changements apportés aux dossiers commerciaux vraisemblablement opaques). VARIABLES, selon les systèmes de dossiers.	ÉLEVÉES au sein du système d'enregistrements (changements généralement opaques pour l'organisme statistique); VARIABLES entre les systèmes d'enregistrements.	TRÈS FAIBLES, en ce sens que les fournisseurs (p. ex. Twitter) peuvent ajouter/soustraire des caractéristiques ou abandonner complètement un produit; changements généralement opaques pour l'organisme statistique; les auteurs des interactions peuvent avoir des cadres de référence très différents.
Exactitude (composantes de l'erreur)*				
Erreur de base de sondage	VARIABLE, possibilité d'un sous-dénombrement ou d'un surdénombrement important.	La base de sondage est habituellement bien définie par une loi, un règlement ou une politique; le problème en cas d'utilisation par un organisme statistique est que la base de sondage pourrait ne pas être exhaustive.	La base de sondage est mal définie pour les besoins d'un organisme statistique, en ce sens qu'elle représente quiconque a eu un achat scanné par un vendeur spécifié ou a utilisé une carte de crédit particulière pour un achat durant une période spécifiée; pose un grand défi à l'organisme statistique en ce qui concerne la détermination de l'usage approprié.	La base de sondage est mal définie pour les besoins d'un organisme statistique, en ce sens qu'elle représente quiconque a décidé, par exemple, de créer un compte Twitter ou d'effectuer une recherche dans Google durant une période spécifiée; pose un grand défi à l'organisme statistique en ce qui concerne la détermination de l'usage approprié.
Non-réponse (totale et partielle)	VARIABLE; peut être importante.	VARIABLE (p. ex. les dossiers de la sécurité sociale couvrent vraisemblablement presque toutes les personnes admissibles, mais les dossiers fiscaux reflètent vraisemblablement la fraude fiscale sous forme d'omission de produire une déclaration de revenus ou de non-déclaration de certains revenus).	SANS OBJET, en ce sens que les « répondants » sont autosélectionnés; le défi pour l'organisme statistique consiste à déterminer l'utilisation appropriée qui ne requiert pas l'hypothèse d'un mécanisme probabiliste.	SANS OBJET, en ce sens que les « répondants » sont autosélectionnés; le défi pour l'organisme statistique consiste à déterminer l'utilisation appropriée qui ne requiert pas l'hypothèse d'un mécanisme probabiliste.

Exactitude (composantes de l'erreur)* (SUITE)				
Dimension/ Source de données	**Recensement/enquête probabiliste (p. ex. CPS/ASEC, ACS, NHIS - voir tableau 2.1)**	**Dossiers administratifs (p. ex. impôt sur le revenu, sécurité sociale, chômage, paye)**	**Enregistrements de transactions commerciales (p. ex. données de scanneurs et de cartes de crédit)**	**Interactions des particuliers avec Internet (p. ex. Twitter; volumes de termes de recherche dans Google)**
Erreur de mesure	VARIABLE, au sein des enquêtes, par question, et entre les enquêtes pour des questions comparables; souvent mal évaluée, même pour les enquêtes réalisées par les organismes statistiques.	VARIABLE entre les systèmes de dossiers et au sein des systèmes de dossiers, par question, selon la mesure dans laquelle la question joue un rôle central dans le fonctionnement du programme (p. ex. une question sur le versement de prestations est vraisemblablement plus exacte que des éléments de données obtenus auprès des bénéficiaires, comme la situation d'emploi).	SANS OBJET pour la source de données en tant que telle, quoique toute caractéristique ajoutée par le vendeur en provenance d'une autre source peut ou non être valide; le défi pour l'organisme statistique consiste à ne pas introduire d'erreur de mesure en utilisant les données de manière inappropriée.	SANS OBJET pour la source de données en tant que telle, quoique toute caractéristique ajoutée par le vendeur en provenance d'une autre source peut ou non être valide; le défi pour l'organisme statistique consiste à ne pas introduire d'erreur de mesure en utilisant les données de manière inappropriée.
Erreur de traitement des données	VARIABLE (p. ex. possibilité d'erreur de saisie des données ou de recodage), mais fait habituellement l'objet d'un bon contrôle statistique, bien que cela soit plus difficile à évaluer pour les enquêtes réalisées par des organismes privés.	VARIABLE (p. ex. possibilité d'erreurs de saisie-clavier ou de codage), vraisemblablement mieux contrôlée pour les variables clés (p. ex. versements de prestations) que pour d'autres variables, mais difficile pour l'organisme statistique de l'évaluer.	VARIABLE (p. ex. possibilité d'erreurs lors de l'attribution des codes à barres ou des prix), vraisemblablement bien contrôlée, mais difficile pour l'organisme statistique de l'évaluer.	SANS OBJET, en ce sens que l'erreur n'est pas définie, quoiqu'il puisse y avoir à l'occasion des problèmes tels que, disons, l'écrasement et la perte d'une journée complète de messages Twitter.
Erreur de modélisation/ estimation	Biais découlant de processus tels que la pondération et l'imputation VARIABLE; souvent, l'organisme statistique déploie d'intenses efforts afin de bien concevoir l'enquête au départ, mais ne procède pas à un réexamen pour s'assurer que les procédures continuent d'être valides.	SANS OBJET (habituellement), en ce sens que les dossiers sont des données « brutes », sauf peut-être dans le cas de certaines variables recodées, mais un biais peut être introduit par l'organisme statistique durant le retraitement.	SANS OBJET (habituellement), en ce sens que les enregistrements sont des données « brutes », sauf peut-être dans le cas de certaines variables recodées ou résumées, mais un biais peut être introduit par l'organisme statistique durant le retraitement.	SANS OBJET (habituellement), en ce sens que les enregistrements sont des données « brutes », mais le retraitement par l'organisme statistique peut introduire un biais important (p. ex. en considérant que le terme « licencié » est toujours indicateur de chômage dans l'analyse des messages Twitter).

