L'ANALYSE DE LA NON REPONSE POUR ELIPA 2

Florent Domergue (*), Virginie Jourdan (*)

(*)Département des statistiques, des études et de la documentation, service statistique du secrétariat général à l'immigration et à l'intégration

Document en version provisoire

Introduction

ELIPA signifie "enquête longitudinale sur l'intégration des primo-arrrivants". Une première interrogation a eu lieu au printemps 2010. Le questionnaire passe en moyenne en une heure et 6 107 personnes ont été interrogées. Ces personnes devaient être retrouvées puis interrogées au printemps 2011 pour une seconde vague, avec un questionnaire toujours calibré pour passer en une heure, mais évidemment construit pour compléter l'information obtenue en vague 1.
Le taux de réponse global est de 77,88 %. A rebours, le taux d'échec est de 22,12 %. Celui-ci se décompose en 15,51 % d'échecs dus à diverses impossibilités de contacter les enquêtés (déménagements, mauvais suivi, …) et 6,62% de refus d'enquête. On a donc considéré que les sorties de champ n'étaient pas discernables d'autres comportements conduisant à l'échec du contact. Les deux chiffres sont tout à fait dans la norme d'enquêtes de ce type, soit comparables aux résultats d'autres enquêtes du système statistique publique ou les enquêtes modèles, ELIC par exemple (L'ELIPA canadienne). Il faut avoir à l'esprit qu'ELIPA n'est pas obligatoire, notamment. On peut rappeler encore le taux d'échec d'une enquête assez semblable (PPM, parcours et profils des migrants, en deux vagues, 2006 puis 2007), de l'ordre de 38 %. L'analyse quantitative de la non-réponse n'a donc pas pour objectif premier d'identifier des dysfonctionnements de collecte.
Cette analyse présente au moins un triple intérêt :

- Au contraire d'enquêtes en une vague, où l'on ne dispose que d'assez peu d'information pour l'ensemble des échantillonnés, on dispose de toute l'information de la première vague. On peut donc comparer l'impact de certaines caractéristiques usuellement connues (habitat, position par rapport à l'emploi, statut matrimonial, ….) avec celui d'un grand ensemble d'autres caractéristiques ;
- Les thèmes d'ELIPA sont assez spécifiques. Outre une partie de questionnaire sur les caractéristiques sociodémographiques, il y a des parties sur le parcours des primo-arrivants et une petite partie sur "l'attitude de l'enquêté" vis-à-vis de l'enquête. Précisément, ces caractéristiques s'avèrent importantes comme déterminants de la réponse ;
- Une question usuelle est de savoir en quoi la non réponse peut fausser les estimations, du fait de l'interaction entre le thème d'enquête et la propension des personnes à y participer. La richesse et la variété des données d'ELIPA permettront un bon éclairage de cette question. Ce dernier point n'est toutefois pas abordé dans ce document.

Les données de base sont les suivantes :

Tableau 1 : effectifs et taux pour l'échec du contact et le refus.

	Non contact	Refus	Répondant	Total
Effectif	947	404	4 756	6 107
%tage	15,51	6,62	77,88	*100*

1. La méthode d'analyse (synthèse)

Première étape : analyses descriptives pour toutes les variables et modalités.

La variable "coderes" est créée, avec les trois modalités : non-contact, refus, répondant. Il s'agit alors de croiser cette variable avec toutes les variables et modalités de la vague 1. Le détail des analyses ne sera pas reproduit, il y a 500 variables environ et pour certaines de très nombreuses modalités (nationalités par exemple). Une exception, toutefois, les variables du module "test linguistique" n'ont pas été mobilisée. Cependant, sur le thème des compétences en langue française, on dispose de plusieurs variables. Cette phase fait apparaître les modalités et variables pour lesquelles les taux (non contact / refus) sont différents des taux moyens. Elle fait également apparaître des regroupements de modalités possibles et les modalités pour lesquelles les effectifs sont insuffisants pour peser dans une modélisation par régression logistique.

Deuxième étape : recodages.

L'analyse descriptive conduit à détecter des modalités déterminantes. On crée ainsi des variables dérivées si possibles binaires (regroupement des modalités particulières par rapport aux échecs du contact ou aux refus, *versus* les autres modalités). Dans certains cas, une variable peut être créée à partir de plusieurs variables de la vague 1 ou de variables dérivées. Par exemple, on a créé une variable synthétique des raisons pour venir en France (il y en a 10), à partir de méthodes d'analyse de données (classifications ascendantes hiérarchiques pour créer des modalités pertinentes). On a encore utilisé d'autres travaux (exploitations de la vague 1) pour reprendre des variables intéressantes dans ces analyses et assez synthétiques.

Environ 130 variables ont ainsi été créées. A nouveau, on fait des analyses descriptives en croisant les nouvelles variables avec "coderes". Par construction, chacune d'entre elles fait apparaître des groupes de personnes dont les comportements, soit pour le contact, soit pour l'acceptation de l'enquête, sont différents des taux moyens et d'effectifs assez importants. Ces résultats ne sont pas détaillés ici. Une synthèse est proposée en partie 4.

Les variables construites sont issues de tous les modules du questionnaire, et relèvent ainsi des thèmes de l'enquête :

- informations sociodémographiques usuelles (sexe, âge, nationalité, …) ;
- informations sociodémographiques spécifiques (durée de présence en France, …) ;
- informations concernant l'emploi (avant migration, au moment de la vague 1) et le logement (type, niveau de confort, y compris informations sur les personnes vivant dans le même logement que l'enquêté, …) ;
- informations sur les motifs de la migration ;
- informations sur les conditions de l'arrivée en France, puis du séjour durant les premiers temps (notamment passages hors légalité) ;
- informations sur les intentions (rester en France, demander la nationalité, …) ;
- informations issues des bilans de la demi-journée d'accueil (pour signature du contrat d'accueil et d'intégration) ;
- informations concernant les qualifications et le niveau de français ;
- informations concernant les conditions de vie (niveaux de revenus, difficultés financières, ….) ;
- informations concernant la famille et les circonstances du mariage ;
- informations concernant le réseau social, notamment sa composition en termes d'origines nationales des membres de ce réseau ;
- informations concernant les pratiques culturelles de l'enquêté ;
- réseau social, notamment sa composition en termes d'origines nationales des membres de ce réseau.

A cela s'ajoutent quelques variables "de gestion" : période de collecte de la vague 1, lieu d'enquête (région et dans / hors domicile), sentiment de l'enquêteur durant la vague 1, acceptation du principe d'une post-enquête ….

Troisième étape : analyses logistiques partielles.

Il s'agit de chercher à faire une synthèse par thème. Pour chaque thème, une analyse logistique ne prenant que quelques variables (de 5 à 20), toutes du thème donné, est faite.
La modélisation est faite séparément pour le non contact et le refus. Cela permet de faire quelques recodages supplémentaires destinés à faciliter la compréhension des ressorts du non contact et du refus et à améliorer certaines variables.

