

Chapitre 4 : Besoins des ménages et probabilité de migration

Dans ce chapitre, nous présentons le volet quantitatif de ce mémoire. Ce chapitre vise à répondre à nos deux premiers objectifs de recherche. Le premier objectif est de *situer l'analyse des comportements migratoires dans leur contexte familial (du ménage), dans une perspective longitudinale*. Le deuxième objectif est *d'étudier l'effet des besoins du ménage sur les comportements migratoires de ses membres*. L'utilisation des ensembles de données de trois passages de l'Enquête sur les niveaux de vie des ménages au Vietnam (VHLSS) nous permet de situer les comportements migratoires dans leur contexte familial, tout en tenant compte de la durée et du parcours du ménage. Dans cette analyse, les besoins du ménage sont mesurés à partir de trois indicateurs : 1) la composition du ménage (rapport de dépendance), 2) le niveau de dépenses reliées à la santé et 3) le niveau de dépenses reliées à l'éducation. L'unité d'analyse est le ménage. Le rapport de dépendance et les niveaux de dépenses sont calculés par rapport au ménage dans son entier. Selon la définition de VHLSS, les membres d'un ménage sont ceux qui ont partagé le logement et les repas pendant un minimum de six mois au cours des douze mois précédant l'enquête, et qui partagent un ensemble de revenus et de dépenses. Les données de VHLSS ne nous permettent pas de tenir compte de la migration internationale dans notre analyse. Ainsi, faute de données plus complètes, ce volet de notre analyse porte exclusivement sur les migrations internes de travail au Vietnam.

Tel que discuté dans le chapitre précédent, au Vietnam, le passage d'une économie planifiée, de type socialiste et contrôlée par l'État, à une économie de marché contrôlée fait en sorte qu'à l'échelle locale, une part non négligeable du financement a été transférée vers les ménages. Ces transformations économiques ont augmenté de beaucoup le fardeau financier au niveau du ménage. Parmi d'autres, les charges aux usagers et les paiements officieux sont des sources de dépenses importantes pour les ménages vietnamiens. Ainsi, afin de vérifier les hypothèses selon lesquelles 1) le stade du parcours du ménage influence les probabilités de migration et 2) que de plus grands besoins au niveau du ménage augmentent les probabilités de migration de travail au niveau du ménage, la prise en compte de la structure du ménage et des niveaux de dépenses en santé et en éducation du ménage est tout indiquée.

L'analyse qui suit se démarque des autres études de trois façons. D'abord en utilisant le ménage comme unité d'analyse. Ensuite, en incluant les besoins pré migration en matière de santé et d'éducation et relativement à la structure du ménage en tant que déterminants de la migration de travail interne au Vietnam. Et finalement, en utilisant un ensemble de données de type panel pour tenir compte des changements qui se produisent au sein des ménages au fil du temps. Dans la section suivante, un bref portrait de la situation concernant la migration interne au Vietnam sera fait. Suivra ensuite la description des données et de la méthode utilisée. Les résultats et une discussion seront par la suite présentés.

Migrations internes au Vietnam

L'augmentation de la mobilité géographique liée au travail et une rapide croissance économique vont de pair (Coxhead, Cuong, et Vu 2015, 3). Au Vietnam, la mise en place du *Đổi Mới* s'est caractérisée par un accroissement des migrations à l'échelle nationale (Nguyen et Locke 2014, 856; Nguyen, Raabe, et Grote 2015, 79). Selon les données des recensements nationaux (Tableau 2), en 1989, un peu plus de 2,5 millions d'individus (4,5% de la population de 5 ans et plus) étaient considérés migrants²². En 1999, ils sont environ 4,5 millions, soit 6,4% des cinq ans et plus, et en 2009, à l'échelle du pays, ils étaient un peu plus de 6,7 millions, soit 8,6% de la population âgée de cinq ans et plus.

²² Au Vietnam, aux fins de recensement, un individu est considéré migrant s'il a changé de lieu de résidence au cours des cinq années précédant le recensement. Cette définition pose différents problèmes : 1) elle n'inclut pas les déplacements précédents, 2) les migrants qui ne s'enregistrent pas à destination ne sont pas comptés, et 3) les déplacements à court terme et circulaires ne sont pas représentés dans les recensements. Ce dernier point est sans doute le plus important (Dang, Tacoli, et Thanh 2003, 1).

Tableau 2 : Proportion de migrants internes et internationaux au Vietnam, population âgée de 5 ans et plus, 1989-2009

| Migrations¹ | 1989 (%) | 1999 (%) | 2009 (%) |
|----------------------------------|--------------------|--------------------|--------------------|
| Aucune | 95,3 | 93,4 | 91,3 |
| Intradistricts | n.d. | 1,9 | 2,1 |
| Interdistricts | 2,0 | 1,6 | 2,2 |
| Interprovinciales | 2,5 | 2,9 | 4,3 |
| Total migrants internes | 4,5 | 6,4 | 8,6 |
| Taille population 5 ans + | 54 717 181 | 69 059 455 | 78 513 966 |

Sources : IPUMS International²³ (Minnesota Population Center 2015), microdonnées des échantillons (%) des recensements nationaux du Vietnam de 1989 (5%), de 1999 (3%) et de 2009 (15%); Calculs de l'auteur, données pondérées.

