



Conseil économique et social

Distr. générale
19 août 2009
Français
Original: anglais

Commission économique pour l'Europe

Conférence des statisticiens européens

Groupe d'experts des recensements de la population et des habitations

Douzième réunion

Genève, 28-30 octobre 2009

Point 6 de l'ordre du jour provisoire

Collecte des données en ligne

Analyse théorique et pratique des recensements par Internet

Note de l'Institut national de la statistique de l'Espagne

I. Introduction

1. Les nouvelles technologies font partie de la vie de tous les jours depuis un certain temps. Il est désormais difficile de l'imaginer sans téléphone mobile, ordinateur portable ou connexion Internet.
2. Internet peut être utilisé à des fins très différentes, par exemple envoyer un courriel (au lieu d'une lettre), lire les journaux, rechercher des informations sur toutes sortes de questions techniques ou consulter ses données bancaires.
3. La présente note examine comment le recours à Internet peut offrir de nombreuses possibilités (et un extraordinaire potentiel de croissance) pour le prochain recensement qui aura lieu aux alentours de 2011.
4. Les recensements de la population et des habitations figurent parmi les activités statistiques les plus importantes, non seulement à cause de leur ampleur, mais aussi par tradition. L'Espagne, par exemple, procède à des recensements modernes depuis 1768 (recensement d'Aranda) ou 1787 (recensement de Floridablanca). Le recensement de 2011 sera le vingtième de son histoire:

XVIII^e siècle (3): 1768, 1787, 1797;

XIX^e siècle (5): 1842, 1857, 1860, 1877, 1897;

XX^e siècle (10): 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1981, 1991;

XXI^e siècle (2): 2001, 2011.

5. Les habitudes de vie, en particulier celles qui ont trait aux technologies et aux appareils électroniques, se sont considérablement modifiées au fil des années, le temps ne faisant qu'accélérer les changements.

6. L'Espagne a été l'un des premiers pays au monde à donner aux gens la possibilité de répondre par Internet au recensement de 2001. Pourtant, même si les occupants de moins de 1 % des logements (13 768 sur 14 187 169) ont participé par Internet au recensement de 2001, l'expérience est utile pour la planification de celui de 2011.

7. Plusieurs raisons expliquent pourquoi la situation sera bien plus prometteuse pour le recensement de 2011 que pour celui de 2001. Le nombre de personnes qui utilisent Internet chaque jour a sensiblement augmenté, la largeur de bande offerte par les sociétés spécialisées dans les technologies de l'information est plus grande qu'en 2001, les applications Web sont bien plus nombreuses et leurs possibilités moins restrictives (les protocoles de sécurité Internet ayant par exemple été considérablement renforcés) et le volume de stockage a connu une croissance exponentielle ces dernières années. Cela explique pourquoi le recensement de 2011, dans lequel Internet jouera un rôle important, aura une configuration très différente de celle de 2001, l'existence d'Internet étant alors purement symbolique.

II. Analyse théorique

8. La présente section traite essentiellement des caractéristiques théoriques des réponses à un recensement faites par Internet et insiste en particulier sur les points forts et sur les faiblesses de ce mode.

A. Est-il justifié de développer une application Web?

9. D'un point de vue économique, le développement d'une application Web qui donne la possibilité de répondre par Internet (y compris une page Web, des protocoles de sécurité et un serveur capable de prendre en charge un nombre adéquat de connexions) suppose uniquement des frais fixes. Par contre, toutes les autres méthodes qui permettent de répondre de manière exhaustive au recensement (réponse à un questionnaire papier, entretien téléphonique ou individuel à l'aide d'un ordinateur de poche) présentent des frais fixes et variables.

10. Les solutions autres qu'Internet comportent des frais variables car le nombre de personnes qui répondent au recensement influe directement sur le coût. Par exemple, le volume de documents papier imprimés ou numérisés et le nombre d'agents recenseurs intervenant dans les entretiens téléphoniques ou individuels assistés par ordinateur dépendent du nombre de personnes qui font usage des différents modes de réponse, ce qui entraîne une série de frais variables.