Exactitude (composantes de l'erreur)* (SUITE)				
Dimension/ Source de données	**Recensement/enquête probabiliste (p. ex. CPS/ASEC, ACS, NHIS - voir tableau 2.1)**	**Dossiers administratifs (p. ex. impôt sur le revenu, sécurité sociale, chômage, paye)**	**Enregistrements de transactions commerciales (p. ex. données de scanneurs et de cartes de crédit)**	**Interactions des particuliers avec Internet (p. ex. Twitter; volumes de termes de recherche dans Google)**
Erreur de spécification	VARIABLE (p. ex. l'état de santé autodéclaré peut indiquer validement la perception du répondant, mais pas nécessairement l'état de santé physique ou mental diagnostiqué); peut évoluer au cours du temps (p. ex. à mesure que l'usage des mots évolue parmi le public).	VARIABLE; peut être importante si les concepts dans les dossiers administratifs diffèrent de ceux dont l'organisme statistique a besoin (p. ex. les règles concernant la déclaration des revenus sur les formulaires de déclaration peuvent ne pas tenir compte de composantes telles que des avantages de cafétéria).	VARIABLE; peut être faible ou élevée en fonction de la mesure dans laquelle les données correspondent aux besoins de l'organisme statistique.	VARIABLE, mais vraisemblablement importante dans l'état actuel des moyens d'acquérir, d'évaluer et d'analyser ces types de données émanant de choix relativement libres effectués par des individus autonomes.
Fardeau*	VARIABLE, peut être élevé.	PAS DE FARDEAU SUPPLÉMENTAIRE imposé par l'organisme statistique à la population pertinente (p. ex. bénéficiaires), mais fardeau imposé à l'organisme administratif.	PAS DE FARDEAU SUPPLÉMENTAIRE imposé par l'organisme statistique à la population pertinente (p. ex. acheteurs), mais fardeau imposé au fournisseur.	PAS DE FARDEAU SUPPLÉMENTAIRE imposé par l'organisme statistique à la population pertinente (p. ex. afficheurs de messages sur Twitter), mais fardeau imposé au fournisseur.
Coût*	VARIABLE, peut être élevé; l'organisme statistique assume la totalité des coûts de conception, de collecte, de traitement et d'estimation.	VARIABLE, mais peut être plus faible que pour une enquête comparable, parce que l'organisme administratif assume les coûts de collecte des données, mais l'organisme statistique assume vraisemblablement des coûts de manipulation/ traitement spécial.	VARIABLE comme pour les dossiers administratifs, mais le fournisseur souhaite vraisemblablement un paiement; l'organisme statistique assume vraisemblablement des coûts de traitement spécial/ manipulation/ analyse.	VARIABLE comme pour les dossiers administratifs, mais le fournisseur souhaite vraisemblablement un paiement; les coûts supplémentaires assumés par l'organisme statistique pour le traitement/ analyse de données non structurées peuvent être élevés.

*La direction de l'échelle change; autrement dit « élevée » est indésirable et « faible » est désirable.
Note : N'inclut pas l'erreur de révision comprise dans la classification de Biemer et coll. (2014).
Source : Évaluation grossière de l'auteure.

5.3 Utilisations des dossiers administratifs dans les programmes fondés sur des enquêtes-ménages

Les participants aux enquêtes-ménages ont prouvé maintes fois que leurs réponses à de nombreuses questions importantes sur le revenu, la richesse, les dépenses et d'autres sujets ne sont pas très exactes. Dans de nombreux cas, l'utilisation de dossiers administratifs offre la possibilité de remédier à cette

situation. Une autre stratégie adoptée par de nombreux programmes d'enquêtes-ménages aux États-Unis consiste à inviter les répondants eux-mêmes à consulter leurs propres dossiers, comme les déclarations de revenus, lorsqu'ils répondent aux questions sur le revenu ou des sujets similaires. Sans aucun doute, les réponses sont vraisemblablement plus exactes lorsque les dossiers sont consultés, comme Johnson et Moore (pas de date) le constatent dans une comparaison de dossiers fiscaux aux réponses à la SCF pour l'exercice 2000. Cependant, la stratégie proprement dite semble être en grande partie un exercice futile. Selon la même étude de la SCF réalisée par Johnson et Moore, seulement 10 % des ménages dont le revenu brut ajusté est inférieur à 50 000 $ consultent leurs dossiers et seulement 22 % des ménages à revenu élevé le font. Voir National Research Council (2013a, p. 89-91) ainsi que Moore, Marquis et Bogen (1996) pour des constatations similaires au sujet des difficultés à obtenir que les répondants consultent leurs dossiers.

En me penchant maintenant sur les stratégies que les organismes statistiques peuvent adopter pour travailler directement avec des données administratives, je cerne huit façons selon lesquelles les dossiers administratifs peuvent contribuer à la qualité des données des enquêtes-ménages, à savoir 1) aider à évaluer la qualité des données d'enquête, en les comparant à des estimations agrégées, ajustées comme il convient pour tenir compte des différences d'univers et de concepts entre les populations, et par appariement exact des enregistrements de l'enquête et des enregistrements administratifs; 2) fournir des totaux de contrôle pour l'ajustement des poids de sondage afin de tenir compte des erreurs de couverture; 3) fournir des bases de sondage supplémentaires pouvant être utilisées dans un plan à bases de sondage multiples; 4) fournir des renseignements supplémentaires à annexer aux enregistrements d'enquête appariés pour améliorer la pertinence et l'utilité des données; 5) fournir des covariables pour les estimations fondées sur un modèle pour des régions géographiques plus petites que celles pour lesquelles des estimations peuvent être produites directement d'après l'enquête; 6) améliorer les modèles pour l'imputation des données manquantes dans les enregistrements de l'enquête; 7) remplacer « non » pour les participants à l'enquête qui auraient dû répondre à une question, remplacer « oui » pour les participants à l'enquête qui n'auraient pas dû répondre à une question, et remplacer les valeurs déclarées pour les participants à l'enquête qui ont fourni une réponse erronée à une question; et 8) remplacer les questions de l'enquête et utiliser les valeurs des dossiers administratifs directement. Dans une version non publiée plus longue du présent article, je donne des exemples actuels et possibles de chaque type d'utilisation et énumère les avantages, les problèmes de confidentialité et de perception du public, ainsi que les limites et les problèmes de faisabilité pour chaque utilisation, de manière générique et en particulier pour les enquêtes-ménages américaines portant sur des sujets tels que le revenu, les actifs et les dépenses. Ce qui importe, en ce qui me concerne, est que les avantages doivent surpasser les inconvénients, étant donné un programme soutenu, pour intégrer des systèmes de dossiers administratifs à des programmes statistiques.

