

BIAIS DE SELECTION ET PRISE EN COMPTE DES MIGRANTS DANS L'ANALYSE DES BIOGRAPHIES

Philippe BOCQUIER, ORSTOM-CERPOD ¹

Le temps est à l'origine de bien des confusions et erreurs en statistique. Cette dimension est souvent prise en compte implicitement dans l'analyse, mais l'on "oublie" parfois d'en préciser les limites. Pourtant, la réflexion sur le temps devrait faire partie intégrante du raisonnement scientifique, car le temps est à la base de la perception que l'on a de la causalité.

Plus que dans d'autres sciences, cette réflexion sur le temps est indispensable en sciences sociales où il n'est guère possible de contrôler l'analyse à un autre stade que celui de la définition conceptuelle. L'expérimentation est quasiment impossible en sciences sociales, et par conséquent la recherche doit se baser sur une solide théorie, c'est-à-dire une explication causale qui met en relation des faits entre eux.

Cette communication tente d'identifier les biais possibles quand le temps et l'événement ne sont pas suffisamment bien définis lors de l'analyse biographique. Elle a pour but de contribuer à une théorie des événements sociaux qui prenne explicitement en compte le temps dans lequel s'inscrivent ces événements. Nous verrons en particulier que la migration constitue le biais principal dans les enquêtes biographiques, et que le problème ne peut se résoudre que sur le plan théorique et pas seulement technique.

Enfin, nous terminerons cette communication par un exemple d'analyse de l'accès à l'emploi qui prend en compte les migrants et les non-migrants dans un même modèle.

¹ Cette communication reprend de large extrait d'un manuel publié par l'auteur (Bocquier, 1996) et d'un article à paraître (LeGrand et Bocquier, 1997).

Migration et événement

A chaque fois que l'on a à évaluer un problème, il faut se poser la question du moment d'observation et du rapport au temps de la caractéristique étudiée dans la population. Chaque fois qu'une caractéristique variant dans le temps est analysée par un modèle qui, lui, ne tient pas compte du temps, il faut questionner la validité des résultats qui sont présentés.

Lorsqu'on travaille sur des données transversales, l'analyse est généralement conjugué au présent : on mesure la probabilité pour que tel individu (ou groupe) ait telle caractéristique au moment de l'observation de l'ensemble de l'échantillon. Le temps est concentré en un point donné.

Cependant, il est plus intéressant de considérer un événement dans le temps. La caractéristique est acquise par une partie de la population au cours d'un intervalle de temps, et la population en début d'intervalle est appelée le groupe à risque de connaître l'événement. Étudier une caractéristique acquise (un événement qui survient dans une population) a quelques conséquences sur la collecte des données : il faudra situer les événements dans le temps et non pas simplement demander à chaque individu s'il a connu ou non l'événement dans sa vie.

Les biais liés à la troncature à droite

Lorsqu'on travaille avec des données d'enquête, les individus ne sont pas observés sur toute leur vie, mais seulement jusqu'au moment de l'enquête. Lorsque l'individu n'a pas encore connu l'événement étudié, on appelle cette interruption d'observation une troncature à droite, pour signifier que l'individu sort d'observation à la date d'enquête, et qu'il n'est plus soumis au risque.

On fait l'hypothèse que les troncatures dues à la date d'enquête se répartissent uniformément au cours de l'unité de temps choisie pour l'analyse. Il suffira alors de choisir cette unité de façon à ce que cette hypothèse d'uniformité sur l'intervalle de temps reste raisonnable tout au long de l'observation. Une première source de biais lié à la troncature à droite réside dans le mauvais choix de l'unité de temps pour l'analyse. Généralement, en sciences sociales, on choisit l'année comme unité de temps, mais si les événements ont lieu sur quelques mois, l'année n'est pas un intervalle raisonnable et on préférera le mois ou la semaine : ce serait le cas, par exemple, de la durée du chômage avant la reprise d'une activité rémunérée.

Une seconde source de biais lié à la troncature à droite réside dans les sorties temporaires ou définitives de la population soumise au risque. En effet, diverses raisons ont pu mener à la sortie de la population soumise au risque :

1. d'abord, l'événement lui-même,
2. ensuite, des cas d'incapacité temporaire de connaître l'événement malgré la présence physique de l'individu dans l'échantillon : il s'agit en principe de **sorties temporaires**,
3. et enfin, l'émigration, c'est-à-dire la sortie physique de l'échantillon soumis au risque, que l'on considère comme une **sortie définitive**.

