

PROCOLES MULTIMODES EFFET DE MESURE ET AGREGATION

Colloque francophone sur les sondages

13/10/2016

Christophe Barret - Christophe Dzikowski

■ ■ ■ ■ ■ • Plan de la présentation

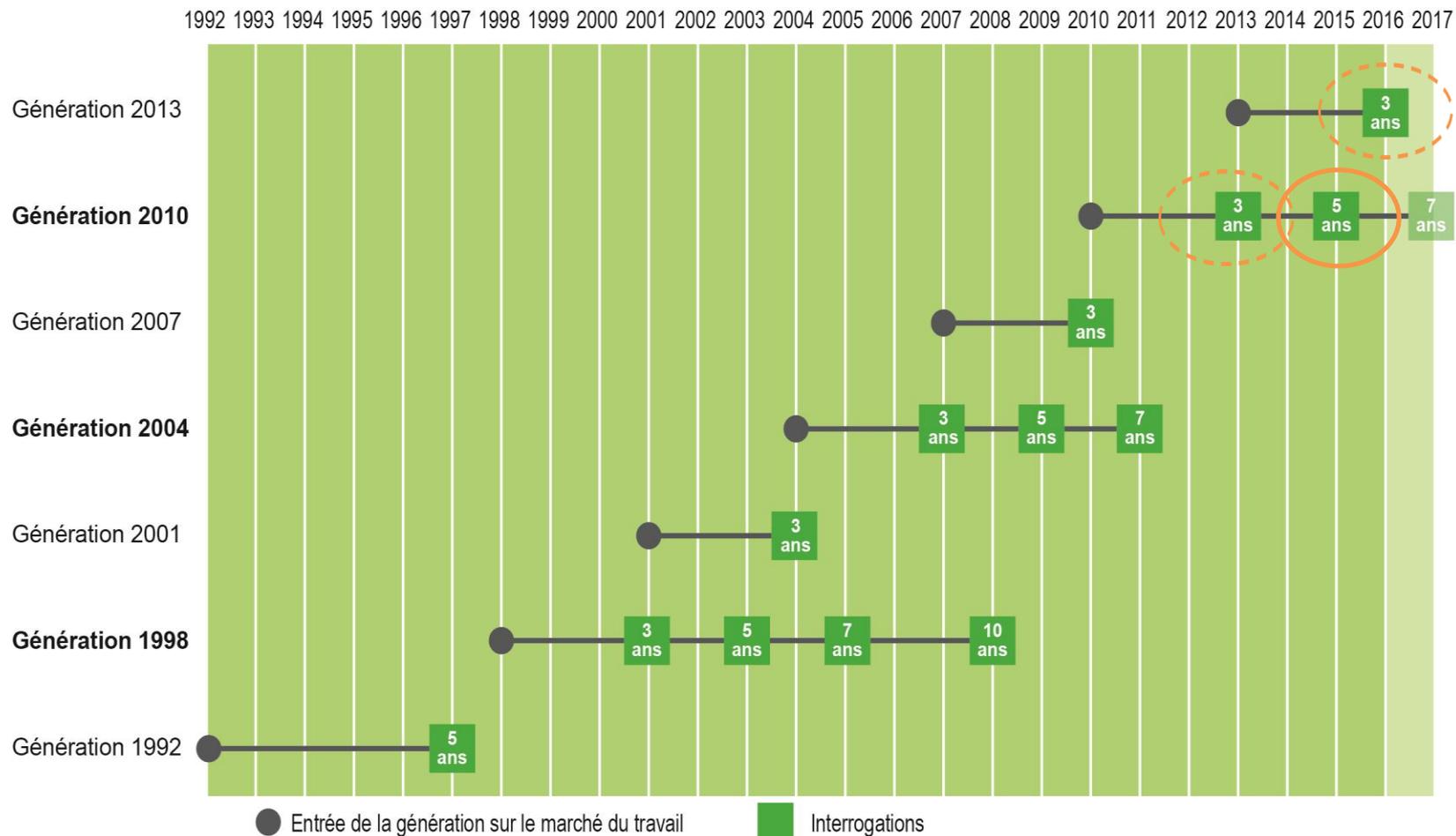
- Les enquêtes Génération
- Protocole de l'expérimentation multimode
- Mise en évidence d'un effet de mesure
- Plusieurs tentatives d'agrégation des données
- *Limites, Préconisations?*