11. Le recours à Internet fait disparaître tous les frais variables puisque le développement de l'application Web ne dépend pas du nombre de personnes qui font appel à cette option¹.

12. On peut donc conclure que si les frais fixes engagés pour développer une application Web ne sont pas élevés et que le taux de réponses sur le Web est important, il est

¹ Sauf en ce qui concerne la capacité du système qui doit assurer le traitement (plus les personnes qui ont recours à cette option sont nombreuses, plus les besoins en matière de capacité et de budget sont élevés). Cela ne suppose cependant pas des coûts supplémentaires considérables.

intéressant d'utiliser Internet. Cette observation sera expliquée chiffres à l'appui dans la section III de la présente note.

13. Autre conclusion, si on part de l'hypothèse que le taux de réponses sur le Web n'est pas connu au départ, le développement d'une application qui permet de répondre par Internet pourrait parfois ne pas être rentable car les frais fixes que suppose un recensement par Internet sont plus élevés que la somme des frais fixes et des frais variables encourus si un autre moyen est utilisé.

14. L'actuelle situation économique mondiale a mis en évidence un point capital, à savoir qu'il importe que tous les recensements soient conçus selon le principe du moindre coût.

B. Informations: qualité et possibilités

15. L'un des plus grands avantages du recensement en ligne est qu'il améliore la qualité des informations reçues de cette façon. Il est possible d'éviter des inconvénients comme la difficulté de reconnaître (ou de numériser) les caractères et la présence de valeurs non valides ou de blancs pour certaines caractéristiques. Par exemple, la réponse concernant la caractéristique «pays/lieu de naissance» sera acceptée uniquement si la valeur inscrite figure dans un menu déroulant.

16. Les recensements en ligne peuvent aussi permettre de supprimer les incohérences entre différentes réponses figurant sur le même formulaire, voire des formulaires différents.

17. La numérisation et la reconnaissance des caractères prennent du temps et demandent des ressources importantes. En outre, la qualité des renseignements figurant sur des formulaires papier n'est pas toujours garantie.

18. Les recensements en ligne présentent un autre avantage, à savoir que, selon le profil d'une personne, l'application pourrait lui poser uniquement les questions nécessaires.

19. Par exemple, il serait absurde de demander à un enfant de 8 ans son état matrimonial légal. Il n'est pas aisé de définir quel genre de question devrait être posée selon le profil d'une personne.

20. Autre avantage du recensement en ligne, il est possible d'utiliser des réponses préremplies concernant les questions pour lesquelles l'on connaît déjà probablement les réponses exactes. En Espagne, ces réponses préremplies ont pu être obtenues grâce au recensement de 2001 et aux registres administratifs.

21. Par exemple, si une question a été posée lors du recensement de 2001, la réponse pourrait être indiquée sous la forme d'une réponse prédéfinie sur le formulaire de 2011 (ou il est possible de ne pas poser directement la question à l'utilisateur et de prendre pour hypothèse le renseignement obtenu lors du recensement de 2001 ou dans les registres). Si la personne qui remplit le formulaire de recensement n'accepte pas la réponse déjà inscrite, il lui sera loisible de la modifier, mais dans le cas contraire, elle n'aura qu'à appuyer sur un bouton de confirmation et la question suivante apparaîtra.

22. Pour conclure, le recensement par Internet sera bien plus efficace qu'avec d'autres méthodes comme l'utilisation de formulaires papier car le flux de questions sera adapté à chaque individu et seules les questions qui exigent une réponse seront posées (par exemple, la question se rapportant au nombre total d'enfants nés vivants ne sera posée qu'aux femmes d'au moins 16 ans).