Certains cas de sortie temporaire peuvent à la limite être considérés comme des sorties définitives : un individu peut sortir de la population pour un temps, mais il peut aussi en sortir pour la vie. Mais le problème se pose plus sérieusement pour les cas de sorties définitives, c'est-à-dire des sorties qui interdisent le retour dans la population soumise au risque.

Si les cas d'incapacité définitive ne posent pas trop de problèmes (il suffit de les considérer comme physiquement hors de la population soumise au risque), l'émigration est un phénomène plus délicat, car elle peut être suivie d'un retour dans la population soumise au risque. En fait, l'échantillon n'est en général constitué que de personnes présentes au moment de l'enquête, et donc, parmi les émigrés, de personnes ayant effectué une migration de retour. Par définition, les émigrés qui ne sont pas retournés au lieu d'enquête ne figurent pas dans l'échantillon, qu'ils aient connu l'événement étudié ou non.

En ce sens, tout échantillon d'enquête rétrospective est nécessairement biaisé. Deux catégories de population ne peuvent être interrogées dans ce type d'enquête : les morts et les émigrants qui ne sont pas revenus au lieu d'enquête. Cela semble trivial, mais il est utile de le rappeler : **aucune enquête rétrospective ne peut se soustraire à des biais de sélection du fait de la survie ou de la sédentarisation** des individus, c'est une limite inhérente à ces enquêtes. On peut seulement s'assurer que la procédure de sondage donne une image représentative des biographies des personnes présentes au moment de l'enquête. On peut espérer, dans un accès d'optimisme, que les itinéraires ne sont pas considérablement affectés par la mortalité (l'itinéraire des vivants représenterait assez bien celui des morts), mais il est en revanche très fréquent que l'itinéraire migratoire soit inextricablement lié au phénomène étudié.

Dans le cas des enquêtes prospectives longitudinales (de suivi, ou à passages répétés) ou bien dans le cas des registres de population, on évite ces biais de sélection du fait de la survie ou de la sédentarisation, puisque l'on dispose en principe des informations sur les individus sortis de l'échantillon. Le lecteur qui aurait la chance de disposer de données prospectives ou de registres, se trouve ainsi allégé d'un problème important en analyse des biographies. Mais le coût de ces opérations de collecte est en général assez élevé et les résultats longs à obtenir.

Nous nous en tiendrons donc ici aux problèmes que posent les enquêtes rétrospectives à un seul passage, plus communes.

Dans les enquêtes rétrospectives, non seulement l'observation des immigrants et des émigrants est sélective, mais la migration de retour (émigration suivie d'un retour avant la date d'enquête) est aussi sélective. Dès lors, l'enquête ne peut donner une mesure exacte ni de l'immigration vers, ni de l'émigration hors de la ville, et par conséquent **il est impossible d'analyser la migration vers et hors la population soumise au risque en tant qu'événement biographique, même si le statut migratoire (défini au moment de l'entrée dans la population soumise au risque) peut être introduit dans l'analyse comme variable explicative.**

Il faut en particulier être très vigilant pour **traiter les cas où l'émigration est motivée par un événement du même type que l'événement étudié.** Considérons par exemple l'émigration suite à une offre d'emploi à l'extérieur : comment traiter ce type d'événement ? Comme une entrée dans la vie active ou comme une émigration ? Pour savoir que répondre, il nous faut mettre en rapport la définition de la population soumise au risque et celle de l'événement. Par exemple, si l'offre concerne un emploi à l'extérieur, donc un marché de l'emploi plus vaste, l'émigration constitue bien une sortie d'observation. Mais si l'offre concernait un emploi dans la zone d'observation, et qu'immédiatement après l'employé a été affecté ailleurs, l'émigration n'est plus considérée comme une sortie d'observation puisqu'elle a eu lieu après l'événement (l'embauche). On conviendra cependant que ces cas sont très difficiles à distinguer les uns des autres : il faudrait disposer d'informations très précises sur le moment et les conditions de l'événement.

