

Utilité des parodonnées dans les grandes enquêtes nationales : défis et avantages

Brady T. West

Survey Research Center, Institute for Social Research
University of Michigan-Ann Arbor
bwest@umich.edu

Frauke Kreuter

Joint Program in Survey Methodology
University of Maryland-College Park
fkreuter@survey.umd.edu

Aperçu de la présentation

- Sommaire de la recherche effectuée pour examiner la qualité et l'utilité des parodonnées dans la *National Survey of Family Growth* (NSFG) aux États-Unis
- Recherche sur l'utilité des parodonnées dans des enquêtes par panel importantes aux États-Unis et en Allemagne :
 - *Medical Expenditure Panel Survey* (MEPS)
 - *Labor Market and Social Security* (PASS)

NSFG

- L'enquête majeure au niveau national sur la fécondité aux États-Unis
- Source importante de données sur l'activité sexuelle, le comportement sexuel et la santé reproductive à l'intention des décideurs
- Population cible (jusqu'en septembre 2015) : 15 à 44 ans
- Population (de septembre 2015 à aujourd'hui) : 15 à 49 ans
- Plan d'échantillonnage continu : quatre échantillons trimestriels nationaux sont diffusés et traités chaque année
- Interviews sur place (60-80 minutes) avec une personne de chaque ménage; AAIAO pour les questions délicates

Paradonnées dans la NSFG

- Observations des intervieweurs
 - Niveau du segment (région) (p. ex., préoccupations en matière de sécurité?)
 - Niveau de l'unité de logement (p. ex., jeunes enfants présents?)
 - Niveau du répondant (p. ex., le répondant sélectionné est-il sexuellement actif?)
 - Observations postérieures à l'enquête (p. ex., comportements de l'AAIAO?)
- Données du registre des appels
 - Nombre de tentatives d'appel
 - Preuve de refus, préoccupations, etc.

Paradonnées dans la NSFG

- Résultats des décisions concernant les cas
 - Répondant, refus définitif, hors-échantillon, etc.
- Données découlant des frappes au clavier
 - Demandes d'aide des intervieweurs, sauvegarde, etc.
 - Répondants qui changent leurs réponses
 - Moment de la réponse pour des questions d'enquête particulières

MEPS

- *Medical Expenditure Panel Survey* aux États-Unis
- Sous-échantillon d'interviews sur place de la *National Health Interview Survey*, avec cinq rondes sur 2 ans
- Paradoxonnées :
 - Données du registre des appels
 - Nombre de tentatives d'appel
 - Preuve de refus, préoccupations, etc.
 - Résultats des décisions relatives aux cas
 - Répondant, refus définitif, hors-échantillon, etc.

PASS

- L'étude *Labour Market and Social Security Study* (PASS) en Allemagne
- Enquête annuelle à mode mixte auprès des ménages fondée sur deux échantillons aléatoires (bénéficiaires de l'aide sociale et ménages d'une enquête sur les bâtiments résidentiels); actuellement à la vague du 10^e panel
- Paradoonnées :
 - Données du registre des appels
 - Résultats des décisions relatives aux cas
 - Observations des intervieweurs (étude expérimentale)

Utilité n° 1 : Ajustement pour tenir compte de la non-réponse

- Les observations des intervieweurs recueillies au sujet de toutes les unités échantillonnées sont incluses dans les modèles de propension à répondre, qui servent à ajuster les poids
- Les observations relatives à la fois aux résultats clés et à la propension à répondre peuvent réduire le biais dû à la non-réponse

NSFG (West, 2013a)

Activité sexuelle actuelle du répondant sélectionné

Présence de jeunes enfants

Empêchements physiques d'accès aux unités de logement

PASS (West et coll., 2014)

Fourchette de revenus (faible, moyenne, élevée)

Membre du ménage touchant des prestations d'aide sociale

Défi n° 1 : Qualité des observations

- Qu'arrive-t-il si les observations sont sujettes à des erreurs?
- Par exemple (West, 2013a; West et coll., 2014) :
 - Activité sexuelle : 78 % « d'exactitude »
 - Jeunes enfants : 72 % « d'exactitude »
 - Prestations reçues : 78 % d'exactitude
 - Fourchette de revenus : 55 % d'exactitude
 - L'exactitude varie aussi de façon substantielle parmi les intervieweurs (West et Kreuter, 2013; Sinibaldi et coll., 2013; West et coll., 2014); **pourquoi?**
- Les observations sujettes à des erreurs nuisent à l'efficacité des ajustements pour tenir compte de la non-réponse (West, 2013a; West, 2013b)