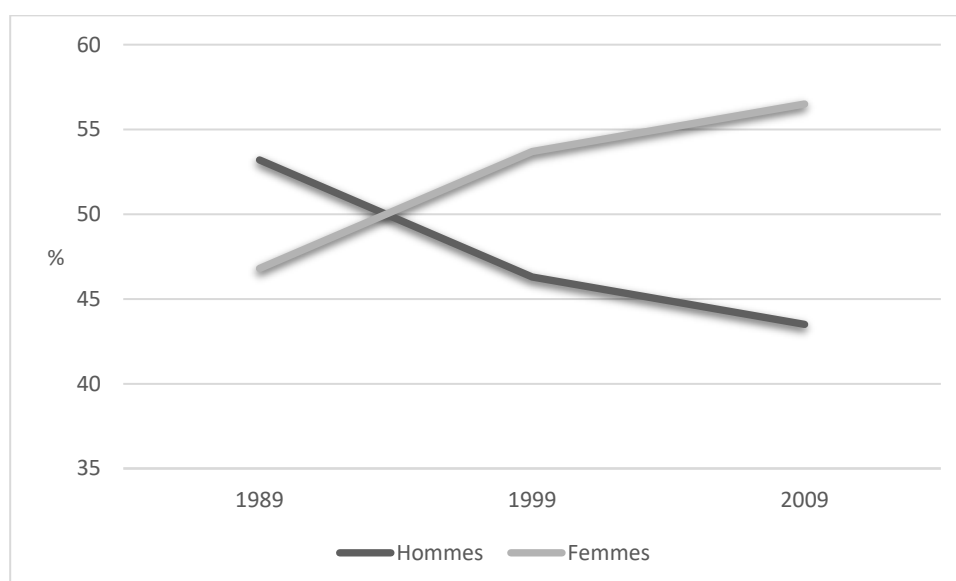
¹La catégorie *inconnue/manquante* n'est pas présentée; elle ne dépasse jamais 0,1% des individus.

Les recherches précédentes montrent que les migrations interprovinciales rurales urbaines représentent la plus grande proportion des flux migratoires au Vietnam (GSO 2015, 90). Plus des deux tiers de ceux ayant changé de province ont également changé de région au sein du Vietnam. Ceci souligne la prédominance des migrations internes de longue distance. Cette tendance est associée aux changements de localisation, de statuts et de revenu des emplois (Dang, Tacoli, et Thanh 2003, 5). Les migrants se déplacent principalement vers les villes et les centres urbains (GSO 2015, 92). Hô Chi Minh-Ville en tête, suivie de Hanoï, ont toutes deux attiré le plus grand nombre de migrants internes en raison des opportunités économiques et des possibilités d'études (Dang, Tacoli, et Thanh 2003, 5; Coxhead, Cuong, et Vu 2015, 9; GSO 2015, 92). Les provinces où le taux d'émigration est le plus élevé se retrouvent dans la région du Delta du fleuve Rouge (très haute densité de population) et les régions côtières centrales (faibles revenus). Après Hô Chi Minh-Ville et Hanoï, ce sont les régions du Sud-Est et des Montagnes centrales qui attirent le plus grand nombre de migrants. Da Nang a également vu sa population augmenter en raison de l'immigration (Dang, Tacoli, et Thanh 2003, 5).

La période 2004-2009 est marquée par la croissance de l'économie et l'accélération des migrations. Cependant, par rapport à cette période, selon les données du Bureau national des statistiques du Vietnam (*General Statistics Office of Vietnam - GSO*), le nombre de

²³ *IPUMS-International* est un projet dédié à la collecte et à la distribution de données de recensement. À l'heure actuelle, leur base de données rend possible l'accès aux données de 277 recensements provenant de 82 pays. *IPUMS-International* collecte et préserve les données et la documentation, harmonise les données et les dissémine gratuitement.

migrants a diminué lors de la période 2009-2014 (GSO 2015, 90). Toutefois, la majorité des migrants (59,2%) se concentrent en zones urbaines et c'est dans le sud-est du pays que se retrouve la plus forte proportion (41,3%) d'entre eux. Les migrants sont également plus jeunes que ceux qui ne migrent pas (Dang, Tacoli, et Thanh 2003, 6; Nguyen *et al.* 2008, 13; Tu, Dao, et Hoang 2008, 6; Coxhead, Cuong, et Vu 2015, 11; GSO 2015, 93). Toujours selon le Bureau national des statistiques du Vietnam, en 2014, il y avait environ 5,7 millions de migrants au Vietnam. De ce nombre, 59% sont des femmes (GSO 2015, 93). Ceci confirme la tendance observée dans les données des trois derniers recensements nationaux (Figure 5) et il est ainsi possible d'affirmer qu'au Vietnam, la mobilité liée à la migration interne des femmes augmente comparativement à celle des hommes.



Source : IPUMS International, microdonnées des échantillons (%) des recensements nationaux du Vietnam de 1989 (5%), de 1999 (3%) et de 2009 (15%). Calculs de l'auteur, données pondérées.

Figure 5 : Répartition des migrants internes au Vietnam selon le sexe, 1989 à 2009

Au Vietnam, la féminisation des migrations internes est un fait important. La plus grande présence de femmes migrantes au Vietnam pourrait être liée à la plus grande prévalence des activités du secteur populaire²⁴, qui dans ce pays, sont plus accessibles aux femmes de tous âges, en particulier pour les femmes mariées (Nguyen et Locke 2014, 867). Plusieurs travaux soulignent d'ailleurs que les travailleuses migrantes, dont une grande part

²⁴ L'expression « secteur populaire » fait référence aux activités économiques qui peuvent être catégorisées de plus ou moins informelles. Par exemple, les vendeurs itinérants, les petits commerçants ou les petites entreprises familiales sont généralement identifiés comme faisant partie du secteur populaire.

s'adonne au petit commerce itinérant, représentent le même poids analytique dans les migrations rurales urbaines que les migrants qui occupent les zones industrielles (Higgs 2003; Resurreccion et Ha 2007; Mitchell 2008, 2009; Truong 2009; Turner et Schoenberger 2012; Arnold 2013).