5.4 Utilisations possibles de sources de données non habituelles

Ayant indiqué antérieurement que les données provenant d'autres sources que les enquêtes et les dossiers administratifs posent un certain nombre de problèmes pour la production de statistiques officielles, il serait négligent de ma part de ne pas discuter brièvement des raisons pour lesquelles ces données semblent si intéressantes. Les entreprises privées ont des fonctions de perte très différentes de celles des organismes statistiques - elles cherchent à avoir un avantage sur leurs concurrents. Des données qui sont plus à jour et qui indiquent des moyens d'accroître les ventes et les profits sont

vraisemblablement utiles à l'entreprise privée, même si elles ne couvrent pas entièrement une population ou qu'elles ont d'autres inconvénients pour les statistiques officielles. Dans cette perspective, les types d'expériences que réalise une entreprise telle que Google, en utilisant ses propres « mégadonnées », afin de trouver des moyens d'augmenter les publicités visionnées sont de bons investissements (voir, par exemple, McGuire, Manyika et Chui 2012). De même, les organismes chargés des programmes, à tous les paliers de gouvernement, souvent en collaboration avec des centres universitaires, regroupent et analysent leurs propres données et d'autres de façons novatrices afin de déceler des tendances, « points chauds », etc., non seulement pour améliorer leurs programmes et planifier de nouveaux services, mais aussi pour classer les ressources par ordre de priorité et améliorer la réponse en temps réel (voir, par exemple, le *Center for Urban Science and Progress* à l'Université de New York (http://cusp.nyu.edu/), ainsi que le *Urban Center for Computation and Data* à l'Université de Chicago (https://urbanccd.org)).

Les organismes statistiques ont besoin, avant tout et par-dessus tout, de sources de données qui couvrent une population connue et présentent des propriétés d'erreur qui sont raisonnablement bien comprises et qui ne sont pas susceptibles de changer sans qu'on s'y attende, c'est-à-dire exemptes de caractéristiques qui sont inhérentes à des sources comme les interactions autonomes avec des sites Web sur Internet. Les programmes fondés sur les enquêtes-ménages des organismes statistiques disposent toutefois d'au moins deux moyens qui pourraient leur permettre de tirer un « avantage » de sources de données non habituelles : l'un consiste à améliorer l'actualité des estimations provisoires des statistiques clés, et l'autre consiste à fournir des indicateurs avancés de l'évolution sociale (p. ex. l'émergence de nouveaux domaines de formation et professions) qui avertissent les organismes statistiques qu'il est nécessaire de modifier leurs concepts et leurs mesures.

6 Des besoins de données aux sources de données : deux exemples aux États-Unis

Afin d'illustrer concrètement mes propos, voici deux exemples relevés aux États-Unis - revenu des ménages et caractéristiques des logements - pour lesquels, selon moi, les organismes statistiques peuvent et doivent transformer leurs programmes d'enquête en programmes à sources de données multiples afin de mieux répondre aux besoins des utilisateurs. Le *U.S. Office of Management and Budget* (2014) a fait un pas dans cette direction dans un mémoire récent où il affirme que les utilisations statistiques des dossiers administratifs des organismes fédéraux représentent un bien positif et énonce les étapes à suivre pour institutionnaliser leur usage.

6.1 Revenu des ménages

Les statistiques officielles sur la répartition des revenus des ménages comptent parmi les indicateurs les plus importants du bien-être économique produits régulièrement par les bureaux nationaux de la statistique, et elles sont encore plus importantes à la lumière des débats actuels concernant les inégalités croissantes et d'autres sujets apparentés. Pourtant, il existe une abondance de preuves que la qualité des mesures du revenu des ménages obtenues d'après les réponses aux enquêtes menées aux États-Unis est altérée considérablement par l'erreur de couverture, la non-réponse totale, la non-réponse partielle et les

erreurs de déclaration. En outre, le concept de revenu monétaire ordinaire appliqué dans les enquêtes américaines est périmé étant donné les moyens complexes et continuellement en évolution par lesquels les ménages obtiennent des ressources pour soutenir leur consommation quotidienne et leur épargne. Il semble impératif que le système statistique des États-Unis améliore ses estimations vedettes du revenu calculées d'après les données de la CPS/ASEC, de la SIPP et, dans la mesure du possible, de l'ACS, en passant d'une approche fondée en grande partie sur les réponses aux enquêtes à une approche visant à intégrer les données de dossiers administratifs à celles de ces enquêtes. Le *Census Bureau* met en œuvre de nouvelles questions et des questions modifiées afin de mieux mesurer le revenu de pension et d'autres sources dans la CPS/ASEC, à la suite d'un examen majeur de la mesure du revenu dans cette enquête réalisée par Czajka et Denmead (2012) et d'un rapport sur les essais cognitifs des modifications apportées au questionnaire de l'ASEC (Hicks et Kerwin 2011). Récemment, le *Census Bureau* a également procédé à une refonte importante de la SIPP, en faisant appel à des méthodes axées sur l'utilisation de calendriers biographiques et d'interviews annuelles à la place d'interviews tous les quatre mois afin de réduire le fardeau et les coûts, et dont les effets sur la qualité doivent être évalués (voir https://www.census.gov/programs-surveys/sipp/about/re-engineered-sipp.html [November 2014]). Un processus a été mis en place pour examiner les questions de l'ACS, mais jusqu'à présent, les questions sur le revenu n'ont pas été abordées. En plus de poursuivre la recherche classique sur les questionnaires afin de trouver des moyens de réduire le fardeau de réponse, de clarifier la signification des questions et de faciliter la réponse aux questions sur le revenu dans la mesure du possible, les enquêtes vedettes seraient améliorées considérablement si :

(1) Le *U.S. Census Bureau* et le *Bureau of Economic Analysis* (BEA) se mettaient d'accord - et réexaminaient périodiquement la situation et la mettaient à jour au besoin - sur un concept contemporain du revenu ordinaire du ménage sur lequel fonder les estimations d'après les données de la CPS/ASEC, de la SIPP et de l'ACS, et la série d'estimations du revenu personnel dans les NIPA, qui sont établies en grande partie d'après des dossiers administratifs. Il existe à l'heure actuelle des différences conceptuelles entre les enquêtes et les NIPA, tel que le traitement des prestations de retraite, qui devraient être conciliées. L'utilisation d'un concept intégré du revenu du ménage rendrait les comptes des revenus personnels et les enquêtes-ménages plus utiles pour analyser les tendances sous les angles macro et microéconomiques.

