

LES EFFETS DU MODE DE COLLECTE DES DONNÉES SUR LA MESURE DE L'EMPLOI : UNE COMPARAISON ENTRE LE WEB ET LE TÉLÉPHONE

Johann Neumayr ¹ & Joachim Schork ² & Guillaume Osier ³

¹ *STATEC (Institut National de la Statistique et des Etudes Economiques du Grand-Duché de Luxembourg) - Johann.Neumayr@statec.etat.lu*

² *STATEC et Luxembourg Income Study (LIS) - Joachim.Schork@statec.etat.lu*

³ *STATEC (Institut National de la Statistique et des Etudes Economiques du Grand-Duché de Luxembourg) - Guillaume.Osier@statec.etat.lu*

Résumé. Cette étude cherche à évaluer les effets de mode impactant les variables tirées de l'enquête sur les forces de travail (EFT) au Luxembourg. Cette enquête repose sur un système de collecte mixte web/téléphone. Une technique de matching (coarsened exact matching) permet d'apparier le sous-échantillon des répondants par Internet avec ceux par téléphone en contrôlant le mécanisme de sélection des sous-échantillons. On peut ainsi évaluer l'ampleur de l'effet de mode qui affecte les résultats.

L'application de cette approche sur les variables de l'EFT ne révèle aucun effet de mode significatif pour ce qui est des variables de statut dans l'emploi, tandis que des effets de mode importants apparaissent avec des questions subjectives comme la satisfaction avec l'emploi ou la rémunération.

Mots-clés. Enquêtes en modes mixtes, enquêtes par Internet, effets de mode

Abstract Based on the Luxembourgish Labor Force Survey, the largest recurring household survey in Luxembourg, we investigate if reports on participants employment status differ in web and telephone data. Analysis of the raw data reveals significant differences in sample composition (e.g. respondents personal characteristics such as age or nationality) as well as in key variables for measuring employment status.

In order to investigate whether differences in employment status are caused by sample composition or survey mode effects, we match web and telephone samples according to variables that lead to dissimilarities in sample composition. Based on the selected variables, we then apply a Coarsened Exact Matching to enable a comparison among web and telephone samples. After matching, we show that most differences between web and

telephone samples disappear, which leads to the conclusion that survey mode has a minor effect on measuring employment status.

Keywords Mixed-mode surveys, web surveys, mode effects

1 Introduction

Dans de nombreux domaines, les enquêtes par sondage sont une source importante de données pour l'évaluation et la décision. Des modes de collecte efficaces et adaptés sont indispensables pour espérer obtenir des données de bonne qualité. Pendant longtemps, les enquêtes ont été réalisées par l'entremise d'enquêteurs dont la mission était justement de recueillir les informations sur le terrain auprès des individus. Les entretiens étaient généralement réalisés en face-à-face ou par téléphone. Ces modes traditionnels de recueil de l'information font face aujourd'hui à de nombreux problèmes tels que le recul depuis maintenant plusieurs années de la participation aux enquêtes, le nombre croissant de ménages n'ayant plus de ligne téléphonique fixe ou encore la pression de plus en plus forte au niveau politique pour obtenir rapidement et à moindre coût des statistiques fiables, ce que les modes traditionnels de collecte ne permettent généralement pas d'atteindre (Blanke and Luiten, 2014; Groves, 2005).

Dans le même temps, les progrès de l'informatique et le nombre de plus en plus élevé de ménages possédant un accès à Internet autorise aujourd'hui le recours à des questionnaires en ligne. A la différence des modes de collecte traditionnels, les enquêtes par Internet offrent de nombreux avantages (Bianchi et al., 2017; Fan and Yan, 2010; Greene et al., 2008; de Leeuw, 2005): elles sont moins coûteuses, plus rapides à traiter, plus commodes pour les répondants et plus faciles à soumettre à un grand nombre de participants. En outre, elles permettent de toucher des segments de la population qui sont habituellement beaucoup plus difficiles à contacter via les canaux traditionnels, comme les jeunes ou les personnes d'origine étrangère. Pour toutes ces raisons, les questionnaires en ligne sont devenus un outil particulièrement intéressant pour les concepteurs d'enquêtes.