■ ■ ■ ■ ■ • Les enquêtes « Génération »

- Enquête de la statistique publique sur l'insertion professionnelle des jeunes à l'issue du système éducatif
- Une « Génération » : l'ensemble des jeunes sortants du système scolaire une année donnée (tous niveaux d'études)
- Enquête téléphonique (CATI) à 3, 5 et 7 ans après la sortie qui retrace notamment le parcours sur le marché du travail au mois le mois
- Entre 20000 et 60000 répondants (à 3 ans) selon les Générations
- Nécessité de rénover le dispositif après 2017
 - une des orientations : passer à une collecte multimode internet (Cawi) et téléphone (Cati)

Le calendrier des enquêtes Génération

-- Expérimentations internet



■ ■ ■ ■ ■ • Protocoles de l'expérimentation internet

Enquête 2015 auprès de la Génération 2010 (environ 15')
(2e interrogation après celle de 2013)

- Les répondants à l'expérimentation 2013 en monomode internet sont réinterrogés en monomode internet
- Un sous-échantillon des répondants CATI 2013 (enquête principale EP) est interrogé selon deux protocoles
 - Un multimode « classique » (MM) *(avec ajout de deux questions pour mieux modéliser la probabilité de répondre sur internet)*
 - Un échantillon embarqué avec affectation aléatoire du mode de collecte (EE)
- Beaucoup d'efforts sur l'ergonomie du questionnaire web pour limiter les abandons en cours de collecte (calendrier, autocomplétion, ...)

■ ■ ■ ■ ■ • Protocole multimode « classique »

Quatre phases

0. envoi lettre avis (papier, mail ou les deux)
1. relance mail uniquement
2. Démarrage du plateau tel avec invitation à répondre par internet
3. Plateau tel avec choix du mode (internet ou tel) laissé à l'enquêté
4. Plateau tel pour faire l'enquête par tel

Tout au long de la collecte,

Relances mail programmées des non répondants (rythme hebdo environ)

Relances mail personnalisées des questionnaires commencés (puis relance tel)

Relance SMS ponctuelle

Quelques remarques :

- Possibilité de passer d'un mode à l'autre en cours d'enquête
- Au téléphone, alternance appels masqués/démasqués
- Plusieurs versions (affectation aléatoire) de la même question (NSP, QCM)

■ ■ ■ ■ ■ • Protocole échantillon embarqué

- Protocole de contact proche :
On contacte l'individu par tous les moyens (papier, mail, tel, sms) (pour éviter un biais de sélection lié au mode de contact)
- Si l'individu « accepte » de répondre au questionnaire, affectation aléatoire du mode de collecte
 - poursuite du questionnaire s'il est sur le bon mode (1 cas sur 2)
 - bascule « imposée » pour les autres (1 cas sur 2)