Défi n° 1 : Qualité des observations

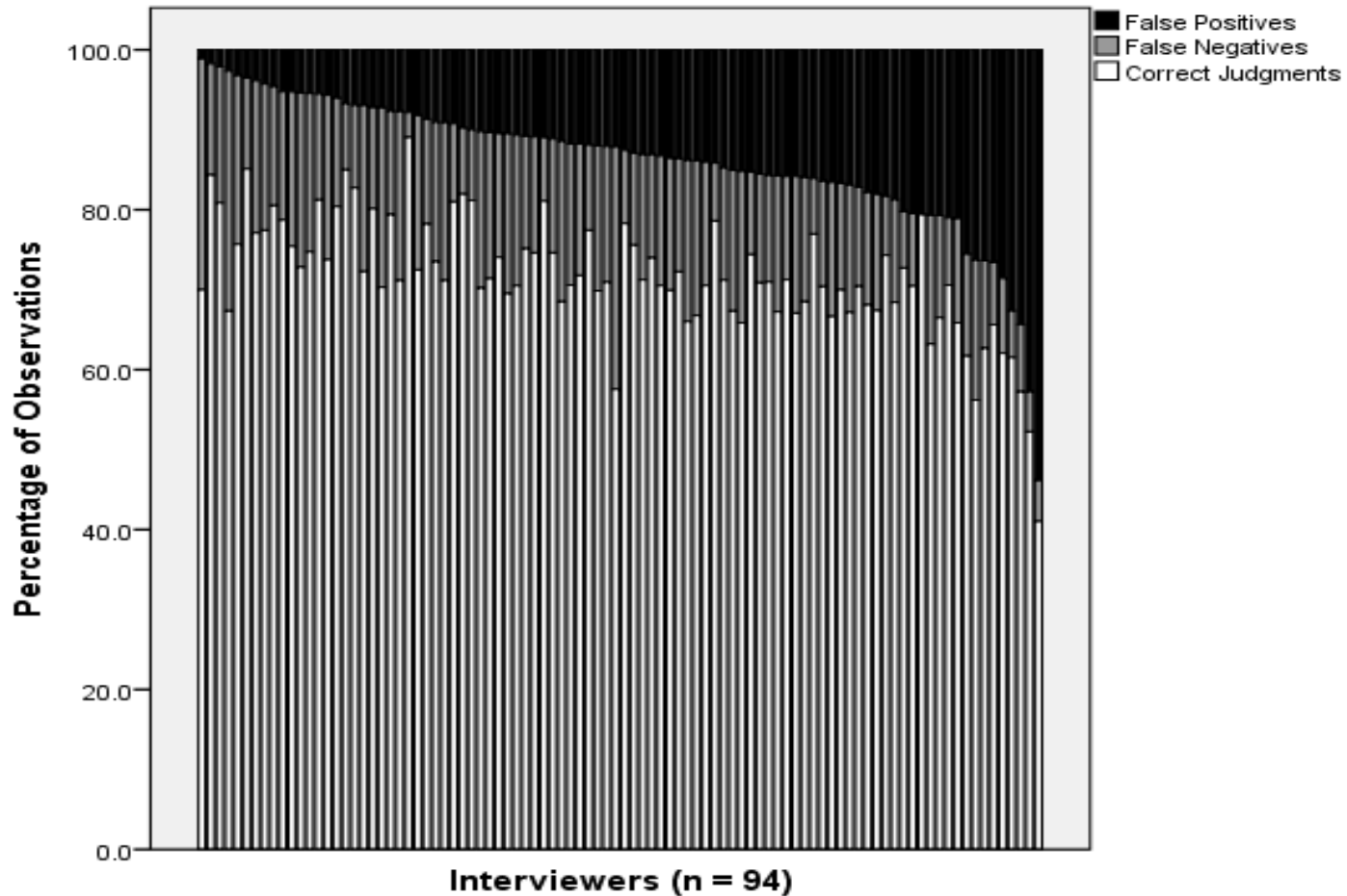


Figure 1: Variance in Observation Accuracy Among 94 NSFG Interviewers.

Défi n° 1 : Qualité des observations

		Unemployment benefit: Self-reported	
		On UB	Not on UB
UB:			
Interviewer Observed		n=1866	n=1347
On UB	(n=1906)	72.8%	27.2%
Not on UB	(n=1234)	21.9%	78.1%
Missing	(n=73)	43.8%	56.2%

		Income: Self-reported		
		Low	Medium	High
Income:				
Interviewer Observed		n=1961	n=684	n=568
Low	(n=1511)	82.3%	13.7%	4.0%
Medium	(n=1362)	45.2%	29.2%	25.6%
High	(n=267)	19.1%	24.7%	56.2%
Missing	(n=73)	69.9%	17.8%	12.3%

Défi n° 1 : Qualité des observations

- **Que pouvons-nous faire à ce sujet?**
- **Une idée** : Fournir aux intervieweurs des prédicteurs importants des caractéristiques qu'ils tentent d'observer (West et Kreuter, 2015)
- Comprendre les indices et les stratégies que les différents intervieweurs utilisent pour faire leurs observations, et **uniformiser la formation** selon les stratégies les plus efficaces (West et Kreuter, 2011; West et coll., soumis)

