

Effet du plan de sondage dans des enquêtes emploi : les enquêtes 1-2-3 en Afrique de l'ouest

Eloi Ouedraogo
Aude Vescovo¹

Cet article établit un diagnostic sur les plans de sondage des phases 1 des enquêtes 1-2-3, enquêtes en phase sur l'emploi et le secteur informel, réalisées entre 2001 et 2003 dans 7 capitales économiques d'Afrique de l'Ouest. Dans la première partie de l'article sont discutées les implications des plans de sondage stratifiés et des plans de sondage à deux degrés en termes d'estimation de la précision. La seconde partie de l'article discute des choix opérés des enquêtes étudiées et des recommandations sont formulées en conclusion. Dans un contexte où les bases de sondage sont souvent anciennes, il est notamment recommandé de se limiter à une stratification à 1 degré, sur la base du découpage administratif des communes. Il est fortement souligné la nécessité de documenter les plans de sondage utilisés, pratique malheureusement rarement mise en œuvre. Enfin, l'importance de la prise en compte des effets de grappe dans le calcul de précision des estimateurs est démontrée. Les applications sont effectuées sur les données de Niamey et Ouagadougou. C'est ainsi que, dans le cas de Niamey, la prise en compte de l'effet de grappe des plans de sondage conduit à une variance du taux de chômage 2,3 fois supérieure à la variance sans prise en compte de cet effet.

Introduction

Les Etats d'Afrique subsaharienne et les partenaires techniques et financiers ont orienté, dans le cadre de la poursuite des objectifs du millénaire pour le développement, l'essentiel de leur politique économique vers la réduction de la pauvreté. L'immense majorité des pauvres des pays africains tirent leurs revenus du travail, facteur le plus abondant. Il s'avère donc nécessaire de mettre en place des systèmes performants de collecte d'informations sur le chômage et l'emploi.

Les enquêtes sur l'emploi auprès des ménages répondent au besoin de statistiques pertinentes, fiables et comparables sur ces phénomènes. La qualité de ces statistiques du marché de l'emploi dépend des choix méthodologiques opérés. Pourtant peu d'enquêtes effectuées en Afrique ont fait l'objet

d'évaluations méthodologiques, d'une part par manque de temps ou de ressources humaines, mais d'autre part aussi du fait de la focalisation sur les résultats au détriment de la méthode. Malgré la publication de plusieurs guides d'évaluation de la qualité des enquêtes auprès des ménages², peu d'études de qualité sont faites sur les enquêtes ménages en Afrique subsaharienne, que ces enquêtes soient mises en œuvre par les instituts nationaux de statistique ou par des organisations internationales.

La qualité d'une enquête peut être définie selon de nombreux aspects méthodologiques allant de la fixation des objectifs de l'enquête à l'analyse et la dissémination des résultats, en passant évidemment par la précision des estimations. Les questions de plan de sondage se posent tout au long du processus d'enquête. Le plan de sondage d'une enquête doit

¹ Eloi Ouedraogo, AFRISTAT, Département des Statistiques Sociales et des Stratégies de Réduction de la Pauvreté. Aude Vescovo, IRD, Centre de Recherche Appliquée (CERA) d'AFRISTAT.

² Par exemple, *Lignes directrices concernant la qualité*, Statistique Canada, 4^{ème} édition, Octobre 2003.

être dessiné en fonction des objectifs analytiques de celle-ci et de façon à assurer des estimateurs sans biais et précis. Il peut être nécessaire de redresser l'échantillon lors de la phase d'analyse. Enfin il doit faire l'objet d'une documentation, à publier ou à fournir en même temps que les autres fichiers de l'enquête.

Cet article est consacré à l'analyse de la précision des estimations dans les enquêtes auprès des ménages. L'analyse de la précision consiste en l'estimation de la variance des statistiques d'intérêt estimées. Plusieurs méthodes d'estimation de variance peuvent être utilisées. Par exemple, le *bootstrap*, méthode de rééchantillonnage, peut servir à estimer un résultat et sa précision. Cependant, il ne prend pas en compte le mode de tirage des ménages interrogés. Aussi la variance estimée par *bootstrap* est potentiellement biaisée, menant à une sous ou sur-estimation de la précision du résultat. L'analyse de la précision doit donc tenir compte de tous les choix méthodologiques relatifs au plan de sondage.

La plupart des estimateurs sont sans biais et convergents s'ils sont calculés sur des échantillons tirés selon un sondage aléatoire simple (SAS). Un sondage aléatoire simple est très souvent irréalisable dans la pratique d'enquêtes, d'une part pour des questions de coût, d'autre part car l'échantillon issu d'un sondage aléatoire simple peut, par hasard, ne pas comporter certaines catégories de la population. Il est donc d'usage de dessiner des plans de sondage complexes, résultats d'arbitrages entre une bonne couverture de l'échantillon, la précision des estimations et le coût financier du travail de terrain.

La variance d'un estimateur, calculée à partir d'un plan de sondage complexe, est souvent sous-estimée lorsqu'elle est calculée naïvement (i.e. en supposant SAS). Dans de tels cas, l'effet du plan de sondage donne une idée de l'ampleur de cette sous-estimation. Bien que beaucoup d'études empiriques aient montré que l'effet du plan de sondage n'est pas négligeable³, cette dimension est très souvent ignorée dans les calculs de précision. Dans cet article, nous présentons les implications en termes d'estimation de la précision de quelques options du plan de sondage, puis nous en étudions l'impact sur la mesure de la précision d'indicateurs tels que les taux de chômage et d'activité, à partir de la phase 1 des enquêtes 1-2-3, enquêtes en phase sur l'emploi et le secteur informel, réalisées entre 2001 et 2003 dans 7 capitales économiques d'Afrique de l'Ouest.