Différents facteurs expliquent les tendances migratoires au Vietnam. Les données des recensements de 1989, 1999 et 2009 montrent que les provinces ayant un plus haut niveau de développement économique attirent un plus grand nombre de migrants (Dang, Goldstein, et McNally 1997, 332; Phan et Coxhead 2010, 105). De plus, la migration peut être fortement affectée par les caractéristiques des ménages ou des communes. On retrouve généralement plus de migrants chez les ménages plus peuplés (Coxhead, Cuong, et Vu 2015, 18) et chez ceux ayant une plus grande proportion de membres sur le marché du travail (Nguyen *et al.* 2008, 13). Les probabilités de migration de travail sont plus élevées pour les individus dont le ménage a pour chef une femme (Coxhead, Cuong, et Vu 2015, 18). L'âge du chef de ménage est également associé positivement, mais jusqu'à un certain point (relation en U inversé), avec les probabilités de migration (Coxhead, Cuong, et Vu 2015, 18). Et un plus haut niveau d'étude augmente également les probabilités de migration (Nguyen *et al.* 2008, 14; Coxhead, Cuong, et Vu 2015, 18). Les résultats de Coxhead *et al.* (2015) apportent cependant une nuance : les probabilités de migration interne pour les individus ayant atteint un niveau moyen d'éducation sont plus élevées que pour ceux ayant très peu ou énormément d'éducation (Coxhead, Cuong, et Vu 2015, 12). Et bien que ces caractéristiques individuelles soient importantes, les réseaux influencent significativement les probabilités de migration et les choix de destination (Tu, Dao, et Hoang 2008, 7).

Différentes études montrent également que les migrations rurales urbaines s'inscrivent dans une logique précise pour les migrants. Si les représentations généralement faites au sujet des migrants ruraux urbains ont tendance à renforcer l'image de personnes objectivement pauvres et marginalisées qui se déplacent passivement, faisant tout travail qui se présente à eux, dans les faits, la plupart des migrants vont en ville avec des idées spécifiques sur ce qu'ils y feront et avec un réseau de soutien facilitant leur entrée et leurs activités en milieu urbain (Resurreccion et Ha 2007; Agergaard et Thao 2011). D'ailleurs,

beaucoup s'en sortent économiquement mieux que les résidents urbains de la classe ouvrière (Nguyen et Locke 2014, 867).

Si quelques travaux cités ci-hauts tiennent compte de certaines caractéristiques de base des ménages et de certains éléments macrostructurels, ils analysent les déterminants de la migration au Vietnam à partir de l'individu et de ses caractéristiques. Dans notre analyse, tel que déjà mentionné, l'unité d'analyse est le ménage.

L'Enquête sur les niveaux de vie des ménages au Vietnam

L'Enquête sur les niveaux de vie des ménages au Vietnam (*Vietnam - Household Living Standards Survey (VHLSS)*) est menée par le Bureau national des statistiques du Vietnam (GSO) avec une assistance technique de la part de la Banque Mondiale (BM). L'enquête VHLSS est une enquête biennale et elle comprend un échantillon restreint et un grand échantillon. Les deux échantillons se veulent représentatifs à l'échelle nationale. L'échantillon restreint, comprenant 9 402 ménages, procure des informations sur les individus, les ménages et les communes. Au niveau individuel, des données démographiques et des données sur l'éducation, l'emploi, la santé et la migration sont disponibles. Les données sur les ménages portent sur les biens durables, les actifs, la production, les revenus, les dépenses et la participation à des programmes gouvernementaux. Le grand échantillon comprend 37 000 ménages additionnels. Par contre, les données concernant les dépenses des ménages ne sont pas collectées pour les ménages du grand échantillon. En conséquence, cette recherche utilise l'échantillon restreint.

L'enquête VHLSS comprend un panel²⁵ rotatif, dans lequel environ la moitié des unités d'échantillonnage primaires sont remplacées à chaque cycle d'enquête, tandis que la moitié restante est retenue pour permettre de réinterroger les mêmes ménages (Bales 2013, 10). En 2012, 4173 (44,4%) des 9402 ménages de 2010 ont été interrogés à nouveau. En 2014, c'est 4183 ménages, soit 44,5% des ménages questionnés en 2012 qui demeurent présents dans l'enquête. Au cours des cinq années sur lesquelles s'échelonnent ces trois

²⁵ Le terme panel fait référence au fait que des données sont collectées à plusieurs reprises dans le temps pour le(s) même(s) ménage(s) ou le(s) même(s) individu(s) (ISM et SINFONICA 2015). Dans les enquêtes statistiques, les données panel comprennent essentiellement des identifiants uniques pour les ménages et/ou les individus. Chaque identifiant est lié à un seul ménage et/ou un seul individu. Dans VHLSS, chaque ménage et chaque individu au sein de chaque ménage possède un identifiant unique permettant de lier l'information les concernant à chaque passage subséquent de l'enquête.

passages de VHLSS, certains ménages demeurent présents et sont à chaque fois (3) interrogés. Il est ainsi possible de reconstruire un panel plus petit, mais d'une durée plus intéressante pour l'analyse des déterminants de la migration au niveau des ménages. Ainsi, de par son caractère longitudinal, VHLSS offre la possibilité de retracer les migrants individuels au sein des ménages suivis par cette enquête. Par opposition à une approche transversale qui peine à tenir compte du temps et de la durée, l'approche longitudinale peut saisir une part de la complexité des comportements migratoires au niveau des ménages.

