

Les effets de la collecte par internet sur le taux
de réponse et sur les résultats d'une enquête
auprès d'un public précaire.
Résultats à partir d'une expérimentation aléatoire.

Sandrine FIRQUET

Dares

14 juin 2018



Sommaire

- 1 Introduction
- 2 Expérimentations
- 3 Bilan des collectes
- 4 Un effet de mode ?

Sommaire

- 1 Introduction
- 2 Expérimentations
- 3 Bilan des collectes
- 4 Un effet de mode ?

Introduction

- Enquête mise en place en 2010 pour les sortants de CUI et fin 2014 pour les sortants d'EAv

Introduction

- Enquête mise en place en 2010 pour les sortants de CUI et fin 2014 pour les sortants d'EAv
- Objectifs de l'enquête :
 - ▶ Savoir ce qu'ils sont devenus six mois après la sortie du contrat aidé
 - ▶ Recueillir des informations sur le déroulement du contrat : formation, tutorat, accompagnement par la mission locale
 - ▶ Connaître leur opinion sur l'utilité de leur passage en contrat aidé

Introduction

- Enquête mise en place en 2010 pour les sortants de CUI et fin 2014 pour les sortants d'EAv
- Objectifs de l'enquête :
 - ▶ Savoir ce qu'ils sont devenus six mois après la sortie du contrat aidé
 - ▶ Recueillir des informations sur le déroulement du contrat : formation, tutorat, accompagnement par la mission locale
 - ▶ Connaître leur opinion sur l'utilité de leur passage en contrat aidé
- Protocole et mode de collecte :
 - ▶ Enquête papier courte (recto/verso)
 - ▶ Réalisée en continue 6 mois après la sortie du contrat aidé
 - ▶ Exhaustive jusqu'en 2015

Introduction

- Taux de réponse faibles : 22% pour l'enquête EAv et 27% pour l'enquête CUI en 2014

Introduction

- Taux de réponse faibles : 22% pour l'enquête EAv et 27% pour l'enquête CUI en 2014
- ☞ Souhait d'intégrer Internet comme mode de collecte
- ☞ Réflexion préalable nécessaire : mise en place de deux expérimentations pour analyser l'impact d'un changement du mode de collecte

Sommaire

- 1 Introduction
- 2 Expérimentations**
- 3 Bilan des collectes
- 4 Un effet de mode ?

Expérimentations : protocoles de collecte

- Tirage aléatoire des deux échantillons parmi les sortants ayant une adresse e-mail
- Deux collectes monomodes en parallèle
- Questionnaires similaires (filtres automatisés sur internet)

Expérimentations : protocoles de collecte

- Tirage aléatoire des deux échantillons parmi les sortants ayant une adresse e-mail
- Deux collectes monomodes en parallèle
- Questionnaires similaires (filtres automatisés sur internet)
- Envois des questionnaires et relances :

Expérimentations : protocoles de collecte

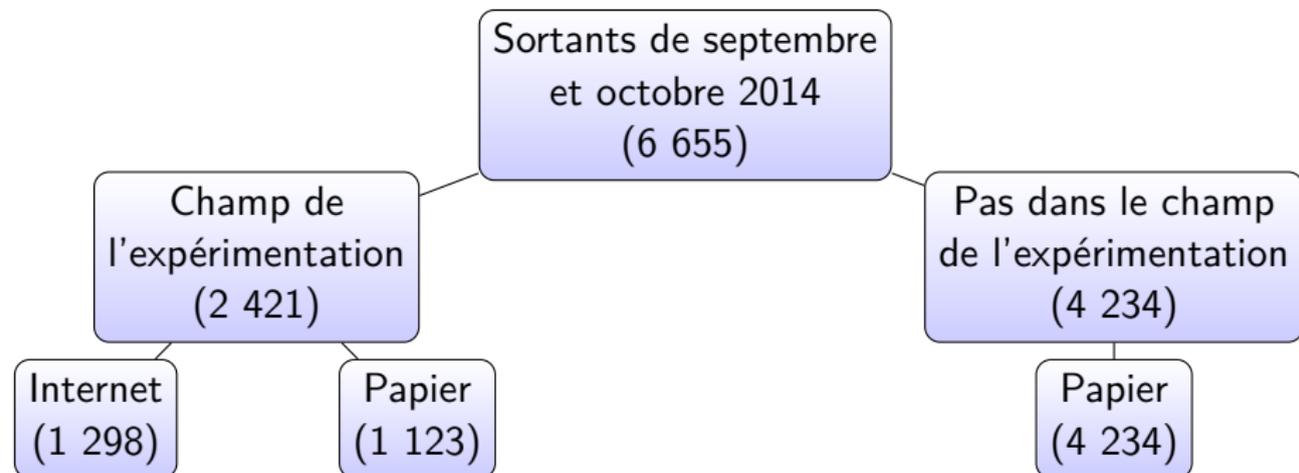
- Tirage aléatoire des deux échantillons parmi les sortants ayant une adresse e-mail
- Deux collectes monomodes en parallèle
- Questionnaires similaires (filtres automatisés sur internet)
- Envois des questionnaires et relances :
 - ▶ Enquête internet : envoi direct du lien vers le questionnaire par e-mail puis deux relances par e-mail et SMS à une semaine d'intervalle chacune