Objectif : gommer les effets de sélection liés au choix du mode

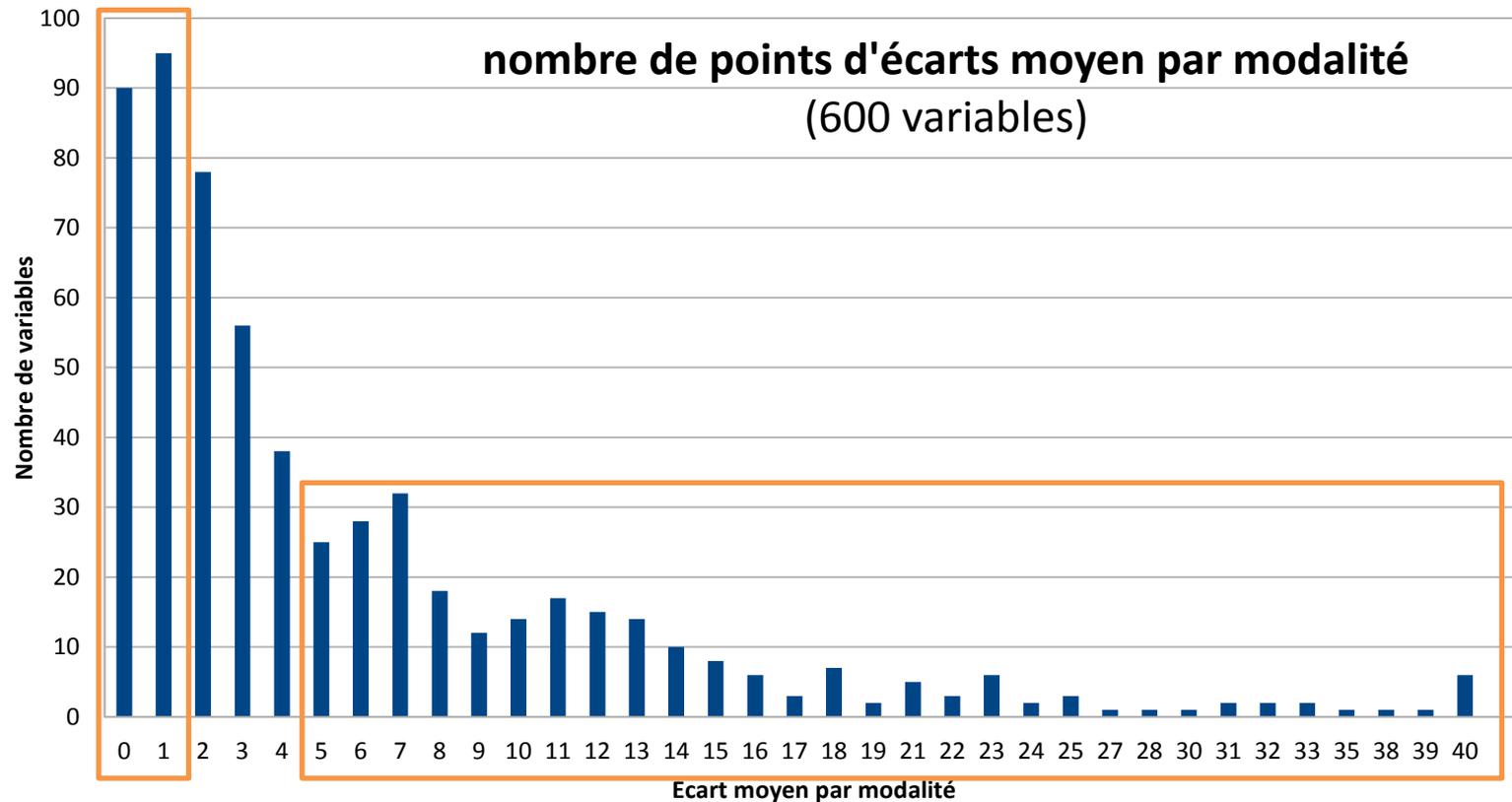
Taux de réponse par protocole et par mode

Taux de réponse

	Multimode classique	Echantillon embarqué	Enquête Principale (Monomode internet)
Enquêtes Réalisées CATI	23%	29%	60%
Enquêtes Réalisées CAWI	23%	13%	0%
TOTAL	46%	42%	60%
Volume fichier	4450	1000	23482

- Un protocole multimode qui a permis d'obtenir 23% de taux de réponse sur internet (*)
- Hypothèse : $TR\ MM \geq TR\ EP$ si protocole téléphonique identique (nombre de tentatives d'appels, dimensionnement)
- À deux réserves près :
 - l'impact potentiellement négatif de la multiplication des moyens de contacts...
 - Le refus déguisé (non rattrapable)
- L'échantillon embarqué : faible taux de réponse sur internet

Les écarts entre les collectes internet et téléphone



Source : Échantillon MM (tel recalé, web recalé)

- 30% de variables avec un écart moyen par modalité de 1 point ou moins
- 30% de variables avec un écart moyen par modalité de 5 points ou plus

■ ■ ■ ■ ● Les écarts entre les collectes internet et téléphone

- L'effet de mode se décompose en
 - un effet de sélection
 - un effet de mesure
- Effet de sélection → une aubaine, mais ...
- Effet de mesure
 - NSP, QCM, Parenthèses
 - Moindre implication de l'enquêté sur internet
 - Désirabilité sociale
 - ... ?
- Mais ces deux effets sont difficiles à démêler
 - Méthode de matching
 - Échantillon embarqué avec affectation aléatoire du mode de collecte

■ ■ ■ ■ ● Mise en évidence de l'effet de mesure par EE

Comment voyez-vous votre avenir professionnel?