À l'échelle des ménages, et en utilisant les données des passages de 2010, 2012 et 2014, il est possible de suivre un panel de 1914 ménages. Ainsi, de par son design, l'enquête VHLSS offre la possibilité de suivre l'évolution des ménages dans le temps en ce qui a trait aux migrations, et en lien avec d'autres aspects du parcours du ménage; c'est-à-dire l'éducation, la santé, le travail, les revenus et la structure du ménage. Cette enquête permet d'opérationnaliser les besoins au niveau du ménage (*i.e.* santé, éducation et structure des ménages), et par rapport aux données des recensements nationaux, les données de VHLSS contiennent beaucoup plus d'informations démographiques, sociales et économiques sur les individus, les ménages et les communes (Barbieri 2009, 141-42; Coxhead, Cuong, et Vu 2015, 9).

L'enquête VHLSS est implantée selon un plan d'échantillonnage complexe. L'échantillonnage se fait en deux étapes et selon les zones d'énumérations du recensement de la population. D'abord, la stratification de l'échantillon (*strate*) est faite en fonction des zones rurales et urbaines de chaque province. Ensuite, les unités d'échantillonnage primaires (*primary sampling unit – PSU*) (*i.e.* les communes en zones rurales et les quartiers en zones urbaines) sont sélectionnées parmi l'ensemble des communes et quartiers du pays. Finalement, les unités d'échantillonnage secondaires (*secondary sampling unit – SSU*), c'est-à-dire les ménages, sont sélectionnées à raison d'un ménage par zone d'énumération sélectionnée. Techniquement, ce plan d'échantillonnage en est un à trois niveaux, mais il est opérationnellement équivalent à un plan à deux niveaux (ISM et SINFONICA 2015). Selon ce plan, l'échantillonnage transversal (*cross-sectional sample*) de l'échantillon restreint de chaque passage de l'enquête se compose de trois ménages sélectionnés dans chacune des 3 063 unités d'échantillonnage primaires choisies (*i.e.* communes/quartiers) afin de s'assurer

que l'échantillon soit représentatif à l'échelle nationale, régionale, provinciale et également au niveau des zones rurales et urbaines du pays (GSO 2013, 2; Bales 2013, 10).

Forces et faiblesses des données

L'enquête VHLSS s'inscrit dans la série des *Living Standards Measurement Study (LSMS)* (Enquêtes sur les niveaux de vie) de la Banque Mondiale. Selon la Banque Mondiale (2016), ces études constituent un programme d'enquête auprès des ménages axé sur la production de données de haute qualité, l'amélioration des méthodes d'enquête et la promotion du développement (*building capacity*). Ainsi, l'objectif des LSMS est, toujours selon la Banque Mondiale, de faciliter l'utilisation des données d'enquêtes auprès des ménages afin d'élaborer des politiques fondées sur des données probantes (Banque Mondiale 2016). Dans cette optique, l'utilisation des enquêtes LSMS offre plusieurs avantages. Un des aspects les plus pertinents dans le cas présent est que ce type d'enquête permet d'analyser la migration par rapport à d'autres variables, et ce, de façon nettement supérieure à d'autres enquêtes (Carletto, de Brauw, et Banerjee 2012, 209). Selon Funkhouser (2012), un des aspects les plus prometteurs des enquêtes longitudinales, mais qui demeure peu exploré, est celui concernant l'inclusion de caractéristiques prémigratoires en tant que déterminants de la migration (2012, 198).

Néanmoins, la prudence est de mise lorsque les données de ces enquêtes sont utilisées. Si les LSMS se veulent longitudinales, étant donné que la majorité de ces enquêtes visent l'élaboration de politiques, les agences statistiques mettent beaucoup plus de soins dans ce qui concerne directement l'élaboration des politiques proprement dite, que dans l'aspect longitudinal des enquêtes (Funkhouser 2012, 195), c'est-à-dire que la qualité de l'aspect longitudinal est parfois négligée. Cela a pour conséquences que des compromis doivent être faits lors de l'élaboration des modèles d'analyse. De plus, la collecte d'informations détaillées concernant la migration n'est pas une priorité de la plupart des enquêtes, incluant les Enquêtes sur les niveaux de vie de la Banque Mondiale (Lucas 2000, 49-82; Carletto, de Brauw, et Banerjee 2012, 208-9). Selon Carletto, de Brauw, et Banerjee (2012), un des principaux inconvénients serait que la taille de la population migrante est trop petite par rapport à la population totale et que cette population migrante a tendance à être regroupée (*clustered*) (2012, 209). Dans cette optique, la prudence est de mise lors de l'interprétation des résultats.

Toujours selon ces derniers, il demeure possible que ce genre d'enquête ne comprenne pas suffisamment de migrants pour en faire de l'inférence statistique fiable (Carletto, de Brauw, et Banerjee 2012, 209). Cependant, tel qu'abordé plus haut et discuté par Lohr (2010), la prise en compte du plan d'échantillonnage de l'enquête (*survey design*) et l'utilisation de logiciel permettant de l'intégrer au modèle d'analyse permettent de contourner cette mise en garde (2010, Chap.11).