Les cas d'émigrations suivies de retour peuvent rester ambigus, même dans le cas où l'émigration n'a pas été motivée par un événement du même type que l'événement étudié. Prenons l'exemple d'un individu qui va faire ses études hors de la zone d'observation, pour revenir ensuite y chercher du travail. Il a été soumis au risque d'entrer dans la vie active à l'extérieur pendant la durée de son émigration, mais il est à nouveau soumis au risque depuis le moment de son retour. Doit-on l'inclure dans l'analyse, et de quelle façon ?

C'est dire si **le problème de la migration ne peut être évacué dans la plupart des analyses de données biographiques : au contraire, il doit fait partie intégrante du raisonnement.** Le fait qu'on ne puisse mesurer la migration vers et hors de l'échantillon ne doit pas nous empêcher d'en stigmatiser l'importance, et d'identifier en conséquence les biais que peuvent introduire la sortie temporaire ou définitive de la population soumise au risque.

Les biais liés à la troncature à gauche

Malgré les difficultés conceptuelles mentionnées ci-dessus, le traitement des troncatures à droite (généralement provoquées par le moment de l'enquête) pose moins de problème que celui des troncatures à gauche. Dans la mesure où la majorité des troncatures à droite sont dues à la date d'enquête et que les sorties de la population soumise au risque ne sont pas trop fréquentes ni trop liées au phénomène étudié, on se contente généralement de déplorer ces biais, inhérent au recueil rétrospectif.

Le problème des troncatures à gauche est plus délicat car il est indispensable de le résoudre pour constituer le sous-échantillon soumis à l'analyse : de la même façon qu'un individu peut sortir d'observation avant les autres, un individu peut rentrer dans le champ d'observation au cours du même intervalle de temps. Ainsi, un immigré rentrera dans la population après sa naissance. L'immigré n'est soumis au risque de connaître l'événement qu'à partir du moment de son immigration dans la population étudiée (représentée par l'échantillon), alors qu'il était soumis au risque dans une autre population (non étudiée) avant son immigration.

Maintenant, on dispose d'outils statistiques qui permettent de traiter ces troncatures à gauche, aussi bien que les troncatures à droite. Mais ce n'est pas seulement une question technique : il ne suffit pas, par exemple, de tenir compte des observations tronquées à gauche pendant la moitié de l'intervalle de temps (comme on l'a fait pour les troncatures à droite). En effet, le problème est avant tout conceptuel, car on ne peut pas considérer *a priori* pour toutes les analyses que l'ensemble des individus de l'échantillon, dont le temps d'observation est tronqué à gauche ou non, forment une population homogène.

En effet, on ne peut mesurer le risque de connaître un événement qu'à partir du moment où l'individu entre dans la population soumise au risque. Or, pour que les individus tronqués à gauche (les immigrés) soient parvenues dans la population soumise au risque, il faut qu'ils n'aient pas subi l'événement dans leur population d'origine avant le moment de l'immigration. On introduit donc un **biais de sélection par l'immigration** si l'on tient compte des immigrants au même titre que les autres individus de l'échantillon. On peut de plus imaginer (c'est en fait un cas fréquent en sciences sociales) que la troncature à gauche soit un événement (l'immigration) lui-même corrélé avec l'événement étudié. La prise en compte des troncatures à gauche n'est pas impossible, mais il n'y a pas de solution unique parce que cela touche au problème plus général de l'hétérogénéité de la population soumise au risque par rapport à une définition de l'événement que l'on veut précise.