2. Le redressement, principaux résultats.

2.1. La modélisation de l'échec du contact

Tableau 2 : variables / modalités « explicatives" pour la réussite du contact

Variable	Référence	Modalité	Odd-ratio	*IC (à 5%)*
Age (3)	Personnes âgées	Jeunes	**0,49**	*0,37-0,66*
		Moyennement jeunes	**0,69**	*0,55-0,86*
Nationalité	Groupe de "bons répondants" (2)	Groupe 1 (1)	**0,72**	*0,58-0,89*
		Cas général	**0,73**	*0,61-0,88*
Durée de présence en France	Toutes sauf…	2-3 ans	**0,76**	*0,61-0,95*
Venu parce que beaucoup d'autres le font	Très important	Refus de répondre	**1,25**	*1,01-1,53*
		Autres modalités	**1,34**	*1,11-1,61*
Bilan du CAI	Mitigé	Bon	**1,22**	*1,01-1,48*
Utilité bilan "OP" (4)	Non	Oui	**0,80**	*0,66-0,96*
Equivalence diplôme	Pas demandée	Autres modalités	**1,36**	*1,13-1,64*
Permis	Pas validé	Validé	**1,35**	*1,10-1,64*
		Pas de permis	**1,14**	*0,96-1,36*
Crédits	Pas nécessaires	Autres modalités	**0,57**	*0,43-0,76*
Bénéfices d'aide (5)	Autres modalités	Via organismes mal identifiés	**0,68**	*0,50-0,93*
		Via organismes bien identifiés	**1,39**	*0,83-2,31*
Couverture maladie	Complémentaire	Ni CMU ni CMC	**0,78**	*0,64-0,94*
Région d'enquête	PACA et IdF	Rhône-Alpe-Alsace	**1,48**	*1,19-1,82*
Période d'enquête (6)	Fin	Début	**1,34**	*1,14-1,58*
Langue d'enquête	Français	Pas en français	**0,71**	*0,60-0,84*
Formation linguistique (lors demi-journée)	Oui et s'en souvient	inutile	**1,11**	*0,89-1,38*
		Oui et ne s'en souvient pas	**0,71**	*0,54-0,92*
Habitat	HLM ou propriétaire	En foyer ou chez des amis	**0,70**	*0,57-0,86*
		Autres modalités		
Réseau : rencontres d'amis …	Durant formations	Ailleurs	**0,80**	*0,68-0,94*
Réseau : rencontres d'amis …	Via l'école (des enfants)	Pas d'enfants en âge d'aller à l'école	**0,75**	*0,62-0,92*
Relations au pays d'origine	Ne plus en avoir	En avoir	**0,61**	*0,40-0,94*
Citer un titre de presse (en France)	Capable	Incapable	**0,83**	*0,69-0,99*

(1) Chine-Turquie-Mali-Tunisie
(2) Cameroun, Congo, Cote d'Ivoire, Madagascar, Maroc, Russie, Sri-Lanka.
(3) Voir partie 4-1.

(4) Considérer que le bilan concernant une réorientation professionnelle est utile ou non.
(5) Aides financières : organismes connus : aides publiques, associations reconnues.
(6) Il s'agit de l'enquête de première vague

On utilise les variables créées pour modéliser l'échec du contact pour "expliquer" celui-ci. Les (environ) 130 variables construites sont introduites, quelques variantes sont encore testées. Le meilleur modèle obtenu fait intervenir 20 variables (donc un peu moins d'une sur six). Celles-ci peuvent s'analyser en :

- trois variables socio-démographiques, plus la variable « durée de présence en France » ;
- trois variables en relation avec les circonstances de la demi-journée d'accueil, plus une variable sur les motifs de migrations. Le sentiment général peut être qu'un manque d'implication dans le projet migratoire se ressent (négativement) sur le contact ;
- quatre variables qui traduisent encore le défaut d'investissement dans la vie au pays d'accueil, voire dans la vie sociale en général et qui se ressentent encore négativement sur le contact ;
- trois variables qui associent mauvaises conditions (plutôt financières) et difficulté du contact ;
- trois variables « de gestion » (dont une associée au niveau linguistique).

Cela permet encore d'obtenir pour chaque enquêté de vague 1 une probabilité (et un poids) d'être contacté. Dans le calcul des pondérations, le poids de vague 1 sera multiplié par ce poids.

2.2. La modélisation du refus d'enquête

On utilise ensuite les variables créées pour modéliser le refus pour "expliquer" celui-ci. Bien évidemment, si certains recodages sont spécifiques pour le refus, ce sont ces derniers qui sont utilisés. Le meilleur modèle obtenu fait intervenir 21 variables (donc un peu moins d'une sur six). Celles-ci peuvent s'analyser en :

- quatre variables socio-démographiques, plus des variables « diplômes » et « PCS » ;
- trois variables en relation avec les circonstances de la demi-journée d'accueil, plus une variable sur les motifs de migrations et une variable sur le parcours administratif. Le sentiment général est assez semblable à ce qui a été relevé pour le contact ;
- trois variables qui traduisent aussi le défaut d'investissement dans la vie au pays d'accueil, voire dans la vie sociale en général et qui se ressentent encore négativement sur le contact ;
- le capital culturel intervient mais de façon peu facilement interprétable : outre deux variables citées par ailleurs qui vont plutôt dans le sens d'une corrélation positive entre niveau global ou aisance en français et acceptation de l'enquête, ce sont les personnes reconnaissant des difficultés pour téléphoner et écrire une lettre qui ont un taux de refus particulièrement bas ;
- quatre variables « de gestion » (dont une associée au niveau linguistique), les deux variables « sentiment de déroulement de l'entretien selon l'enquêteur » et refus (voire refus de répondre à la question) de post-enquête interviennent ici.

Cela permet aussi d'obtenir pour chaque enquêté de vague 1 contacté lors de la vague 2 une nouvelle probabilité (et un poids) d'accepter l'enquête. Le poids précédent sera multiplié par ce poids.

Les odds-ratios (ou leurs inverses) sont plus élevés, quasiment tous entre 1,5 et 2, soit :

Faire tableau plus joli et plus précis sur modèle tableau 2

Tableau 3 : variables / modalités « explicatives, odds ratio »

Variable	Effet négatif sur l'acceptation dû à
Sexe	Etre une femme

Nationalité	Chine-Turquie-Mali-Tunisie (très fort)
	Tous sauf 7 pays
Enfants	En avoir un ou deux (donc ni aucun, ni beaucoup)
Choix de venir en France	Refuser de répondre à l'item "parce que liens historiques" (très fort)
Passage en préfecture	Etre critique (juste significatif)
Bilan CAI	Pas intéressé par la documentation
CAI utile pour nationalité	Penser que non
Utilité form. "Vivre en France"	Toutes modalités sauf "oui mais"
Diplômes	Pas ou niveau certificat
Période enquête	Fin de période (vague 1)
Langue d'enquête	Autre que Français (et Russe) (vague 1)
Déroulement d'enquête	Autre que bon accueil selon enquêteur (vague 1)
Post-enquêtes	Refus ou refus de répondre
PCS	Autres que employés, ouvriers qualifiés
Travail	Ne pas en avoir changé depuis l'arrivée
Aisance Français (1)	Autre que "difficultés au téléphone"
Aisance Français (2)	Autre que "difficultés pour écrire une lettre"
Aisance Français	Dans les deux cas précédents, c'est en fait les personnes en difficulté qui ont des taux de réponses très élevés.
Logement	Autre que ne pas avoir demandé de HLM
Réseau social	Très centré sur habitudes et personnes du pays d'origine
Recours au réseau	Autre que recours modéré aide communautaire
Vacances au pays	Oui

3. L'analyse descriptive et les recodages : méthodes.

3.1. Première étape : analyses descriptives

La variable "coderes" est créée, avec les trois modalités : non-contact, refus, répondant. Il s'agit alors de croiser cette variable avec toutes les variables et modalités de la vague 1.

Pour faciliter l'analyse, on calcule si nécessaire un indicateur qui tient compte de l'anormalité du taux (de non contact ou de refus) et de l'effectif de la modalité. Il faut en effet que l'écart soit suffisant pour un nombre suffisant d'enquêtés pour que l'effet résultant intervienne significativement. Cet indicateur est simple : on calcule d'abord le ratio des taux : taux pour la modalité / taux pour l'ensemble. On arrange le résultat sous forme d'écart (valeur absolue) à 1. On multiplie le résultat par la proportion de personnes dans la modalité (puis par 100).