³ Voir Ganninger Häder et Gabler (2007) qui citent une dizaine de ces études empiriques.

Plan de sondage complexe et mesure de la précision

Chacun des choix méthodologiques relatifs au plan de sondage a potentiellement un impact, qu'il convient d'évaluer, sur la précision des estimations. Ignorer ces éléments lors du calcul de précision mène à des sur ou sous-estimations de variance. Parmi ces choix méthodologiques, la stratification apporte un gain de précision qu'il est nécessaire de prendre en compte sous peine de sous-estimer la précision de l'estimation (sur-estimer la variance). Un tirage à deux degrés et/ou à probabilités inégales induit des effets dits de grappe qui conduisent à une sous-estimation de la variance si le type de plan de sondage n'est pas introduit dans l'analyse de précision.

La stratification

La stratification réduit la variance, par rapport à un échantillon de même taille issu d'un sondage aléatoire simple. Elle permet aussi d'assurer un nombre suffisant d'observations issues de sous-populations d'intérêt. Une strate sera sur-représentée dans l'échantillon si elle présente un intérêt particulier ou le risque que les individus de cette strate répondent moins. Par exemple, il est intéressant de sur-représenter une strate de haut standing si une sous-déclaration des hauts revenus est anticipée, ou une strate à fort taux d'emploi privé formel.

La stratification apporte un gain de précision d'autant plus grand que les variables auxiliaires sur lesquelles elle est basée sont corrélées aux variables d'intérêt. Même si elle est basée sur des critères peu corrélés aux variables d'intérêt, la précision sera inférieure ou égale à celle qui aurait été obtenue sans stratification. Il est donc préférable d'y recourir, même sans conviction sur les critères de stratification. A priori, les divisions administratives des villes constituent des strates tout à fait acceptables, le niveau de vie des habitants étant relativement homogène au sein des strates et différent entre les strates, selon le standing des quartiers. Les stratifications basées sur des données antérieures de la variable d'intérêt (par exemple le niveau d'informalité) sont également pertinentes, dès lors que la qualité des données antérieures est attestée.

Ignorer la stratification conduit à surestimer la variance, en espérance, puisque la variance entre les éléments de strates différentes, la variance inter-strate (par opposition à la variance entre des éléments d'une même strate, la variance intra-strate), qui devrait être ignorée, est ajoutée dans le calcul de précision. La stratification doit être prise en compte même lorsqu'elle est proportionnelle

(même si la part d'une strate dans l'échantillon est égale à la part de cette même strate dans la population).

Sondage à deux degrés

La population d'intérêt est découpée en groupes de ménages qui sont les unités primaires de sondage (UP). Le tirage à deux degrés consiste à tirer tout d'abord un échantillon d'UP, puis de tirer un échantillon des ménages au sein de chaque UP tirée.

Le tirage de zones de dénombrement n'a aucun intérêt du point de vue de la précision statistique. Il se justifie par l'abaissement du coût de collecte (limitation des déplacements des enquêteurs), mais engendre souvent des effets dits de grappe. La présence d'un effet de grappe s'explique par une homogénéité des individus ou ménages d'une même UP supérieure à celle entre des individus ou ménages d'UP différentes. La positivité de la covariance intra-UP signifie une corrélation entre les observations d'une même UP. Cette corrélation peut-être interprétée comme le reflet de l'homogénéité des enquêtés d'une même UP et/ou le reflet de l'homogénéité due à l'enquêteur. Si cet effet de grappe n'est pas pris en compte dans le calcul de la variance, c'est-à-dire si la variance intra-UP est oubliée, la variance totale est sous-estimée.

Il est d'usage de procéder à un tirage des UP par probabilités proportionnelles à leur taille (PPT) si leur taille est corrélée aux phénomènes que l'analyste veut étudier. Supposons que le phénomène d'intérêt soit le chômage, ce tirage est approprié si le taux de chômage est positivement corrélé à la taille des UP.

Dans le cas d'enquêtes auprès des ménages, les UP sont fréquemment les zones de dénombrement (ZD) issues de recensements passés, correspondant à un découpage géographique fait selon trois critères. Le premier est l'unité, la continuité géographique de la zone, le deuxième la facilité du repérage de son contour par les agents sur le terrain. Le troisième est l'uniformité des tailles, en termes de population, des zones. Selon l'arbitrage fait entre ces critères, les zones de dénombrement peuvent être hétérogènes entre elles ou non. Elles peuvent être de taille très variable, en termes de population, alors que de superficie constante, auquel cas elles seront de densité de population variable. Mais elles peuvent également être de taille variable bien que de densité égale. Selon le cas, la taille des UP sera ou non corrélée à des variables d'intérêt (à supposer que la densité de population soit elle-même corrélée aux variables d'intérêt). Ceci ne peut se vérifier qu'après l'enquête.