Expérimentations : protocoles de collecte

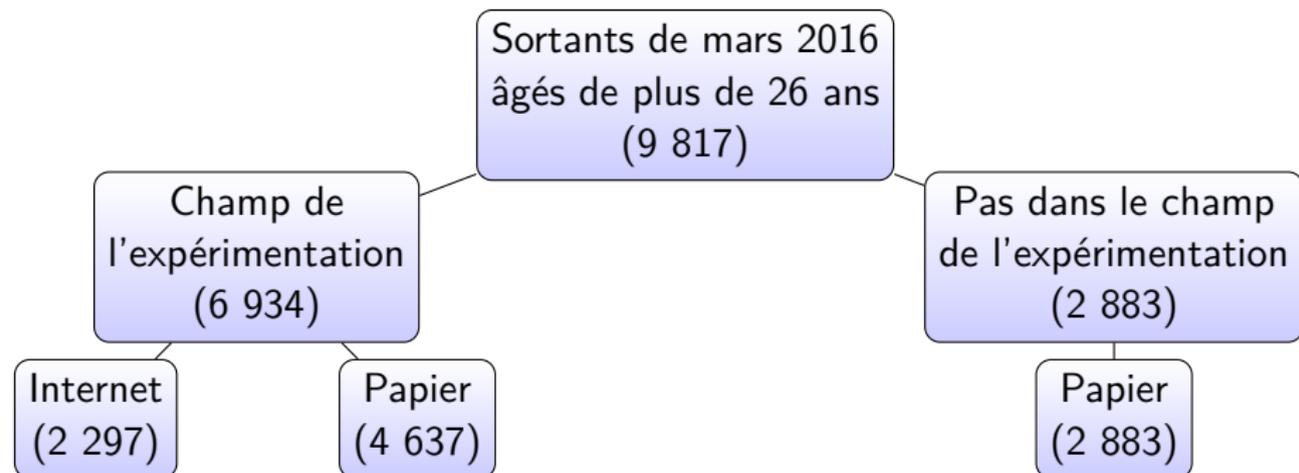
- Tirage aléatoire des deux échantillons parmi les sortants ayant une adresse e-mail
- Deux collectes monomodes en parallèle
- Questionnaires similaires (filtres automatisés sur internet)
- Envois des questionnaires et relances :
 - ▶ Enquête internet : envoi direct du lien vers le questionnaire par e-mail puis deux relances par e-mail et SMS à une semaine d'intervalle chacune
 - ▶ Enquête papier : envoi initial du questionnaire par courrier puis une ou deux relances (CUI, EA_v) à un mois d'intervalle chacune

Expérimentation EAv (2015)

Expérimentation EA_v (2015)



Expérimentation CUI (2016)



Sommaire

- 1 Introduction
- 2 Expérimentations
- 3 Bilan des collectes**
- 4 Un effet de mode ?

Impact du mode de collecte sur les taux de réponse

Impact du mode de collecte sur les taux de réponse

- Taux de réponse à l'enquête internet supérieur à celui de l'enquête papier

	EAv_2015			CUI_2016		
	Taux de réponse (%)	Taux de NPAI (%)	Taux de réponse hors NPAI (%)	Taux de réponse (%)	Taux de NPAI (%)	Taux de réponse hors NPAI (%)
Internet	38	5	40	47	1	48
Papier	22	21	28	35	10	39

Tableau – Taux de réponse aux enquêtes selon le mode de collecte
Source : ASP, traitement Dares

Impact du mode de collecte sur les taux de réponse

- Taux de réponse à l'enquête internet supérieur à celui de l'enquête papier

	EA _v _2015			CUI_2016		
	Taux de réponse (%)	Taux de NPAI (%)	Taux de réponse hors NPAI (%)	Taux de réponse (%)	Taux de NPAI (%)	Taux de réponse hors NPAI (%)
Internet	38	5	40	47	1	48
Papier	22	21	28	35	10	39