Tableaux 4 : exemple de calcul de l'indicateur

Variable "thlnb" : il s'agit du nombre d'habitants du logement, y compris l'enquêté, ce nombre est donc compris entre 1 (enquêté vivant seul) et 12.

Exemple de calcul :

Si thlnb = 1 (enquêté vivant seul), l'effectif est de 567, soit 428 répondants, 103 échecs du contact et 36 refus d'enquête (tableau 4-a).
Le taux de non contact est de 18,2 % (le taux de refus est de 6,3 %) (tableau 4-b).
On calcule 18,2 / 15, 5 (= 1,17), d'où un écart absolu (multiplié par 100) de 17 (tableau 4-c).
(15,5 est le taux d'échec du contact moyen)
De même, si thlnb = 5, le taux de non contact est de 14,5 %, 14,5 / 15,5 = 0,93, d'où un écart absolu multiplié par 100 de 7.
Indices : l'écart précédent est multiplié par l'effectif de la modalité et divisé par 6107 ("coefficient"). Dans cet exemple, 9,3 % des enquêtés sont dans la modalité 1, d'où un indice de 17*0,093 soit environ 1,6.

Tableau 4-a : effectifs

thlnb	NC	refus	répondant	Total
1	103	36	428	567
2	339	163	1666	2 168
3	205	74	1041	1 320
4	140	53	723	916
5	82	43	442	567
6	48	17	235	300
7	19	12	123	154
8	6	4	52	62
9	4		27	31
10		1	11	12
11	1	1	4	6
12			4	4
total	947	404	4 756	6 107

Tableau 4-b: taux

thlnb	NC	refus	répondant	Total
1	18,2	6,3	75,5	*100*
2	15,6	7,5	76,8	*100*
3	15,5	5,6	78,9	*100*
4	15,3	5,8	78,9	*100*
5	14,5	7,6	78,0	*100*
6	16,0	5,7	78,3	*100*
7	12,3	7,8	79,9	*100*
8	9,7	6,5	83,9	*100*
9	12,9	0,0	87,1	*100*
10	0,0	8,3	91,7	*100*
11	16,7	16,7	66,7	*100*
12	0,0	0,0	100,0	*100*
total	15,5	6,6	77,9	*100*

Tableau 4-c : écarts des taux

	Ecarts				Indices				coefficient
1	17,1	4,0	3,1		1,6	0,4	0,3		*0,092844*
2	0,8	13,7	1,3		0,3	4,8	0,5		*0,355002*
3	0,2	15,3	1,3		0,0	3,3	0,3		*0,216145*
4	1,4	12,5	1,4		0,2	1,9	0,2		*0,149992*
5	6,7	14,6	0,1		0,6	1,4	0,0		*0,092844*

6	3,2	14,3	0,6		0,2	0,7	0,0		*0,049124*
7	20,4	17,8	2,6		0,5	0,4	0,1		*0,025217*
8	37,6	2,5	7,7		0,4	0,0	0,1		*0,010152*
9	16,8	100,0	11,8		0,1	0,5	0,1		*0,005076*
10	100,0	26,0	17,7		0,2	0,1	0,0		*0,001965*
11	7,5	151,9	14,4		0,0	0,1	0,0		*0,000982*
12	100,0	100,0	28,4		0,1	0,1	0,0		*0,000655*

3.2. Deuxième étape : recodages.

Des tableaux comme le tableau 4 permettent de détecter des modalités déterminantes : si l'indicateur est en-dessous de 1, il n'est pas utile de poursuivre l'analyse. Entre 1 et 2, ça se discute, au-delà de 2, on tente une recodification pour faire ressortir une ou plusieurs modalités.
Dans l'exemple du nombre d'habitants du logement, le recodage est différent suivant que l'on s'intéresse au "non-contact" ou au refus. Une variable "thlnb_nc", qui oppose les personnes seules (thlnb=1) à toutes les autres, est créée. De même, une variable "thlnb_ref" est créée en trois modalités : thlnb=3, thlnb=2, autres modalités (tableau 5). Dans ce dernier cas, l'analyse du tableau 4-b suggère que d'autres regroupements de modalités auraient été possibles. Ils ont été testés (en partie), au final, on choisit si possible un recodage simple et des regroupements de modalités ayant une logique, mais ce n'est pas toujours le cas (regroupements de nationalités par exemple).
Dans certains cas, une variable peut être créée à partir de plusieurs variables de la vague 1 ou de variables dérivées. Par exemple, dans le module A, on a créé une variable synthétique des raisons pour venir en France (il y en a 10), à partir de méthodes d'analyse de données (classifications ascendantes hiérarchiques pour créer des modalités pertinentes).

Dans l'exemple du nombre d'habitants du logement, le recodage est différent suivant que l'on s'intéresse au "non-contact" ou au refus.

Tableau 5 : taux pour les variables dérivées de "thlnb"

thlnb_nc	NC	refus	répondant	Total
1	18,2	6,3	75,5	*100*
2	15,2	6,6	78,1	*100*
total	15,5	6,6	77,9	*100*

thlnb_ref	NC	refus	répondant	Total
1	15,4	6,4	78,2	*100*
2	15,6	7,5	76,8	*100*
3	15,5	5,6	78,9	*100*
total	15,5	6,6	77,9	*100*

3.3. Troisième étape : analyses logistiques partielles.

Il s'agit de chercher à faire une synthèse par module. L'exemple qui suit est l'analyse faite à partir du premier module, soit le module "ego". Une vingtaine de variables ont été examinées. Après analyses comme décrites précédemment, on recode quelques variables car :

- Le genre ressort (c'est-à-dire que l'indicateur prend des valeurs suffisamment élevées pour que l'on se demande s'il va intervenir significativement dans les modèles). Il joue de façon opposée sur le non contact (plus élevé chez les hommes) et le refus (plus élevé chez les femmes).
- Le nombre d'habitants dans le logement ressort, les personnes seules étant plus difficiles à contacter, les couples refusant davantage (tableau 4).
- L'âge joue un peu différemment entre non contact et refus, mais ce sont essentiellement les jeunes qui répondent moins.

- L'analyse par nationalité (le pays de naissance donne des résultats identiques) est reprise de travaux préliminaires avec quatre groupes :
    - Un groupe de "bon répondants" : Cameroun, Congo, Cote d'Ivoire, Madagascar, Maroc, Russie, Sri-Lanka ;
    - Un groupe de mauvais répondants, et surtout de "refusants" : Chine et Turquie
    - Un groupe de mauvais répondants, et surtout de "impossible à contacter" : Mali et Tunisie.
    - Le groupe des "autres".
- Les variables de durée de présence / âge à l'arrivée sont étudiées via la durée de présence, avec une durée (deux-trois ans) moins favorable à la réponse.
- La variable d'activité fait apparaître un comportement spécifique des personnes en emploi, plus difficiles à contacter, et les élèves / étudiants[1], qui refusent davantage. Des analyses préliminaires laissent prévoir que cette variable est colinéaire au genre et à l'âge.
- La variable "stabilité du logement" est binarisée. Les personnes déclarant un logement durable ou stable sont plus faciles à contacter.

On procède ensuite à quelques analyses logistiques, en introduisant les variables si possible recodées, avec différents essais.

Au final, on fait apparaître deux modèles (pour le non contact et pour le refus) qui paraissent les meilleurs possibles pour prendre en compte l'information d'un module donné.