Dans les pays de l'UEMOA comme dans beaucoup d'autres pays africains, la taille des UP à la date des enquêtes est mal connue. Le dernier recensement peut dater de 10 ans ou plus encore. Si les divergences entre les tailles d'UP issues du recensement et les tailles à la date de l'enquête sont supposées faibles, elles peuvent être ignorées et le tirage à PPT utilisé.

Si les divergences entre les tailles du recensement et les tailles à la date de l'enquête sont importantes, il est possible de traiter le problème de façon formelle. La taille du recensement est considérée comme un proxy de la taille à l'année de l'enquête (la superficie de la ZD pourrait également être utilisée si elle est supposée corrélée au nombre de ménages dans la ZD). Ce traitement est appelé tirage à probabilités proportionnelles à la taille estimée (PPTE). Avec un tirage à PPTE, les UP sélectionnées sont dénombrées lors de l'enquête. Le dénombrement permet de disposer d'une base de tirage actualisée des ménages au sein des UP et de connaître la taille exacte des UP sélectionnées.

Ensuite, le ratio entre taille réelle et taille approchée est utilisé pour re-pondérer les observations, par exemple, pour donner des poids plus importants aux observations des UP dont la taille réelle est supérieure à la taille anticipée, et qui sont ainsi sous-représentées dans l'enquête. Le désavantage est qu'alors les poids diffèrent entre les UP d'une même strate. Cette méthode permet cependant d'obtenir des estimateurs non biaisés même si la base de sondage est vieille, puisque le poids d'un ménage reste l'inverse de sa probabilité d'inclusion.

Mesurer l'effet du plan de sondage

Le calcul de l'effet de plan de sondage a deux utilités : calculer la taille effective de l'échantillon lors de la préparation d'une enquête et corriger les variances et les écarts-type estimés naïvement. L'effet du plan de sondage n'est pas le même pour toutes les variables de l'enquête, d'où la difficulté d'avoir une précision convenable dans les enquêtes ayant pour objectif de mesurer des indicateurs très variés.

Etant donné un plan de sondage complexe, l'effet du plan de sondage est dû à l'inégalité des probabilités d'inclusion et à l'utilisation d'un tirage à deux degrés (Gabler et al. 1999). Il se définit comme le rapport entre la variance de l'estimateur $\hat{\theta}$ sous le plan de sondage réel complexe ($Var_{réel}(\hat{\theta})$) et la variance du même estimateur sous un hypothétique sondage aléatoire simple ($Var_{SAS}(\hat{\theta})$).

$$deff = \frac{Var_{réel}(\hat{\theta})}{Var_{SAS}(\hat{\theta})} \quad (1)$$

L'effet de plan de sondage le plus important et le plus intuitif est l'effet de grappe (Kish 1965). Il provient des tirages à deux degrés et existe lorsque les réponses des individus d'une même UP se ressemblent davantage entre elles qu'elles ne ressemblent à celles des individus des autres UP⁴. Dans un sondage à deux degrés et à probabilités d'inclusions égales, le tirage d'UP introduit une homogénéité qui réduit la taille nette de l'échantillon d'un facteur multiplicatif, qui est l'effet du plan de sondage dû au tirage d'UP. Le *design effect* est alors une mesure de la sous-estimation de la variance causée par l'utilisation d'un plan de sondage à deux degrés plutôt qu'un SAS (le *design effect* est supérieur à 1). Dans le cas d'un plan de sondage stratifié, avec un sondage aléatoire simple dans chaque strate, le *design effect* mesure le gain de précision permis par la stratification (le *design effect* est inférieur à 1).

Plusieurs méthodes peuvent être utilisées pour estimer ce *design effect* correspondant aux méthodes d'estimation de variance⁵. Le numérateur, la variance de la statistique sous le plan de sondage réel complexe, peut être estimé par la méthode de linéarisation de Taylor ou la méthode du jackknife. La linéarisation de Taylor est une méthode d'estimation de la variance fréquemment utilisée pour des statistiques ou des plans de sondage complexes. Elle permet d'obtenir une estimation non biaisée de la variance d'estimateurs linéaires. La méthode de jackknife consiste à calculer l'estimateur en supprimant de l'échantillon les ménages d'une UP. Elle nécessite donc autant d'itérations qu'il y a d'UP. A chaque itération, elle tient compte du plan de sondage pour le calcul de l'estimateur et de sa variance.

Les données des enquêtes 1-2-3, enquêtes sur l'emploi

L'impact du plan de sondage sur l'estimation de la précision des indicateurs du marché de l'emploi issus d'enquêtes auprès des ménages se doit d'être étudié empiriquement. Pour ce faire, les enquêtes 1-2-3 réalisées en milieu urbain ouest-africain sont utilisées.

⁴ Cette corrélation peut-être interprétée comme le reflet de l'homogénéité des enquêtés d'une même UP et/ou le reflet de l'homogénéité due à l'enquêteur.

⁵ Pour une présentation plus détaillée des techniques de linéarisation, voir Ardilly (1994), chapitre 5.

Les enquêtes 1-2-3 du PARSTAT

La mise en œuvre des enquêtes étudiées dans les capitales économiques de sept des huit pays de l'Union Économique et Monétaire Ouest Africaine (UEMOA) entre dans le cadre du programme régional d'appui statistique à la surveillance multilatérale (PARSTAT) mis en place par la Commission de l'UEMOA. Ce programme s'inscrit dans l'objectif de convergence des politiques et des économies des pays membres à travers la surveillance multilatérale. Cet objectif nécessite la disponibilité de statistiques harmonisées et fiables.