	Modalités	Répondants EEtel	Répondants EEweb	test d'égalité des proportions ($\alpha=5\%$)
Génération 2010 à 3 ans (tél.)	vous êtes plutôt inquiet?	23,3	23,0	h0 acceptée
	vous êtes plutôt optimiste?	72,7	74,4	h0 acceptée
	Ne sait pas	4,0	2,5	h0 acceptée
Génération 2010 à 5 ans (tél. ou web)	vous êtes plutôt inquiet?	12,1	24,7	h0 rejetée
	vous êtes plutôt optimiste?	53,0	44,5	h0 rejetée
	Ne sait pas	34,9	30,8	h0 acceptée

Sur échantillon embarqué

- EE joue bien son rôle de contrôle de la sélection et donc permet de mettre en évidence un effet de mesure (ici, « désirabilité sociale »)
- La méthode de matching (avec utilisation de variables webographiques) donne des résultats cohérents (significatif ou pas, signe identique)
 - Evaluation par EE de l'effet de mesure sur « inquiet » : +12,6
 - Evaluation par matching de l'effet de mesure sur « inquiet » (ATE) : +7

■ ■ ■ ■ ● Agréger les données en présence d'effet de mesure ?

- Objectif :
 - Agréger en corrigeant le biais lié à l'effet de mesure
 - Fournir aux chercheurs un fichier de diffusion simple à utiliser
- Méthodes proposées
 - Ne rien faire ! (= agrégation simple)
 - Omettre un des deux modes de collecte
 - Imputer un des deux modes de collecte
 - Transformation de réponse d'un des deux modes
- Nécessité de définir un mode de référence (pour une variable donnée)
 - Mode historique pour limiter la rupture de série
 - Mode majoritaire
 - Mode jugé plus fidèle de la réalité (ex : autoadministré pour désirabilité sociale)

■ ■ ■ ■ ● Agréger les données en présence d'effet de mesure ?

Estimation de la part de jeunes déclarant être inquiet quant à leur avenir professionnel

	agrégation simple (multimode complet)			Omission d'un mode	
	tel+web	tel (sans recalage)	web (sans recalage)	tel recalé	web recalé
Génération 2010 à 3 ans (tél.)	26,9	25,6	28,3	25,2	27,2
Génération 2010 à 5 ans (tél. ou web)	15,7	11,7	20,0	12,5	17,9

	agrégation après imputation (tel=ref et web=imputé)			agrégation après imputation (web=ref et tel=imputé)			transformation de réponse (modèle sur web et tel=transformé) mesure = +12,56 (à partir EE)			transformation de réponse (modèle sur web et tel=transformé) mesure = +7,0 (à partir Matching)		
	tel+web	tel	web	tel+web	tel	web	tel+web	tel	web	tel+web	tel	web
Génération 2010 à 3 ans (tél.)	22,7	25,6	19,6	25,4	22,6	28,3						
Génération 2010 à 5 ans (tél. ou web)	10,8	11,7	9,8	23,0	25,9	20,0	21,9	23,8	20,0	19,3	18,7	20,0

■ ■ ■ ■ ● Limites

- Choix du mode de référence arbitraire
- Problème de robustesse dans la détection des effets de mesure
 - échantillon embarqué trop petit mais qui semble remplir sa fonction
 - méthode de matching : dépendance du modèle
- Omission, imputations
 - problème du volume de données omises et/ou imputées
 - problème de gestion des filtres
- Transformation de réponse
 - Nécessite une évaluation de l'effet de mesure
 - Qualité du modèle (*utilisation de variables non sujettes à un effet de mesure*)
 - Nécessite une (trop?) bonne capacité économétrique en production
- Ces méthodes conservent-elles les corrélations, les modélisations ?
(cf. présentation de Cyril Favre-Martinoz)

■ ■ ■ ■ ● Préconisations ?