L'usage grandissant des enquêtes en ligne soulève cependant un certain nombre de difficultés, en particulier la question de savoir si les réponses données par Internet sont systématiquement différentes de celles que les individus auraient communiquées s'ils avaient été interrogés en utilisant d'autres modes de collecte. On parle alors d'effet de mode. Ce problème des effets de mode dans les enquêtes sociales fait l'objet d'intenses recherches au niveau du Système Statistique Européen (SSE), en particulier dans le cadre de l'enquête sur les Forces de Travail (EFT). Des études comparatives entre les données EFT obtenues

par web et celles obtenues par téléphone ont montré qu'il n'y a pas d'effet de mode sur les variables liées à l'emploi et que les différences peuvent en fait s'expliquer par des effets de couverture et de non-réponse (Luiten and Blanke, 2015). D'un autre côté, toujours en travaillant sur les données de l'EFT, des effets de mode significatifs ont été observés sur les questions liées aux heures de travail, au statut dans l'emploi et au niveau d'éducation (Körner and Liersch, 2014; Schouten and van der Laan, 2014).

Ces résultats montrent que l'ampleur des effets de mode dépend à la fois de la nature de l'enquête et de la variable que l'on étudie. C'est pourquoi l'analyse des effets de mode dans une enquête doit se faire séparément pour chaque groupe de variables. L'étude qui suit va porter sur l'EFT qui est conduite par le STATEC, l'Institut National de la Statistique du Luxembourg. Par rapport aux recherches déjà réalisées sur les effets de mode, cette étude présente trois particularités intéressantes:

- 1) La population luxembourgeoise a une structure très originale, avec presque la moitié des résidents qui sont de nationalité étrangère. Le Luxembourg possède trois langues officielles (Luxembourgeois, Français et Allemand) et le questionnaire de l'EFT est administré en cinq différentes langues.
- 2) A l'opposé des variables classiques de l'EFT sur le statut dans l'emploi ou le revenu, nous examinons aussi des questions subjectives comme la satisfaction avec l'emploi actuel ou avec la rémunération.
- 3) L'approche que nous proposons ici utilise une technique de matching (coarsened exact matching) entre les sous-échantillons des différents modes afin de les rendre comparables (Iacus et al., 2008; Iacus et al., 2012).

2 Données

Afin d'analyser les effets de mode, nous utilisons l'enquête sur les Forces de Travail (EFT) qui est conduite par le STATEC auprès de la population résidente du Luxembourg. Afin d'améliorer la précision de nos estimations, les données de trois années consécutives (2015, 2016 et 2017) ont été mises en commun. Ainsi, on obtient un échantillon de $n = 57,570$ individus âgés entre 15 et 74 ans.

Jusqu'en 2015, l'EFT a été conduite uniquement par téléphone en utilisant la méthode "Random Digit Dialing", qui consistait à générer aléatoirement des numéros de téléphone. La chute de la participation à l'enquête a amené le STATEC à revoir le plan de sondage et à mettre en place un système mixte CATI/CAWI. L'échantillon initial est sélectionné

à partir du Registre National des Personnes Physiques (RNPP), qui couvre l'ensemble de la population résidente au Grand-Duché.

- a) Une fois l'échantillon tiré dans le RNPP, il est vérifié si l'individu possède ou non un numéro de téléphone fixe. Ceux pour lesquels un numéro de téléphone est trouvé sont approchés par téléphone, le reste l'étant par Internet *via* une lettre de contact sur laquelle figure un code d'accès (login) leur permettant de se connecter et d'accéder au questionnaire en ligne. Il faut donc bien garder en tête que l'allocation CATI/CAWI ici n'est pas aléatoire: elle dépend de l'existence ou non d'une ligne de téléphone fixe.
- b) Les individus qui n'ont pas rempli le questionnaire en ligne après deux semaines reçoivent une lettre de rappel. Quant au sous-échantillon "téléphone", jusqu'à 30 tentatives de contact sont réalisées.