Les enquêtes 1-2-3 du PARSTAT portent sur chacune des sept agglomérations suivantes : Niamey, Ouagadougou, Dakar, Bamako, Cotonou, Lomé et Abidjan. Le champ de l'enquête correspond à la ville économique, c'est-à-dire la "ville administrative" et toutes les petites localités qui y sont directement rattachées et avec lesquelles il y a des échanges fréquents.

Il s'agit d'enquêtes dites en phases, ou emboîtées. La phase 1 de ces enquêtes porte sur l'emploi et sert de filtre pour obtenir un échantillon représentatif d'unités de production informelle interrogées en phase 2 et un sous échantillon de ménages interrogés sur leur consommation en phase 3. En effet, la technique classique qui consiste à recenser les établissements et d'en extraire ceux du secteur informel a échoué à fournir un échantillon représentatif d'unités de production informelle (UPI). Aussi, une stratégie alternative a été mise au point, celle des enquêtes mixtes (ménages/établissements) ou en phases. Elle consiste à identifier un échantillon représentatif d'UPI à partir d'une enquête sur l'activité des ménages pour ensuite réaliser une seconde enquête spécifique sur ces dernières. L'enquête 1-2-3 est une enquête en phases, ayant fait l'objet de recommandations par les organisations internationales de statisticiens du travail, développée au début des années quatre-vingt-dix. Aux deux phases initiales a été ajoutée une troisième phase dont l'objectif est de quantifier la demande adressée au secteur informel par les ménages (phase 3).

Le plan de sondage retenu

Phase 1

Le tirage de l'échantillon de la phase 1 est effectué selon un sondage stratifié à deux degrés. Seul le plan de sondage de Lomé n'est pas stratifié. Après une stratification des UP sur la base de critères économiques ou administratifs (excepté à Lomé), 125 UP sont tirées selon un tirage à probabilité proportionnelle à leur taille à partir de travaux cartographiques d'enquêtes ou de recensements de

la population antérieurs⁶. Les UP échantillonnées sont dénombrées exhaustivement pour pallier à la vétusté des bases de sondage. Au sein de chaque UP tirée sont alors échantillonnés 20 ménages (25 à Cotonou) selon un sondage aléatoire simple. A Dakar cependant, les ménages ont été échantillonnés après une stratification selon le sexe du chef de ménage. Les échantillons de la phase 1 sont donc de 2 500 (20*125) ménages par ville (3 000 (25*125) à Cotonou). Les échantillons enquêtés sont de taille similaire puisqu'il a vraisemblablement été procédé à des remplacements des ménages non trouvés sur le terrain, même si cela n'a pas été documenté.

Les documents relatifs au plan de sondage des villes de Bamako et Abidjan n'ont pas été archivés. Par ailleurs de gros problèmes d'échantillonnage et de problèmes de poids ont été décelés dans les documents de Lomé⁷. Enfin, certaines questions de redressement de l'échantillon de Cotonou n'étant pas résolues (redressement de la taille de Cotonou, de la taille des ménages, choix entre une post-stratification basée sur les résultats de l'EDS et un redressement basé sur le RGPH 2002), l'effet du plan de sondage est étudié seulement dans cet article dans les cas de Niamey, Ouagadougou et Dakar.

Vétusté des bases de sondage

Le plan exposé ci-dessus a été appliqué avec plus ou moins de réussite du fait de la vétusté des bases de sondage (jusqu'à 10 ans avant l'enquête). Pour pallier à l'obsolescence des bases de sondage, toutes les UP tirées au premier degré sont intégralement dénombrées avant l'enquête.

Le nombre de ménages dénombrés a souvent été inférieur à celui figurant dans la base de sondage. Le dénombrement n'est pas seul en cause. D'une part, les travaux de cartographie et de dénombrement des RGPH ne sont pas exempts d'approximations. D'autre part, la population des

ZD a pu varier. Il est tout à fait possible que la population de certaines UP tirées et dénombrées, soit réellement inférieure à la population de l'année du recensement, d'autant plus si le recensement date de 10 ans. Cela n'implique pas que la population totale de la ville ait diminué, mais que les quartiers du centre, denses (et donc plus probablement tirés), se soient "vidés".

Pour deux des trois villes étudiées ici (Niamey et Ouagadougou), un coefficient correcteur a été appliqué aux coefficients d'extrapolation issus du plan de sondage.

A Niamey, les coefficients d'extrapolation ont été multipliés par un facteur propre à chaque strate (commune), obtenu, en 2001, en divisant la population par strate issue des résultats provisoires du RGPH 2001 par la population estimée par les coefficients d'extrapolation. Ce type de redressement, différencié selon les strates, est adéquat si les strates sont hétérogènes entre elles. Les résultats définitifs du RGPH diffèrent des résultats provisoires. Nous avons recalculés les facteurs de redressement à partir des résultats définitifs.

A Ouagadougou, un facteur multiplicatif de 1,3 pour les 5 strates, calculé par une projection linéaire des résultats des RGPH de 1985 et 1996, a été appliqué uniformément aux coefficients d'extrapolation. Cela revient à faire l'hypothèse implicite que la croissance démographique de la ville de Ouagadougou a été uniformément répartie dans tous les quartiers, ce qui est faux⁸. D'autre part, cette technique de projection, assez fruste car linéaire, sous-estime la population de 2006 de 14,6 %. Le redressement uniforme est donc peu utile (n'affecte que les totaux) et peu fiable (basé sur peu d'information et des hypothèses démographiques réductrices).