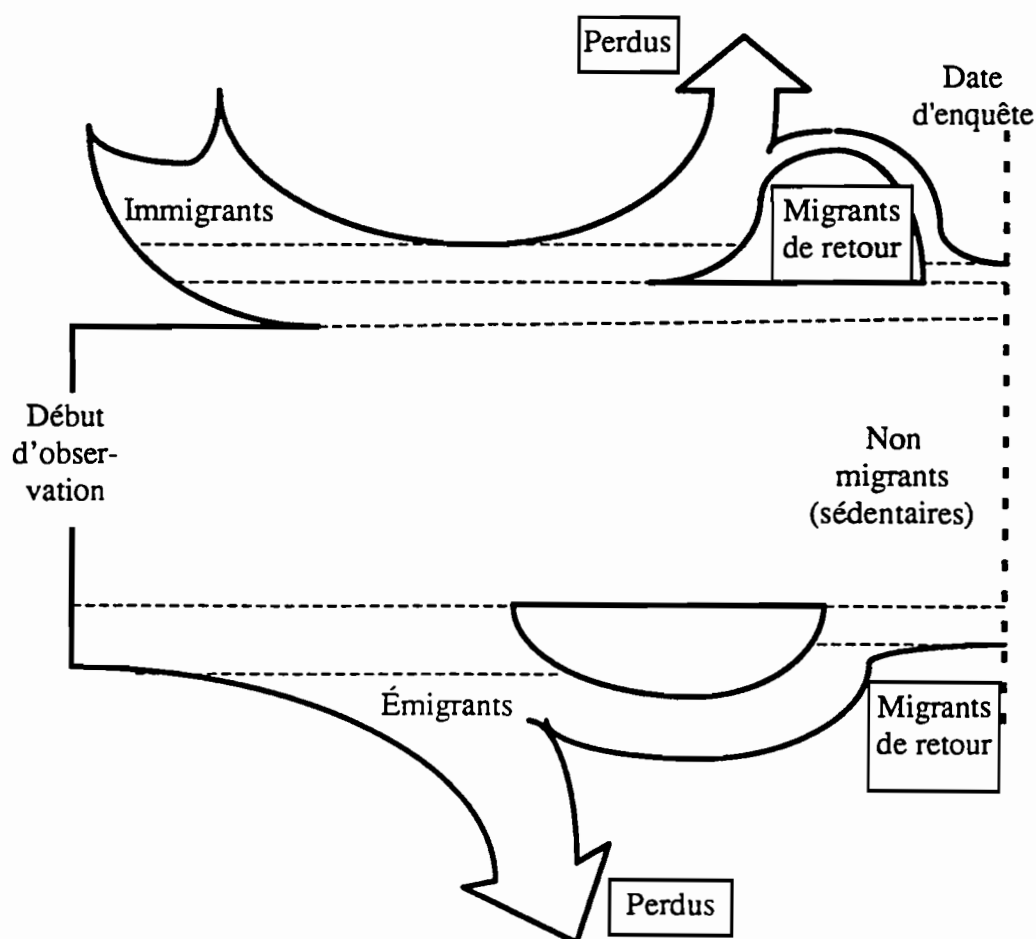


Figure 1 : Entrées et sorties d'une population soumise au risque

En résumé, on se trouve confronté à deux problèmes majeurs en traitant des données d'enquête : d'abord, les enquêtés peuvent **sortir par émigration** de la population soumise au risque pour un temps plus ou moins long, et pour une raison plus ou moins liée au phénomène que l'on étudie ; ensuite, les enquêtés peuvent **entrer par immigration** dans le lieu d'enquête après le début d'observation et connaître ainsi l'événement étudié.

La figure 1 schématise les différentes possibilités d'entrée et de sortie de la population soumise au risque étudiée, et la population finale qui en résulte. Parmi les enquêtés, on peut distinguer différents groupes selon le statut migratoire :

- les **sédentaires** forment une population relativement homogène du point de vue de l'unité de lieu et de l'unité de temps (ils n'ont pas migré). La date d'enquête, qui n'est pas en principe corrélée avec l'événement étudié, est la troncature à droite par excellence.

- les émigrants peuvent aussi être considérés dans la population soumise au risque dans la mesure où ils ne se distinguaient pas des sédentaires jusqu'au moment de la migration. L'émigration est alors traitée comme une troncature à droite. Cependant, l'émigration peut aussi être corrélée avec l'événement étudié. De plus, il faut être conscient qu'un **biais de sélection** s'opère au moment de l'enquête : parmi les émigrés, on n'interroge que les migrants de retour. Une partie des émigrés est perdue au moment de l'enquête, surtout parmi les émigrés de fraîche date qui ont peu de temps pour éventuellement revenir au lieu d'enquête.

- une partie de l'échantillon a pu venir au lieu d'enquête après la date fixée pour le début de l'observation. Ces **immigrants** (comme d'ailleurs les migrants de retour) ont pu connaître l'événement en concurrence avec les sédentaires.