Par exemple : module "ego"

Tableau 6 : modèle explicatif du succès du contact, explicatives issues du premier module (“ego”)

Variable	**Référence**	**Modalité**	**Odd-ratio**	***IC (à 5%)***
Sexe	Femme	Homme	**0,86**	*0,74-0,99*
Age (3)	Ages intermédiaires	Jeunes	**0,76**	*0,62-0,94*
		Plus âgés	**1,35**	*1,08-1,67*
Nationalité	Groupe général	(1)	**0,85**	*0,72-1,00*
		(2)	**1,30**	*1,08-1,55*
Durée de présence en France	Toutes sauf…	2-3 ans	**0,76**	*0,62-0,94*
Logement	stable	incertain	**0,84**	*0,72-0,98*
Variables ne ressortant pas significativement (mais entrées dans le modèle) : - Nombre d’habitants dans le logement - Position sur le marché du travail au moment de l’enquête				

(1) Chine-Turquie-Mali-Tunisie
(2) Cameroun, Congo, Cote d'Ivoire, Madagascar, Maroc, Russie, Sri-Lanka.
(3) Voir partie 4-1.

[1] Les étudiants primo-arrivants ne sont pas dans le champ de l'enquête (en 2009, environ 55 000 étudiants sont venus des pays tiers) Toutefois, une personne peut se déclarer étudiante au moment de l'enquête. Voir encore [1].

Tableau 7 : modèle explicatif de l'acceptation de l'enquête, explicatives issues du premier module ("ego")

Variable	Référence	Modalité	Odd-ratio	*IC (à 5%)*
Sexe	Femme	Homme	**1,25**	*1,01-1,54*
Age (3)	Ages intermédiaires	Jeunes	**0,61**	*0,40-0,93*
		Plus âgés	**1,17**	*0,83-1,64*
Nationalité	Groupe général	(1)	**0,59**	*0,47-0,75*
		(2)	**1,61**	*1,20-2,15*
Nombre d'habitants dans le logement	Toutes sauf…	nombreux	**0,76**	*0,62-0,94*
Variables ne ressortant pas significativement (mais entrées dans le modèle) : - Durée de présence en France - Logement (stable /incertain) - Position sur le marché du travail au moment de l'enquête				

4. Les résultats intermédiaires

On résume ici les variables / modalités qui sont spécifiques soit de l'échec du contact, soit du refus d'enquête. Seules les variables recodées sont présentées. On les regroupe par module d'enquête, ou si les thèmes de ces modules le permettent, par groupe de modules.

4.1. Module "Ego"

Les informations sociodémographiques usuelles font l'objet de ce module.

L'analyse démarre de façon intéressante, la première variable testée, le genre, montre des effets assez marqués et différents :

Tableau 4-1-a : effets du genre sur la non-réponse

sexe	NC	refus	répondant	Total
Hommes	16,7	6,0	77,2	*100*
Femmes	14,3	7,2	78,5	*100*
total	15,5	6,6	77,9	*100*

Hommes : contact difficile – refus faible
Femmes : contact facile – refus fort

Nombre de personnes dans le logement (voir partie 3)

Personne seule : difficile à contacter, acceptation normale
Deux personnes (couples, autres situations)Couples : contact normal, refus plus fréquent
Trois ou quatres personnes (pas forcément de la même famille) : contact normal, acceptation plus fréquente.
Plus de 5 personnes : situations normales.

Age : deux regroupements, l'un pour le non contact, l'autre pour le refus (l'enquête ne porte que sur des personnes de 18 ans et plus). Cependant, les logiques sont similaires : non contact et refus pour les plus jeunes.

Tableau 4-1-b : effets de l'âge sur la non-réponse

Age_nc		NC	refus	répondant
19 à 23 ans	1	19,1	8,4	72,5
Autres âges	2	15,7	6,5	77,8
Plus de 41 ans	3	12,1	5,9	82,1
	total	15,5	6,6	77,9

Age_ref		NC	refus	répondant
18-21 ans	1	16,5	10,8	72,7
Autres âges	2	15,8	6,1	78,1
Plus de 44 ans	3	12,5	6,8	80,6
	total	15,5	6,6	77,9

Tableau 4-1-c : effets des nationalités sur la non-réponse
Les nationalités

	NC	refus	répondant
Chine et Turquie	17,2	12,0	70,8
Mali et Tunisie	19,0	8,2	72,9
Autres	15,7	6,4	77,9
Cameroun, Congo, Cote d'Ivoire, Madagascar, Maroc, Russie, Sri-Lanka,	12,7	4,1	83,2
total	15,5	6,6	77,9

On remarque des effets marqués, mais il n'y a pas de logique géographique.

La durée de présence en France : Les primo-arrivants peuvent être présents en France depuis plusieurs années. Cette durée avant obtention du titre de séjour est notamment liée au motif d'obtention. En l'occurrence la durée singulière correspond assez bien au motif "réfugié" (ou "famille de réfugié")

Tableau 4-1-d : effets de la durée de présence sur la non-réponse

	NC	refus	répondant
Présents en France depuis 2 ou 3 ans	18,9	6,2	74,9
Autres durées de présence	15,1	6,7	78,2
total	15,5	6,6	77,9

Tableau 4-1-e : effets de la situation par rapport à l'emploi (au moment de l'enquête) sur la non-réponse

	NC	refus	répondant
En emploi	16,1	6,2	77,7
Autres situations	15,0	7,0	78,0
total	15,5	6,6	77,9

Cette variable est en fait peu pertinente, elle est largement colinéaire au sexe

Tableau 4-1-f : effets de caractéristiques de logement sur la non-réponse

	NC	refus	répondant
Logement stable	14,8	6,7	78,5
Logement "précaire"	17,4	6,4	76,2
total	15,5	6,6	77,9

Le logement : caractère stable ou incertain : un petit effet guère surprenant de la précarité sur le non contact.

4.2. Modules sur les autres habitants du logement

Il s'agit des conjoints, enfants voire autres cohabitants.

La présence, le sexe et la localisation du conjoint : la variable construite résume des informations plus finement disponibles par ailleurs, sur le logement et le sexe notamment.

Tableau 4-2-a : effets « de conjoints » sur la non-réponse

	NC	refus	répondant
Pas de conjoint	17,0	8,0	75,0
Un conjoint, mais hors logement	17,7	6,3	76,0
Le conjoint vit dans le logement, c'est un homme (la personne est donc une femme)	13,8	7,1	79,1
Le conjoint vit dans le logement, c'est une femme (la personne est donc un homme)	15,8	5,3	78,9
total	15,5	6,6	77,9

La présence d'enfants joue à peu près dans les mêmes conditions pour non contact et refus, bien que la différence se fasse à partir du deuxième pour la difficulté de contact, du troisième pour le refus. On notera encore que ces résultats sont cohérents (et colinéaires) aux résultats issus de la variable donnant le nombre d'habitants. La variable indiquant s'il y a d'autres cohabitants est par contre intéressante car elle apporte une information originale et pertinente. Le fait que la personne interrogée ait des cohabitants autres signifie le plus souvent qu'elle habite chez des parents ou amis (la belle-famille éventuellement). La difficulté de contact est très nette dans ce cas.

Tableau 4-2-b : effets «d'enfants» et d'autres cohabitants sur la non-réponse

Enfants (non contact)	NC	refus	répondant
Plusieurs	16,2	6,8	77,0
Aucun ou un seul	12,4	5,8	81,9
total	15,5	6,6	77,9

Enfants (refus)	NC	refus	répondant
Aucun ou plus de 3	16,0	7,4	76,6
Un ou deux	14,6	4,9	80,5
total	15,5	6,6	77,9

Autres cohabitants	NC	refus	répondant
Oui	19,8	7,6	72,5
Non	15,0	6,5	78,5
total	15,5	6,6	77,9

4.3. Modules sur les données de la base et de gestion d'enquête.

Un certain nombre de variables de ces modules s'avèrent particulièrement pertinentes.