Résultats

La stratification au premier degré

A Dakar, la stratification du premier degré a été faite selon la commune de l'agglomération croisée avec le degré d'informalité issu d'une enquête

⁶ Cotonou RGPH de 1992 ; Ouagadougou RGPH de 1996 ; Abidjan RGPH de 1998 ; Bamako RGPH de 1998 ; Niamey RGPH de 2001 ; Dakar RGPH de 2002 ; Lomé Enquête Démographie et Santé de 1998.

⁷ Le tirage des 125 UP a été fait à partir d'un échantillon de 129 UP (celui de l'enquête togolaise sur les migrations et l'urbanisation, ETOMU) et non à partir d'une base exhaustive d'UP. Le nombre total d'UP de la base utilisée pour l'ETOMU et le mode de tirage de ces 129 UP sont inconnus. Une probabilité unitaire de tirage a ensuite été affectée à chaque UP ! Celà a conduit à sous-estimer la taille de Lomé, "redressée" en multipliant tous les poids par un facteur proche de 7. L'absence d'information sur le premier degré du sondage rend impossible la correction des poids, à moins de disposer de la base de sondage de l'enquête ETOMU.

⁸ La croissance de Ouagadougou est celle des extensions périphériques. Si cette croissance des quartiers périphériques était principalement due, avant les années 1980, à l'exode rural, les difficultés financières des ménages urbains y contribuent. Face à la cherté des loyers et des parcelles loties, plus au centre, les ménages sont obligés de solliciter des terrains auprès des propriétaires terriens des villages périurbains pour y construire leurs logements. Voir Ouattara (2001).

antérieure. Certes les variances à l'intérieur des strates sont élevées et, dans une des communes, l'ordre des strates en termes de degré d'informalité est inversé par rapport à l'information auxiliaire issue de l'enquête antérieure. Cependant, la situation est au moins aussi bonne que si aucune stratification n'avait été opérée.

Nous présentons ici la précision du taux de chômage (chômage élargi, c'est-à-dire incluant les chômeurs découragés, voulant travailler mais ne cherchant plus activement), à Niamey (Graphique 1). L'objectif est de montrer les différences entre les taux de chômage et écarts-type calculés par des méthodes d'estimation de variance différentes et leurs conséquences sur les comparaisons entre strates. La présentation des résultats de cette seule capitale suffit à mettre en garde sur la significativité des différences entre les strates. Les taux de chômage issus de *bootstrap*⁹ et du calcul sans pondération sont égaux, et également soumis au biais d'échantillonnage. Ceux tenant compte du plan de sondage (linéarisation et *jackknife*) sont sans biais et identiques puisqu'il s'agit de l'estimateur d'Horvitz et Thompson. Les variances diffèrent cependant selon les quatre méthodes.

La prise en compte du plan de sondage dans le calcul de la variance (linéarisation et *jackknife*) augmente la précision de l'estimation, comparée à la variance estimée par *bootstrap*. Si, lors du calcul de précision, la stratification est ignorée, cela revient à supposer que l'échantillon est issu d'un SAS. Cela biaise donc les estimations de la précision, en surestimant la variance. L'erreur commise en ignorant le plan de sondage est mesurée par l'effet du plan de sondage, ou *design effect*, dont la valeur est inscrite sur le graphique 1, sous l'intervalle de confiance estimé. Si le plan de sondage avait été ignoré dans le calcul de précision du taux de chômage de Niamey, la variance estimée aurait été 2,20 fois plus petite que celle estimée par linéarisation.

⁹ La méthode de rééchantillonnage, en anglais "bootstrap", consiste à tirer aléatoirement et avec remise, au sein de l'échantillon initial de taille n , B "nouveaux" échantillons de même taille n . Cette opération est répétée un nombre B de fois, autant que nécessaire pour satisfaire un critère de convergence. L'estimateur est calculé sur chaque nouvel échantillon. Les variations entre les différentes valeurs de l'estimateur obtenues fournissent la variance. Il est également possible d'adapter le tirage du bootstrap au plan de sondage. Cependant, peu de résultats théoriques attestent la validité du bootstrap en sondages (D'Haultefoëuille 2007). Le bootstrap utilisé ici est donc le bootstrap simple, sans utilisation du plan d'échantillonnage.

Ces sous-estimations peuvent conduire à des interprétations erronées. Supposons que le *design effect* est ignoré, c'est-à-dire que la variance est estimée comme si les observations étaient issues d'un SAS. Des tests d'égalité de moyenne concluent que deux à deux, les taux de chômage par strate sont inégaux, au seuil de 5 %¹⁰. Si ces mêmes tests sont menés sur les estimations tenant compte du plan de sondage, les p-values augmentent et l'égalité entre les taux de chômage des communes I et II d'une part, et des communes II et III d'autre part est acceptée.

