

Impact comparé d'un questionnaire rétrospectif par rapport à des relevés budgétaires quotidiens

Christophe Muller¹

Cet article étudie l'impact comparé sur l'estimation de la dépense de consommation, de l'utilisation d'un questionnaire budgétaire rétrospectif par rapport à un questionnaire de relevés quotidiens. Les données d'une enquête nationale au Rwanda montrent que l'utilisation du questionnaire rétrospectif entraîne une sous-estimation de la consommation moyenne. Ceci est dû en partie à la sous-estimation des dépenses récurrentes d'un faible montant. Le niveau de valeur des transactions individuelles apparaît comme un élément essentiel dans le choix du questionnaire budgétaire à utiliser. Les omissions avec ces questionnaires sont susceptibles de conduire à des biais importants dans les estimations, non seulement de consommation agrégée mais également, de modèles linéaires estimés par les moindres carrés ordinaires où la consommation du ménage apparaît comme une variable dépendante ou indépendante. Dans des cas favorables, la direction de ces biais est connue grâce à la connaissance de la corrélation négative des omissions et des valeurs de transaction.

Introduction

De nombreuses publications traitent de la méthodologie des enquêtes sur la consommation auprès des ménages et en particulier de l'impact de l'usage de différents types de questionnaires sur la qualité de la mesure du phénomène étudié, notamment dans les pays en voie de développement². Par ailleurs, les préoccupations au sujet des erreurs de mesures occupent une part importante dans la littérature statistique³. Dans ce contexte, divers auteurs ont étudié les sous-déclarations et mauvaises réponses liées à l'emploi de questionnaires budgétaires spécifiques durant les entretiens d'enquêtes auprès des ménages⁴.

Le fait que plusieurs méthodes de collecte (pesage de la nourriture, entretiens journaliers et rétrospectifs, bulletins de comptes) produisent des informations de qualité inégale est généralement reconnu. Toutefois, la mesure de cette qualité pour un indicateur donné, notamment pour la valeur de la consommation agrégée des ménages, n'est pas bien connue.

Entre autres auteurs, Glaude (1983) pour la France, Grootaert et Cheung (1985), Deaton et Grosh (1998), Deaton et Zaidi (1999), Pradhan (2001) pour divers pays en voie de développement, comparent des questionnaires journaliers et rétrospectifs pour divers biens de consommation. Ce type de comparaisons de questionnaires est généralement organisé en distinguant les produits selon qu'ils sont a priori bien observés à l'aide de

¹ Professeur à l'Université d'Alicante, chercheur associé à DIAL. Je suis reconnaissant pour le support financier du Ministère Espagnol des Sciences et Technologies, Project No. SEJ2005-02829/ECON.

² E.g., Winter (1970), OCDE (1978), Booker et al. (1989), Scott et al. (1980), Casley et Lury (1987), United Nations (1982, 1986), World Fertility Survey (1980), Blaizeau et Dubois (1989), Deaton (1994), Grosh et Munoz (1996), Grosh et Glewwe (1996).

³ Par exemple: Morgenstern (1963), Rossi, Wright et Anderson (1983), Fuller (1987).

⁴ Bradburn (1983), Fowler et Mangione (1990), Silberstein (1989, 1990), Biemer et al. (1991), Cannari et d'Alessio (1991), Maki et Nishiyama (1996), Philipson (1997), Deaton et Grosh (1998), Deaton et Zaidi (1999), Pradhan (2001).

l'un ou l'autre des questionnaires comparés. On considère aussi les fréquences plausibles de transactions pour différents biens afin de choisir entre un questionnaire basé sur des relevés quotidiens et un autre basé sur des relevés rétrospectifs.

Dans cette note nous présentons un élément empirique intéressant à propos des propriétés de collectes de relevés budgétaires quotidiens et rétrospectifs dans un pays Africain. L'originalité de cet élément est de montrer comment les propriétés de chaque questionnaire pouvant dépendre de la valeur de la transaction enregistrée dans le questionnaire, et pas seulement du produit concerné ou de la fréquence de transaction.

Nous présentons les données dans une première partie, puis nous discutons la comparaison de questionnaires budgétaires rétrospectifs et quotidiens pour une enquête au Rwanda dans la partie suivante.