- agrégation simple avec un système d'alerte sur les variables présentant un effet de mesure (par EE ou matching)
- en amont, supprimer les questions à problème (brutal) ou jouer sur l'ergonomie ou la formulation pour limiter certains effets de mesure (generalized design).
- ~~● en amont, ne poser les questions que dans un seul des deux modes (mais problème de sélection)~~
- ou en aval supprimer un des deux modes (coûteux de poser des questions dont on ne se sert pas, mais on le sait pas forcément a priori sauf dans le cas d'enquêtes répétées)
 - Dans ce cas, fournir trois jeux de pondérations (MM, 100%web, 100%tel)
 - Problème lorsque l'on croise deux variables...
- Si imputation ou transformation, doubler les variables (bruts et corrigés) et indicateurs de qualité sur chaque variable (part de données imputées)

PROTOCOLES MULTIMODES EFFET DE MESURE ET AGREGATION

Colloque francophone sur les sondages

13/10/2016

Christophe Barret - Christophe Dzikowski

DIAPOS COMPLEMENTAIRES

Colloque francophone sur les sondages

13/10/2016

Christophe Barret - Christophe Dzikowski

Exemple d'amélioration de l'ergonomie

CALENDRIER D'ACTIVITE

Merci de sélectionner vos différentes périodes d'activité par un cliquer-glisser

SITUATION		2013				2014								2015								
		SEPT.	OCT.	NOV.	DEC.	JANV.	FEVR.	MARS	AVRIL	MAI	JUIN	JUIL.	AOUT	SEPT.	OCT.	NOV.	DEC.	JANV.	FEVR.	MARS	AVRIL	
Périodes d'emploi	INTÉRIM																					
	ENTREPRISE ou administration																					
Périodes de non emploi	RECHERCHE D'EMPLOI																					
	REPRISE D'ÉTUDES à temps plein dans un établissement scolaire																					
	FORMATION autre qu'une reprise d'étude																					
	AUTRES SITUATIONS																					

MARS 2014 - JUILLET 2014
ENTREPRISE
CEREQ, 13002 MARSEILLE 2E ARRONDISSEMENT

VALIDATION DU CALENDRIER



TABLEAU RÉCAPITULATIF

Début	Fin	Situation	Entreprise	
OCTOBRE 2013	FEVRIER 2014	ENTREPRISE	INSEE	75014 PARIS 14E ARRONDISSEMENT
MARS 2014	JUILLET 2014	ENTREPRISE	CEREQ	13002 MARSEILLE 2E ARRONDISSEMENT
AOUT 2014	AVRIL 2015	RECHERCHE D'EMPLOI		

Exemple d'amélioration de l'ergonomie

Calendrier d'activité

A partir de **Avril 2016**, vous avez travaillé ou vous n'avez pas travaillé ?

Oui ⇒ Etait-ce en intérim ou en Entreprise ?

Non ⇒ Etiez-vous en recherche d'emploi, en reprise d'études à temps plein dans un établissement scolaire, en formation ou dans une autre situation ?

Jusqu'à quand ?

Merci de sélectionner vos différentes périodes d'activité en cliquant une première fois sur une case pour marquer le début d'une période et une deuxième fois pour marquer la fin de cette période.

SITUATION	2013								2014								2015								2016										
	juin	juil.	août	sept.	oct.	nov.	déc.	janv.	févr.	mars	avr.	mai	juin	juil.	août	sept.	oct.	nov.	déc.	janv.	févr.	mars	avr.	mai	juin	juil.	août	sept.	oct.	nov.	déc.	janv.	févr.	mars	
Périodes d'emploi	Intérim ⓘ																																		
	Entreprise ou administration ⓘ																																		
Périodes de non-emploi	Recherche d'emploi ⓘ																																		
	Reprise d'études à temps plein dans un établissement scolaire ⓘ																																		
	Formation autre qu'une reprise d'études ⓘ																																		
	Autres situations ⓘ																																		

Valider le calendrier

Récapitulatif de la saisie

	Début	Fin	Type de période	Nom	Commune
	juin 2013	février 2014	Recherche d'emploi		
	mars 2014	février 2015	Emploi - Entreprise	cereq	MARSEILLE 6E ARRONDISSEMENT
	mars 2015	mars 2016	Formation		MARSEILLE 6E ARRONDISSEMENT

Calendrier d'activité

A partir de **Avril 2016**, vous avez travaillé ou vous n'avez pas travaillé ?