Utilité n° 2 : Évaluation des intervieweurs

- Les différents types de parodonnées recueillies éclairent les modèles de propension à l'admissibilité, au contact (quotidien) et à la coopération (quotidienne) (Krueger et West, 2014)
- Ces modèles sont utilisés pour calculer les attentes au chapitre de la propension à _____ pour un point donné dans le temps
- Le rendement des intervieweurs peut alors être évalué en comparant les résultats quotidiens réels aux attentes et en faisant une moyenne des écarts pour un intervieweur donné (West et Groves, 2013)

Utilité n° 2 : Évaluation des intervieweurs

- Utiliser les données découlant des frappes au clavier pour identifier les intervieweurs qui ont des tendances inhabituelles à corriger les réponses ou à progresser trop rapidement
- Intervenir auprès des intervieweurs en question pour améliorer leur rendement pendant les interviews

Utilité n° 2 : Évaluation des intervieweurs

- Facteur 1 : Trop rapide (score Z fondé sur toutes les questions)
- Facteur 2 : Grand nombre de vérifications d'erreur
- Facteur 3 : Grand nombre de réponses « Ne sait pas » et « Refus »

Average of Zscore		Column Labels								
		<input type="checkbox"/> C8					<input type="checkbox"/> Q02			
		<input type="checkbox"/> Y2								
		<input type="checkbox"/> Q01								
Row Labels	<input type="checkbox"/> IWER 3	W08	W10	W12	W06	W08	W10	W12		
factor1		-0.36	-0.52	-0.54	-0.79	-0.83	-0.74	-0.82		
factor2		3.48	3.41	3.25	-0.83	-0.36	0.26	0.16		
factor3		0.53	0.74	1.47	0.08	0.50	1.20	1.44		

Défi n° 2 : Spécification du modèle

- Comment savons-nous si un type donné de modèle de propension a été spécifié correctement?
- Les observations des intervieweurs sujettes à erreurs peuvent encore une fois jouer un rôle...
- Les effets aléatoires des intervieweurs devraient-ils être inclus dans les modèles (afin qu'ils puissent être évalués)?
- Qu'arrive-t-il si des parodonnées sont manquantes?