(2) Le *Census Bureau* effectuait une étude sur les avantages probables de la mise en œuvre d'ajustements des pondérations des enquêtes socioéconomiques en plus des ajustements des pondérations démographiques. En supposant qu'il existe un avantage, le *Census Bureau* déterminerait ensuite quelles sont les sources appropriées, qui pourraient être les dossiers fiscaux ou la SCF, pour ajuster les pondérations dans la CPS, la SIPP et l'ACS afin de tenir compte des différences de couverture par grande catégorie socioéconomique.

(3) Le *Census Bureau* progressait stratégiquement, source par source, afin d'améliorer les imputations des montants de revenu dans la CPS/ASEC et la SIPP en utilisant les valeurs des dossiers administratifs. Le *Census Bureau* a déjà accès à de nombreux dossiers et entreprend des démarches pour en obtenir d'autres (p. ex. dossiers du SNAP provenant des États) dans le cadre de la planification du recensement de 2020.

(4) Le *Census Bureau* évoluait - prudemment, étant donné les obstacles supplémentaires à l'utilisation des dossiers administratifs aux États-Unis - vers le modèle de Statistique Canada, qui permet aux répondants de sauter des blocs entiers de questions sur le revenu en autorisant l'accès à leurs dossiers administratifs (voir http://www.statcan.gc.ca/eng/survey/household/5200 [November 2014]).

Je ne sous-estime aucunement les difficultés que présentent les étapes susmentionnées pour la production des statistiques sur le revenu aux États-Unis. Ces difficultés, sans ordre particulier, comprennent 1) les obstacles juridiques et bureaucratiques à l'obtention d'un accès facile aux dossiers administratifs, lesquels sont considérablement plus importants pour les dossiers détenus par les organismes d'État en raison de différences entre les lois, les politiques et les normes et systèmes de données des États; 2) les considérations concernant l'obtention du consentement des répondants, particulièrement si les valeurs tirées des dossiers sont substituées à des questions; 3) les perceptions de type « Big Brother » et de menace pour la vie privée, qui peuvent limiter l'accessibilité des microdonnées pour la recherche et l'analyse des politiques; 4) le manque de ressources permettant aux organismes statistiques d'entreprendre des activités telles que le remaniement des systèmes d'imputation; 5) les effets indésirables sur l'actualité des données dans la mesure où les dossiers sont mis à la disposition des organismes statistiques avec un certain décalage temporel, problème qui pourrait être résolu en diffusant des estimations provisoires suivies par des estimations définitives quand suffisamment de données administratives deviennent disponibles; 6) la connaissance insuffisante des structures d'erreur des dossiers, qui pourrait donner lieu à des surprises désagréables; 7) les différences conceptuelles entre les mesures appliquées dans les dossiers et dans les enquêtes, dont il n'est pas tenu compte facilement (p. ex. les revenus déclarés à l'IRS ne sont pas les revenus bruts, mais les revenus imposables); 8) le fardeau supplémentaire imposé aux employés déjà fortement sollicités des bureaux centraux des organismes statistiques; 9) la nécessité de réécrire les systèmes de traitement afin de relier de multiples flux de données et d'exécuter toutes les tâches nécessaires d'appariement, de réconciliation et d'estimation dans les délais prévus; 10) la méfiance de nombreux utilisateurs de microdonnées aux États-Unis, qui semblent préférer un ensemble de données provenant d'une source unique, comme une enquête, indépendamment des inexactitudes dans les données, à un ensemble de données provenant de sources multiples, qui pourraient contenir des valeurs fondées sur un modèle pour certaines variables; et 11) l'hésitation des employés de l'organisme statistique, qui semblent souvent croire qu'il n'est pas approprié d'utiliser, disons, des dossiers administratifs pour imputer un revenu à un répondant qui n'a pas fait état d'un revenu ou d'utiliser des dossiers administratifs pour remplacer certaines questions ou pour améliorer certaines imputations, à moins que cela ne puisse être fait pour toutes les questions. Dans sa planification du recensement de 2020, le *Census Bureau* envisage d'utiliser de manière limitée des dossiers administratifs pour le suivi des cas de non-réponse, et cela pourrait servir de modèle pour l'usage sélectif de dossiers administratifs dans les enquêtes-ménages. Même si elles sont formidables, ces difficultés ne sont nullement insurmontables. Un plan stratégique par étapes, bien formulé, en vue d'adopter une approche à sources de données multiples pourrait donner aux organismes statistiques la possibilité de travailler à l'amélioration des estimations du revenu et de réussir simultanément à réduire le fardeau de réponse et, peut-être, les coûts des programmes d'enquêtes clés.

6.2 Caractéristiques des logements, y compris les installations de plomberie

En réponse aux préoccupations concernant la mauvaise qualité des logements au pays, exprimées dans le *New Deal*, le recensement décennal des États-Unis de 1940 comprenait quelques questions sur les caractéristiques des logements. Cette préoccupation était bien fondée - le recensement de 1940 a révélé, par exemple, que 45 % des logements n'étaient pas équipés d'une installation de plomberie complète (eau courante chaude et froide, toilette avec chasse d'eau, douche ou baignoire). Voir https://www.census.gov/hhes/www/housing/census/historic/plumbing.html [November 2014]. Les questions sur le logement se sont multipliées et ont été incluses dans les recensements jusqu'en 2000. Quand l'*American Community Survey* est entrée en vigueur, elle comprenait les questions sur le logement qui figuraient antérieurement dans le questionnaire complet du recensement. L'*American Housing Survey* (AHS), qui est une enquête bisannuelle de beaucoup plus petite portée, recueille une gamme encore plus vaste de renseignements sur les logements et les quartiers.