Une autre mise en garde concerne le type de migrations répertoriées par ces enquêtes. C'est-à-dire que les migrants non enregistrés ne sont pas comptés dans l'échantillon (Nguyen, Raabe, et Grote 2015, 80). Au Vietnam, bien que la réforme économique du milieu des années 1980 ait donné lieu à un relâchement du contrôle de la mobilité, les entraves au changement d'enregistrement du lieu de résidence persistent (Rigg, Nguyen, et Luong 2014, 380). Ainsi, la poursuite d'un travail informel est devenue une stratégie répandue (Karis 2013, 258). Ceci a pour résultat qu'une part non négligeable des migrants échappent aux statistiques officielles, du fait qu'ils conservent leur enregistrement résidentiel au sein de leur ménage d'origine. En conséquence, l'interprétation des résultats doit se faire avec prudence puisque la catégorie de migrants répertoriée par le VHLSS est sans doute, voire très certainement, composée presque exclusivement de migrants permanents et avec un enregistrement résidentiel à jour. Ces derniers représentent un sous-groupe parmi la population de migrants vietnamiens. Parmi d'autres, Pincus et Sender (2008) s'inquiètent également de l'étude de la migration à partir des ensembles de données du VHLSS, car seuls les ménages enregistrés officiellement depuis au moins 6 mois précédant l'enquête sont couverts (2008, 115). Les migrants sans statut de résidence permanente sont ignorés dans l'échantillon. De plus, du fait que les migrations temporaires ne sont pas captées par l'enquête, les risques de sous-estimation des migrations internes sont présents (Nguyen, Raabe, et Grote 2015, 81).

Les conclusions pourraient être très différentes si l'ensemble de la population migrante se retrouvait dans l'analyse. Néanmoins, la source d'information que représente VHLSS se veut représentative au niveau national. D'autre part, l'objectif de notre analyse n'est pas de mesurer de manière précise des taux d'émigration, mais bien de mesurer l'effet de certaines variables sur la probabilité qu'un ménage compte un migrant. De ce fait, VHLSS

demeure une des meilleures sources de données disponibles en ce qui concerne la situation des ménages vietnamiens et leurs comportements migratoires.

Le modèle d'analyse

Les modèles de régression présentés dans le cadre de cette analyse visent à décrire les relations observées entre les caractéristiques des ménages et les probabilités de migration. Le but n'est pas de prédire des valeurs précises, mais bien de déterminer si ces différentes caractéristiques ont un effet significatif sur la probabilité de migration et, le cas échéant, d'en déterminer le sens (positif ou négatif) et l'intensité. Le niveau d'analyse est le ménage. Ceux-ci sont suivis sur une période de cinq années de façon à évaluer l'impact des besoins du ménage sur la probabilité de migration interne de travail au Vietnam.

À partir de l'échantillon panel de 1914 ménages, trois modèles de régression logistique ont été produits sur un sous-échantillon de 1810 ménages. Le retrait de 104 ménages s'explique par la sélection de la population à l'étude qui est formée des ménages de deux personnes ou plus. Les ménages d'une seule personne ont été exclus de l'analyse en raison de l'impossibilité d'avoir effectué une migration tout en demeurant dans l'analyse. Ainsi, en raison de ce critère de sélection et de la nature des données panel, les ménages migrants (*i.e.* ménages dont tous les membres migrent) ne sont pas pris en compte dans l'analyse. Par contre, le taux d'attrition est jugé suffisamment faible. Si, le taux d'attrition ne peut être estimé directement à partir de l'ensemble de données en raison de la possibilité de remplacement des ménages sans réponse (Bales 2013, 11), comparer la taille de l'échantillon à partir de la conception (panel rotatif de 50%) avec la taille réelle du panel (*i.e.* les ménages correspondants) indique des taux d'attrition relativement faibles : 5,6% de 2010 à 2012 et 5,5% de 2012 à 2014. La comparaison entre le panel restreint (n=1810) et les autres ménages de l'échantillon de l'année de départ indique que la proportion de ménages ayant pour chef une femme est légèrement plus faible et que les ménages sont légèrement plus populeux. Aucune variable clé de notre analyse (*i.e.* dépenses en santé, dépenses en éducation et rapport de dépendance) n'a montré de différences significatives lors de nos comparaisons. Cela suggère que l'attrition du panel que nous utilisons n'est pas problématique pour cette analyse.

Concernant la qualité d'ajustement des modèles de régression, l'approche utilisée se base sur le plan d'échantillonnage de l'enquête (*design-based*). Si la structure des données

doit être prise en compte, tout autant dans une approche basée sur le plan d'échantillonnage de l'enquête (*design-based*) que dans une approche basée sur le modèle d'analyse (*model-based*), dans une enquête avec plan échantillonnage complexe, cette considération est d'autant plus importante. Les observations peuvent avoir différentes probabilités d'inclusion dans une enquête à échantillonnage complexe, telle que le VHLSS. En ce sens, une analyse qui ne tient pas compte des différentes probabilités de sélection peut conduire à des biais dans les paramètres de régression (Lohr 2010, 434). Si le plan d'échantillonnage de l'enquête implique des regroupements (*clustering*), et c'est le cas de VHLSS, les erreurs standards calculées retournées par un programme qui ne tient pas compte du plan d'échantillonnage ne seront pas exactes (Lohr 2010, 435).