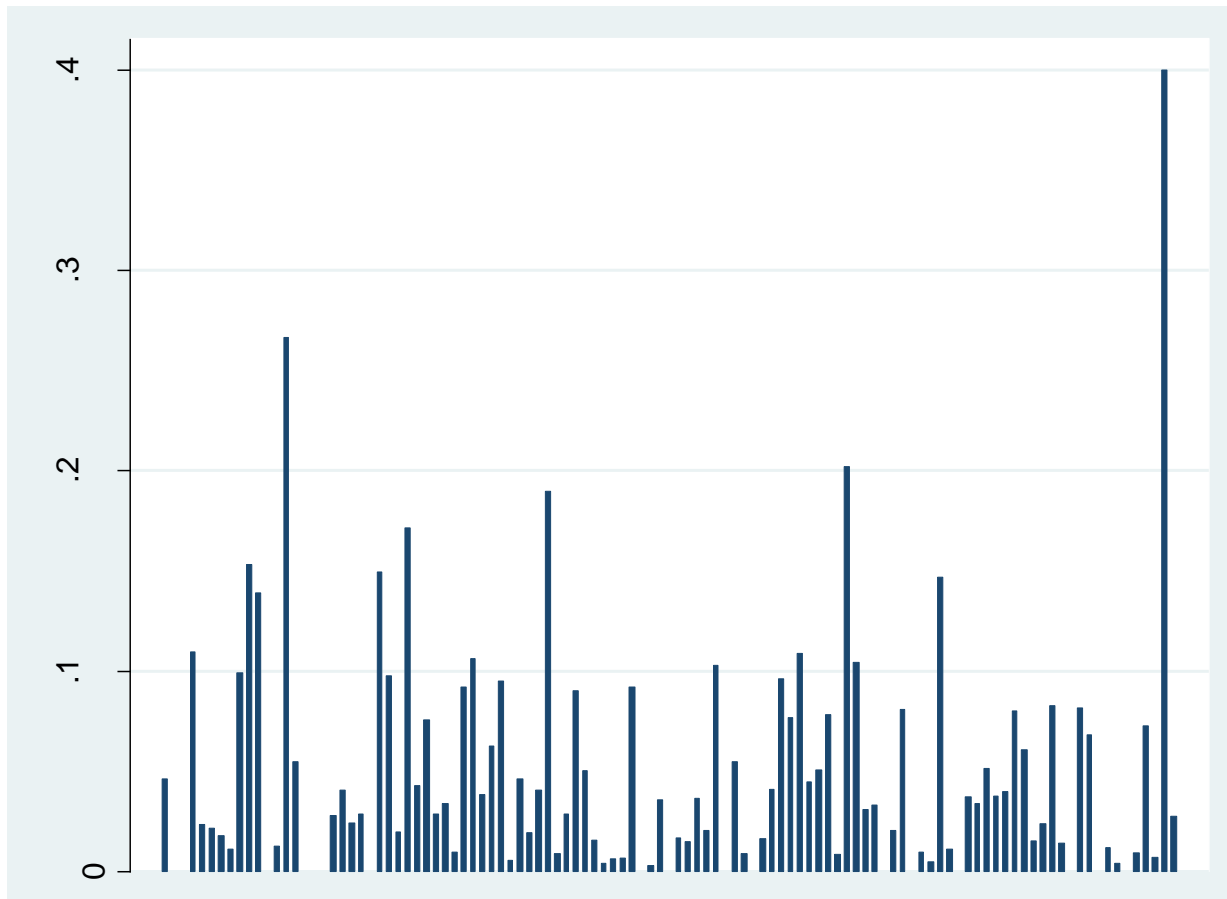
Utilité n° 3 : Qualité de la réponse

- Utiliser les observations postérieures à l'enquête pour identifier les répondants qui peuvent fournir des données de mauvaise qualité
- Évaluer les comportements des intervieweurs qui peuvent avoir des répercussions sur les réponses aux questions délicates pendant l'AAIAO (West et Peytcheva, 2014)

Utilité n° 3 : Qualité de la réponse

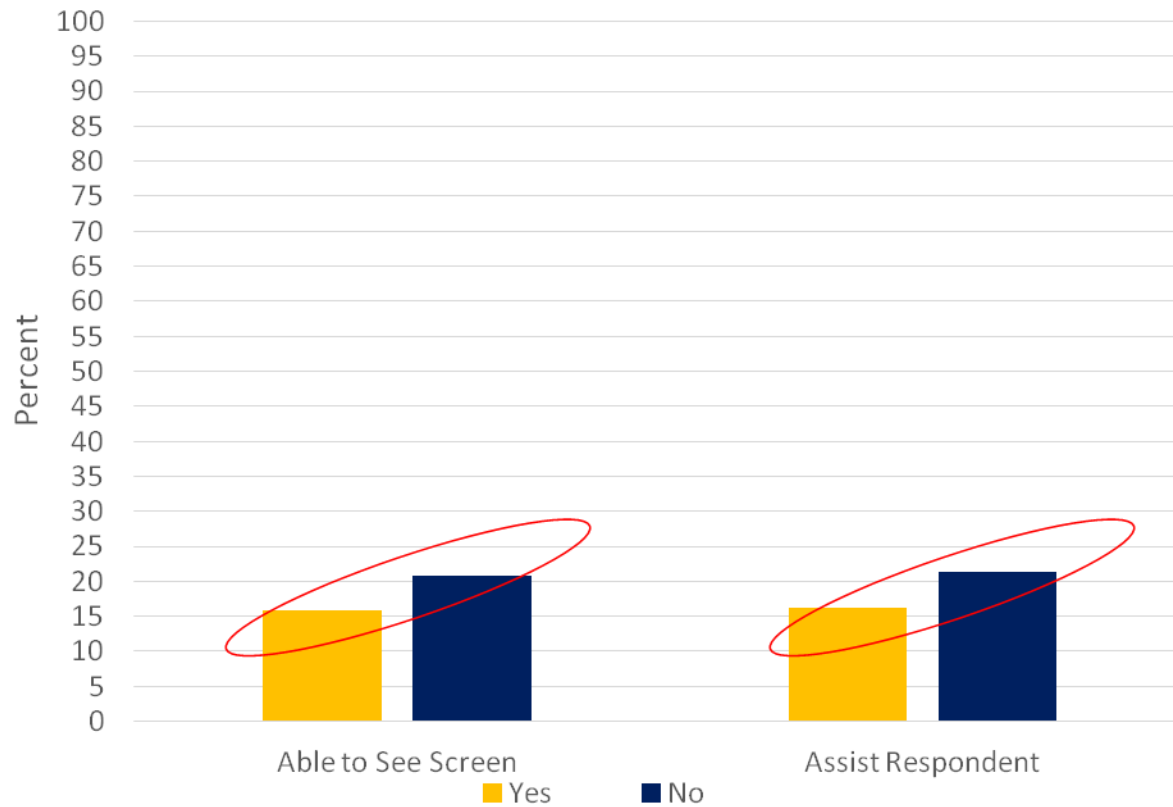
- Les intervieweurs varient substantiellement du point de vue de leur proximité avec l'écran lorsqu'ils sont assis
- Les rapports de l'AAIAO sur les comportements de nature délicate varient selon la vision qu'a l'intervieweur de l'écran