La raison principale de chercher des moyens de transférer les questions sur le logement de l'ACS d'un programme fondé sur une enquête à un programme fondé sur une enquête plus d'autres sources de données est le fardeau de réponse, tant réel que perçu, qui, dans le climat politique actuel aux États-Unis, menace la viabilité de l'ACS. Comme le travail sur le terrain de l'ACS se déroule auprès d'un grand échantillon d'environ 280 000 ménages chaque mois, au lieu de tous les dix ans comme pour l'échantillon du questionnaire complet du recensement qu'elle a remplacé, l'enquête génère un flot modéré, mais continu, de plaintes pour les membres du Congrès, qui ont donné lieu à la tenue d'audiences. Le *Census Bureau* a déterminé les quatre questions de l'ACS qui suscitent le plus de plaintes, à savoir le revenu, l'incapacité, l'heure de départ pour le travail et les installations de plomberie (voir http://www.census.gov/acs/www/Downloads/operations_admin/2014_content_review/ACSContentReviewSummit.pdf[November 2014]). Les questions sur les installations de plomberie figurant dans le questionnaire complet du recensement faisaient aussi régulièrement l'objet de plaisanteries et de plaintes. En fait, les gens répondent de manière assez exhaustive à ces questions (voir http://www.census.gov/acs/www/methodology/item_allocation_rates_data/[October 2014]), mais elles continuent de susciter du ressentiment et sont parfois mal comprises (voir Woodward, Wilson et Chestnut 2007). De surcroît, un examen du questionnaire complet de l'ACS donne à penser que l'ensemble complet d'environ 30 questions sur le logement impose un fardeau considérable à de nombreux ménages, particulièrement les propriétaires ayant un prêt hypothécaire.

En réponse à ces préoccupations au sujet du fardeau de l'ACS, le *Census Bureau* a réduit le nombre d'appels et de visites de suivi (voir Zelenak et Davis 2013), a établi un « champion des répondants » et a fourni aux membres du public des renseignements justifiant les questions. Néanmoins, le 30 mai 2014, la Chambre des représentants des États-Unis a voté une loi de crédits qui, si elle est adoptée, transformera l'ACS en une enquête à participation volontaire plutôt qu'obligatoire. Même si des données de bonne qualité pouvaient vraisemblablement être recueillies avec un suivi suffisant, le coût de l'ACS augmenterait considérablement (voir Griffin 2011). Récemment, le *Census Bureau* a demandé aux organismes fédéraux de fournir une justification législative ou réglementaire pour chaque question, et il est fort possible que certaines questions soient abandonnées (voir http://www.census.gov/acs/www/about_the_survey/acs_content_review/ [November 2014]). Les questions sur les installations de plomberie semblent être de bonnes candidates à l'élimination dans l'ACS, étant donné qu'en 2012 aux États-Unis, seulement 0,4 % des logements n'étaient pas équipés d'une installation de plomberie complète (voir

http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_12_1YR_DP04&prodType=table [November 2014]). Cependant, ce faible pourcentage est concentré dans des régions particulières, comme les réserves habitées par les Autochtones et les régions rurales. En outre, l'élimination de toute question de l'ACS semble être une mesure radicale sans d'abord examiner si d'autres sources pourraient fournir les données.

En fait, les réponses à certaines questions sur le logement du questionnaire de l'ACS pourraient fort vraisemblablement être tirées d'une gamme d'autres sources, annexées au Fichier maître des adresses (FMA) du *Census Bureau*, et être disponibles en vue de leur inclusion dans l'ACS et d'autres enquêtes qui utilisent le FMA comme base de sondage. D'autres sources comprennent les dossiers administratifs des administrations locales sur les impôts fonciers établis, l'année de la construction et d'autres caractéristiques des propriétés, renseignements qui sont de plus en plus fréquemment compilés par des fournisseurs commerciaux, ce qui réduit la nécessité d'interagir individuellement avec les milliers d'administrations publiques aux États-Unis. Elles comprennent aussi des sources comme *Google Street View* pour les caractéristiques extérieures des propriétés, les sites Web des agents immobiliers pour la valeur des logements et les caractéristiques intérieures (p. ex. nombre de pièces), les compteurs intelligents pour les coûts des services publics (installés dans certaines régions et dont l'usage se répandra vraisemblablement), et les bases de données sur les prêts hypothécaires détenues par les organismes fédéraux et les vendeurs commerciaux. Les caractéristiques des logements qui changent rarement peuvent aussi être extraites des réponses fournies par les échantillons antérieurs ayant reçu le questionnaire complet du recensement. Les installations de plomberie en sont un parfait exemple - une fois qu'un logement est doté d'une installation de plomberie, celle-ci n'est presque jamais démantelée (même s'il peut arriver qu'elle ne soit pas fonctionnelle).

Ces sources de données de rechange varient en ce qui concerne la facilité d'acquisition et d'évaluation des données, la menace réelle ou perçue qu'elles font peser sur la vie privée et la confidentialité, et la mesure dans laquelle elles couvrent l'entièreté ou la majorité du pays. L'élaboration d'un Fichier maître des adresses et des logements (FMAL) augmenté, qui peut servir l'ACS et d'autres programmes statistiques du *Census Bureau*, prendra du temps et, pour certains sujets (p. ex. installations de plomberie), il pourrait être nécessaire d'utiliser une version distincte (plus longue) du questionnaire dans des régions géographiques sélectionnées auquel sont ajoutées les questions pertinentes. Tout cela sera complexe et délicat, mais les avantages potentiels à long terme sont considérables. Pour passer à un FMAL augmenté, le *Census Bureau* peut tirer parti des travaux de l'*Office of Policy Development & Research* du *U.S. Department of Housing and Urban Development* pour simplifier le long questionnaire de l'AHS en utilisant d'autres sources de données pour de nombreuses caractéristiques du logement et du quartier en vue de remplacer les questions de l'enquête; voir http://www.huduser.org/portal/datasets/ahs.html#planning [November 2014].