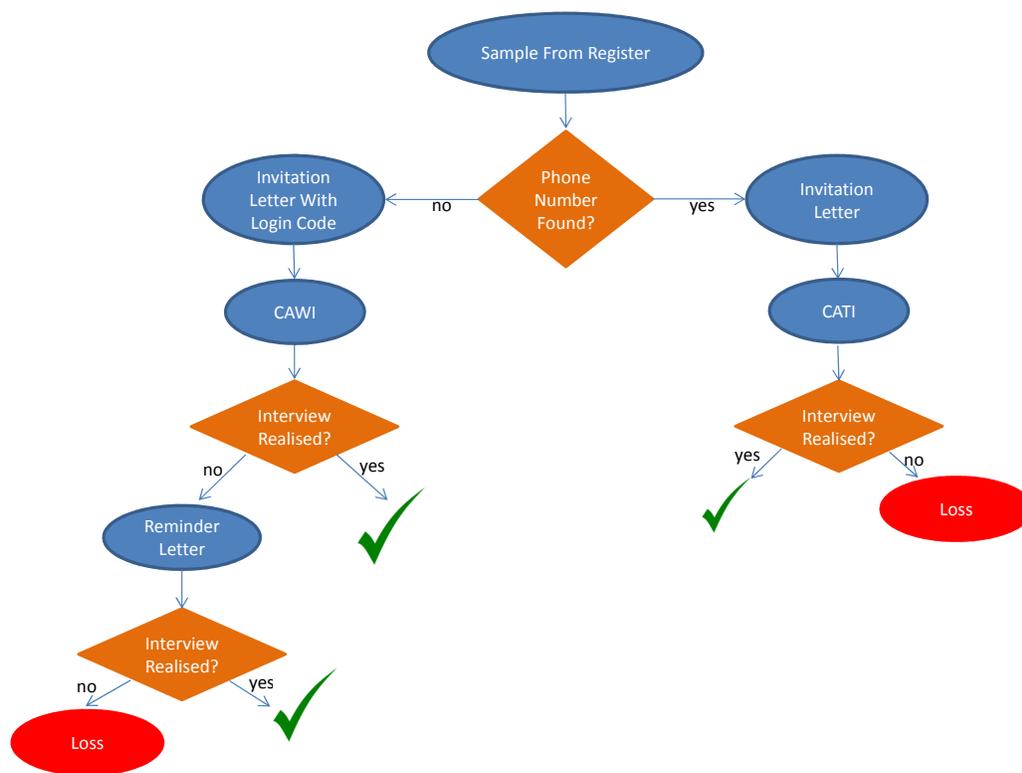


Figure 1: Plan de sondage de l'EFT

3 Différences entre les échantillons web et téléphone

La figure 2 illustre les différences de composition entre les sous-échantillons "web" et "téléphone". Sur un échantillon total de 57,570 individus âgés entre 15 et 74 ans, 60% composent le sous-échantillon "web" et 40% le sous-échantillon "téléphone". Les nationaux luxembourgeois et âgés de 50 ans ou plus apparaissent sur-représentés dans le sous-échantillon "téléphone", tandis que les personnes étrangères de 20 à 49 ans sont relativement plus représentées dans le sous-échantillon "web".