Les définitions du temps et de l'événement doivent être précisées pour justifier l'analyse. De nombreuses erreurs d'interprétation proviennent souvent de la négligence de cette étape essentielle de conceptualisation. On ne saurait trop insister sur ce point : **l'essentiel de l'analyse statistique consiste à réfléchir à l'événement, à la durée avant l'événement (c'est-à-dire au temps d'observation) et aux variables indépendantes.** Le calcul informatique d'estimation est un aspect relativement secondaire dans cette démarche. Il faut souligner que **les définitions temporelles et spatiales sont très liées entre elles** : l'événement ne peut se définir sans l'espace dans lequel il se produit, c'est-à-dire la population soumise au risque à laquelle il est rapporté.

Une étude de cas :

l'effet de la sélectivité sur l'accès au secteur formel à Dakar

Si l'on veut construire un échantillon homogène selon le temps d'observation, on est bien souvent obligé de définir une population présente au lieu d'enquête au même âge et qui n'a pas migré jusqu'à l'événement. C'est ce qu'on est tenu de faire pour une analyse descriptive, par exemple, de l'accès à l'emploi des jeunes dans les enquêtes biographiques de Dakar et de Bamako (voir les « biographies des enquêtes Insertion »). Ainsi, on a sélectionné un échantillon d'individus présents à Dakar ou à Bamako à l'âge de 15 ans, qu'on peut appeler les dakarois et les bamakois, pour les distinguer des migrants. Ce sous-échantillon ne rend pas compte le sort des migrants après l'âge de 15 ans, et l'on exclue les individus dès qu'ils ont émigré hors de la ville (LeGrand et Bocquier, 1997).

Or, la proportion des dakarois et bamakois qui ont émigré avant de connaître l'événement est importante. A Bamako, en particulier, si l'on veut étudier l'accès à l'emploi, on ne peut ignorer l'importance de l'émigration car elle est aussi fréquente que, par exemple, l'accès au salariat formel.

En introduisant une double sélection, en début d'observation (présence dans la ville à 15 ans) et en fin d'observation (troncature due à l'émigration), on rend l'analyse très partielle, puisqu'elle ne rend pas compte de l'ensemble des individus en concurrence sur le marché de l'emploi. Il est donc nécessaire de passer au stade de la modélisation, pour essayer de résoudre le problème de la constitution d'une population soumise au risque homogène et pour minimiser l'importance de l'émigration.

Le modèle de transition à risques proportionnels de type semi paramétrique, avec des variables indépendantes fonction du temps, est utilisé ici pour estimer les déterminants de la durée avant l'accès des hommes au premier emploi salarié dans le secteur moderne. L'originalité du modèle semi paramétrique utilisé ici est de prendre en compte les immigrants et les migrants de retour, aussi que bien que les non-migrants.

Les variables indépendantes utilisées dans les régressions sont : la scolarité et la connaissance du français, le lieu de naissance, le statut migratoire, l'activité courante, l'ethnie, la caste, et les périodes de temps en interaction avec les générations (ce qui capte les effets de l'âge).

Dans cette étude, nous utilisons la durée de présence en ville (c'est-à-dire à Bamako ou à Dakar) pour mesurer le temps d'exposition sur le marché de l'emploi. La migration vers la ville après l'âge de 15 ans est captée par deux variables dichotomiques pour les immigrants et les migrants de retour. Le modèle de régression ne considère que le passage du temps *vécu dans la ville* avant qu'un individu n'obtienne un emploi salarié dans le secteur formel. Le temps d'observation des hommes qui ont quitté la ville avant d'obtenir un tel emploi est interrompu au moment de leur départ. Ainsi, pour eux, le temps passé hors de la ville n'est pas pris en compte dans l'analyse. Leur retour en ville est pris en compte sous certaines conditions que nous allons exposer maintenant.