7 Défis et stratégies pour procéder à un changement de paradigme

J'ai présenté des arguments en faveur d'un nouveau paradigme en vertu duquel les organismes statistiques conçoivent et mettent à jour leurs programmes vedettes en déterminant quelle est la meilleure combinaison de sources de données et de méthodes pour répondre aux besoins des utilisateurs dans un

domaine dont l'importance est croissante. J'utilise les enquêtes-ménages réalisées aux États-Unis comme exemple d'une situation où il existe des preuves solides que s'appuyer uniquement sur les réponses aux enquêtes ne suffira pas pour répondre aux besoins critiques d'information de haute qualité sur le revenu, les dépenses et des sujets apparentés. Je pense qu'il est également vrai que l'utilisation de dossiers administratifs seulement, comme dans certains pays dotés de registres de population détaillés, pourrait ne pas fournir de renseignements suffisamment complets et de haute qualité en l'absence d'efforts réguliers en vue d'examiner la qualité des données des registres et d'augmenter et de corriger ces derniers au moyen d'information provenant d'autres sources, telles que les enquêtes. Par exemple, Axelson, Homberg, Jansson, Werner et Westling (2012) décrivent l'utilité des enquêtes pour évaluer la qualité des données sur les logements et les ménages provenant d'un nouveau registre des logements créé pour le recensement de 2011 en Suède.

Je conclus par une liste de facteurs qui rendent le changement de paradigme difficile, en énonçant aussi des moyens de procéder au changement que je recommande et de l'intégrer dans la culture des organismes statistiques. Les systèmes statistiques des États-Unis et d'autres pays ont fait preuve d'innovation dans de nombreux aspects de leurs programmes, mais changer les paradigmes est toujours un exercice difficile, comme en témoigne la bataille en vue d'introduire l'échantillonnage probabiliste dans la statistique officielle aux États-Unis durant les années 1930. Il est particulièrement difficile de repenser des programmes statistiques permanents, établis depuis longtemps, avec lesquels l'organisme producteur et les utilisateurs se sentent à l'aise.

Les facteurs qui peuvent entraver le changement comprennent 1) l'inertie, particulièrement quand un programme était au départ novateur et très bien conçu, de sorte qu'il peut se reposer sur ses lauriers; 2) le décalage par rapport à l'évolution des besoins des parties prenantes, lequel peut être exacerbé quand un organisme se voit comme la seule source des données nécessaires et sans concurrence; 3) la crainte d'amoindrir les programmes existants conjuguée à la crainte du « non inventé ici »; 4) l'évaluation continue inadéquate de toutes les dimensions de la qualité des données; et 5) la compression des ressources humaines et budgétaires, conjuguée à une hésitation compréhensible du personnel de l'organisme ou de leur base établie d'utilisateurs à réduire l'une ou l'autre des séries statistiques établies de longue date afin de réaliser d'importants progrès dans d'autres séries.

Pourtant, il existe de nombreux exemples remarquables d'innovations importantes mises en place par les organismes statistiques aux États-Unis et dans d'autres pays, de sorte qu'il existe manifestement des moyens de surmonter les obstacles énumérés plus haut pour procéder au changement de paradigme. Selon moi, l'ingrédient essentiel à un changement de paradigme est le ralliement des cadres et le soutien permanent de la haute direction d'un organisme statistique, déployé proactivement de manière à rallier le personnel à tous les niveaux de l'organisme. À titre d'exemple exceptionnel d'un tel leadership, voir dans National Research Council (2010a) la discussion du rôle de Morris Hansen et de ses collègues dans le remaniement de ce qui avait été un recensement effectué par des agents recenseurs en un recensement avec envoi et retour du questionnaire par la poste. Les travaux de remaniement ont été lancés et poursuivis après que l'on ait dégagé des preuves de l'existence d'un biais et d'une variance d'intervieweur considérables pour des éléments de données importants. On craignait également qu'il devienne plus difficile de recruter des agents recenseurs à mesure que les femmes entraient sur le marché du travail.

Des mesures particulières en vue d'obtenir l'appui des cadres de l'organisme dans le but précis d'inculquer l'utilisation de multiples sources de données dans les programmes permanents de statistiques officielles comprennent (voir Prell et coll. 2009, qui ont effectué des études de cas d'utilisation statistique fructueuse de dossiers administratifs aux États-Unis, pour des conclusions similaires) : 1) l'établissement d'attentes et d'objectifs précis pour les employés, par exemple l'attente que les programmes statistiques combineront d'office des sources telles que les enquêtes et les dossiers administratifs afin de produire des données pertinentes, exactes et à jour de manière rentable et en imposant un fardeau de réponse minimal; 2) l'attribution d'un rôle important aux spécialistes du domaine - interactions avec les utilisateurs externes et les producteurs internes des données; 3) la dotation des programmes opérationnels en personnel compétent en ce qui concerne toutes les sources de données pertinentes, ce qui inclut mettre sur un pied d'égalité les spécialistes de la conception des enquêtes et les spécialistes des dossiers administratifs ou d'autres sources de données; 4) la rotation des affectations, y compris des rotations internes, des rotations entre organismes statistiques, des rotations avec les organismes utilisateurs des données et des rotations avec les entités fournissant les sources de données de rechange; 5) la mise en place de ressources pour l'évaluation continue et 6) le traitement des organisations possédant des sources de données de rechange qui jouent un rôle important dans les programmes statistiques comme des partenaires. Sur ce dernier point, voir, par exemple, Hendriks (2012, p. 1473), qui, en décrivant les expériences de Statistics Norway au sujet de son premier recensement fondé sur des registres en 2011, insiste sur le fait que « Les trois C des statistiques fondées sur des registres (afin de produire des données de qualité) sont la coopération, la communication et la coordination » [*Traduction*].