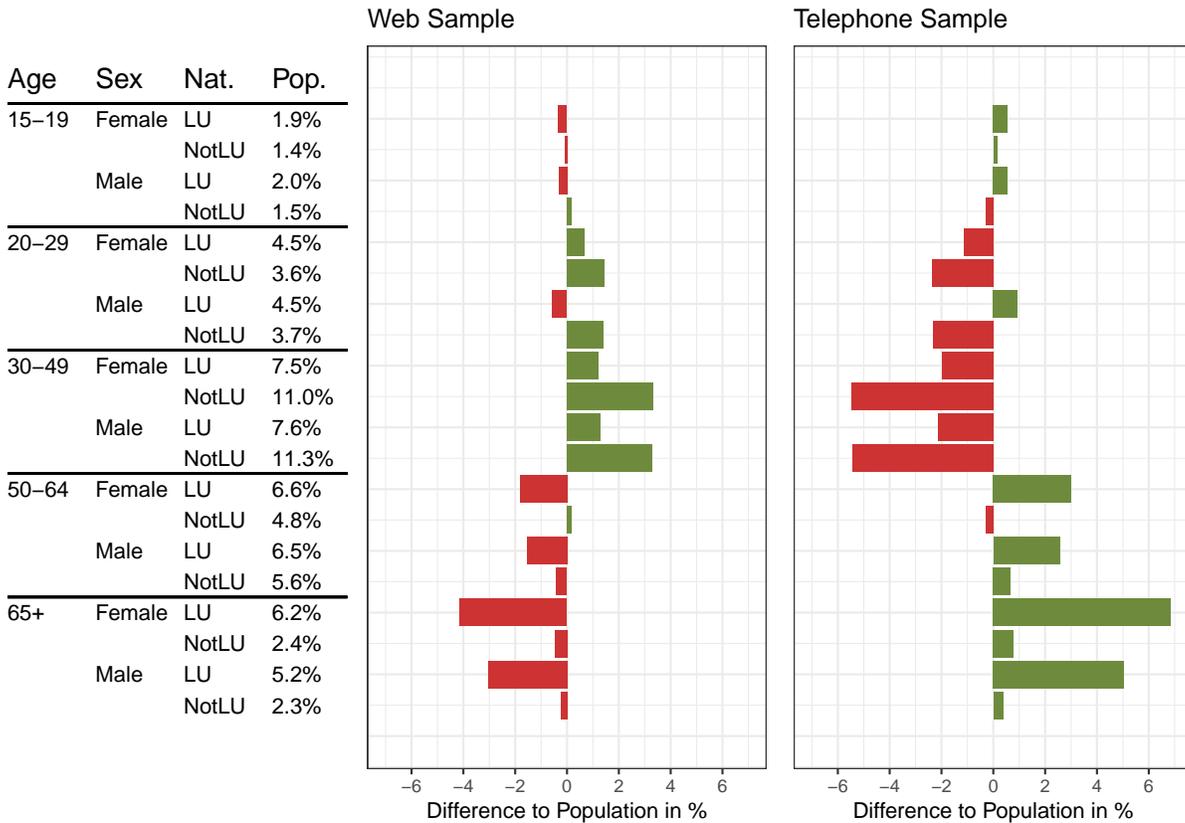


Figure 2: Distribution par âge, sexe et nationalité, comparaison web/téléphone

En général, il existe des écarts significatifs sur de nombreuses variables d'intérêt entre les données collectées par Internet et celles collectées par téléphone (Körner and Liersch, 2014; Lugtig et al., 2011; Sarracino et al., 2017; Schouten and van der Laan, 2014). Les répondants par Internet sont davantage des hommes jeunes, éduqués, célibataires et vivant dans des zones urbaines. Pour ce qui est de l'emploi, ces personnes sont aussi plus souvent en emploi, disposent d'un revenu plus important et déclarent plus souvent travailler plus

de 40 heures par semaine. Si rien n'est fait pour y remédier, ces différences sont des sources potentielles de biais dans les résultats.

Dans notre étude, nous examinons trois différentes variables de l'EFT: le statut d'activité, qui est une information objective et deux variables subjectives, à savoir le degré de satisfaction de la personne avec son travail et avec sa rémunération. Afin de rendre les sous-échantillons "web" et "téléphone" comparables, une technique de matching (coarsened exact matching) a été appliquée en utilisant plusieurs variables de contrôle. L'objectif de l'appariement des deux sous-échantillons est de "gommer" l'effet de sélection des sous-échantillons pour ne conserver que l'effet de mode pur.