Les données

Les données proviennent d'une enquête nationale budget consommation réalisée au Rwanda en 1982-83, bien avant la guerre de 1994⁵. Il s'agit donc de données assez anciennes qui gardent toutefois ici leur intérêt pour une simple note méthodologique. Elles sont aussi intéressantes en raison de leur qualité vraisemblablement exceptionnelle due à l'intensité de la collecte. Ainsi, on peut espérer pouvoir dégager les propriétés de questionnaires différents dans une situation d'observation relativement favorable.

L'échantillon de ménages enquêtés représente la population rurale du Rwanda en 1982-83 (95 % de toute la population à l'époque de l'enquête). Le plan de sondage (Roy, 1983) a quatre niveaux de sondage et trois stratifications géographiques au premier niveau. Un district fut tiré dans chaque commune tirée, et une grappe de trois ménages fut tirée dans chaque district tiré.

La collecte des données fut réalisée de novembre 1982 à décembre 1983 auprès de 270 ménages. Toutefois seulement 256 ménages sont utilisés dans les estimations en raison de valeurs manquantes. Les ménages sélectionnés

⁵ Cette enquête fut conduite par le Ministère du Plan du Rwanda en collaboration avec le Ministère Français de la Coopération et du Développement (Ministère du Plan, 1986 a, b). L'expertise de l'INSEE fut essentielle pour la conception et l'exécution de la collecte et du plan de sondage. Voir Muller (1992, 2005) pour plus de détails sur les indicateurs de consommation de cette enquête.

furent interrogés sur leurs caractéristiques démographiques, leur budget, leur consommation et leurs activités journalières. Le volume des récipients présents dans chaque ménage fut mesuré avant la collecte des informations budgétaires et les unités de mesure traditionnelles furent également étalonnées. De cette façon, des mesures précises de quantités de produits achetés ont pu être obtenues. Chaque ménage fut visité par un enquêteur au moins une fois par jour (souvent deux fois), durant deux semaines à chacun des quatre trimestres de l'enquête. Cet enquêteur posa des questions et effectua des pesées de nourriture dont les résultats furent enregistrés dans les questionnaires de l'enquête. De plus, chaque ménage enquêté remplit des carnets de compte durant l'absence de l'enquêteur qui les révisa à son retour dans le ménage.

L'ensemble des sujets touchés par l'enquête fut très large et plusieurs questions furent traitées de façon redondante par diverses méthodes de collecte. Par exemple, il nous fut possible de confronter les transactions alimentaires avec les pesées de nourriture lors de la préparation des repas, et avec les déclarations d'emploi du temps des membres du ménage. Cette situation nous permit d'entreprendre un nettoyage approfondi des données et d'utiliser des algorithmes sophistiqués de vérification basés sur les redondances présentes dans les questionnaires.

Muller (2005) décrit les procédures complexes de valorisation de l'autoconsommation et la consommation reçue en cadeau, à partir des informations tirées d'une enquête de prix spéciale et des valeurs unitaires de l'enquête. Dans l'ensemble, le prix adopté pour la consommation est très proche du prix d'achat au marché le plus proche pour le même trimestre et le même district.

A partir de ces informations recueillies, nous avons calculé des indicateurs de consommation. Nous décrivons maintenant la comparaison des questionnaires budgétaires quotidiens et rétrospectifs de cette enquête en fonction du niveau de valeur de la transaction.

La comparaison de questionnaires budgétaires journaliers et rétrospectifs

La méthode

Le calcul des indicateurs de consommation finale des ménages commence avec le repérage des enregistrements pertinents dans chaque questionnaire. Pour ce faire les modalités de chaque

variable constituant l'enregistrement sont soigneusement examinées. On sélectionne par exemple uniquement les dépenses strictement de consommations dans les questionnaires décrivant les transactions du ménage. La définition précise des enregistrements à conserver pour l'étude de la consommation peut être basée par exemple sur les définitions de la comptabilité nationale, ce qui correspond généralement aux objectifs des instituts de statistique (INSEE, 1983).

Une approche analytique simple de la comparaison de questionnaires quotidiens et rétrospectifs consiste, pour un produit donné, à comparer les niveaux de consommations enregistrés dans les deux questionnaires, en les ramenant à une période commune. En fonction du résultat de la comparaison, il est possible de combiner les deux mesures provenant de chaque questionnaire, ou de garder la mesure jugée la plus fiable⁶.