Les organismes statistiques ont montré qu'ils étaient capables d'effectuer des changements de grande portée en réponse aux menaces qui pèsent sur les moyens établis de fonctionnement. La deuxième moitié du 20^{e} siècle a donné le paradigme des enquêtes à échantillonnage probabiliste en réponse à la croissance des coûts et du fardeau associée à la réalisation de dénombrements complets et aux défauts des plans d'échantillonnage non probabilistes. Le 21^{e} siècle peut sans aucun doute nous donner le paradigme de l'utilisation des meilleures sources de données, incluant les enquêtes, les dossiers administratifs et d'autres sources, pour répondre au besoin de statistiques officielles pertinentes, exactes, à jour et rentables des décideurs et des membres du public.

Remerciements

Le présent article est fondé sur les années d'expérience acquise par l'auteure auprès du *Committee on National Statistics* (CNSTAT), mais les opinions exprimées sont les siennes et ne doivent pas être interprétées comme représentant celles du CNSTAT ni celles de la *National Academy of Sciences*. L'auteure remercie John Czajka, David Johnson et Rochelle Martinez de leurs commentaires constructifs au sujet d'une version antérieure. Une version plus longue du présent article peut être obtenue sur demande auprès de l'auteure.

Bibliographie

Anderson, M.J. (1988). *The American Census: A Social History*. New Haven. CT: Yale University Press.

Axelson, M., Homberg, A, Jansson, I., Werner, P. et Westling, S. (2012). Doing a register-based census for the first time: The Swedish experience. *Paper presents at the Joint Statistical Meetings*, San Diego, CA (août). Statistics Sweden, Stockholm.

Bee, A., Meyer, B.D. et Sullivan, J.X. (2012). The validity of consumption data: Are the consumer expenditure interview and diary surveys informative? *NBER Working Paper No. 18308*. Cambridge, MA: National Bureau of Economic Research.

Biemer, P., Trewin, D., Bergdahl, H. et Lilli, J. (2014). A system for managing the quality of official statistics, with discussion. *Journal of Official Statistics*, 30(3, septembre), 381-442.

Brackstone, G. (1999). La gestion de la qualité des données dans un bureau de statistique. *Techniques d'enquête*, 25(2), 159-171.

Bradburn, N.H. (1992). A response to the nonresponse problem. 1992 AAPOR Presidential Address. *Public Opinion Quarterly*, 56(3), 391-397.

Citro, C.F. (2012). *Editing, Imputation and Weighting*. Encyclopedia of the U.S. Census: From the Constitution to the American Community Survey, Second Edition, M. J. Anderson, C.F. Citro et J.J. Salvo, eds, 201-204. Washington, DC: CQ Press.

Couper, M.P. (2013). Is the sky falling? New technology, changing media, and the future of surveys. Keynote presentation at the *5th European Survey Research Association Conference*. Ljubliana, Slovenia. http://www.europeansurveyresearch.org/sites/default/files/files/Couper%20keynote.pdf [July 2014].

Czajka, J.L. (2009). SIPP data quality. Appendix A in *Reengineering the Survey of Income and Program Participation*. National Research Council. Washington, DC: The National Academies Press.

Czajka, J.L. et Denmead, G. (2012). Income measurement for the 21st century: Updating the current population survey. Washington, DC: *Mathematica Policy Research*. Disponible au http://www.mathematica-mpr.com/~/media/publications/PDFs/family_support/income_measurement_21_century.pdf [July 2014].

Czajka, J.L., Jacobson, J.E. et Cody, S. (2004). Survey estimates of wealth: A comparative analysis and review of the Survey of Income and Program Participation. *Social Security Bulletin*, 65(1). Disponible au http://www.ssa.gov/policy/docs/ssb/v65n1/v65n1p63.html [July 2014].

Daas, P.J.H., Ossen, S.J.L., Tennekes, M. et Nordholt, E.S. (2012). Evaluation of the quality of administrative data used in the Dutch virtual census. *Paper presents at the Joint Statistical Meetings*, San Diego, CA (août). Methodology Sector and Division of Social and Spatial Statistics, Statistics Netherland, The Hague.

De Leeuw, E.D. et De Heer, W. (2002). *Trends in Household Survey Nonresponse: A Longitudinal and International Comparison*. R.M. Groves, D.A. Dillman, J. L. Eltinge et R.J.A. Little, eds. Survey Nonresponse, 41-54. New York: Wiley.

Duncan, J. W. et Shelton, W. C. (1978). *Revolution in United States Government Statistics 1926–1976*. Office of Federal Statistical Policy and Standards, U.S. Department of Commerce. Washington, DC: Government Printing Office.

Eurostat. (2000). Assessment of the quality in statistics. *Doc. Eurostat/A4/Quality/00/General/Standard report*. Luxembourg (4-5 avril). Disponible au http://www.unece.org/fileadmin/DAM/stats/documents/2000/11/ metis/crp.3.e.pdf [July 2014].

Fixler, D. et D.S. Johnson (2012). Accounting for the distribution of income in the U.S. National Accounts. *Paper prepared for the NBER Conference on Research in Income and Wealth*, 30 septembre. Disponible au http://www.bea.gov/about/pdf/Fixler_Johnson.pdf.

Fricker, S. et R. Tourangeau (2010). Examining the relationship between nonresponse propensity and data quality in two national household surveys. *Public Opinion Quarterly*, 74(5), 935-955.

Griffin, D. (2011). Cost and workload implications of a voluntary American community survey. *U.S. Census Bureau,* Washington, DC (June 23).

Groves, R.M. (2011). Three eras of survey research. *Public Opinion Quarterly*, 75(9), 861-871. Special 75th Anniversary Issue.

Groves, R.M. et Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly*, 72(2), 167-189.

Harris-Kojetin, B. (2012). *Federal Household Surveys*. Encyclopedia of the U.S. Census: From the Constitution to the American Community Survey, Second Edition, M. J. Anderson, C.F. Citro et J.J. Salvo, eds, 226-234. Washington, DC: CQ Press.

Heckman, J. J. et LaFontaine, P.A. (2010). The American high school graduation rate: trends and levels. *NBER Working Paper 13670*. Cambridge, MA, National Bureau of Economic Research. Disponible au http://www.nber.org/papers/w13670 [July 2014].

Hendriks, C. (2012). Input data quality in register based statistics-The Norwegian experience. Proceedings of the *International Association of Survey Statisticians-JSM 2012*, 1473-1480. Article présenté au Joint Statistical Meetings, San Diego, CA (août). Statistics Norway, Kongsvinger, Norway.