La méthode CEM (coarsened exact matching) a été introduite récemment dans la littérature (Iacus et al., 2012). Elle possède de bonnes propriétés statistiques et peut dépasser en performance les méthodes traditionnelles de matching comme celles basées sur des scores de propensité. Dans le CEM, l'appariement se fait de façon exacte. Les observations d'un sous-échantillon qui n'ont pas pu trouver de "jumeau" dans l'autre sous-échantillon sont alors éliminées du fichier final. On obtient donc un échantillon final de taille réduite, donc de variance plus élevée. D'un autre côté, l'appariement étant exact, on espère ainsi diminuer le biais résultant du matching et donc améliorer la qualité globale des estimations.

4 Sélection des variables

Construire un modèle de matching qui soit efficace (*i.e.* un ensemble de variables de contrôle utilisées pour rendre comparables les sous-échantillons "web" et "téléphone") est un préalable essentiel pour que la méthode CEM conduise à des résultats fiables (Iacus et al., 2008; Iacus et al., 2012). Il est courant dans la pratique d'inclure dans le modèle des variables qui expliquent la probabilité d'être interrogé par web ou par téléphone (Imbens, 2004; Rosenbaum and Rubin, 1983). La littérature récente pointe aussi l'importance d'inclure des variables ayant un fort pouvoir prédictif des variables d'intérêt de l'enquête (Brookhart et al., 2006; de Luna et al., 2011; Fowler, 2012). En outre, les modèles de matching doivent être aussi parcimonieux que possible dans la mesure où chaque nouvelle variable rend les appariements plus difficiles (Iacus et al., 2012).

Afin de trouver le meilleur modèle possible, nous devons identifier les variables les plus pertinentes parmi la liste des possibles. Dans une première étape, une sélection automatique des variables a été faite à partir d'un algorithme de segmentation des données de type "Random Forest". Ensuite, cette première liste de variables a été épurée sur la base à la fois de considérations théoriques et empiriques.

5 Résultats

La figure 3 montre les différences avant et après application de la méthode CEM au niveau de la variable de statut d'activité pour les trois catégories actif, chômage et inactif. Avant le CEM, les écarts sont très marqués entre le web et le téléphone. Par rapport au sous-échantillon "téléphone", les répondants par Internet apparaissent davantage actifs et moins au chômage. Cependant, après l'application du CEM, les écarts entre le web et le téléphone ne sont plus statistiquement significatifs. Cela indique que les différences peuvent en fait être expliquées par la composition des deux sous-échantillons. Il ne semble donc pas y avoir d'effets de mode pesant sur la variable de statut d'activité.