Utilité n° 3 : Qualité de la réponse



Utilité n° 3 : Qualité de la réponse

2+ Occasions Using Marijuana



Défi n° 3 : Observations postérieures à l'enquête

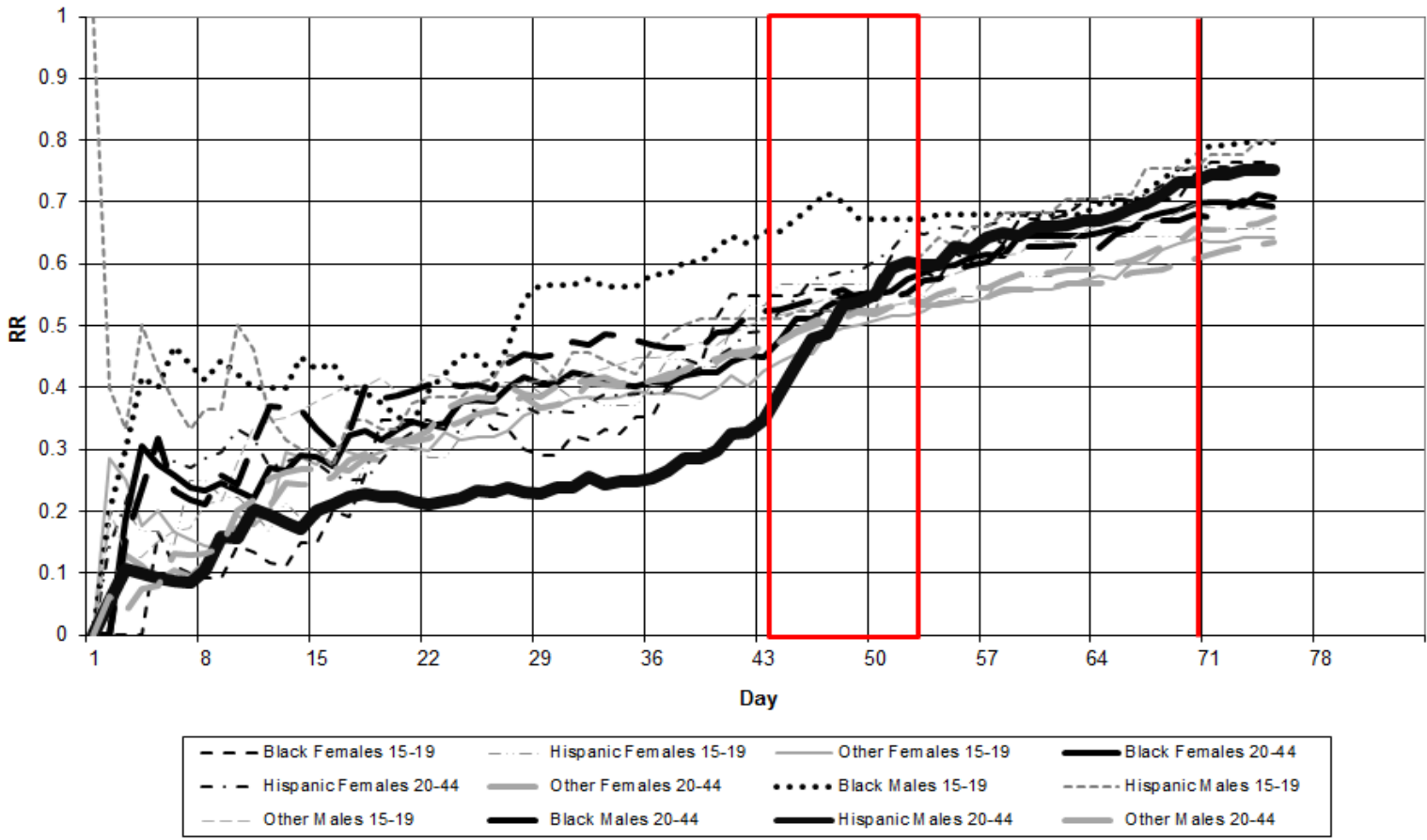
- Les observations des intervieweurs postérieures à l'enquête représentent-elles des indicateurs fiables de la qualité des données (Wang et coll., 2013)?
- Les ouvrages publiés ont montré que cela est fonction des caractéristiques du répondant (plutôt que de la qualité des données), et il existe des preuves uniformes de la variance entre les intervieweurs à cet égard.
- Ces observations peuvent-elles être combinées pour indiquer de façon fiable la qualité des données?
Question ouverte!