Dans cette note, nous faisons une hypothèse supplémentaire qui consiste à supposer que le niveau observé de consommation le plus élevé correspond à la meilleure mesure. Cette hypothèse est cohérente avec la dominance d'erreurs d'omissions lors de la collecte, ce qui est plausible dans de nombreux cas. Toutefois elle peut conduire à négliger les effets télescopiques, c'est-à-dire la possible surestimation de la consommation causée par l'inclusion d'enregistrements correspondant à une période passée hors du champ du questionnaire. La littérature statistique semble indiquer que dans la grande majorité des enquêtes budgétaires, les omissions provoquent généralement des erreurs plus graves que l'effet télescopique, ce qui justifie notre hypothèse.

Dans la littérature sur les enquêtes de consommation la comparaison de relevés budgétaires quotidiens et rétrospectifs est généralement basée sur : (1) la prise en compte des produits échangés lors de la transaction, et (2) les fréquences probables de transaction pour ces produits. Dans cette note nous fournissons une table basée sur la valeur de la consommation agrégée, sans considération du type de produit et des fréquences de transaction. En effet, cette table est tout ce qui subsiste d'archives décrivant le traitement de la comparaison de questionnaires pour cette enquête.

⁶ Signalons qu'une considération importante dans le choix du questionnaire est le coût de l'enquête et pas seulement les performances statistiques qui sont l'objet de cette note. Un arbitrage, similaire à celui du choix de la taille de l'échantillon et plan de sondage, peut être fait entre augmenter la précision et diminuer le coût financier de l'enquête.

Le questionnaire de relevés budgétaires rétrospectifs (QR) de l'enquête budgétaire du Rwanda en 1983 a été utilisé pour enregistrer les transactions réalisées lors des dix semaines précédant le passage de l'enquêteur. Les transactions de valeur inférieure à 250 Frw⁷ furent a priori exclues de ces enregistrements, bien que de nombreuses transactions de cette sorte aient été en fait conservées dans les fichiers de collecte. Ces instructions de collecte étaient justifiées par le fait qu'il paraît peu probable que des ménages, même pauvres, puissent se rappeler convenablement des transactions de faible valeur.

Le questionnaire de relevés budgétaires quotidiens (QQ) a été administré de façon journalière durant deux semaines à l'aide d'un carnet de compte laissé au ménage et de visites journalières de l'enquêteur. Ces questionnaires sont décrits dans Ministère du Plan (1986a). Dans ces conditions le questionnaire QQ est utilisé pour mesurer les dépenses de valeur individuelles inférieures à 250 Frw. La question est : Quel questionnaire utiliser pour mesurer les transactions d'achat de consommation de valeur individuelle de 250 Frw ou plus ?

Le test

Le tableau 1 présente les niveaux cumulés de valeurs d'achat de consommation totale pour les 265 ménages enquêtés, par tranches de 50 Frw pour la valeur de la transaction individuelle, et pour chaque questionnaire. Seules les transactions au-delà de 250 Frw sont présentes dans cette table. Il s'agit d'une description de la queue supérieure de la distribution des transactions d'achats de consommation selon leur valeur, pour chacun des questionnaires étudiés. En pratique la majorité des transactions pour les ménages Rwandais ruraux de cette époque correspondent à des transactions de très faible valeur, en dessous de 250 Frw. Ces transactions sont clairement mal enregistrées à l'aide d'un questionnaire rétrospectif portant sur plusieurs semaines. Il n'y a donc pas d'ambiguïté sur le choix du questionnaire à utiliser dans ce cas : il s'agit du questionnaire QQ basé sur les relevés quotidiens. Pour le reste il est logique de concentrer l'attention uniquement sur la queue de distribution des valeurs de transactions. Aucune transaction ne fut relevée au-delà d'une valeur de 1.300 Frw.

⁷ Frw = Franc Rwandais. En 1983, le taux de change moyen était de 100,17 Frw pour un US \$ de 1983 (sources: IMF, Penn Tables).

Tableau 1 :

Comparaison des niveaux de valeur cumulée des transactions de consommation enregistrées avec les questionnaires budgétaires rétrospectifs (QR) et quotidiens (QQ) par tranches de 50 Frw (au-delà de 250 Frw) :