Le taux de chômage de Niamey est de 23,54 %, compris dans l'intervalle de confiance de [21,89 % ; 25,20 %] selon la méthode de linéarisation de Taylor et l'intervalle de [22,10 % ; 24,99 %] selon la méthode du *jackknife*. La précision de ce taux diffère entre les trois strates. La commune III est caractérisée par un taux de chômage plus élevé que les deux autres communes, une variance plus élevée mais un *design effect* plus faible, quelle que soit la méthode d'estimation. Il y a donc une plus grande variabilité du phénomène de chômage dans la commune III. Si cette plus grande variabilité est confirmée pour d'autres variables d'intérêt, il est souhaitable de sur-représenter la commune III de Niamey dans les enquêtes futures sur l'emploi.

Le tirage à probabilités proportionnelles à la taille au premier degré

A Niamey, l'effet de grappe, ou *clustering effect*, une des composantes principales du *design effect*, est important. Si l'effet de grappe n'est pas pris en compte dans le calcul de la variance, la variance du taux de chômage est sous-estimée. Ainsi la prise en compte de cet effet donne une variance 2,31 fois supérieure à Niamey et 1,76 fois supérieure à Ouagadougou (Tableau 1).

Les individus tirés au sein d'une même UP sont donc davantage similaires que des individus issus de différentes UP. Cela paraît assez intuitif, les UP pouvant correspondre à des quartiers de standing différents. Cette corrélation intra-UP est d'importance variable selon les strates et villes. Par exemple, dans l'arrondissement de Boulmiougou de Ouagadougou, l'effet de grappe est quasi-inexistant (Tableau 1). Les UP de cette strate sont donc homogènes en termes d'importance du chômage. A l'inverse l'arrondissement de Sig-nonghin est constitué d'UP très hétérogènes.

¹⁰ En prenant en compte du plan de sondage complexe : taux de chômage des strates 1 et 2 (resp. 1 et 3 puis 2 et 3) égaux : p-value de 8,0 % (resp. de 0,4 % puis de 12,6 %) ; en supposant un SAS : taux de chômage des strates 1 et 2 (resp. 1 et 3 puis 2 et 3) égaux : p-value de 1,0 % (resp. de 0,0 % puis de 4,2 %).

Graphique 1 :
Précision de l'estimation du taux de chômage à Niamey (%) :

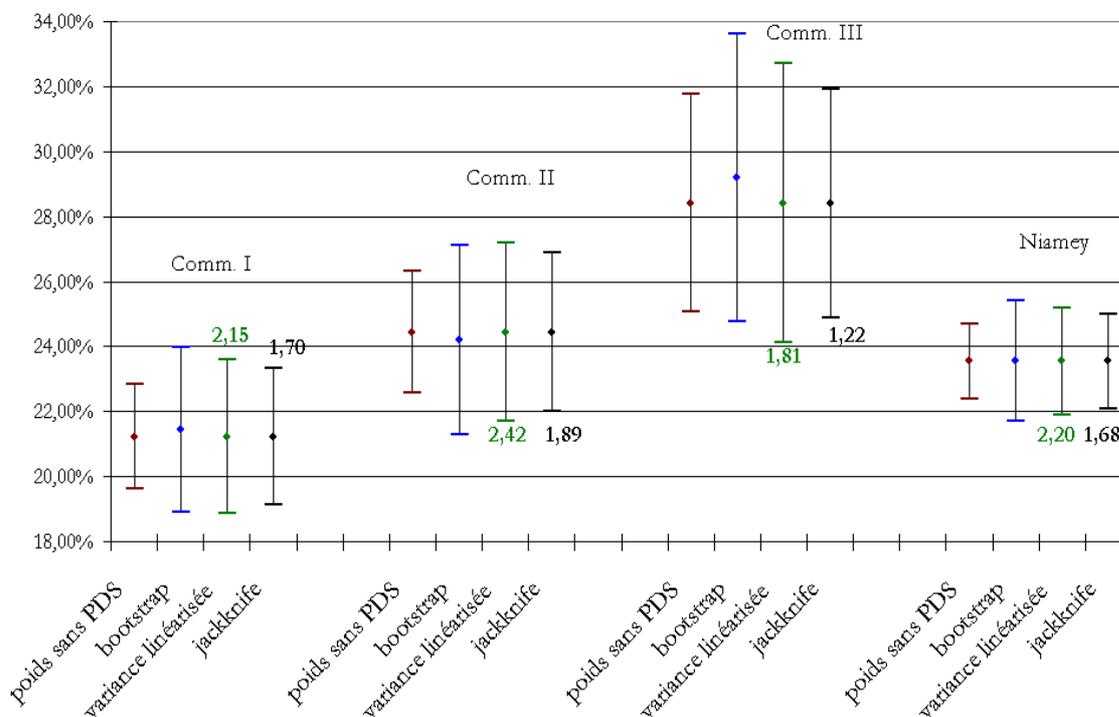


Tableau 1 :
Effet de grappe sur l'estimation de la variance du taux de chômage à Niamey et Ouagadougou (linéarisation de Taylor) :

Niamey		Ouagadougou	
Commune	Effet de grappe	Arrondissement	Effet de grappe
Commune I	2,45	Baskuy	1,33
Commune II	2,51	Bogodogo	2,30
Commune III	1,62	Boulmiougou	1,09
	--	Nongremassom	2,27
	--	Sig-nonghin	2,54
Ensemble	2,31	Ensemble	1,76

Sources : Enquêtes 1-2-3, phase 1 emploi, Instituts Nationaux de la Statistique, AFRISTAT, DIAL, calculs des auteurs.