Ainsi la difficulté de contact plus encore que le refus est plus grande si la personne a été enquêtée en vague 1 en fin de période (les cinq dernières semaines sur les trois mois de collecte). On peut assez naturellement interpréter cela comme une difficulté à joindre les personnes, se traduisant par un contact plus tardif : plus de tentatives, les enquêteurs cherchent d'abord à réaliser les enquêtes "faciles", …. La question sur les post-enquêtes donne aussi des indications très nettes : un refus de répondre s'associe à une difficulté de contact et un refus élevé, le refus franc manifesté lors de la

vague 1 permet de prévoir un refus de la vague 2 (mais l'effet sur la difficulté de contact est non significatif). Le sentiment général de l'enquêteur sur le déroulement de l'enquête est encore un bon indicateur tant pour l'échec du contact que pour le refus.

Tableau 4-3-a : effets de période d'enquête, comportement pendant l'enquête, ….sur la non-réponse

Période d'enquête	NC	refus	répondant
Début et milieu	14,6	6,4	79,0
Fin (juin 2010)	18,2	7,3	74,5
total	15,5	6,6	77,9

Post enquêtes	NC	refus	répondant
Oui	15,2	4,8	80,0
Non	15,1	8,8	76,1
Refus de répondre	19,2	8,3	72,4
total	15,5	6,6	77,9

Sentiment de l'enquêteur	NC	refus	répondant
Autres que bon	19,0	11,0	70,0
Bon	15,3	6,3	78,4
total	*15,5*	*6,6*	*77,9*

D'autres indicateurs semblent pertinents : la langue d'enquête, l'attitude de tiers et le lieu d'enquête. Les effets sont sans surprises si ce n'est par leur ampleur.

Tableau 4-3-b : autres effets de circonstances de l'enquête, …. sur la non-réponse

Langue d'enquête	NC	refus	répondant
Autres	20,0	8,6	71,4
Français	13,6	5,7	80,7
Russe	11,3	5,1	83,6
total	15,5	6,6	77,9

Comportement des tiers	NC	refus	répondant
Neutre	15,8	5,4	78,8
Pas de tiers	15,0	6,8	78,1
Interventionniste	16,7	7,4	75,9
total	15,5	6,6	77,9

Lieu d'enquête	NC	refus	répondant
Hors domicile usuel	19,1	7,6	73,3
Dans le domicile usuel	14,8	6,4	78,8
total	15,5	6,6	77,9

On peut enfin récupérer la région détaillée d'enquête :

Tableau 4-3-c : effets de la région d'enquête, …. sur la non-réponse

Région (effectif)	Contact	Refus	Global
Alsace (268)	Très bon	Très bon	Très bon
Ile de France (4029)	Mauvais	Mauvais	Mauvais
Paca (937)	Très mauvais	Excellent	Moyen
Rhône-Alpes (873)	Très bon	Mauvais	Très bon

On peut constater que les taux alsaciens sont conformes aux taux obtenus en général (plutôt bons) ce qui s'interprète comme une bonne intégration des primo-arrivants (qui adoptent vite le comportement local) ou comme une efficacité locale des enquêteurs.
Les résultats en Rhône Alpes sont un peu inhabituel (meilleurs que sur d'autres enquêtes), tandis que les résultats dans les deux autres régions sont assez conformes à ce qui est obtenu en général…

4.4. Module sur le motif de venue en France.

Ce module permet de savoir quelles ont été les raisons du départ du pays d'origine, puis des motifs importants pour la venue en France. On y insère également les motifs du titre de séjour obtenu et certains éléments sur les liens antérieurs avec la France (y avoir déjà séjourné par exemple).
Les motifs du titre de séjour donnent lieu à des différences intéressantes :

- Un regroupement binaire pour l'échec du contact oppose des motifs "favorables au contact" : migration de travail, autres membres de familles de français, réfugiés et apatrides, aux autres.
- Un regroupement ternaire pour le refus regroupe les motifs "favorables à l'acceptation" : conjoints de Français, réfugiés et apatrides et membres de leurs familles, les motifs " défavorables" (divers), en plus d'une catégorie neutre.

Tableau 4-4-a : effets des motivations de la migration…. sur la non-réponse

Motif	NC	refus	répondant
Défavorables au contact	17,3	6,3	76,4
Favorables au contact	14,6	6,8	78,6
total	15,5	6,6	77,9

Motif	NC	refus	répondant
Favorables à l'enquête	15,1	6,1	78,8
Défavorables à l'enquête	14,5	8,4	77,0
Neutres	16,6	6,6	76,8
total	15,5	6,6	77,9

Il n'y a pas de véritable cohérence, globalement le taux de réponse est un peu inférieur pour le motif de travail, surtout en raison des difficultés de contact, un peu supérieur pour les motifs de conjoint de Français, car facilité du contact et pour le motif de réfugié (et apatride), davantage en raison d'une bonne acceptation. Ces résultats sont en cohérence avec d'autres variables (travail, raisons du départ) et constitue plus un élément de synthèse que des déterminants "primaires".

Le fait d'avoir souvent fait des séjours en France avant de s'y installer est assez favorable au contact (mais sans doute lié à la facilité linguistique).

L'analyse des raisons du choix de la France comme pays d'accueil est assez complexe, il y a dix raisons proposées et les quatre possibilités de réponses donnent lieu à des différences parfois significatives : raison très / assez / pas importante, refus de répondre.

Ainsi, refuser souvent de répondre est plutôt associé à des difficultés de contact et des refus :

Tableau 4-4-b : effets synthétiques des motivations de la migration…. sur la non-réponse

Refus sur les 10 raisons	NC	refus	répondant
Moins de deux	15,3	6,5	78,3
Deux ou plus	18,9	8,7	72,3
total	15,5	6,6	77,9

Une tentative de synthèse par analyse de données donne quelques résultats, mais peu facilement interprétables. On procède à une classification ascendante hiérarchique sur les dix questions et quatre modalités, ce qui conduit à définir six classes. Cela permet en fait de distinguer un petit quart

de l'effectif comme difficiles à contacter et à interroger, mais cette classification est moins pertinente que de sélectionner quelques raisons bien déterminantes.

Il ressort deux raisons : "liens historiques" et le fait de parler le français.

Tableau 4-4-c : effets les plus marquants des motivations de la migration…. sur la non-réponse

Liens historiques	NC	refus	répondant
Très important	15,8	6,4	77,8
Important	14,8	6,0	79,2
Pas important	15,3	6,5	78,2
Refus	21,0	13,1	65,9
Total	15,5	6,6	77,9

Parler le Français	NC	refus	répondant
Très important	14,9	5,8	79,2
Important	16,2	5,5	78,3
Pas important	15,4	7,9	76,6
Refus	18,8	10,7	70,5
Total	15,5	6,6	77,9

A cela s'ajoute la modalité "très important" pour la raison "beaucoup de personnes du pays vont en France", défavorable au contact mais facilitant l'acceptation.

4.5. Module sur l'arrivée en France et bilan du séjour

On regroupe les modules concernant l'accueil et l'arrivée en France d'une part, le bilan au jour de l'enquête sur la vie en France d'autre part.

Le type de visa à l'arrivée : il ressort une différence (pour non contact et refus) entre ceux qui n'ont pas de visa (mauvais taux) et les autres.

Tableau 4-5-a : effets de la possession d'un visa à l'arrivée ... sur la non-réponse

visa à l'arrivée	NC	refus	répondant
Non	18,0	7,3	74,7
Oui	15,0	6,5	78,6
total	15,5	6,6	77,9

La variable décrivant le maintien (ou non) en situation régulière entre l'arrivée et la demande de visa est synthétisée : les personnes arrivées régulièrement sont un peu plus difficiles à contacter mais acceptent l'enquête comme ceux restés dans la régularité. On les regroupe, contre les autres situations.