Valeur de la transaction	QR	QQ	5 x QQ	Ratio	P-value du test de Wilcoxon	P-value du test des signes	Résultat du test des signes de H_0
251-300	97000 (372)	42000 (162)	210000	0,46	0,0067	0,0013	Rejet à 1 %
301-350	235000 (768)	104000 (341)	520000	0,45	0,0105	0,0022	Rejet à 1 %
351-400	129000 (366)	41000 (113)	205000	0,62	0,0184	0,0038	Rejet à 1 %
401-450	259000 (640)	65000 (160)	325000	0,79	0,0035	0,0023	Rejet à 1 %
451-500	104000 (225)	34000 (73)	170000	0,61	0,0060	0,0021	Rejet à 1 %
501-550	274000 (544)	98000 (195)	490000	0,55	0,0104	0,0037	Rejet à 1 %
551-600	60000 (107)	22000 (40)	110000	0,54	0,0182	0,0065	Rejet à 1 %
601-650	270000 (446)	64000 (105)	320000	0,84	0,0325	0,0112	Rejet à 2 %
651-700	58000 (88)	17000 (25)	85000	0,68	0,0587	0,0193	Rejet à 2 %
701-750	173000 (245)	40000 (57)	200000	0,86	0,091	0,0327	Rejet à 4 %
751-800	93000 (123)	24000 (31)	120000	0,78	0,142	0,054	Non Rejet à 5 %, Rejet à 6 %
801-850	200000 (249)	41000 (51)	205000	0,98	0,230	0,0898	Non Rejet à 5 %, Rejet à 9 %
851-900	47000 (55)	11000 (13)	55000	0,85	0,285	0,144	Non Rejet
901-950	138000 (153)	28000 (24)	110000	1,25	0,361	0,226	Non Rejet
951-1000	35000 (36)	14000 (15)	70000	0,50	0,149	0,109	Non Rejet
1001-1050	399000 (39)	82000 (82)	410000	0,97	0,291	0,187	Non Rejet
1051-1100	40000 (38)	18000 (17)	90000	0,44	0,414	0,312	Non Rejet
1101-1150	81000 (73)	18000 (16)	90000	0,90	0,852	0,50	Non Rejet
1151-1200	28000 (24)	7000 (6)	35000	0,80	0,781	0,75	Non Rejet
1201-1250	184000 (153)	34000 (28)	170000	1,08	0,317	1	Non Rejet
1251-1300	50000 (40)	10000 (8)	50000	1	-	-	Non Rejet

Le nombre de transactions relevées est présenté entre parenthèses. Les tests sont effectués sur la base des tranches supérieures ou égales à la rangée considérée.

Le test de Wilcoxon et le test des signes sont calculés à partir des distributions approximées avec les tranches de valeur. L'hypothèse nulle est H_0 : QR = 5 QQ (pour la valeur centrale cumulée) qui est testée contre l'hypothèse alternative H'_a : QR \neq 5 QQ pour le test de Wilcoxon, et H'_a : QR < 5 QQ pour le test des signes.

Pour faciliter la comparaison nous avons également indiqué la valeur obtenue en extrapolant les données de QQ à la même période de référence pour QR (dans la colonne « 5 x QQ »). Nous présentons entre parenthèses le nombre de

transactions relevées pour chaque questionnaire et chaque tranche de valeur. Ces effectifs de transaction d'achats de consommation sont souvent faibles ce qui s'explique par la dominance de l'autoconsommation des ménages ruraux Rwandais

à cette époque⁸. Un problème est que seuls les chiffres indiqués dans la table sont disponibles. En particulier, on ne dispose pas des écart-types qui permettraient de mettre en œuvre des tests de Student. Dans une telle situation on souhaiterait également pouvoir utiliser des tests séquentiels de dominance stochastique du type de ceux proposés par Davidson et Duclos (2000), mais ce n'est pas possible ici.

Nous mettons en œuvre des tests non paramétriques calculables à partir des statistiques descriptives disponibles. Le test de Wilcoxon et le test des Signes (voir par exemple Manoukian, 1986) nous permettent de repérer l'existence d'une différence significative entre les questionnaires QQ et QR en dessous de valeurs assez fortes de transactions individuelles, mais pas au-delà d'un certain seuil.

L'hypothèse nulle à tester est H_0 : « la tendance centrale de la distribution des observations représentées dans la table (i.e. pour les différentes tranches de valeur) est identique pour la colonne QR et pour la colonne QQ ». Nous notons pour simplifier, H_0 : « $QR = 5 QQ$ ». L'hypothèse alternative est H_a : « $QR \neq 5 QQ$ » dans le cas du test de Wilcoxon dont la statistique est la somme des rangs des différences positives des deux colonnes « QR » et « 5 QQ ». L'hypothèse alternative dans le cas du test des signes, dont la statistique est le nombre de signes positifs de la différence des deux colonnes, est H_a : « $QR < 5 QQ$ » afin de prendre en compte la forte conviction qu'un questionnaire quotidien permet de collecter plus d'information. En effet, ce qui compte vraiment est de mesurer si le degré d'omissions relatives avec QR, par rapport à QQ, est significatif. Nous donnons donc plus d'importance au résultat du test des signes, le test de Wilcoxon étant utilisé comme un simple contrôle.