Utilité n° 4 : Plan de collecte adaptatif

- Les parodonnées recueillies dans le cadre de la NSFG sont examinées quotidiennement dans le contexte d'un PCA, afin de contrôler la production sur le terrain et l'efficacité
- Des interventions sont mises en œuvre lorsque les parodonnées laissent supposer que certains processus peuvent entraîner un biais ou de l'inefficacité
- **Exemple** : Contrôle des taux de réponse entre les différents sous-groupes sociodémographiques, et accent plus grand mis par les intervieweurs sur les sous-groupes qui tirent de l'arrière (Wagner et coll., 2012)

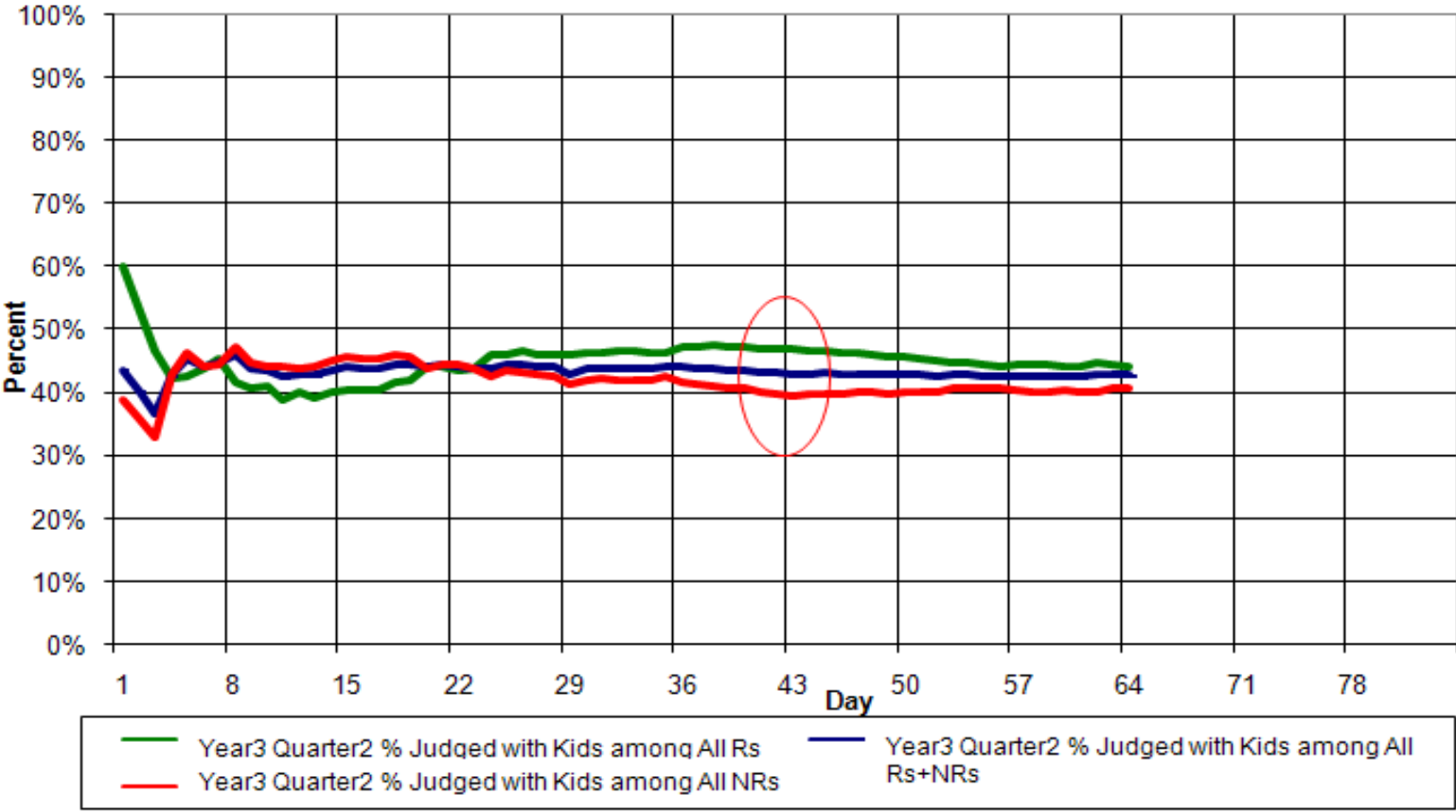
Utilité n° 4 : Plan de collecte adaptatif