Tableau 4-5-b : effets du parcours (maintien en situation régulière ou non)... sur la non-réponse

Maintien en situation régulière	NC	refus	répondant
Arrivée régulière, passage par irrégularité	16,1	6,4	77,5
Situations diverses (donc passages par irrégularité)	17,3	7,3	75,3
Toujours en situation régulière	14,6	6,4	79,0
total	15,5	6,6	77,9

Pour les différentes démarches (au pays, à l'arrivée en France) ne ressortent que la question sur l'attente en préfecture (avec des résultats différents pour contact et refus) :

Tableau 4-5-c : effets du parcours (attente en préfecture)... sur la non-réponse

Attente en préfecture	NC	refus	répondant
Non	14,0	7,6	78,3
Oui	16,0	6,3	77,7
total	15,5	6,6	77,9

Et la question finale de bilan (des démarches en préfectures)

Tableau 4-5-d : effets du parcours (bilan des démarches en préfecture)... sur la non-réponse

Bilan	NC	refus	répondant
Bien	14,9	6,2	78,9
Autres modalités	16,5	7,4	76,1
total	15,5	6,6	77,9

On note bien un taux d'échec du contact plus élevé pour ceux qui ont déclaré que la décision de venir en France n'était pas bonne (et qui sont peut-être repartis ?). Le refus de répondre à cette question est lié à un taux de refus très élevé. Mais les effectifs sont très modestes.
Le refus de répondre à la question : "encourageriez vous d'autres personnes à venir en France" ou à la question analogue (limitée aux membres de la famille) ressort dans le taux de refus. Dans ce dernier cas, il est probable que la réponse « non » peut résulter du fait que la famille est déjà en France, ce qui serait cohérent avec un taux de contact un peu meilleur et à l'inverse un taux de refus plus élevé, du fait de femmes venues avec motif familial par exemple.

Tableau 4-5-e : effets des bilans tirés du parcours (recommander à d'autres de venir en France) ... sur la non-réponse

Recommander de venir	NC	refus	répondant
Refuser de répondre	18,4	6,5	75,1
Répondre	15,0	6,6	78,3
total	15,5	6,6	77,9

Une analyse plus fine des raisons montre que les difficultés administratives sont liées à un taux de réponse faible tandis que les raisons "économiques" (vie chère, différentes difficultés à trouver du travail) joue en sens inverse. Les personnes qui savent déjà qu'elles vont probablement repartir sont elles aussi difficiles à contacter. Le fait de vouloir devenir français se traduit par des taux de contact meilleurs.

Tableau 4-5-f : effets des bilans tirés du parcours (intention de rester ou repartir, demander la nationalité) ... sur la non-réponse

Intentions	NC	refus	répondant
Repartir (peut-être ou sûr)	18,4	6,5	75,1
S'installer	15,0	6,6	78,3
total	15,5	6,6	77,9

Demander la nationalité	NC	refus	répondant
Non, pas sûr, …	17,5	7,8	74,7
Oui	15,1	6,4	78,6
total	15,5	6,6	77,9

4.6. Modules sur le contrat d'accueil et d'intégration.

Cette demi-journée consiste à présenter des informations (film, documentation écrite), proposer et faire des bilans dans différents domaines. Elle peut être l'occasion d'établir le programme de cours de français (jusqu'à 400 heures). Elle se conclut par la signature du contrat d'accueil et d'intégration et la délivrance du titre de séjour. L'enquête reprend les différentes phases, pour évaluer la perception des enquêtés de cette demi-journée.

- Le film (réactions à propos d'un film d'informations) : la première question n'apporte pas grand-chose, mais la seconde distingue les personnes intéressées (ou ayant vraiment compris … ?) des autres.
- Le livret : distinction entre ceux qui l'ont consulté et les autres, sensible sur le refus.
- L'assistante sociale : le fait d'être orienté vers celle-ci est lié à de meilleurs taux de réponse, particulièrement si cela est jugé utile (taux de refus divisé par 3 !). Hélas, les effectifs concernés sont faibles.
- L'utilité : taux plus défavorables si l'on est filtré par la formation linguistique (ce qui renvoie à cette variable). Taux favorable pour ceux qui pensent que le CAI peut servir à trouver un emploi ou acquérir la nationalité française, taux défavorable pour ceux qui refusent de répondre à la question sur le logement. Enfin, la question finale (et bilan du module) distingue les personnes qui répondent "plutôt bien passé" avec un bon taux de contact mais un taux élevé de refus. Les taux sont très défavorables pour les réponses encore plus négatives, mais avec des effectifs quasi insignifiants. Il semble donc que les personnes critiques ou peu coopératives se sont réfugiées dans la modalité citée…Notons enfin un taux de refus fort parmi ceux qui considèrent que le CAI ne sert à rien ou presque.

L'analyse logistique sur les variables précédemment décrites ne permet pas de retenir quoi que ce soit pour la modélisation du non contact (le film serait retenu à un seuil de l'ordre de 7 %).
Par contre, on obtient des modèles du refus assez convenables, le meilleur faisant intervenir les explicatives « film », « livret », « assistante sociale » et « utilité pour la nationalité ».

On ajoute quelques autres informations plus spécifiques sur cette journée.

- La visite médicale : petits effets (dans la lignée du module précédent). Le fait d'avoir besoin et d'aller voir un médecin depuis la journée CAI est lié à un meilleur taux de réponse.
- La formation civique (module H) : trouver la formation intéressante, même schéma que les questions sur l'utilité du CAI. Les trois questions concernant l'intérêt ou l'utilité sont fortement corrélées. Le fait de déclarer connaître toutes les valeurs de la République est associé à un taux de refus plus élevé, ce qui est paradoxal[2]. Déclarer être plutôt (mais pas complètement) d'accord avec la laïcité de la France est associé à un taux très modeste d'échec de contact. Ne pas être d'accord concerne très peu de monde, mais les taux de réponses sont mauvais. L'ajout de variables dérivées de ces indicateurs ne joue pas dans les analyses logistiques pour le modèle de refus. Par contre, elles sont toutes deux assez significatives et influent sur le contact.
- La formation "Vivre en France" : taux de réponse un peu supérieur à la moyenne si elle est prescrite, alors que ce taux est mauvais si elle est prescrite mais pas suivie. Dans de dernier cas, les effectifs sont trop faibles. Un effet (positif) sur la réponse pour un comportement déjà constaté : les personnes qui répondent avoir presque tout compris (et non pas "tout") ont peut-être plus de recul… On constate le même phénomène sur les questions d'évaluation de la formation, en particulier la qualité de la formation. Les personnes qui adhèrent sans réserve à l'utilité de la formation sont un peu moins répondantes que les autres. Cela est particulièrement le cas de l'utilité de la formation en rapport avec le logement.
- La formation "Bilan de compétence" : taux de réponse d'autant plus supérieur à la moyenne que si elle est prescrite, faite, comprise. On créé une variable qui synthétise cela. Les variables concernant l'utilité de ce bilan pour une réorientation professionnelle, une reconversion et une orientation vers une formation sont assez discriminantes :

[2] Répondre aux enquêtes fait partie de ces valeurs...

Tableau 4-6 : Taux suivant les conclusions du bilan de compétence

réorientation professionnelle	NC	refus	répondant
non	16,4	7,2	76,4
oui	12,2	4,6	83,2
total	15,5	6,6	77,9

reconversion	NC	refus	répondant
Hors sujet	16,9	7,5	75,6
Pas utile	14,9	5,9	79,2
Utile	12,6	5,2	82,3
total	15,5	6,6	77,9

orientation vers une formation	NC	refus	répondant
Hors sujet	16,9	7,5	75,6
Pas utile	15,3	5,8	78,9
Utile	12,8	5,3	81,9
total	15,5	6,6	77,9

Les analyses logistiques conduisent à simplifier la variable de bilan de compétences en deux modalités. Pour le non contact, cette dernière et la variable de réorientation jouent sensiblement et significativement, tandis que la variable dérivée de l'appréciation de la formation "vivre en France" et la variable de réorientation jouent de même sur le refus.