Les P-values des deux tests sont montrées dans la table pour chaque classe de valeur de transactions individuelles. Elles correspondent aux tests effectués à partir des observations des tranches correspondant à cette rangée et les rangées successives. Ainsi plus on descend dans la table, moins de transactions de faible valeur sont considérées pour les tests, étant donné qu'on pense que QR fonctionne moins bien pour ces transactions.

Les résultats du test des signes montrent que H_a est rejeté au niveau de 5 % (et même à 4 %), dès qu'on prend en compte des valeurs de transaction individuelles inférieures ou égales à 750 Frw

(850 Frw au niveau de 10 %). Par conséquent, si on se limite à des valeurs individuelles de transaction supérieures à 750 Frw, le test des signes ne rejette pas l'hypothèse H_a , que le questionnaire QR fournit des estimations plus élevées que le questionnaire QR. En ce qui concerne les résultats du test de Wilcoxon, l'identité des résultats des questionnaires QQ et QR n'est pas rejetée à 5 % dès qu'on inclut des transactions de valeur individuelle inférieure ou égale à 650 Frw.

Ainsi, les résultats des tests montrent que QR correspond à des omissions, relativement à QQ, pour les transactions de faible valeur individuelle, mais que ces omissions ne sont plus significatives lorsque l'on considère seulement les transactions de valeur individuelle suffisamment élevée.

Discussion des résultats

Les résultats obtenus montrent que le questionnaire QQ permet d'enregistrer une valeur d'achats de consommation significativement plus élevée jusqu'à une valeur individuelle de transaction de 750 Frw (correspondant à environ 7,45 Dollars Américains de 1983, ce qui est considérable pour des ménages aussi pauvres). Le ratio des colonnes « QR » sur « 5xQQ » est de 0,45 à 0,86 pour ces tranches de valeurs. Au-delà de cette valeur individuelle de transaction, QR est généralement préférable dans la mesure où ses enregistrements correspondent à un niveau cumulé de valeurs similaire à QQ, avec des ratios entre 0,78 et 1,25, pour ces tranches de valeur, excepté pour les tranches [951, 1.000] et [1.051, 1.100], où ce ratio est plus faible.

Ces résultats confirment la stratégie de collecte des concepteurs de cette enquête qui était de n'utiliser que le QQ pour relever des informations sur les transactions de faible valeur inférieures à 250 Frw. Toutefois, l'approche fut encore trop ambitieuse, et il semble qu'un seuil de 750 Frw aurait du être adopté pour éviter des problèmes d'omissions massives de transactions avec le questionnaire QR.

Une interprétation de ces résultats est d'abord que le questionnaire rétrospectif QR est caractérisé par des omissions presque systématiques pour des valeurs de transaction inférieure à 750 Frw. Ensuite pour des valeurs de transaction plus élevées, QQ et QR deviennent plus ou moins comparables.

Le fait que QQ, basé sur une procédure d'enquête intensive, puisse être jugé équivalent à QR pour des grandes valeurs de transaction, qui correspond à une procédure légère, peut provenir de plusieurs sources. D'abord, les statistiques basées sur les erreurs relevées grâce à QQ et QR sont sensibles (plus pour QQ) aux fluctuations aléatoires dues aux

⁸ Dans cette enquête, l'autoconsommation était relevée directement à l'aide de pesées alimentaires qui ont été enregistrées dans un autre questionnaire.

petits échantillons des transactions considérées. Ainsi, l'absence de différence significative entre les performances de QQ et de QR à partir d'une valeur de transaction de 750 Frw est peut-être en partie attribuable à de telles fluctuations dans une situation où les échantillons de transactions sont réduits⁹. D'autre part, l'importance des transactions relevées avec QR au-delà de 750 Frw est peut-être en partie causée par l'existence d'effets rétrospectifs affectant particulièrement les transactions de grande valeur. C'est-à-dire que certaines transactions de grande valeur enregistrées dans QR peuvent provenir d'une période antérieure à la période de référence de dix semaines pour ce questionnaire. Toutefois, cette interprétation est peu crédible ici en raison de la structure en données de panel de la collecte qui nous a permis de détecter de tels cas (chaque ménage est enquêté successivement durant quatre trimestres). En effet, une transaction enregistrée par erreur avec QR et correspondant à un retard de onze semaines par exemple, apparaîtrait également avec le questionnaire QQ du passage précédent et aurait pu être facilement corrigée.