Hicks, W. et Kerwin, J. (2011). Cognitive testing of potential changes to the Annual Social and Economic Supplement of the Current Population Survey. *Report to the U.S. Census Bureau*, Westat, Rockville, MD (25 juillet).

Holt, D.T. (2007).The official statistics Olympics challenge: Wider, deeper, quicker, better, cheaper. *The American Statistician*, 61(1, février), 1-8. Avec les commentaries de G. Brackstone et J.L. Norwood.

Horrigan, M.W. (2013). Big data: A BLS perspective. *Amstat News*, 427(janvier), 25-27.

Hoyakem, C., Bollinger, C. et Ziliak, J. (2014). The role of CPS nonresponse on the level and trend in poverty. *UKCPR Discussion Paper Series*, DP 2014-05. Lexington, KY: University of Kentucky Center for Poverty Research.

Iwig, W., Berning, M., Marck, P. et Prell, M. (2013). Data quality assessment tool for administrative data. Prepared for a subcommittee of the *Federal Committee on Statistical Methodology*, Washington, DC (février).

Johnson, B et Moore, K. [pas de date]. Consider the source: Differences in estimates of income and wealth from survey and tax data. Disponible au http://www.irs.gov/pub/irs-soi/johnsmoore.pdf [July 2014].

Keller, S.A., Koonin, S.E. et Shipp, S. (2012). Big data and city living - what can it do for us? *Statistical Significance*, 9(4), 4-7, août.

Kennickell, A. (2011). Look again: Editing and imputation of SCF panel data. *Paper prepared for the Joint Statistical Meetings*, Miami, FL (août).

Laney, D. (2001). 3-D data management: Controlling data volume, velocity and variety. *META Group [now Gartner] Research Note*, 6 février. Voir http://goo.gl/Bo3GS [July 2014].

Manski, C.F. (2014). Communicating uncertainty in official economic statistics. *NBER Working Paper No. 20098*. Cambridge, MA: National Bureau of Economic Research.

McGuire, T., Manyika, J. et Chui, M. (2012). Why big data is the new competitive advantage. *Ivey Business Journal* (juillet-août).

Meyer, B. D. et Goerge, R.M. (2011). Errors in survey reporting and imputation and their effects on estimates of Food Stamp Program participation. Working Paper. *Chicago Harris School of Public Policy*, University of Chicago.

Meyer, B.D., Mok, W. K.C. et Sullivan, J.X. (2009). The under-reporting of transfers in household surveys: Its nature and consequences. *NBER Working Paper No. 15181*. Cambridge, MA: National Bureau of Economic Research.

Moore, J.C., Marquis, K.H. et Bogen, K. (1996). The SIPP cognitive research evaluation experiment: Basic results and documentation. *SIPP Working Paper No. 212*. U.S. Census Bureau, Washington, DC (janvier). Disponible au http://www.census.gov/sipp/workpapr/wp9601.pdf [July 2014].

Morganstein, D. et Marker, D. (2000). A conversation with Joseph Waksberg. *Statistical Science*, 15(3), 299-312.

Mule, T. et Konicki, S. (2012). *2010 Census Coverage Measurement Estimation Report: Summary of Estimates of Coverage for Housing Units in the United States*. U.S. Census Bureau, Washington, DC.

National Research Council (1995). *Measuring Poverty: A New Approach*. Washington, DC: The National Academies Press.

National Research Council (2004). *The 2000 Census: Counting Under Adversity*. Washington, DC: The National Academies Press.

National Research Council (2010a). *Envisioning the 2010 Census*. Washington, DC: The National Academies Press.

National Research Council (2010b). *The Prevention and Treatment of Missing Data in Clinical Trials*. Washington, DC: The National Academies Press.

National Research Council (2013a). *Measuring What We Spend: Toward a New Consumer Expenditure Survey*. Washington, DC: The National Academies Press.

National Research Council (2013b). *Nonresponse in Social Science Surveys: A Research Agenda*. Washington, DC: The National Academies Press.

National Research Council (2013c). *Principles and Practices for a Federal Statistical Agency*. Washington, DC: The National Academies Press.

Nelson, N. et West, K. (2014). Interview with Lars Thygesen. *Statistical Journal of the IAOS*, 30, 67-73.

Passero, B. (2009). The impact of income imputation in the Consumer Expenditure Survey. *Monthly Labor Review* (août), 25-42.

Prell, M., Bradsher-Fredrick, H., Comisarow, C., Cornman, S., Cox, C., Denbaly, M., Martinez, R.W., Sabol, W. et Vile, M. (2009). Profiles in success of statistical uses of administrative records. Report of a subcommittee of the *Federal Committee on Statistical Methodology*, U.S. Office of Management and Budget, Washington, DC.

Shapiro, G.M. et Kostanich, D. (1988). High response error and poor coverage are severely hurting the value of household survey data. *Proceedings of the Section on Survey Research Methods*, 443-448, American Statistical Association, Alexandria, VA. Disponible au http://www.amstat.org/sections/srms/Proceedings/papers/1988_081.pdf [July 2014].

Steeh, C.G. (1981). Trends in nonresponse rates, 1952-1979. *Public Opinion Quarterly*, 45, 40-57.

Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology*, 55, 775-801.

U.S. Office of Management and Budget. (2014). *Guidance for Providing and Using Administrative Data for Statistical Purposes*. Memorandum M-14-06. Washington, DC.

Waksberg, J. (1978). Sampling methods for random digit dialing. *Journal of the American Statistical Association*, 73, 40-46.

Woodward, J., Wilson, E. et Chesnut, J. (2007). *Evaluation Report Covering Facilities - Final Report. 2006 American Community Survey Content Test Report H.3.U.S.* Census Bureau. Washington, DC: U.S. Department of Commerce. Janvier.

Zelenak, M.F. et M.C. David (2013). *Impact of Multiple Contacts by Computer-Assisted Telephone Interview and Computer-Assisted Personal Interview on Final Interview Outcome in the American Community Survey*. U.S. Census Bureau, Washington, DC.