C. Importance des méthodes d'incitation

23. Il a été démontré (par des études effectuées dans d'autres pays comme le Canada ou la Suède) que la possibilité de répondre en recourant simultanément à plusieurs moyens (questionnaire papier, entretien téléphonique assisté par ordinateur ou téléphone) a un effet négatif sur le pourcentage de réponses sur le Web.

24. Si l'objectif recherché est d'obtenir un pourcentage élevé de réponses par Internet, il y a lieu de prendre en compte un ou plusieurs des points suivants:

a) La lettre envoyée à chaque logement pour encourager les intéressés à répondre par Internet devrait être rédigée avec discernement. Il faudrait envisager de charger des experts externes d'envoyer des lettres adéquates à un grand nombre de personnes. Par ailleurs, il convient de recourir à des techniques cognitives pour évaluer l'efficacité des lettres pour la réalisation des objectifs;

b) L'option Internet devrait être offerte aux personnes interrogées avant tout autre moyen;

c) C'est seulement lorsque l'occupant d'un logement indique qu'il préfère ne pas répondre par Internet que l'on pourra recourir à d'autres moyens tels que la réponse sur papier, la visite d'un agent pour un entretien individuel assisté par ordinateur ou l'entretien téléphonique assisté par ordinateur.

25. Au cours du recensement pilote réalisé en Espagne du 13 avril au 31 juillet 2009, les deux méthodes ci-après ont été appliquées:

a) Méthode 1: Une lettre et des formulaires de recensement ont été envoyés simultanément à chaque logement. Les utilisateurs pouvaient répondre sur papier, par Internet ou par téléphone;

b) Méthode 2: Seule une lettre a été envoyée à chaque logement. Premièrement, les utilisateurs pouvaient répondre par Internet ou par téléphone. Deuxièmement, les informations concernant tous les logements dont les occupants n'avaient pas répondu avant une date donnée ont été collectées par des agents lors d'un entretien individuel assisté par ordinateur.

26. Dans les deux cas de figure, en raison des faibles taux de réponses, une lettre supplémentaire a été envoyée un mois après pour rappeler aux gens l'importance qu'il y avait à participer au recensement pilote.

27. On trouvera dans la section III de la présente note des observations concernant les différences entre les taux de réponses par Internet obtenus avec les deux méthodes.

D. Autres points clefs du recensement par Internet

28. Les réponses fournies par Internet aux différentes questions pourraient être diffusées bien plus rapidement que celles obtenues par d'autres voies puisqu'il n'est pas nécessaire de les convertir au format électronique. Par ailleurs, les valeurs incohérentes ou non valides seront évitées si l'application Web programmée est assortie de contraintes appropriées.

29. La conception de l'application est un autre point essentiel destiné à garantir la réussite du recensement sur le Web, le système devant être complet et facile à comprendre. Les facteurs comme la vitesse, une présentation attrayante et la robustesse sont également très importants.

30. Par ailleurs, la sécurité et la confidentialité doivent être sérieusement étudiées. L'application devra traiter des données personnelles et privées sensibles. Il faudra veiller à éviter les attaques informatiques de la part de pirates ou d'autres utilisateurs non autorisés. C'est pourquoi les éléments suivants sont indispensables:

- a) Mots de passe pour authentification (y compris éventuellement une signature électronique);
- b) Pare-feux bloquant les accès non autorisés;
- c) Conception des pages Web à l'aide de protocoles de sécurité http (https).

31. Si l'occupant d'un logement répond au recensement par différents moyens (par exemple Internet, questionnaire papier, entretien téléphonique ou individuel assisté par ordinateur), l'application Internet doit être conçue de façon à synchroniser les réponses. Si l'intéressé a répondu par Internet par exemple, il serait illogique de lui envoyer des formulaires papier. Autre exemple, si la réponse a été faite sur papier, la visite d'un agent pour un entretien assisté par ordinateur ne serait pas indiquée.

32. La mise en œuvre d'une base de données centrale coordonnant les réponses collectées de différentes façons pourrait faciliter la gestion de la question.