Tableau – Taux de réponse aux enquêtes selon le mode de collecte
Source : ASP, traitement Dares

- Résultat qui se retrouve en travaillant à caractéristiques sociodémographiques comparables et pour un même type de contrat aidé

Impact du mode de collecte sur les taux de réponse

- Taux de réponse à l'enquête internet supérieur à celui de l'enquête papier

	EA _v _2015			CUI_2016		
	Taux de réponse (%)	Taux de NPAI (%)	Taux de réponse hors NPAI (%)	Taux de réponse (%)	Taux de NPAI (%)	Taux de réponse hors NPAI (%)
Internet	38	5	40	47	1	48
Papier	22	21	28	35	10	39

Tableau – Taux de réponse aux enquêtes selon le mode de collecte
Source : ASP, traitement Dares

- Résultat qui se retrouve en travaillant à caractéristiques sociodémographiques comparables et pour un même type de contrat aidé
- Taux de non-réponse partielle plus faible pour l'enquête réalisée par internet

Impact du mode de collecte sur le biais de sélection

Impact du mode de collecte sur le biais de sélection

	EA _v _2015		CUI_2016	
	Internet	Papier	Internet	Papier
Sexe				
Homme	35*	40*	31	31
Femme	65*	60*	69	69
Age				
Moins de 21 ans	45	46	n.c.	n.c.
22 ans et plus	55	54	n.c.	n.c.
26 - 35 ans	n.c.	n.c.	38*	24*
36 - 49 ans	n.c.	n.c.	35*	40*
Plus de 50 ans	n.c.	n.c.	27*	36*
Niveau de formation				
Inférieur au CAP	16*	12*	6*	9*
BEP-CAP	52*	61*	28*	34*
Bac	25*	20*	26	25
Supérieur au Bac	8*	6*	40*	32*

Tableau – Caractéristiques des répondants selon le mode de collecte (en pourcentages)

Seuil de significativité : * 5% ; n.c. : non concernés

Source : ASP, traitement Dares

Impact du mode de collecte sur le biais de sélection

- Déterminants du taux de réponse :

Impact du mode de collecte sur le biais de sélection

- Déterminants du taux de réponse :
 - ▶ Etude par le biais de régressions logistiques binaires
 - ▶ Facteurs communs aux deux modes de collecte : sexe, niveau de formation
 - ▶ Facteurs propres à chaque mode de collecte, et à chaque expérimentation (populations aux caractéristiques différentes)

Sommaire

- 1 Introduction
- 2 Expérimentations
- 3 Bilan des collectes
- 4 Un effet de mode ?**

Qu'est-ce qu'un effet de mode ?

- Effet de mode : impact du mode de collecte sur les réponses données
- Décomposition en biais de sélection et biais de mesure
- Plusieurs méthodes possibles pour le mesurer, pas de consensus

Des différences de réponses, même après calage sur marges

Des différences de réponses, même après calage sur marges

	Insertion dans l'emploi	Formation réalisée	Formation qualifiante réalisée	Accompagnement interne	Accompagnement externe
EA_v_2015					
Internet	39*	43*	14	56*	78
Papier	44*	49*	19	72*	76
CUI_2016					
Internet	52	33*	3*	27*	8*
Papier	54	40*	5*	38*	11*

Tableau – Estimations des variables d'intérêt après calage sur marges
(en pourcentages)

Seuil de significativité : * 5% ; n.c. : non concernés

Source : ASP, traitement Dares

Contrôle des caractéristiques observables

Contrôle des caractéristiques observables

- Régressions logistiques binaires :
 - ▶ Taux d'insertion dans l'emploi
 - ▶ Avoir suivi une formation
 - ▶ Avoir suivi une formation qualifiante
 - ▶ Avoir eu un tuteur
 - ▶ Avoir bénéficié d'un accompagnement externe

Contrôle des caractéristiques observables

- Régressions logistiques binaires :
 - ▶ Taux d'insertion dans l'emploi
 - ▶ Avoir suivi une formation
 - ▶ Avoir suivi une formation qualifiante
 - ▶ Avoir eu un tuteur
 - ▶ Avoir bénéficié d'un accompagnement externe
- Régression logistique multinomiale :
 - ▶ Opinion du salarié sur sa situation avant son passage en contrat aidé