Response Rates by Subgroup



Utilité n° 4 : Plan de collecte adaptatif

% Respondents and Nonrespondents Judged to Have Kids by Day



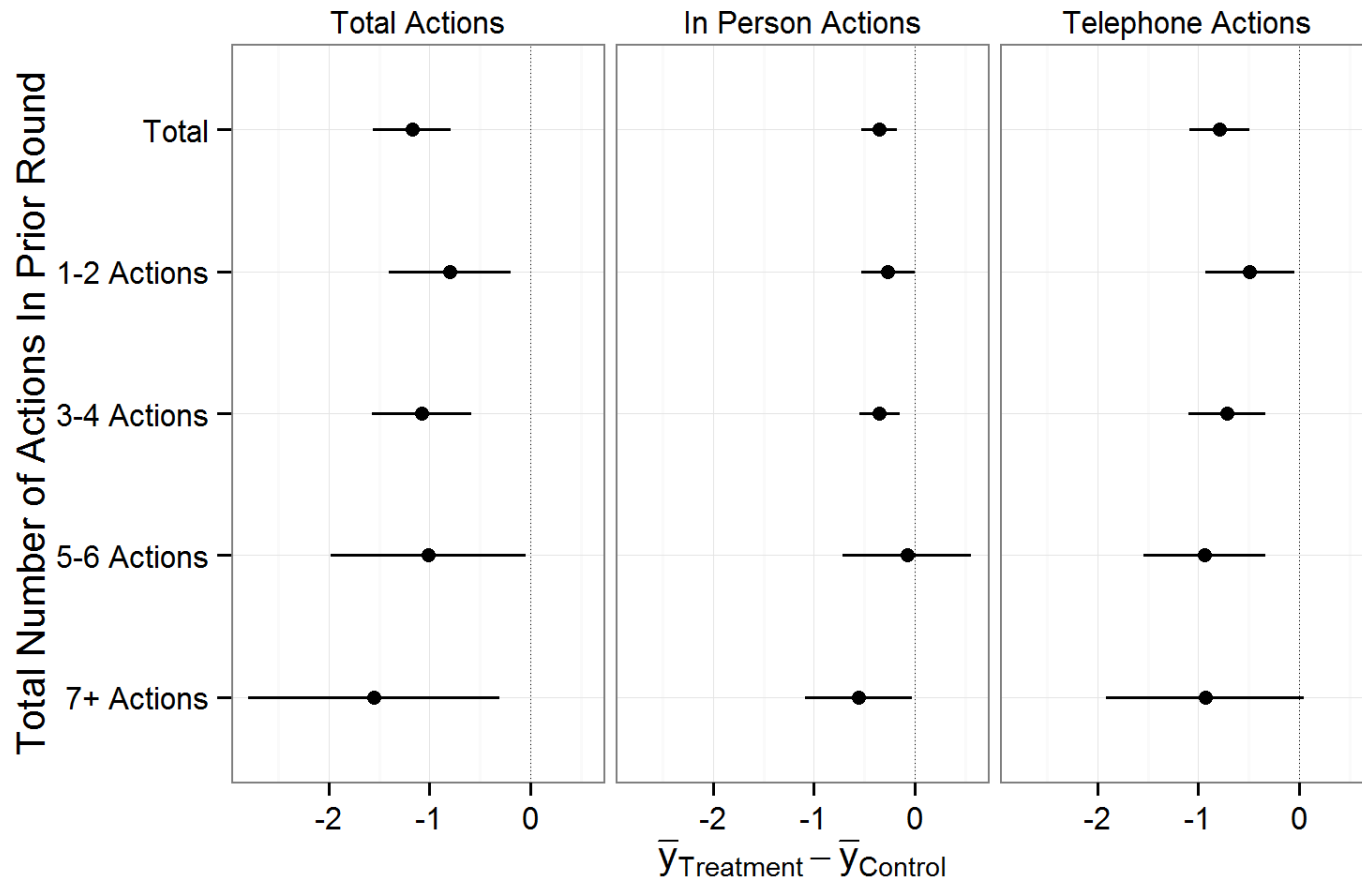
Défi n° 4 : Indicateurs de biais

- Les taux de réponse dans les différents sous-groupes sociodémographiques sont-ils les meilleurs indicateurs du biais lié à la non-réponse?
- Devrait-on réellement contrôler la variabilité du taux de réponse parmi les autres sous-groupes plus étroitement par rapport aux résultats clés (p. ex., présence d'enfants)?
 - Qu'arrive-t-il si les variables du sous-groupe (p. ex., observations) sont sujettes à des erreurs?
- Les paradonnées pourraient éclairer une gamme variée d'indicateurs possibles du biais lié à la non-réponse (Nishimura et coll., à paraître; voir aussi Krueger et West, 2014)

Utilité n° 5 : Efficacité des appels

- Les paradosnées des MEPS et PASS ont servi à adapter les tentatives de contact aux « meilleurs » moments, selon les données historiques.
- Dans la MEPS, des cartes postales ont été envoyées pour indiquer qu'une visite serait effectuée au même moment que l'année précédente (Kreuter et coll., 2014).
- Dans la PASS, on a utilisé des fenêtres de tentatives d'appel réussies pour réduire le temps nécessaire pour le premier contact et la première interview dans les vagues subséquentes (Kreuter et Mueller, 2015)