La stratification au deuxième degré : le cas dakarois

A Dakar, le tirage des ménages au sein des UP a été effectué après une stratification selon le sexe du chef de ménage. Cette variable avait été collectée lors du dénombrement des UP échantillonnées. C'est la seule ville dans laquelle cette méthode a été utilisée, n'étant nullement justifiée par les objectifs de l'enquête. Cette complication augmente pourtant les risques d'erreur dans le calcul des poids d'échantillonnage, et pose le problème de la cohérence entre les informations collectées lors du dénombrement et celles provenant de la collecte de données sur l'emploi. En effet, 19 % des ménages de la base de données ne sont pas dirigés par une

personne du même sexe que celui utilisé pour le tirage. Enfin, de nombreux remplacements, ne tenant pas compte du sexe, ont été opérés sur le terrain, mais sans être ni justifiés ni documentés.

Les gains à cette stratification ne sont pas visibles (Tableau 2). La précision du taux d'activité des chefs de ménages féminins est très proche de celle obtenue à Ouagadougou sans stratification, et toujours largement inférieure à celle obtenue pour les hommes. De plus, les variances intra-UP (non présentées ici) sont similaires, à Dakar, avec les deux configurations (stratifié au 2nd degré ou non), parfois celles du deuxième degré stratifié sont même supérieures.

Tableau 2 :
L'absence de gains à la stratification au second degré à Dakar :

	Ouagadougou	Dakar
Ménages dirigés par une femme, non pondérée (%)	13,51	27,11
Ménages dirigés par une femme, pondérée (%)	13,20	26,60
Taux d'activité des chefs de ménage hommes (erreur-type)	88,71 (0,87)	78,56 (1,07)
Taux d'activité des chefs de ménage femmes (erreur-type)	75,21 (2,83)	60,26 (2,07)

Sources : Voir Tableau 1.

Conclusion et tableau récapitulatif des recommandations

Les sept enquêtes du projet PARSTAT ont pour atout la comparabilité des méthodes. Cependant, la méthodologie prévue n'a pas été strictement respectée dans toutes les capitales. La définition et l'application d'un plan de sondage commun à plusieurs pays sont tout à fait possibles dans le cadre de l'UEMOA. Les difficultés rencontrées sont partagées par la majorité des capitales. Les variantes du plan initial tiennent donc plus à des équipes de travail différentes qu'à des difficultés différentes.

Une stratification doit être opérée au premier degré. Des critères administratifs tels que les arrondissements sont préférables, pour leur simplicité, et parce que l'hétérogénéité entre les strates ainsi définies-constatée dans la majorité des cas- est ainsi prise en compte. En revanche, la stratification au 2nd degré, selon le sexe du chef de ménage, n'est ni justifiée, ni efficace. D'une part, les enquêtes 1-2-3 n'avaient pas, dans le cadre du projet PARSTAT, pour objectif spécifique d'analyser les ménages dirigés par des femmes. D'autre part, une stratification au second degré nécessite que l'information sur les ménages constituant le critère de stratification soit recueillie lors du dénombrement précédant l'enquête de façon fiable. Cela alourdit le dénombrement, qui peine déjà parfois à fournir des tailles d'UP correctes.

Dans la majorité des villes, des remplacements de ménages ont été faits, en cas de refus, de disparition du ménage ou de mauvais repérage du ménage lors du dénombrement. Dans aucunes des villes concernées, ces remplacements n'ont été documentés. Ceci pose un problème de biais des estimateurs, impossible à corriger sans information sur ces remplacements. Cela pose un problème supplémentaire dans le cas de stratification au second degré. A Dakar, les remplacements ont été faits indépendamment du sexe du chef de ménage.

Si les zones de dénombrement utilisées sont de taille variable, nous estimons préférable d'utiliser un sondage avec une probabilité de tirage proportionnelle à la taille (estimée). Il est nécessaire de conserver le fichier de la liste des zones de

dénombrement sur lequel a été effectué le tirage. En effet, comme dans la majorité des enquêtes, tous les tirages ont été faits selon un tirage systématique, qui ne peut être assimilé à un tirage aléatoire que sous certaines conditions. Si l'ordre de la liste des ZD n'est pas complètement aléatoire (si les ZD voisines se suivent dans la liste par exemple), alors le tirage systématique crée une stratification implicite qu'il est possible de prendre en compte par une méthode exposée par Kish (1965), à la condition de disposer de cette liste exhaustive des ZD et de leur taille.

Les redressements des coefficients en cas de sous-estimation de la population ne doivent pas être faits de façon uniforme, d'après la taille estimée (souvent grossièrement) à partir du dernier recensement. Il est préférable de vérifier si la sous-estimation concerne toute la population ou certaines zones, certains groupes d'âge, certaines tailles de ménages, afin de procéder par la suite à une post-stratification. Cette post-stratification doit être basée sur des résultats censitaires plutôt que sur les résultats d'enquêtes par sondage précédentes. En effet, les résultats des enquêtes par sondage précédentes sont eux aussi basés sur des estimations et des redressements.

Enfin, nous réitérons la recommandation de documenter chaque étape de la constitution de l'échantillon. Depuis la liste exhaustive des zones de dénombrement et du fichier de tirage, jusqu'à la post-stratification ou redressement des poids après la collecte des données. Les logiciels statistiques permettent de tenir compte de plans de sondage de plus en plus complexes et proposent de plus en plus de méthodes de correction de divers biais. Cependant, pour les mettre en œuvre, il faut disposer de toute l'information nécessaire sur la constitution de l'échantillon et des poids.