Dans l'ensemble, les résultats, notamment ceux pour les tranches inférieures à 750 Frw, indiquent qu'il n'existe pas d'évidence en faveur d'effets rétrospectifs *systématiques* lors de l'usage de QR. Ces résultats révèlent plutôt des omissions massives causées par l'utilisation de QR pour des transactions de faible valeur. Bien entendu, des omissions ont certainement eu lieu également avec QQ, et d'autres types d'erreurs de mesure peuvent aussi intervenir à la fois pour QQ et QR¹⁰.

⁹ En dépit de la faible taille d'échantillon et du faible nombre de transactions relevées, un calcul rapide d'erreur standard basé sur des écart-types réalistes et un tirage aléatoire simple montre qu'il est peu plausible que les différences de moyennes considérées soient systématiquement non-significatives. Cette remarque couvre les erreurs dues au sondage et le bruit dû à la plupart des erreurs aléatoires centrées, dus à la saisie et aux autres erreurs de mesure. Ce sont donc des erreurs systématiques de mesure qui apparaissent largement dans les différences de moyennes dans la table.

¹⁰ En effet, l'estimation des dépenses n'est pas uniquement affectée par la méthode d'enquêtes (QQ, QR) mais aussi par le niveau de détail de la définition des produits considérés. Plus le module des dépenses du questionnaire est désagrégé, plus importante sera la dépense totale estimée par rapport aux questionnaires plus courts. Ceci a été démontré dans le cas de l'Indonésie par Pradhan (2001). Dans le 1^{er} cas, un questionnaire long capte 27 % de dépenses alimentaires et 40 % des dépenses non alimentaires en plus que dans un questionnaire court. Pradhan trouve que le questionnaire court sous-estime la dépense totale d'entre 11,7 % et 19,6 % en moyenne (Pradhan, 2001). Il est aussi avéré que le choix de l'informateur a aussi une incidence sur le montant estimé des dépenses. Si au lieu d'enquêter uniquement le chef du ménage on interroge

Une autre possibilité, impossible à vérifier ici, est que des omissions apparemment plus importantes avec QR pour ces faibles valeurs de transaction correspondent en fait à une plus mauvaise performance de ce questionnaire pour certains produits ou certaines fréquences d'achat qui seraient corrélées avec ces faibles valeurs.

Finalement, il est encore possible que les différences observées entre QR et QQ soient dues au léger décalage des périodes de références des questionnaires. Toutefois, cette hypothèse paraît peu vraisemblable pour une enquête ayant porté sur quatre trimestres complets, et couvrant l'ensemble des phénomènes saisonniers, les plus susceptibles de différencier les périodes de références des questionnaires.

Conséquence des résultats pour les estimations

Nous avons interprété nos résultats en termes d'erreurs d'omissions dépendant de la valeur de la transaction enregistrée dans le questionnaire QQ. A partir de cette perspective les résultats suggèrent qu'utiliser des questionnaires budgétaires rétrospectifs peut entraîner une sous-estimation des estimations de consommation agrégée (ou moyenne) dans les enquêtes auprès des ménages même lorsque leur usage est réduit à la mesure de transactions de valeur relativement élevée.

Mais ce n'est pas tout. Les données de telles enquêtes sont souvent utilisées pour estimer des modèles économétriques dans lesquelles la consommation de chaque ménage intervient comme variable dépendante ou indépendante. Quelle est la conséquence d'erreurs de mesure sur cette variable lorsque seul un questionnaire budgétaire rétrospectif a été utilisé pour la mesure comme c'est souvent le cas ? Les résultats de la Table 1 peuvent être utilisés pour explorer la réponse à cette question dans la mesure où ils fournissent une représentation approchée des erreurs d'omission de ce questionnaire.

aussi toutes les membres du ménage ayant une autonomie dans les dépenses, le montant total sera supérieur à celui déclaré par le chef du ménage.