Comme l'activité, la migration est un choix des individus et de leurs familles, et les perceptions des possibilités de réussite sur le marché de l'emploi influencent certainement les décisions de migration. De ce fait, il existe une sélectivité parmi les migrants : ceux qui ont une probabilité de réussite élevée sont vraisemblablement aussi ceux qui ont une tendance plus grande à migrer. S'il était possible d'inclure dans les régressions toutes les caractéristiques des individus qui affectent sa mobilité sur le marché de l'emploi, nous pourrions alors contrôler les effets de cette sélectivité. Ce n'est pas possible parce que beaucoup de caractéristiques ne sont pas observées (par exemple le « dynamisme » d'un individu). Il est aussi impossible d'inclure explicitement dans le modèle une analyse des décisions migratoires car les enquêtes ne contiennent pas d'information sur ceux qui n'ont pas migré (et qui résident donc hors de Dakar ou de Bamako) et sur ceux qui ont décidé de partir ailleurs (à l'étranger par exemple).

Cependant pour tenir compte de la situation du migrant avant son arrivée dans la ville, nous avons associé aux variables contrôlant le statut migratoire, des variables définissant l'activité simplifiée: on distingue ainsi les migrants qui avaient fait des études hors de la ville (souvent à l'étranger), des migrants occupés ou chômeurs avant leur arrivée en ville. On peut donc comparer les chances d'accès au salariat formel pour les étudiants revenus de leurs études, qui sont souvent à la recherche d'un emploi de ce type, et pour les hommes déjà actifs, pour qui la migration est synonyme de mobilité professionnelle.

Les problèmes de sélectivité sont plus aigus pour ceux qui ont trouvé un emploi *avant* leur migration via les visites de courte durée en ville, les réseaux familiaux et sociaux, ou tout simplement par affectation (par exemple dans le cas des fonctionnaires mutés). Leur devenir immédiatement après l'immigration en ville ne devraient même pas faire partie de l'observation parce qu'ils n'ont pas été soumis au risque de trouver un premier emploi salarié moderne *après* leur arrivée en ville. Afin de les exclure de l'échantillon, nous avons omis les premiers six mois de vie des immigrants et des migrants de retour après leur arrivée en ville; seule l'expérience des hommes toujours sans emploi salarié dans le secteur moderne six mois après l'arrivée est prise en considération.

Cette procédure vaut pour les nouveaux immigrants aussi bien que pour les migrants de retour. Pour tenir compte de l'expérience déjà vécue dans la ville, ces deux catégories d'immigrants ont cependant été distinguées par une variable indépendante dans l'analyse de régression.

On notera que cette procédure exclue aussi ceux qui ont pu chercher un tel emploi et le trouver dans la demi-année après leur arrivée. Ce sont les migrants qui ont généralement les caractéristiques les plus favorables pour pénétrer dans ce secteur d'emploi. Pour cette raison, la procédure adoptée devrait aussi réduire les effets de sélectivité sur les coefficients estimés.

Les résultats de la régression sont à cet égard éloquentes. A Bamako comme à Dakar, les immigrants et les migrants de retour n'ont ni plus ni moins de chances que les non migrants de trouver un emploi salarié dans le secteur moderne. Leur statut d'activité avant la migration ne joue même pas sur la probabilité d'accès à ce type d'emploi.

Le choix des migrants qui n'avaient pas obtenu d'emploi formel dans les six mois suivant leur arrivée, a certainement réduit le phénomène de sélectivité. Les migrants ainsi choisis semblent se retrouver dans les mêmes conditions que les non migrants sur le marché de l'emploi moderne. Même si l'on exclue les migrants les plus favorisés, on constate que le statut migratoire n'est pas pertinent.

Cependant, la sélectivité des migrants a pu jouer par "exclusion" : les données sont recueillies auprès d'un échantillon de personnes présentes dans la ville au moment de l'enquête. Comme il est expliqué plus haut, des personnes ont pu émigrer hors de la ville, avant ou après avoir connu l'événement étudié. Cette émigration, dont l'issue ne nous est pas connue, peut entraîner des biais dans les estimations sur la population qui est restée dans la ville. L'émigration peut être révélatrice d'un recours à des réseaux hors de la ville, ou de capacités d'adaptation en ville différentes. On ne sait cependant dans quel sens joue le biais pour l'estimation de la probabilité d'accès à l'emploi formel.