Contrôle des caractéristiques observables

- Pour chaque variable à expliquer, deux modèles : $Y = f(X, \text{mode})$ et $Y = f(X, \text{mode}, \text{mode} * X)$ avec $X =$ variables sociodémographiques

Contrôle des caractéristiques observables

- Pour chaque variable à expliquer, deux modèles : $Y = f(X, \text{mode})$ et $Y = f(X, \text{mode}, \text{mode} * X)$ avec $X =$ variables sociodémographiques
- Exemple des modèles pour le taux d'insertion dans l'emploi :

$$EAv : \text{logit}(p_{\beta}(x_i)) = \alpha + \beta_1 \mathbb{1}_{CDI_i} + \beta_2 \mathbb{1}_{marchand_i} + \beta_3 \mathbb{1}_{internet_i}$$

$$\begin{aligned} CUI : \text{logit}(p_{\beta}(x_i)) = & \alpha + \beta_1 \mathbb{1}_{inf_CAP_i} + \beta_2 \mathbb{1}_{BAC_i} + \beta_3 \mathbb{1}_{sup_BAC_i} \\ & + \beta_4 \mathbb{1}_{minima_sociaux_i} + \beta_5 \mathbb{1}_{rupture_i} + \beta_6 \mathbb{1}_{CDI_i} \\ & + \beta_7 \mathbb{1}_{marchand_i} + \beta_8 \mathbb{1}_{26h_35h_i} + \beta_9 \mathbb{1}_{plus_35h_i} + \beta_{10} \mathbb{1}_{internet_i} \end{aligned}$$

$$\text{avec } \text{logit}(p_{\beta}(x_i)) = \log\left(\frac{p_{\beta}(x_i)}{1-p_{\beta}(x_i)}\right), \quad p_{\beta}(x_i) = P_{\beta}(Y_i = 1 | X_i)$$

Contrôle des caractéristiques observables

	Insertion dans l'emploi	Formation réalisée	Formation qualifiante réalisée	Accompagnement interne	Accompagnement externe
EA_v_2015					
Internet	-13,54**	-8,18**	n.s.	-16,82***	n.s.
Papier	Réf.	Réf.	Réf.	Réf.	Réf.
CUI_2016					
Internet	n.s.	-6,73***	-1,62**	-10,67***	n.s.
Papier	Réf.	Réf.	Réf.	Réf.	Réf.

Tableau – Tableau tronqué des effets marginaux du mode de collecte (en pourcentages)

Seuils de significativité : *** 1% ; ** 5% ; * 10% ; n.s. : non significatif

Source : ASP, traitement Dares

Contrôle des caractéristiques observables

	Insertion dans l'emploi	Formation réalisée	Formation qualifiante réalisée	Accompagnement interne	Accompagnement externe
EA_v_2015					
Internet	-13,54**	-8,18**	n.s.	-16,82***	n.s.
Papier	Réf.	Réf.	Réf.	Réf.	Réf.
CUI_2016					
Internet	n.s.	-6,73***	-1,62**	-10,67***	n.s.
Papier	Réf.	Réf.	Réf.	Réf.	Réf.

Tableau – Tableau tronqué des effets marginaux du mode de collecte (en pourcentages)

Seuils de significativité : *** 1% ; ** 5% ; * 10% ; n.s. : non significatif

Source : ASP, traitement Dares

- Effets résiduels : biais de sélection sur inobservables ou biais de mesure ?

Conclusion

Conclusion

- Internet : taux de réponse plus élevé, taux de non-réponse partielle plus faible mais quid de la qualité des données ?

Conclusion

- Internet : taux de réponse plus élevé, taux de non-réponse partielle plus faible mais quid de la qualité des données ?
- Situation des personnes ayant été interrogées par internet plus défavorable que celles ayant été interrogées par papier

Conclusion

- Internet : taux de réponse plus élevé, taux de non-réponse partielle plus faible mais quid de la qualité des données ?
- Situation des personnes ayant été interrogées par internet plus défavorable que celles ayant été interrogées par papier
- A l'avenir :
 - ▶ utilisation du FHA pour comparer le taux d'insertion dans l'emploi
 - ▶ proposition peut être faite à l'ASP d'une collecte multimode séquentielle (incluant une relance papier)
 - ▶ traitements statistiques spécifiques pour l'agrégation des données ?

Merci pour votre attention

Contact :

Sandrine FIRQUET (Dares)

Tel : 01.44.38.23.70

Courriel : sandrine.firquet@travail.gouv.fr