De nombreux statisticiens ont débattu de la question et de l'utilité des poids d'enquête (DuMouchel et Duncan 1983, 535). Or, dans le cas d'une approche basée sur le plan d'échantillonnage, les quantités d'intérêt sont celles de la population (*finite population characteristics*), peu importe que le modèle soit, ou non, parfaitement ajusté à cette population (Lohr 2010, 448). C'est-à-dire qu'une pondération qui ne tient pas compte du plan d'échantillonnage ne donne pas les mêmes résultats qu'une pondération qui en tient compte. En conséquence, une pondération ajustée selon le plan d'échantillonnage est nécessaire afin d'arriver à des estimations justes de population, et devrait donc être utilisée pour les analyses en régression (Klein et Morgan 1951, 442; Lohr 2010, 448). (Lohr 2010, 449-50). DuMouchel et Duncan (1983) et Pfeffermann et Holmes (1985) soulignent que le fait d'utiliser une pondération lors des régressions peut procurer une certaine robustesse face à la mauvaise spécification (*misspecification*) des modèles (1983, 535; 1985, 275). C'est-à-dire que les estimations pondérées ne sont que très peu affectées si une variable explicative est manquante dans le modèle (Lohr 2010, 450). Dans cette analyse, les analyses multivariées sont effectuées à l'aide du logiciel STATA 14.2. Ce logiciel permet de tenir compte des différents regroupements du plan d'échantillonnage. Ainsi, lors des analyses bivariées et multivariées présentées ici, le plan d'échantillonnage de l'enquête a été respecté.

Le choix de la régression logistique binaire s'est imposé en raison de la variable dépendante dichotomique (ménage avec migrant et ménage sans migrant). L'utilisation de cette méthode permet de modéliser l'influence de la variation d'une variable indépendante

sur la probabilité d'une modalité précise d'une variable dépendante tout en maintenant les autres variables constantes. Ce type de modèle multivarié permet donc d'estimer l'effet de l'augmentation d'une unité ou du passage d'une catégorie à l'autre d'une variable indépendante (par rapport à la catégorie de référence de cette même variable), sur la variable dépendante lorsque toutes choses demeurent égales par ailleurs.

La variable dépendante

Les enquêteurs du VHLSS demandent à chaque passage de l'enquête si tous les membres du ménage sont toujours présents dans le ménage²⁶. Cependant, le fait qu'un membre ne soit plus présent dans le ménage n'implique pas nécessairement une migration. Les individus peuvent être décédés ou peuvent s'être séparés de leur ménage afin d'en former un autre et continuer de vivre à proximité, voire être voisins. Si des individus ont effectivement quitté le ménage (*i.e.* excluant le décès), le répondant doit fournir la raison principale. Les possibilités sont, entre autres, pour le travail, pour le mariage ou pour les études²⁷. Concernant la destination de ces migrations, l'information est récoltée seulement pour les migrations de travail, ce qui représente une importante faiblesse de ces données. De plus, l'enquête ayant recensé très peu de cas de migration internationale, l'analyse présentée dans ce chapitre porte seulement sur les migrants internes au Vietnam.

Ainsi, dans l'analyse qui suit, la migration est définie comme une migration interne de travail. La variable dépendante de l'analyse est d'avoir (ou de ne pas avoir) au moins un membre du ménage qui ait quitté le ménage pour une migration de travail à la fin de la période d'observation en 2014 (Figure 6). En d'autres mots, le migrant doit être présent dans le ménage en 2010 et en 2012, lors des deux premiers passages de l'enquête, mais doit avoir entrepris sa migration avant le dernier passage de l'enquête en 2014.

²⁶ La question est : *In this year, is [name] a household member ?*

²⁷ L'ensemble des choix comprend : travail, mariage, formation d'un nouveau ménage, études, aller vivre avec un autre ménage et autre.

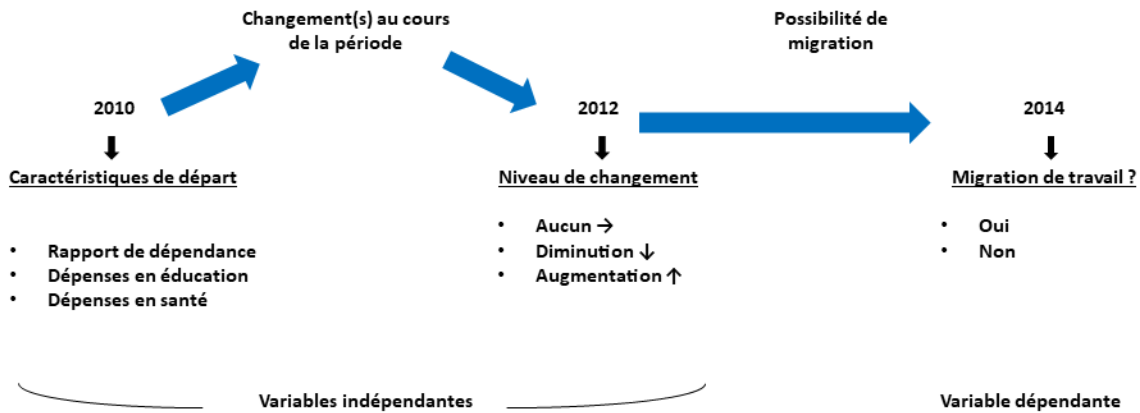


Figure 6 : Schéma des variables et du modèle d'analyse multivariée

Les variables indépendantes : besoins du ménage

Le tableau 3 présente la distribution de l'ensemble des variables utilisées dans les modèles de régression. Les variables indépendantes en relation avec les besoins du ménage comprennent les caractéristiques initiales du ménage au début de la période d'observation en 2010 et le niveau de variation de ces mêmes caractéristiques au cours de la période 2010-2012 (Figure 6). Par opposition à une approche transversale, ce choix se base sur la volonté de mesurer les changements ayant cours au sein des ménages à travers le temps, en amont de la migration. N'ayant pas de données pré 2010, les caractéristiques des ménages en 2010 sont utilisées comme point de départ. Ceci permet ensuite, du fait que les mêmes ménages sont réinterrogés, de mesurer le niveau de changement que chaque ménage expérimente, et ce, pour chacune des variables explicatives.