Les enquêtes du PARSTAT sont des enquêtes en phases. L'échantillon de phase 2 est constitué d'unités de production informelle dont les chefs ont été identifiés en phase 1 tirés selon un sondage stratifié simple parmi les 2 000 à 3 000 unités de production informelle identifiées dans chaque ville. L'échantillon de la phase 3 est un sous échantillon de celui de la phase 1, tiré selon un sondage stratifié simple. Aussi, notre travail se poursuit actuellement

par l'étude de la qualité des échantillons extraits de celui de la phase 1, et avec l'explication d'éventuelles incohérences entre les phases. Il

nécessite de rechercher tous les documents relatifs au sondage rédigés il y a cinq ans environ, et de reconstituer les choix opérés non documentés.

Tableau 3 :

Grille des choix méthodologiques :

Sondage à deux degré	Quasi inévitable pour des raisons de coût et de temps. Augmente la variance et le risque de la sous-estimer : effet de grappe
Stratification au 1 ^{er} degré	Il y a toujours des gains à la stratification. Si les divisions administratives sont des critères tout à fait acceptables, il convient de se méfier des critères très élaborés (qualité de l'information et surnombre de strates de petite taille) et des enquêtes "omnibus".
Sondage à PPT au 1 ^{er} degré	Il est à justifier selon l'objectif de l'enquête et la corrélation entre variables d'intérêt et tailles des UP. Il nécessite de porter attention au calcul des probabilités d'inclusion des UP (probabilités supérieures à 1 à corriger), aux regroupements et segmentations d'UP.
Base de sondage	Il est préférable d'utiliser les résultats du dernier RGPH, exhaustifs, plutôt que les échantillons d'autres enquêtes. Une solution à la vétusté de la taille des UP est le tirage à probabilité proportionnelle à la taille estimée (estimation sans biais). Il faut se méfier des projections démographiques grossières qui supposent souvent à tort une croissance démographique constante et homogène.
Tirage des UP	Le tirage systématique vérifie la plupart des critères exigés d'un algorithme de tirage. De plus, si l'ordre des UP est aléatoire, le tirage systématique peut-être assimilé à un sondage aléatoire et les formules classiques de variance sont applicables. Si l'ordre des ZD n'est pas aléatoire, une stratification implicite des UP peut apporter un gain de précision (même gain que celui de la stratification "explicite") calculé seulement si la base de tirage est disponible. Le seul cas où l'ordre non aléatoire peut nuire à la précision est quand l'ordre des éléments de la liste est périodique, d'une période proche du pas de tirage, cas non pertinent dans le cas d'enquêtes auprès des ménages.
Ménages tirés dans chaque UP	Si les UP sont choisies selon un tirage à PPT, tirer un nombre constant de ménages permet d'avoir un échantillon auto-pondéré. Si le tirage des UP est à PPTE, un nombre constant de ménages reste avantageux pour le travail de terrain, et, si les tailles du RGPH et du dénombrement sont proches, l'échantillon sera presque auto-pondéré.
Stratification au 2 nd degré	La stratification au 2 nd degré doit être justifiée par les objectifs de l'enquête et basée sur une information auxiliaire de qualité dont la collecte n'alourdit pas le dénombrement.
Remplacements	Tous remplacements de ménages échantillonnés doivent être documentés et les ménages concernés identifiés dans la base de données. Ils doivent respecter la stratification si stratification il y a au 2 nd degré.
Post-stratification	Attention aux redressements brutaux consistant à multiplier tous les coefficients d'extrapolation par un facteur de redressement global. Il est préférable d'identifier les types de ménages sous-représentés et de procéder à une post-stratification.
Archivage et extrapolation	Le plan de sondage doit faire l'objet d'un document complet et publié. Les fichiers de sondage doivent être conservés, documentés et les choix justifiés. Il est possible de corriger beaucoup de biais lorsque toute l'information sur chacune des étapes du sondage est disponible.

Références Bibliographiques

Amegashie F., Brilleau A., Coulibaly S., Koriko O., Ouédraogo E., Roubaud F. et Torelli C. (2005), « La conception et la mise en oeuvre des enquêtes 1-2-3 en UEMOA. Les enseignements méthodologiques », *STATECO n°99*, 2005, 21-41.

Ardilly P. (1994), *Les techniques de sondage*, Edit Technip-Paris.

D'Haultfoeuille X. (2007), *Cours de sondages approfondis*, ENSAE.

Gabler S., Häder S. and Lynn P. (2005), « Design effects for multiple design samples », ISER Working Paper 2005-12

Ganninger M., Häder S. et Gabler S. (2007), « Design Effects and Interviewer Effects in the European Social Survey: Where are we now and where do we want to go tomorrow? », Quality Enhancement Meeting, European Social Survey.

Howes S. et Lanjouw J.O. (1997), « Poverty Comparisons and Household Survey Design », LSMS Working Paper n° 129, 1997.

Kish L. (1965), *Survey Sampling*. Wiley, New York.

Ouattara A. (2004), « Les enjeux de la dynamique des extensions périurbaines à Ouagadougou », contribution présentée lors de l'atelier Dynamiques périurbaines : population, habitat et environnement dans les périphéries des grandes métropoles, organisé par le CEPED, 2004.