Utilité n° 5 : Efficacité des appels



Défi n° 5 : Efficacité des appels

- L'utilisation des données sur les appels de la vague précédente peut entraîner un biais en faveur des unités répondantes stables → d'autres recherches sont nécessaires pour voir si une prédiction du changement peut être intégrée
- Le fait de prescrire des stratégies de contact peut perturber les procédures établies sur le terrain; l'adhésion des représentants sur le terrain est nécessaire
- La répartition optimale des temps de contact joue un rôle clé

Sommaire

- Dans le cadre de ces trois enquêtes, on s'est engagé à utiliser une gamme variée de paradonnées pour améliorer les opérations et les produits de données finaux
- Un programme actif de recherche sur les paradonnées est nécessaire pour bien comprendre (et améliorer) les propriétés d'erreur de mesure de ces données
- Les trois enquêtes comportent un long historique de collaboration avec d'autres chercheurs intéressés à ces domaines de recherche!

Bibliographie

- Kreuter, F. et Mueller, G. (2015). A Note on Improving Recruitment Efficiency Using Paradata Information in Panel Surveys, *Field Methods*, 27(1), 55-65.
- Kreuter, F., Mercer, A., et Hicks, W. (2014). Increasing Fieldwork Efficiency through Pre-Specified Dates, *Journal of Survey Statistics and Methodology*, 2(2), 210-223.
- Krueger, B.S. et West, B.T. (2014, Authors Alphabetical). Assessing the Potential of Paradata and Other Auxiliary Information for Nonresponse Adjustments. *Public Opinion Quarterly*, 78(4), 795-831.
- Nishimura, R., Wagner, J., et Elliott, M. (2015). Alternative Indicators for the Risk of Nonresponse Bias: A Simulation Study. *International Statistical Review*. DOI:10.1111/insr.12100.
- Sinibaldi, J., Durrant, G.B., et Kreuter, F. (2013). Evaluating the Measurement Error of Interviewer Observed Paradata. *Public Opinion Quarterly*, 77, 173-193.
- Wagner, J., West, B.T., Kirgis, N., Lepkowski, J.M., Axinn, W.G., et Kruger-Ndiaye, S. (2012). Use of Paradata in a Responsive Design Framework to Manage a Field Data Collection. *Journal of Official Statistics*, 28(4), 477-499.

Bibliographie

- Wang, Y., West, B.T., et Liu, M. (2013). Interviewer perceptions of Survey Data Quality. *Presented at the 38th Annual conference of the Midwest Association for Public Opinion Research (MAPOR), Chicago, IL, 11/23/2013.*
- West, B.T. (2013a). An Examination of the Quality and Utility of Interviewer Observations in the National Survey of Family Growth. *Journal of the Royal Statistical Society, Series A (General)*, 176(1), 211-225.
- West, B.T. (2013b). The Effects of Error in Paradata on Weighting Class Adjustments: A Simulation Study. Chapitre 15 dans *Improving Surveys with Paradata: Analytic Use of Process Information*. Wiley Publishing.
- West, B.T. et Groves, R.M. (2013). The PAIP Score: A Propensity-Adjusted Interviewer Performance Indicator. *Public Opinion Quarterly*, 77(1), 352-374.
- West, B.T., Li, D., et Ma, Y. (*Under Review*). The Identification of Effective Observational Strategies for Survey Interviewers. *Submitted to Sociological Methods and Research, January 2016.*

Bibliographie

- West, B.T. et Kreuter, F. (2011). Observational Strategies Associated with Increased Accuracy of Interviewer Observations: Evidence from the National Survey of Family Growth. Dans les *JSM Proceedings, AAPOR 2011*. Alexandria, VA : American Statistical Association, pp. 5646-5658.
- West, B.T. et Kreuter, F. (2015). A practical technique for improving the accuracy of interviewer observations of respondent characteristics. *Field Methods*, 27(2), 144-162.
- West, B.T. et Kreuter, F. (2013). Factors Affecting the Accuracy of Interviewer Observations: Evidence from the National Survey of Family Growth (NSFG). *Public Opinion Quarterly*, 77(2), 522-548.
- West, B.T., Kreuter, F., et Trappmann, M. (2014). Is the Collection of Interviewer Observations Worthwhile in an Economic Panel Survey? New Evidence from the German Labor Market and Social Security (PASS) Study. *Journal of Survey Statistics and Methodology*, 2(2), 159-181.
- West, B.T. et Peytcheva, E. (2014). Can Interviewer Behaviors During ACASI Affect Data Quality? *Survey Practice*, 5(7).

Merci!

- N'hésitez pas à envoyer vos questions à bwest@umich.edu ou fkreuter@umd.edu.
- Nous sommes prêts à répondre à vos questions.