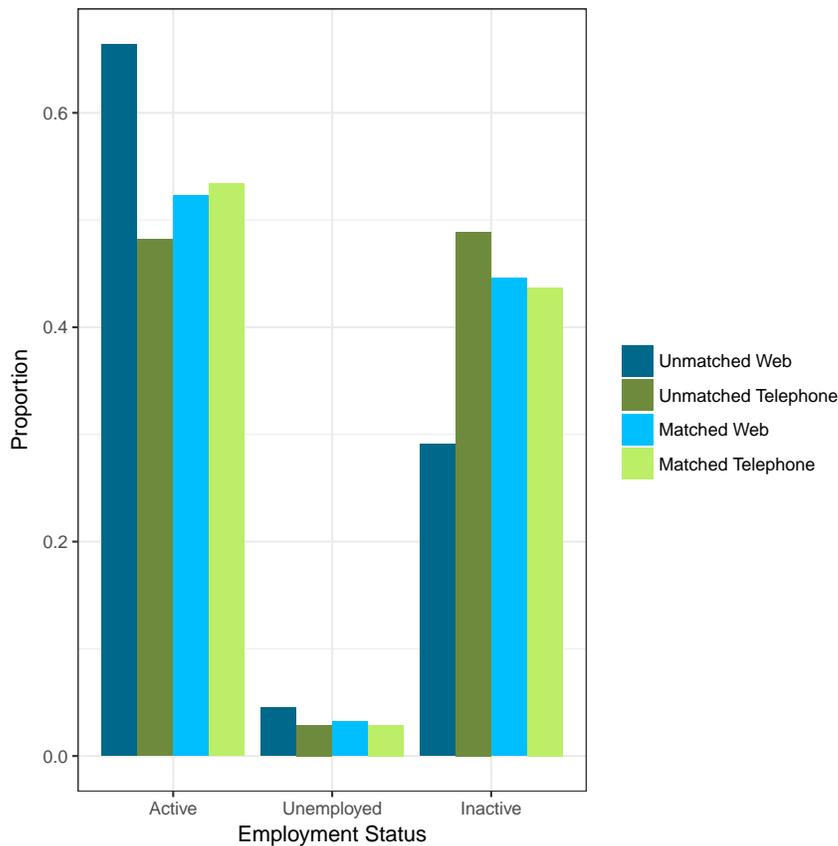


Figure 3: Effets de mode sur le statut d'activité

La figure 4 visualise les écarts avant et après matching pour les variables de satisfaction avec l'emploi et de satisfaction avec la rémunération. Avant le CEM, les répondants par Internet perçoivent moins souvent leur salaire comme étant juste. Un constat similaire est

fait avec la variable de satisfaction avec l'emploi. Ces écarts persistent après l'application du CEM, mettant ainsi en évidence des effets de mode pour ces deux variables.

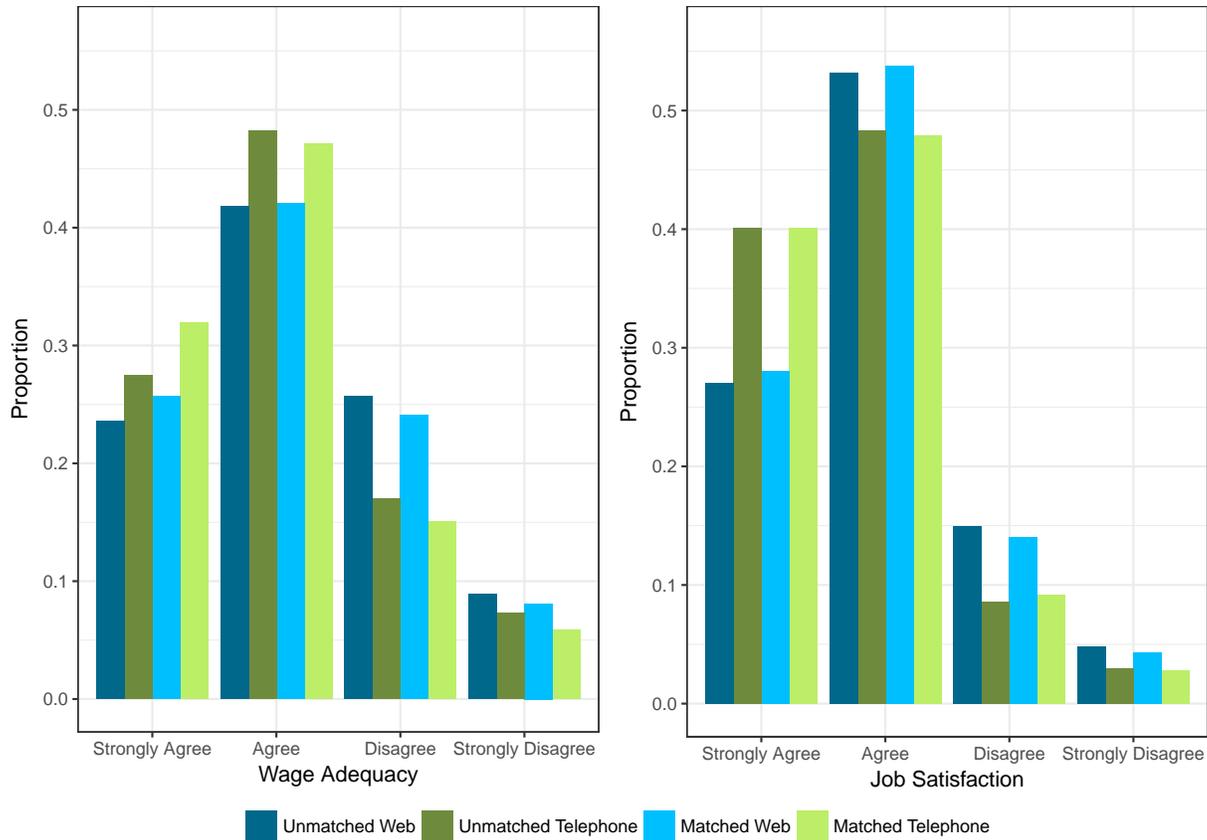


Figure 4: Effets de mode pour les variables de satisfaction avec l'emploi et avec la rémunération