33. Il y a lieu d'étudier attentivement le traitement des réponses incohérentes et incompatibles données par Internet.

34. Une application comportant un grand nombre de vérifications et de messages pourrait être difficile à comprendre par les personnes interrogées tout en causant des difficultés de programmation pour l'équipe technique. Cependant, une application dépourvue de vérifications produira probablement des données incohérentes et il convient donc d'arriver à un équilibre en la matière, ce qui n'est pas chose aisée.

35. L'application conçue par l'Espagne pour le recensement pilote de 2009 accepte deux types d'erreur:

- a) Erreurs graves (incompatibilité): les utilisateurs doivent modifier leurs réponses, sinon il est impossible de continuer à répondre aux questions;
- b) Autres types d'erreur qui ne sont pas graves: les utilisateurs peuvent confirmer ou modifier leurs réponses, puis continuer à répondre aux questions suivantes.

E. Diagramme SWOT

36. Pour conclure, le diagramme SWOT (forces, faiblesses, opportunités et menaces) reproduit ci-après tente de résumer les caractéristiques du recensement par Internet.

37. L'analyse SWOT consiste à définir l'objectif du projet et de déterminer les facteurs internes et externes qui sont favorables ou défavorables à la réalisation de l'objectif.

| <i>Forces</i> | <i>Faiblesses</i> | |
|--|--|--------------------------------------|
| <ul style="list-style-type: none"> • Existence de frais fixes uniquement • Expérience: il était possible de répondre par Internet au recensement de 2001 et au recensement pilote de 2009 • Qualité de l'information • Diffusion rapide des données • Moins de contraintes pour les personnes interrogées | <ul style="list-style-type: none"> • L'application Web doit être sûre, robuste et de conception attrayante • Intégration à différents modes de réponse • Sécurité de l'application Web • Possibilité de répondre simultanément sur papier et par Internet | I N T E R N E S |
| <ul style="list-style-type: none"> • Un nombre croissant de personnes ont une connexion Internet • Compte tenu de la situation économique mondiale: réduction des coûts au minimum • L'ensemble de la population saura qu'un recensement aura lieu en 2011 | <ul style="list-style-type: none"> • Il n'est pas possible de prévoir exactement le taux de réponses sur le Web • Les personnes interrogées doivent avoir des connaissances techniques et être familiarisées avec les applications Web • Internet n'est pas disponible dans 100 % des logements | E X T E R N E S |
| <i>Opportunités</i> | <i>Menaces</i> | |

III. Données numériques sur la viabilité et l'utilité du recensement par Internet en Espagne

38. On trouvera dans la présente section des données numériques qui témoignent de la véracité des différentes constatations expliquées dans la section II.

A. Utilisation d'Internet

39. Selon le pourcentage de logements disposant d'une connexion Internet (tableau 1), il y a lieu de penser que d'ici à 2011, 60 % au moins des logements espagnols auront leur propre connexion.

40. En outre, de nombreuses personnes peuvent utiliser Internet au travail ou chez un parent ou un voisin, ce qui porte les possibilités d'utilisation de ce moyen en 2011 à un niveau plus élevé que la proportion de logements disposant d'une connexion Internet.

41. Le tableau ci-après montre clairement les différences entre 2011, époque où Internet sera quelque chose de familier pour presque tout le monde, et 2001, époque où ce n'était pas encore le cas.

Tableau 1
Pourcentage de logements disposant d'une connexion Internet

| | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
|--------------|-----------------|-----------------|-----------------|------|-----------------|-----------------|-----------------|
| Espagne | 17 | 28 | 34 | 36 | 39 | 45 | 51 |
| UE (moyenne) | 39 ² | 43 ² | 42 ³ | 48 | 51 ³ | 54 ⁴ | 60 ⁴ |

B. Aspect économique du recensement par Internet

42. Dans l'expérience menée par l'Espagne en 2001, le coût du développement d'une application Internet pour le recensement (y compris le centre d'appel) s'est élevé à 2 millions d'euros. Si les conditions sont les mêmes en 2011, le budget requis sera en principe de 3,5 millions d'euros.