La variable *rapport de dépendance* du ménage est introduite dans l'analyse sous l'hypothèse que la structure du ménage, tout comme le stade de la vie des membres du ménage (jeunes enfants, aînés, membres en âge de travailler...), influencent les besoins et la charge à laquelle le ménage est soumis, et par ricochet, encouragent ou limitent les possibilités de migration au niveau du ménage. Le rapport de dépendance se définit comme le nombre de membres dépendants dans le ménage sur la taille totale du ménage²⁸. Les

²⁸ Par rapport à la définition plus classique du rapport de dépendance, soit le rapport de la population combinée de jeunes et de personnes âgées à la population en âge de travailler (jeunes + personnes âgées pour 100 travailleurs), la définition retenue ici présente l'avantage de ne pas générer de données manquantes. Par exemple, selon la définition classique, les logiciels statistiques calculent une valeur manquante pour les ménages dont tous les membres sont également dépendants. Il n'est pas possible de calculer une proportion en divisant par 0. Conserver cette définition exclurait de tels ménages des analyses.

membres dépendants sont ceux âgés de moins de 16 ans et ceux âgés de 60 ans et plus. Ce choix concernant les limites d'âge est fait en fonction des définitions données dans le questionnaire de VHLSS et en fonction de l'âge de la retraite au Vietnam. Concernant les 16 ans et moins, selon le questionnaire de VHLSS, ces personnes sont considérées « à charge »²⁹, tandis que pour les 60 ans et plus (55 ans pour les femmes³⁰), ils sont techniquement à la retraite, selon le système qui a cours au Vietnam (VBN 2017).

La variable *dépenses en éducation* est introduite dans les modèles sous l'hypothèse que des dépenses élevées en matière d'éducation puissent nécessiter une diversification des sources de revenus ou simplement une augmentation de ceux-ci. La migration peut être une option. Concernant les données utilisées, il s'agit des dépenses comparables du ménage. C'est-à-dire que les données de VHLSS sont ajustées afin de permettre une comparaison entre les différents passages de l'enquête (dépenses comparables).

Concernant la variable *dépenses de santé*, l'hypothèse est la même que pour les dépenses en éducation. Des dépenses plus élevées en matière de santé peuvent nécessiter une diversification des sources de revenus, ou simplement une augmentation de ces derniers. Encore une fois, la migration fait partie des avenues possibles. Il s'agit, ici aussi, des dépenses comparables au niveau du ménage.

Au niveau des variables de dépenses des ménages, il est important de noter que si les prix ont été soumis à un contrôle de l'État, il est tout de même possible qu'une fluctuation des prix ait eu lieu selon les régions et les mois de l'année. Cependant, les données comparables de VHLSS représentent la meilleure source d'information disponible à ce sujet. Il convient toutefois de garder en tête que certains biais peuvent être présents dans les résultats.

²⁹ Cette limite est également fonction du parcours scolaire. L'éducation de base, c'est-à-dire jusqu'à la fin du secondaire de base ou de la 9^e année, se termine généralement à 15 ans.

³⁰ Dans notre analyse, nous n'avons pas différencié l'âge de la retraite selon le sexe afin de ne pas surestimer le rapport de dépendance. Dans les faits, l'âge de la retraite ne s'applique qu'aux employés de l'État et du secteur formel. Ainsi, nous préférons être prudents et ne pas surestimer le nombre de personnes à charge au niveau du ménage.

Les autres variables indépendantes des modèles sont celles liées au niveau de changement dans les besoins du ménage entre 2010 et 2012. Ces variables sont catégorielles et les modalités introduites dans les modèles sont 1) aucun changement, 2) une diminution ou 3) une augmentation. La catégorie de référence de chacune des variables relatives aux changements des besoins du ménage est la catégorie *aucun changement*.

Les variables intermédiaires

Dans les modèles, une série de variables intermédiaires sont introduites afin de tenir compte des caractéristiques du chef du ménage, des caractéristiques des ménages autres que celles relatives aux besoins et des caractéristiques des communes. Ces variables sont utilisées afin de mesurer l'impact des besoins du ménage, toutes choses égales par ailleurs.

Des travaux récents sur la migration individuelle ont montré que les jeunes migrent plus, que l'âge du chef de ménage a une relation en U inversé avec les probabilités de migration dans le ménage et que les individus dont le chef de ménage est une femme, ou dont le chef est plus éduqué, ont plus de chances de migrer (Coxhead, Cuong, et Vu 2015). Ainsi, le *sexe du chef* (homme/femme) de ménage, son *âge* (variable continue) et son *niveau d'éducation* (variable continue) sont introduits dans les modèles.

Certaines caractéristiques des ménages sont également prises en compte dans l'analyse. Toutes les variables intermédiaires au niveau des ménages sont dans le modèle à la fois pour 2010 et 2012. Au niveau du ménage, les variables intermédiaires sont :

La taille des ménages

Cette variable répond à la logique voulant qu'un ménage comptant un plus grand nombre de membres ait plus de risques de connaître la migration d'un de ses membres.

La migration d'un membre entre 2010 et 2012 (oui/non)

Les réseaux familiaux ou de connaissances peuvent influencer de façon significative les probabilités de migration. L'hypothèse est qu'une migration très récente ne provoque pas nécessairement une migration subséquente immédiate; les réseaux à long terme sont plus solides et mieux établis. La question des ressources est également à considérer ici. La mobilisation de ressources suite à une migration récente peut être difficile.