Il n'en reste pas moins que l'on a pu montrer, avec un modèle approprié, que l'effet de sélectivité à l'immigration ne perdure pas. Au-delà de 6 mois de résidence, les migrants ne se distinguent pas des autres sur le marché de l'emploi moderne. De plus, les résultats montrent que leur origine géographique et ethnique n'a quasiment pas d'effet, de même que leur appartenance de caste. D'une manière générale, on peut dire que le travailleur sur le marché de l'emploi moderne perd en quelque sorte la "mémoire", par rapport à son origine sociale. Même l'activité des migrants (leur expérience sur d'autres marchés de l'emploi) avant leur installation en ville ne joue pas sur leurs chances d'accès au salariat formel.

Ce qui frappe à l'issue de ces analyses, c'est la similitude des conclusions entre Dakar et Bamako. Ce sont *a priori* deux capitales forts différentes : le niveau technologique est plus élevé à Dakar, qui bénéficie de sa situation de ville portuaire, qu'à Bamako, ville enclavée et précisément très dépendante des villes portuaires d'Abidjan et de Dakar. Cependant, le secteur formel est à peine plus important à Dakar qu'à Bamako, et les mécanismes d'entrée dans ce secteur y sont très proches : dans les deux villes, ce sont les facteurs économiques qui l'emportent sur l'origine sociale.

En fait, l'accession à ce type d'emploi dépend essentiellement du capital humain, de l'activité courante et de la conjoncture macro-économique. D'abord, le bénéfice de l'instruction en français est indéniable. La fréquentation scolaire est la voie d'accès privilégiée à la "formalité" sur le marché de l'emploi. Ensuite, le chômage est très révélateur d'une situation de blocage du marché de l'emploi moderne. Ce sont effectivement les chômeurs qui accèdent le plus souvent au salariat formel : il s'agit essentiellement d'un chômage d'attente (d'un emploi dans le secteur formel) ou d'insatisfaction (face à d'autres emplois possibles, dans le secteur informel).

Enfin, la conjoncture, avec l'instruction et l'activité courante, est le troisième facteur principal qui explique l'accès à l'emploi formel. La crise du secteur moderne a touché toutes les catégories de travailleurs mais ce sont les jeunes qui en subissent le plus les conséquences à court et à long terme. Le manque d'emplois salariés formels maintient une grande partie d'entre eux au chômage. Or les

tendances des années 80 se poursuivent dans les années 90 : les perspectives d'embauche dans le secteur moderne sont donc très faibles pour les jeunes de moins de 30 ans. Le chômage des jeunes, plus que la migration, constitue l'enjeu fondamental sur la scène des villes africaines.

Conclusion

Nous avons tenté dans cette communication, de stigmatiser l'importance des biais de sélection par la migration sur les analyses des enquêtes biographiques rétrospectives. Cela nous semble être le biais principal de ces enquêtes, même si d'autres biais peuvent intervenir à différents niveaux de la collecte ou de l'analyse.

Il nous semble important de rappeler que le biais de sélection par la migration ne peut être évacué comme simplement inhérent à l'enquête. Il doit en fait être considéré comme *inhérent à l'analyse*. On ne peut s'y soustraire, ne serait-ce que pour constituer un sous-échantillon en conformité avec l'événement qu'on étudie. Plus généralement, le temps, l'événement, et les événements perturbateurs tels que la migration, doivent être intégrés dans un système de relations causales, c'est-à-dire une théorie explicative.

Pour illustrer notre propos, nous avons pris un exemple qui concerne la dynamique du marché de l'emploi. L'apparente complexité du modèle statistique utilisé n'est pas un artifice : elle provient de la complexité même du phénomène étudié. La technique statistique nous permet en fait d'être en conformité avec la théorie : la technique n'est pas là pour faire écran, mais au contraire elle est un révélateur de la théorie.