4.7. Module sur l'emploi.

Ce module est très complexe, long et avec beaucoup de filtres. On y relève des effets intéressants, mais qui sont manifestement des effets indirects. Certaines variables seront également testées dans les modules concernés (langue, réseau social).

Effet assez marqué d'un groupe de PCS (PCS avant migration), soit l'ensemble « agriculteur, professions intermédiaires, employés, ouvriers qualifiés », favorable à la réponse. Cet effet capte assez largement les différences suivant le statut et le secteur d'activité, par ailleurs peu marquées. Effet également marqué (et favorable) pour les personnes qui ont trouvé leur emploi en France avant de venir, mais elles sont trop peu nombreuses. Trouver rapidement son premier emploi en France est associé un peu étonnamment à davantage de refus, mais l'effectif est insuffisant. On binarise une variable "avoir trouvé son premier emploi grâce au réseau social", la réponse "oui" étant associé à un moindre refus. Petit effet favorable du temps partiel (temps disponible ? / progression grâce à emploi ultérieur ?). On binarise également une variable "avoir trouvé son emploi actuel grâce au réseau social", la réponse … "non" étant associé à davantage de difficulté du contact. Le fait de déclarer changer de type de travail entre le pays d'origine et la France est associé à un refus moindre. C'est l'inverse pour les personnes qui se déclarent "manœuvre / ouvriers spécialisés", en cohérence avec la PCS (avant migration). Des difficultés de contact si le secteur d'activité est « construction, transport, commerce, hôtels, café, restaurants ». Il y a un effet très fort, de refus, mais plus que compensé par la facilité du contact pour les personnes à temps partiel mais ne souhaitant pas travailler davantage. L'effectif est toutefois très faible (et ce sont des femmes). Un phénomène analogue se produit pour les personnes qui pratiquent des horaires type 3*8, mais ce sont des hommes. De façon globale les conditions de travail semblent influer assez peu, c'est plus le type d'activité qui paraît déterminant. Cela incluant la satisfaction au travail ou encore le type d'environnement (communautaire ou non). Les personnes (peu nombreuses) qui travaillent usuellement dans une autre langue que le français sont un peu plus difficiles à contacter. On a encore testé une variable qui indique si la personne ne cherche pas de travail, soit qu'elle donne comme raison des soins à donner à des proches, soit des raisons de santé, qui semble associé à un meilleur contact. Une variable issue du module rassemblant des informations sur la famille (migrations des parents, circonstances du mariage, emploi du conjoint) est prise en considération ici. Il s'agit de l'emploi du conjoint. Le fait que le conjoint soit ouvrier ou employé est associée à de meilleurs taux. C'est même la seule variable jouant significativement sur ces taux.

4.8. Modules sur les connaissances en Français et les diplômes

Beaucoup de variables jouent sur les taux (mais elles sont logiquement assez corrélées). Ainsi le fait que l'on ne se souvienne plus qu'une formation en Français a été prescrite est très corrélé à de mauvais taux.

Tableau 4-8-a : effets des bilans faits lors de la demi-journée d'accueil (formation en français) ... sur la non-réponse

Formation prescrite	NC	refus	répondant
Pas de formation prescrite	14,1	6,4	79,5
Oui mais l'ignore	21,5	7,5	71,0
Oui et s'en souvient	17,5	6,8	75,7
total	15,5	6,6	77,9

A l'inverse, le fait d'avoir parlé le Français en bas âge est corrélé à de bons taux, de même que le fait de parler le Français avec son conjoint, avec des amis ou encore d'avoir eu une partie (au moins) de son enseignement en Français.

Tableau 4-8-b : effets des pratiques linguistiques ... sur la non-réponse

Langue(s) parlée(s) enfance	NC	refus	répondant
Autres modalités	16,3	7,2	76,6
Français et / ou autre	13,6	5,2	81,2
total	15,5	6,6	77,9

Langue parlée avec le conjoint	NC	refus	répondant
Autres modalités	16,3	6,9	76,7
Français seulement	13,2	5,7	81,1
total	15,5	6,6	77,9

Tableau 4-8-c : effets du niveau déclaré de français ... sur la non-réponse

Déclaré, téléphone	NC	refus	répondant
Autres situations	15,5	6,8	77,8
difficultés	16,1	4,6	79,3
total	15,5	6,6	77,9

Déclaré, chemin	NC	refus	répondant
Autres situations	15,6	6,7	77,8
difficultés	14,5	5,3	80,2
total	15,5	6,6	77,9

Les 4 indicateurs sur le niveau auto-évalué de français apportent une information difficile d'analyse : déclarer des difficultés à téléphoner ou à écrire une lettre est corrélé à un taux de refus faible.

Un effet un peu étonnant et difficile d'interprétation : les personnes n'ayant jamais suivi d'études ou n'ayant que le certificat d'études se signalent par un fort taux de refus, mais ce n'est pas le cas des personnes sans diplômes. Toutefois l'effet n'est pas très fort en raison des effectifs.
Le fait de ne pas demander d'équivalence du diplôme obtenu est associé à un taux de non contact élevé. La variable indiquant si des cours de formation ont eu lieu en français est significative. De même que l'équivalence de diplôme, le fait d'avoir un permis étranger et de le valider pour la France est associé à un meilleur taux de réponse. On distingue pour les autres ceux qui ont le permis mais ne valident pas, et ceux qui ne l'ont pas. Les résultats sont assez proches de ce que l'on observe pour le genre, d'où l'hypothèse non testée que les hommes sont plutôt dans le premier cas, les femmes dans le second.

Tableau 4-8-d : effets du niveau de formation initiale ... sur la non-réponse

Formation initiale	NC	refus	répondant
Autres modalités	15,5	6,2	78,2
Niveau CEP, pas d'études	15,2	9,8	75,0
total	15,5	6,6	77,9

Tableau 4-8-e : effets des demandes d'équivalence ... sur la non-réponse

Equivalence diplômes	NC	refus	répondant
Autres modalités	15,1	6,8	78,2
Pas demandée	17,5	5,9	76,6
total	15,5	6,6	77,9

Equivalence permis	NC	refus	répondant
Mis à jour	13,1	5,7	81,3
Pas de permis	16,0	7,2	76,7
Permis, non valable	17,2	6,6	76,2
total	15,5	6,6	77,9

4.9. Modules sur conditions de vie et difficultés financières

Le niveau de vie par unité de consommation n'est pas déterminant, par contre la modalité de réponse "refus" l'est. On aborde le niveau de vie par les restrictions : difficulté à payer les fluides, un crédit, le loyer ou la cantine. Pour ce dernier cas, il y a un effet net, mais clairement associé à la structure du ménage. On crée une variable supplémentaire pour détecter le refus de répondre à ces questions. L'effet principal (hors refus) est le fait de ne pas être concerné par les crédits, ce qui est associé à un contact très facile. Le fait de bénéficier d'une aide n'est globalement pas significatif, mais suivant que celle-ci provient d'associations ou d'organismes publics d'une part, d'autres origines (famille, proches ?) d'autre part, on observe des taux de réponses très différents (meilleurs dans le premier cas).
Le fait de juger sa situation financière en évolution est aussi significatif, ce sont ceux qui la jugent en dégradation qui répondent le mieux.

Il y a un effet (négatif) sur le taux de réponse si la personne n'a pas accès à un compte bancaire, principalement si cet accès n'a pas lieu d'être par manque de rentrées régulières (ce ne semble pas être un effet de genre, mais l'effet est de toute façon peu marqué).
De même un effet peu marqué favorise l'acceptation si les aides au logement sont perçues, le contact si l'enquêté n'y a pas droit (ou le pense). Les variables sur les allocations familiales, garde d'enfant, … sont manifestement colinéaires aux variables conjoint, enfant, …
Le fait d'avoir une couverture maladie complémentaire, soit de type CMU, soit autre, ou non crée des catégories distinctes. Les analyses logistiques confirment les analyses descriptives.