43. Par ailleurs, selon des études menées en Espagne, le coût par habitant du recensement de 2011 serait à peu près le suivant pour les différentes options:

- a) Réponses sur papier: 5 euros par questionnaire;
- b) Entretien téléphonique assisté par ordinateur: 12 euros par questionnaire;
- c) Entretien individuel assisté par ordinateur: 25 euros par questionnaire.

44. La corrélation des deux montants (budget prévu pour le recensement par Internet de 2011 et coût par habitant des différents modes de réponse) donne les conclusions ci-après:

a) Avec plus de 700 000 réponses par Internet (soit 1,5 % environ de la population), la stratégie du recours à Internet serait justifiée (par rapport à la stratégie du recours au questionnaire papier);

b) Avec plus de 300 000 réponses par Internet (soit 0,7 % environ de la population), la stratégie du recours à Internet serait justifiée (par rapport à la stratégie de l'entretien téléphonique assisté par ordinateur);

c) Avec moins de 140 000 réponses par Internet (soit 0,3 % environ de la population), la stratégie du recours à Internet serait justifiée (par rapport à la stratégie de l'entretien individuel assisté par ordinateur).

45. La viabilité et l'utilité du recensement par Internet ne font donc aucun doute.

C. Recensement pilote de 2009 en Espagne

46. Ainsi qu'il a été observé dans la section II, l'Office national de statistique espagnol a lancé le 13 avril 2009 un recensement pilote, dont le but était de tester diverses entreprises de technologie ainsi que de nouvelles formes et méthodes de collecte de données avant le prochain recensement total de 2011.

47. Le recensement pilote est très utile. Les enseignements tirés faciliteront la prise de décisions ayant trait aux questions à poser, aux méthodes à appliquer et à d'autres aspects du recensement de 2011.

² UE-15.

³ UE-25.

⁴ UE-27.

48. Soixante zones de recensement (10 zones dans six régions différentes) ont été choisies dans tout le pays pour l'expérience, à laquelle participeront, au total, environ 30 000 ménages sur les quelque 15,6 millions que compte le pays.

49. Ainsi qu'il a été dit dans la section II, deux méthodes différentes sont expérimentées (les deux étant des stratégies mixtes) pour la collecte de données. Chaque méthode est expérimentée dans 30 zones. On trouvera des précisions dans le tableau 2.

Tableau 2

Calendrier et modes de collecte de données selon chaque méthode

| | <i>Questionnaire papier</i> | <i>Entretien téléphonique assisté par ordinateur</i> | <i>Internet</i> | <i>Entretien individuel assisté par ordinateur</i> |
|-----------|---------------------------------|--|------------------------------|--|
| Méthode 1 | Du 13 avril au 31 juillet | Du 13 avril au 31 juillet | Du 13 avril au 31 juillet | Du 13 avril au 31 juillet |
| Méthode 2 | Pas de formulaire papier | Du 13 avril au 31 juillet | Du 13 avril au 31 juillet | De la mi-juin au 31 juillet |

50. Le pourcentage de personnes interrogées selon les différentes approches (méthodes et modes de réponse) varie pour chaque méthode et chaque mode.

51. Comme prévu, le recours au questionnaire papier est le mode qui a produit le plus de réponses pour la méthode 1. Il est cependant étonnant de constater que les occupants de près de 10 % des logements (15 % après ajustement des données en fonction d'un taux de réponses de 100 %) ont décidé de répondre au recensement par Internet même s'ils disposaient des formulaires chez eux.

52. Pour la méthode 2, le pourcentage total de personnes interrogées (40,0) est inférieur à celui enregistré pour la méthode 1 (62,5). Cependant, si les différents modes de réponse de la méthode 2 (entretien téléphonique assisté par ordinateur, Internet) sont considérés conjointement avec le fait que le recensement pilote a été effectué sans aucune publicité, ces taux de réponses ne sont pas surprenants.