Ménage catégorisé pauvre (oui/non)

Au Vietnam, les autorités locales classifient les ménages sous leur autorité selon leur niveau de pauvreté. La variable *ménage considéré pauvre* par les autorités locales est introduite dans les modèles sous l'hypothèse que les ménages ayant moins de ressources risquent de voir leurs chances d'avoir un migrant diminuer; tout projet migratoire demandant un minimum de ressources financières.

Présence ou non de membre(s) salarié(s) (oui/non)

En considérant la migration comme une stratégie de diversification des revenus, il est possible de dire que le fait d'avoir un ou plusieurs membres salariés dans le ménage est une stratégie de diversification des revenus qui diminue possiblement le recours à la migration.

Ménage a des activités non agricoles (oui/non)

Toujours par rapport à la diversification des revenus, le fait d'avoir mis en place des activités non agricoles (*i.e.* avoir sa propre entreprise ou sa propre activité de production non agricole) fait partie des choix de diversification qu'un ménage peut envisager. Ce choix peut influencer le comportement migratoire au niveau des ménages, soit en mobilisant la main-d'œuvre et/ou en fournissant une alternative d'emploi, soit en fournissant un revenu supplémentaire qui influence à la baisse le besoin ou la nécessité de migrer.

La dimension de la terre agricole du ménage per capita

La dimension de la terre agricole d'un ménage peut également influencer à la baisse ou à la hausse les possibilités de migration de ses membres. L'hypothèse est que, d'un côté, les ménages agricoles peuvent avoir recours à la migration dans le but de diversifier leurs sources des revenus, et de l'autre, en raison des besoins en main-d'œuvre que cela amène, les ménages possédant de plus grandes terres risquent de voir les probabilités de migration diminuer pour leurs membres.

Des variables sont également ajoutées afin de caractériser les communes. La localisation géographique et le niveau de pauvreté des communes sont pris en considération selon les modalités suivantes : *commune urbaine, commune rurale, commune rurale pauvre, commune rurale éloignée et commune rurale éloignée et pauvre*. Au Vietnam, les migrations

ont majoritairement lieu des zones rurales vers les zones urbaines. De plus, Coxhead, Cuong, et Vu (2015) ont montré que les probabilités de migration diminuent pour les individus des ménages « coincés » dans les plateaux centraux (2015, 20). En ce sens, la prise en compte du niveau d'éloignement des communes permet un certain contrôle en ce qui a trait à la distance et au niveau d'infrastructure disponible sur les lieux d'origine des migrants potentiels.

Tableau 3 : Distribution des variables des modèles de régression logistique, ménages du panel de 2010 à 2014

| | | | n | % (n=1810) | |
|--|--|--|---------------------------------------|---------------|------|
| Dépendante | • Ménages avec émigrant(s) après 2012 | Non | 1719 | 95,3 | |
| | | Oui | 91 | 4,7 | |
| Besoins du ménage | • Rapport de dépendance en 2010 | | Continue, de 0 à 1 | | |
| | • Dépenses en éducation en 2010 | | Continue, de 0 à 87,7 (Millions VND) | | |
| | • Dépenses en santé en 2010 | | Continue, de 0 à 211,3 (Millions VND) | | |
| | • Changement dans le rapport de dépendance de 2010 à 2012 | Aucun | 1099 | 60,2 | |
| | | Augmentation | 357 | 20,5 | |
| | | Diminution | 354 | 19,2 | |
| | • Changement dans les dépenses en éducation de 2010 à 2012 | Aucun | 480 | 27,2 | |
| | | Augmentation | 868 | 47,3 | |
| | | Diminution | 462 | 25,5 | |
| | • Changement dans les dépenses en santé de 2010 à 2012 | Aucun | 39 | 1,8 | |
| Augmentation | | 842 | 47,3 | | |
| Diminution | | 929 | 50,9 | | |
| Caractéristiques du chef en 2010 | • Le chef est une femme | Non | 1420 | 77,9 | |
| | | Oui | 390 | 22,1 | |
| | • Âge du chef | | Continue, de 21 à 89 ans | | |
| | • Éducation du chef | | Continue, de 0 à 12 années | | |
| Caractéristiques des ménages | • Taille du ménage | 2010 | Continue, de 2 à 12 personnes | | |
| | • Migration d'un membre entre 2010 et 2012 | Non | 1707 | 94,2 | |
| | | Oui | 103 | 5,8 | |
| | • Pauvre? | 2010 | Non | 1597 | 89,4 |
| | | | Oui | 213 | 10,6 |
| | | 2012 | Non | 1566 | 88,1 |
| | | | Oui | 244 | 11,9 |
| | • Présence de salarié(s)? | 2010 | Non | 614 | 34,0 |
| | | | Oui | 1196 | 66,0 |
| | | 2012 | Non | 618 | 33,5 |
| | | | Oui | 1192 | 66,5 |
| | • Activité non agricole? | 2010 | Non | 1198 | 65,3 |
| | | | Oui | 612 | 34,7 |
| | | 2012 | Non | 1196 | 66,1 |
| Oui | | | 614 | 33,9 | |
| • Terre agricole (milliers m ² /pers) | 2010 | Continue, de 0 à 19,3 (milliers m ²) | | | |
| | 2012 | Continue, de 0 à 16,0 (milliers m ²) | | | |
| Caractéristiques des communes | • En 2010 | Urbaine | 479 | 28,0 | |
| | | Rurale | 899 | 51,8 | |
| | | Rurale pauvre | 46 | 2,4 | |
| | | Rurale éloignée | 137 | 6,8 | |
| | | Rurale pauvre et éloignée | 249 | 11,1 | |

Source : ensembles de données de VHLSS 2010, 2012 et 2014; Calculs de l'auteur, données pondérées