Les premières analyses logistiques font apparaître un effet marqué de la variable "crédit" et un effet certain mais avec des intervalles de confiance très large de la variable d'aide pour l'échec du contact. On recode la seconde variable. Pour le refus, seul le fait de refuser de répondre pour l'estimation du niveau de vie intervient.

4.10. Modules sur logement (type, confort …)

Le module est long et assez complexe, il ressort un assez grand nombre d'effets en statistiques descriptives, mais d'intensité modérée.

On intégrera dans les analyses logistiques les variables relatives au logement et figurant puis ressortant dans le module "ego". Il s'agit de variables sur le nombre d'habitants du logement, et une variable sur le caractère "durable" de celui-ci.

Concernant le statut d'occupation du logement, un regroupement en trois modalités est opéré: propriétaires (ou accédant) avec contact facile et refus faible, hébergés en foyer de travailleurs ou chez amis, surtout difficile à contacter. A noter que les personnes hébergées en CADA, etc… sont dans la moyenne. Il apparaît judicieux ensuite de regrouper les personnes logées en HLM avec les …propriétaires, tandis que les personnes qui sont logées par des personnes de même origine sont à rapprocher (et sont largement inclus) dans les personnes logées chez des amis. Ce constat est bien confirmé par le mode d'obtention du logement. La taille du logement joue (modalités favorables : 3-4 pièces). A noter que la possession d'un téléphone, fixe ou portable, ne joue pas. On retient dans les équipements le fait de posséder (ou non) un moyen de transport motorisé (facilitateur du contact, sans doute par effet de second ou troisième tour, travail par exemple). On teste encore la variable permettant de savoir si la personne dispose d'une pièce "à elle", même si on attend des colinéarités (nombre de pièces, nombre d'habitants). Un effet (défavorable) s'observe si les personnes déclarent ne pas avoir les éléments de confort de base, mais c'est assez faible. Une variable est construite, mais si les taux sont différents, l'effectif n'est pas assez important.
Des questions permettent de savoir, lorsque la personne a changé de logement depuis l'arrivée en France, par quel type de logement elle est passée. Les filtres sont un peu complexes et les effectifs souvent faibles dans les modalités de passage. On opère une première synthèse ad hoc : passer par une chambre d'hôtel payée soi-même, une chambre universitaire, structure d'accueil type CADA, camping, squat, lieu précaire comme cave ou voiture. La réponse positive à ces questions (mais seulement pour moins d'un mois pour les logements les plus précaires) est associée à de meilleurs taux de réponse. Inversement, pour les personnes qui ont passé peu de temps chez des amis, le taux d'échec du contact est élevé.

Tableau 4-10 : effets des parcours « résidentiels » ... sur la non-réponse

Transits	NC	refus	répondant
Autres situations	15,4	6,8	77,8
Passage amis	20,3	5,5	74,1
Transits divers	14,1	5,7	80,2
total	15,5	6,6	77,9

Enfin, concernant le désir de changer de logement, deux faits apparaissent : si la personne veut aller dans un logement du parc privé ou changer de logement HLM, les taux de réponses sont meilleurs, de même que si la personne ne veut pas faire une demande (ou n'a pas à la faire). Inversement, les personnes qui devraient faire une demande de logement social et qui ne l'ont pas faite ont de moins bons taux de réponses.

4.11. Modules sur réseau social, habitudes culturelles

- Les parents : on construit une variable indiquant si le père (puis la mère) vit dans la même zone (pays ou regroupement de pays, soit une vingtaine de zones pour l'ensemble du champ) qu'à sa naissance (ou s'il est décédé).
- La modalité "avoir migré" s'associe à un contact plus facile s'il s'agit du père, d'un refus plus important s'il s'agit de la mère, tandis que le décès de celle-ci est associé (faiblement) à un contact plus facile.
- Le conjoint : effets déjà détectés dans le module conjoint. Pas d'effet notable si le conjoint est français. Petit effet favorable si le conjoint a l'intention de venir en France, qui sera capté par le fait que l'enquêté souhaite s'installer. Le pays de la rencontre ou du mariage joue essentiellement de la même façon que la nationalité. Le fait qu'il y ait (eu) des pressions de l'une ou l'autre famille joue négativement, mais les effectifs sont anecdotiques, alors que le fait que le couple ne soit pas accepté par de l'une ou l'autre famille joue positivement. Effet fort sur le refus pour la modalité "membre de la famille (cousins par exemple)" de la question modalité de rencontre. Quelques recodages ont été faits, par acquis de conscience.
- Le niveau d'étude du conjoint conduit à une variable en trois modalités : refus (associé à de mauvais taux), niveau faible (associé à de bons taux), autres niveaux. Le type de travail du conjoint conduit à une variable en trois modalités : cadres et ouvriers spécialisés (associé à de mauvais taux), employés et ouvriers qualifiés (associé à de bons taux), autres types. Le secteur d'activité du conjoint conduit à une variable en trois modalités ad hoc. Les effets de

cette dernière variable sont moins marqués et probablement captés par la variable "type de travail".

On tente quelques analyses logistiques, incluant ou non le sexe, ce qui conduit à binariser certaines variables, mais les modèles sont médiocres.

Le fait de ne pas avoir de membres de la famille déjà en France lors de l'arrivée joue négativement sur le contact. Une modalité particulière à la question sur l'apport des personnes déjà en France, "un peu", qui joue négativement. Par type de rencontre, on relève quelques effets intéressants : faire des rencontres (d'amis) via les cours de langue ou une formation joue négativement, particulièrement sur l'acceptation. Le simple fait d'avoir des enfants d'âge scolaire joue favorablement, paradoxalement, c'est plus marqué hors école. Le fait que les nouvelles connaissances soient seulement des personnes de même origine joue négativement.
Ne plus avoir de connaissances au pays d'origine joue positivement sur le contact et négativement sur l'acceptation. Un effet apparaît à partir de la question sur les visites rendues par la famille restée au pays, essentiellement un effet négatif si ces visites sont inexistantes. Cependant, l'examen des raisons pousse à construire une variable regroupant le motif de ce manque, soit comme raisons principales, le fait d'avoir obtenu les papiers récemment. Effets attendus à partir des questions sur la lecture de la presse (du pays d'origine / autre, citer un titre, non) et plus encore de la télévision (mêmes questions). De même pour les pratiques associatives, culinaires (quotidiennes), fréquentation du café, du restaurant, qu'on les pratique de façon communautaire ou que l'on refuse de répondre. A noter que fréquenter un café uniquement fréquenté par des français joue négativement. On complète les variables binarisées sur ces pratiques par un compteur. De façon synthétique, la coupure se fait avec 3 pratiques communautaires (ou plus). Une analyse de données montre encore que le premier critère discriminant est la pratique quotidienne de la cuisine communautaire. Elle permet d'isoler un groupe aux pratiques communautaires intenses, mais les résultats sont largement corrélés à ceux obtenus par la variable de décompte, que l'on conserve car plus simple d'emploi et d'interprétation. Dernière variable construite, celle qui sépare les personnes qui ne sont pas allées en vacances au pays (durant les 12 derniers mois) de celles qui y sont allées et sont revenues trop épuisées pour répondre à l'enquête.

Bibliographie

[1] F. Domergue et C.Regnard, les nouveaux migrants en 2009, Infos-migrations N° 19, janvier 2011. http://www.immigration.gouv.fr/spip.php?page=dossiers_det_res&numrubrique=232&numarticle=2478