Oui ⇒ Etait-ce en intérim ou en Entreprise ?

Non ⇒ Etiez-vous en recherche d'emploi, en reprise d'études à temps plein dans un établissement scolaire, en formation ou dans une autre situation ?

Jusqu'à quand ?

Merci de sélectionner vos différentes périodes d'activité en cliquant une première fois sur une case pour marquer le début d'une période et une deuxième fois pour marquer la fin de cette période.

SITUATION		2013							2014													
		juin	juil.	août	sept.	oct.	nov.	déc.	janv.	févr.	mars	avr.	mai	juin	juil.	août	sept.	oct.	nov.	déc.		
Périodes d'emploi	Intérim ⓘ																					
	Entreprise ou administration ⓘ																					
Périodes de non-emploi	Recherche d'emploi ⓘ																					
	Reprise d'études à temps plein dans un établissement scolaire ⓘ																					
	Formation autre qu'une reprise d'études ⓘ																					
	Autres situations ⓘ																					

Récapitulatif de la saisie

	Début	Fin	Type de période	Nom	Commune
	  juin 2013	février 2014	Recherche d'emploi		
	  mars 2014	février 2015	Emploi - Entreprise	cereq	MARSEILLE 6E ARRON
	  mars 2015	mars 2016	Formation		MARSEILLE 6E ARRON

Exemple d'amélioration de l'ergonomie



Preview :

Dans quelle commune habitez-vous lorsque vous êtes entré en 6ème ?

Consigne :
Veillez saisir les premières lettres de la **commune** dans le champ à droite du bouton **Libellé**, puis cliquer sur ce bouton **Libellé** pour faire apparaître les communes commençant par ces lettres.

Attention : Veuillez ne pas saisir les accents ni les articles, par exemple pour "La TRINITÉ", merci de saisir "TRINITE". Pour les communes concernées, veuillez saisir SAINT et non pas ST.

Toutes les communes de France sont répertoriées (villages, villes, ... mais pas les lieux dits, les départements, les régions, les quartiers...). Merci de **vérifier les 2 chiffres du département**. Attention, le code indiqué correspond au **code INSEE** et non au code postal.

Pour l'étranger (et les TOM), veuillez chercher **le pays** (et non pas la ville).

3) Menu déroulant
proposé au format
« brut » avec
difficultés dans la
recherche

	Code	Libellé
<input type="radio"/>	40016	AUBAGNAN
<input type="radio"/>	13005	AUBAGNE
<input type="radio"/>	21030	AUBAINE
<input type="radio"/>	30019	AUBAIS
<input type="radio"/>	65044	AUBAREDE
<input type="radio"/>	24014	AUBAS
<input type="radio"/>	43011	AUBAZAT
<input type="radio"/>	19013	AUBAZINES

Exemple d'amélioration de l'ergonomie

 **Enquête 2015** auprès de la **Génération 2010**

Centre d'études et de recherches sur les qualifications

MINISTÈRE DE L'ÉDUCATION NATIONALE, DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

MINISTÈRE DU TRAVAIL, DE L'EMPLOI, DE LA FORMATION PROFESSIONNELLE ET DU DIALOGUE SOCIAL

Progression  93%

 Quitter et reprendre plus tard

Période d'emploi actuelle

Pouvez-vous me dire très précisément quelle est votre profession actuelle ?

Si Profession non trouvée : recherchez AUTRE dans le menu

banq

- AGENT COMMERCIAL DE BANQUE
- CHARGE DE CLIENTELE BANQUES
- EMPLOYE DE BANQUE**
- GRADE DE BANQUE

 Précédent  Suivant 