Il existe d'autres sources d'erreurs d'estimation des dépenses. Pour ne citer que quelques-unes : La nécessité d'utiliser un déflateur spatial de prix afin de tenir compte des disparités régionales des niveaux de prix ; la méthode d'imputation des loyers fictifs ; la prise en compte ou non de la valeur d'usage des biens durables ; le choix des informateurs (chef du ménage ; chaque membre ayant une autonomie dans les décisions d'achat, etc.) ; les imputations de prix pour les produits autoconsommés ; la distinction entre consommation intermédiaire et finale pour les ménages ayant une activité productive à l'intérieur du logement ; la prise en compte des repas hors du ménage.

Le schéma usuel des ‘erreurs classiques de mesure’ dans un modèle économétrique linéaire est le suivant :

$$\begin{aligned} y^* &= X^* b + e ; \\ X &= X^* + u, \\ y &= y^* + v, \end{aligned}$$

où y^* et X^* sont des matrices représentant des variables inobservées (en empilant les observations au niveau ménage d’un échantillon d’enquête), y et X sont leur contreparties observées, u et v sont les erreurs de mesure correspondantes, et e est le terme d’erreur du modèle d’intérêt reliant les variables d’intérêt y^* et X^* . Dans le schéma classique, u et v sont supposées non corrélées avec y^* , X^* et e .

Or, nos résultats indiquent que les omissions causées par l’utilisation de QR sont corrélées avec les valeurs des transactions individuelles considérées. Si on suppose que plus un ménage a un bas niveau de vie, plus ses transactions sont de faible valeur (par exemple à cause de contraintes de liquidité de court terme, ou d’un manque de capacité de stockage), alors utiliser seulement les données provenant des questionnaires rétrospectifs conduit à des biais d’estimation des moindres carrés ordinaires (MCO) dans les modèles linéaires où la consommation totale du ménage est utilisée comme variable dépendante ou variable indépendante (voir par exemple Bound et al, 2005).

Lorsque la consommation est la variable dépendante, les résultats des MCO sont « atténués » par les erreurs d’omission dans QR, conduisant à des coefficients trop peu significatifs. Ceci est dû à la corrélation négative des omissions avec le niveau de la variable de consommation.

Lorsque la consommation est incluse comme variable indépendante dans le modèle et que les

erreurs d’omissions sont corrélées avec les autres variables du modèle, les biais des coefficients estimés dépendent de la direction des corrélations partielles entre les erreurs et ces variables. En général, il existe des biais mais leur direction ne peut être anticipée. Si de plus on sait que les erreurs d’omissions ne sont pas corrélées avec les autres variables du modèle, alors on peut en déduire que l’effet de la consommation sur la variable dépendante est surestimé par les MCO. En effet, ici en raison de la corrélation négative des erreurs d’omission avec la variable mesurée – la consommation – le biais « d’atténuation » typique se traduit par une surestimation au lieu d’une sous-estimation (See Bound et al, 2000).

Résumons maintenant la discussion, en mettant en exergue les principaux enseignements de la table commentée dans cette note. Tout d’abord il apparaît que le niveau de valeur des transactions individuelles est un élément essentiel dans le choix du questionnaire budgétaire à utiliser, ce qui complète les méthodes usuelles de choix de questionnaire basées uniquement sur le type de produit et les fréquences probables des transactions.

Ensuite, même pour des niveaux de valeur de transaction relativement importants au vu de la pauvreté des ménages considérés, les omissions observées avec des questionnaires budgétaires à relevés rétrospectifs sont susceptibles de conduire à des biais importants dans les estimations non seulement de consommation agrégée en moyenne par ménage, mais aussi de modèles linéaires estimés par les MCO où la consommation du ménage apparaît comme une variable dépendante ou indépendante. Dans des cas favorables, la direction de ces biais est connue, grâce à la connaissance de la corrélation négative des omissions et des valeurs de transactions.

Références Bibliographiques

Biemer P.P, Groves R.M., Lyberg L.E., Mathiowetz N.A., Sudman S. (Eds.) (1991), *Measurement errors in surveys*, John Wiley, New York.

Blaizeau, D., Dubois J.-L. (1989), *Connaître les conditions de vie des ménages dans des pays en voie de développement*, Ministère de la Coopération et du Développement.

Booker W., Singh P., Savane L. (1980), « Household survey experience in Africa », *Living Standard Measurement Studies*, n°6.

Bound J., Brown C., Mathiowetz N. (2005), « Measurement errors in survey data », Chapter 59 in J.J. Heckman And E. Leamer (Eds.), *Handbook of Econometrics, Volume 5*, North-Holland.

Bradburn N.M. (1983), « Response Effects », Chapter 8 in Rossi, Wright And Anderson.