53. Il semble évident aussi que la situation en 2011 sera différente de celle qui existait lors de l'expérience pilote de 2009. S'il est possible de motiver les gens ou de les convaincre des avantages d'Internet, les possibilités de développement de ce mode de réponse sont importantes.

Tableau 3
Pourcentage de personnes interrogées⁵

| | Questionnaire papier | Entretien téléphonique assisté par ordinateur | Internet | Entretien individuel assisté par ordinateur | Total |
|--------------|--------------------------------|--|-------------------|---|-------|
| Méthode 1 | 45,6 | 7,4 | 9,5 | Pas de données disponibles | 62,5 |
| Méthode 2 | Pas de questionnaire papier | 18,1 | 21,9 ⁶ | Pas de données disponibles | 40,0 |
| Total | 20,7 | 13,3 | 16,3 | Pas de données disponibles | 50,3 |

Tableau 4
Pourcentage de personnes interrogées⁵ (ajusté en fonction d'un taux de réponses
de 100 % pour chaque méthode)

| | Questionnaire papier | Entretien téléphonique assisté par ordinateur | Internet | Total |
|--------------|-----------------------------|--|----------|-------|
| Méthode 1 | 73,0 | 11,8 | 15,2 | 100 |
| Méthode 2 | Pas de questionnaire papier | 45,2 | 54,8 | 100 |
| Total | 41,2 | 26,4 | 32,4 | 100 |

IV. Conclusions et futures stratégies visant à améliorer le taux de réponses sur le Web

54. Bien que le pourcentage de personnes interrogées par Internet dans le recensement pilote soit plus élevé que prévu, diverses stratégies pourraient être appliquées pour améliorer le taux de réponses sur le Web pour le prochain recensement de 2011.

55. Pour accroître le pourcentage de personnes interrogées par Internet, une ou plusieurs des idées ci-après pourraient être envisagées:

- a) Campagne de publicité dans les «médias traditionnels» tels que la télévision, la radio ou les journaux;
- b) Publicité sur les pages Web. Il pourrait aussi être utile de faire appel à des experts externes spécialistes du «bouche-à-oreille électronique»⁷;
- c) Publicité dans des réseaux sociaux comme Facebook (www.facebook.com), Myspace (www.myspace.com) ou Twitter (www.twitter.com);

⁵ Données provisoires à la date du 21 juin.

⁶ Dans certaines zones de recensement, des pourcentages supérieurs à 42,5 % ont été obtenus.

⁷ Le «bouche-à-oreille électronique» peut s'entendre d'une stratégie mise en place par une entreprise pour attirer l'attention sur elle-même tout en amenant ses supporteurs à diffuser eux-mêmes le message, ce qui crée de nouvelles utilisations du produit de l'entreprise, de nouveaux slogans argotiques, de nouvelles communautés ayant des intérêts en commun et de nouvelles façons d'être qui on pense être.

d) Rédaction judicieuse de la lettre envoyée à chaque logement, cette lettre étant le seul contact direct entre l'Office national de statistique et de nombreuses personnes. Comme nous l'avons indiqué précédemment, il pourrait être envisagé de faire appel à des experts externes;

e) Mesure visant à faciliter l'accès à un ordinateur connecté à Internet pour les personnes qui ne disposent pas d'une connexion à domicile et qui souhaitent répondre par Internet. La collaboration des municipalités et des régions pourrait s'avérer utile dans ce domaine.

56. Même si on ne prévoit pas un taux de 100 % pour les réponses par Internet, l'Office national de statistique est persuadé que si toutes les idées visant à améliorer l'utilisation d'Internet sont efficacement mises en œuvre, on obtiendra un pourcentage élevé de réponses. Les fonds qui pourraient être économisés ainsi que la meilleure qualité des données reçues par Internet seront deux facteurs clefs qui inciteront à obtenir le maximum de réponses par Internet.