A ce stade, on se pose la question pourquoi la satisfaction avec l'emploi ou la rémunération sont affectées par des effets de mode, contrairement au statut d'activité. Cette dernière variable est objective tandis que les deux premières sont plutôt subjectives dans la mesure où l'opinion et le ressenti des individus y jouent un rôle important.

6 Synthèse

- Afin de mesurer l'ampleur des effets de mode dans l'enquête sur les Forces de travail, les sous-échantillons "web" et "téléphone" ont été rendus comparables via une méthode de coarsened exact matching.

- Les résultats suggèrent que les effets de mode n’affectent pas le statut d’activité, qui est une variable objective.
- Au contraire, les variables subjectives telles que la satisfaction avec l’emploi et celle avec la rémunération sont significativement affectées par les effets de mode.

References

- Bianchi, A., Biffignandi, S., and Lynn, P. (2017). Web-face-to-face mixed-mode design in a longitudinal survey: Effects on participation rates, sample composition, and costs. *Journal of Official Statistics*, 33(2):385–408.
- Blanke, K. and Luiten, A. (2014). Query on data collection for social surveys. *ESSnet Project "Data Collection for Social Surveys using Multiple Modes"*.
- Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., and Stürmer, T. (2006). Variable selection for propensity score models. *American Journal of Epidemiology*, 163(12):1149–1156.
- de Leeuw, E. D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, 21(2):233–255.
- de Luna, X., Waernbaum, I., and Richardson, T. S. (2011). Covariate selection for the nonparametric estimation of an average treatment effect. *Biometrika*, 98(4):861–875.
- Fan, W. and Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior*, 26(2):132–139.
- Fowler, P. (2012). Covariate selection and coarsened exact matching in causal inference. a simulation study. *Master Thesis*.
- Greene, J., Speizer, H., and Wiitala, W. (2008). Telephone and web: Mixed-mode challenge. *Health Services Research*, 43(1 P1):230–248.
- Groves, R. M. (2005). *Survey Errors and Survey Costs*. Hoboken, New Jersey: John Wiley & Sons, 2nd edition.
- Iacus, S. M., King, G., and Porro, G. (2008). Matching for causal inference without balance checking. *Available at SSRN 1152391*, (617):1–17.
- Iacus, S. M., King, G., and Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis*, 20(1):1–24.

- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *The review of economics and statistics*, 86(1):4–29.
- Körner, T. and Liersch, A. (2014). Case study on mode effects in the german labour force survey. *Deliverable for work package III of the ESSnet on Data Collection for Social Surveys Using Multiple Modes*.
- Lugtig, P., Lensvelt-Mulders, G. J. L. M., Frerichs, R., and Greven, A. (2011). Estimating nonresponse bias and mode effects in a mixed-mode survey. *International Journal of Market Research*, 53(5):1–16.
- Luiten, A. and Blanke, K. (2015). Conclusions of the ESSnet – DCSS on web and mixed mode data collection in official social surveys. *Conference of New Techniques and Technologies for Statistics (NTTS)*.
- Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.
- Sarracino, F., Riillo, C. A. F., and Mikucka, M. (2017). Comparability of web and telephone surveys for the measurement of subjective well-being. *Survey Research Methods*, 11(2):141–169.
- Schouten, B. and van der Laan, J. (2014). ESSnet deliverable WPIII: Mode effect decompositions for the dutch labour force survey. *Deliverable for work package III of the ESSnet on Data Collection for Social Surveys Using Multiple Modes*.