Cannari L., D’Alessio G. (1978), « Non-reporting and under-reporting behaviour in the bank of Italy’s survey of household income and wealth », OCDE;

- Casley D.J., Lury D.A. (1987)**, *Data collection in developing countries* Clarendon Press, Oxford.
- Davidson R., Duclos J-Y. (2000)**, « Statistical inference for stochastic dominance and for the measurement of poverty and inequality », *Econometrica*, Vol. 68, n 6, 1435-1464, November.
- Deaton A. (1997)**, *The analysis of household surveys: a microeconomic approach to development policy*, John Hopkins University Press.
- Deaton A., Grosh M. (1998)**, « Consumption » in P. Glewwe & M. Grosh (Eds.) *Designing household survey questionnaires for developing countries: lessons from ten years of LSMS experience*, World Bank, Washington DC.
- Deaton A., Zaidi S. (1999)**, « Guidelines for constructing consumption aggregates for welfare analysis », Processed. Princeton University and Development Research Group, World Bank, Washington DC.
- Fowler F.J. Jr., Mangione T.W. (1990)**, *Minimizing interviewer-related error*, Sage Publications Inc.
- Fuller W.A. (1987)**, *Measurement Error Models*, John Wiley & Sons.
- Grootaert C.G., Cheung K.F. (1985)**, « Household expenditure surveys: some methodological issues », *Living Standard Measurement Studies*, n 20.
- Grosh M.E., Glewwe P. (1996)**, « Household survey data from developing countries: progress and prospects », *American Economic Association Papers and Proceedings*, May.
- Grosh M.E., Munoz J. (1996)**, « A manual for planning and implementing the living standards measurement study survey », *Living Standards Measurement Study*, working paper n 126.
- INSEE (1983)**, « Système élargi de comptabilité nationale », *Collection C*, n 44-45.
- Maki A., Nishiyama S. (1996)**, « An analysis of under-reporting for micro-data sets: the misreporting or double-hurdle model », *Economics Letters*, 52, 211-220.
- Manoukian E.B. (1986)**, *Guide de statistique appliquée*, Hermann.
- Ministère du Plan du Rwanda (1986)**, *Enquête nationale sur le budget et la consommation des ménages. Volume 1 : Présentation méthodologique de l'échantillon et de la collecte en milieu rural*, Kigali, Rwanda.
- Ministère du Plan du Rwanda (1986)**, *Enquête nationale budget consommation 1982-83 en milieu rural. Volume 2 : Structure du budget des ménages*, Kigali, Rwanda.
- Morgenstern O. (1963)**, *On the accuracy of economic observations*, Princeton University Press.
- Muller C. (1992)**, « Estimation des consommations de producteurs agricoles d'Afrique centrale », *Economie et Prévision*, n°105.
- Muller C. (2005)**, « The Valuation of non-monetary consumption in household surveys », *Social Indicators Research*, Vol. 72, n 3, 319-341, July.
- O.C.D.E. (1978)**, *Multi-purpose household surveys in developing countries*, Paris.
- Philipson T. (1997)**, « Data Markets And The Production Of Surveys », *Review of Economic Studies*, 64.
- Pradhan M. (2001)**, « Welfare analysis with a proxy consumption measure. Evidence from a repeated experiment in Indonesia », *Cornell Food and Nutrition Policy Program working paper n 112*, Ithaca, New-York.
- Rossi P.H., Wright J.D., Anderson A.B. (Eds.) (1983)**, *Handbook of Survey Research*, Academic Press Inc.

Scott C., De Andre P.T.A., Chander R. (1980), « Conducting surveys in developing countries », *Living Standard Measurement Studies*, n°5, October.

Silberstein A.R. (1989), « Recall effects in the U.S. Consumer expenditure interview survey », *Journal of Official Statistics*, Vol. 5, n 2.

Silberstein A.R. (1990), « Effets du premier cycle d'interviews dans la consumer expenditure survey aux E.U. », *Techniques d'enquêtes*, Vol. 16, No. 2, Décembre.

United Nations (1986), « Manuel des méthodes d'enquêtes sur les ménages », *Etudes Méthodologiques*, Série F, n°31.

World Fertility Survey (1976), *Editing and Coding Manual*, I.S.I, Basic Documentation n 7, London.

United Nations (1962), *National household survey capability programme : survey data processing. A Review of Issues and Procedures*.

Winter G. (1970), *Méthodologie des enquêtes niveau de vie en milieu rural africain*, Orstom.