BIBLIOGRAPHIE INDICATIVE

- ANDERSEN, Per Kragh, BORGAN, Ørnulf, GILL, Richard D. & KEIDING, Neil (1993) Statistical Models based on Counting Processes, Springer Series in Statistics, Springer Verlag, 768p.
- ANTOINE, Ph., BOCQUIER, Ph., A.S. FALL, Y. GUISSÉ & J. NANITELEMIO (1995) Les familles dakaroises face à la crise, co-édition IFAN-ORSTOM-CEPED, 209p.
- BOCQUIER, Ph. (1996) L'analyse des enquêtes biographiques à l'aide du logiciel STATA, préface de Daniel COURGEAU et Eva LELIEVRE, Documents et Manuels du CEPED, N°4, Paris, 208p.
- BOCQUIER, Ph. (1996) L'insertion et la mobilité professionnelles à Dakar, Thèse de Doctorat (nouveau régime) en Démographie, sous la direction du Pr. Yves CHARBIT, Université Paris V - René Descartes - Sorbonne, Editions de l'ORSTOM, coll. Etudes et Thèses, 312p.
- COURGEAU, Daniel & LELIEVRE, Eva (1989) Analyse démographique des biographies, INED, 268 p.
- COURGEAU, Daniel & LELIEVRE, Eva (1995) « L'apport de l'analyse biographique en démo-économie », in L'analyse longitudinale en économie sociale - ADEPS/Université Nancy II/CNRS. (15èmes journées de l'Association d'Economie Sociale, 14-15 septembre 1995).
- DUCHÊNE, Josiane, WUNSCH, Guillaume, VILQUIN, Eric (1987) Chaire Quételet '87 L'explication en sciences sociales. la recherche des causes en démographie, Institut de Démographie, Université Catholique de Louvain, Éditions Ciaco, 478p.
- HECKMAN, James J. & SINGER, Burton (1985) Longitudinal analysis of labor market data, Econometric Society Monographs - Cambridge University Press, 410p.
- LEGRAND, Th, BOCQUIER, Ph (1996) "*Access to formal sector employment in West Africa : a study of migrant and non-migrant men in Bamako and Dakar*", Cahier 1296, Université de Montréal, Centre de Recherche et Développement en Economie, 36p.
- MAYER, Karl Ulrich & TUMA, Nancy Brandon (1990) Event History Analysis in Life Course Research, Life Course Studies - The University of Wisconsin Press, 297p.
- WUNSCH, Guillaume (1988) Causal Theory and Causal Modeling, Leuven University Press, Krakenstraat 3, B-3000 Leuven, Belgium, 200p.

Tableau 1 Résultats des régressions semi paramétriques :
Accès au premier emploi salarié dans le secteur formel à Dakar et à Bamako

Modèle:	DAKAR		BAMAKO	
	Exogène	Complet	Exogène	Complet
Education	***	***	***	***
0: parle français	0.901***	0.639***	1.564***	1.608***
Primaire sans dipl	1.469***	1.149***	1.458***	1.459***
Dipl primaire & +	1.782***	1.266***	2.170***	1.830***
Lieu de naissance				
Autre zone urbaine	0.129	0.161	-0.046	0.036
Zone rurale	0.009	0.141	-0.173	-0.036
Immigration	***		**	
Immigrant (étudiant)	0.091	0.221	-0.020	0.510***
Migrant de retour (étudiant)	0.157	0.277	-0.475	-0.109
occupé avant migr.	-0.220	0.010	-0.820***	-0.214
chômeur avant migr.	-0.570	-0.474	-0.791	-0.176
Activité		***		***
Salarié informel	n.a.	-0.415*	n.a.	-1.190***
Indépendant ou employeur	n.a.	-1.644***	n.a.	-1.946***
Apprenti	n.a.	-0.200	n.a.	-0.348
En quête du premier emploi	n.a.	0.399**	n.a.	0.631**
Chômeur avant déjà travaillé	n.a.	0.354	n.a.	0.635
Période X Generation	***	***	***	***
<1960.gen 1	0.524**	0.651***	-0.096	-0.319
1960-68.gen 1	0.089	0.199	0.196	0.077
1968-75.gen 1	-0.261-	-0.128	0.299	0.230
1975-83.gen 1	-0.670**	-0.505	-2.221**	-2.269**
1983-87.gen 1	-2.738***	-2.642***	-1.433	-1.492
1987-89/91.gen 1	-¥-	¥	-0.625	-0.871
1991-92.gen 1	n.a.	n.a.	-∞	-¥
1960-68.gen 2	-0.080	-0.077	-0.302	-0.519
1968-75.gen 2	-0.218	-0.244	-0.118	-0.224
1983-87.gen 2	-0.585*	-0.489	-0.737	-0.693
1987-89/91.gen 2	-1.982***	-1.901***	-1.508**	-1.531**
1991-92.gen 2	n.a.	n.a.	-0.232	-0.485
1968-75.gen 3	-1.224**	-1.140*	-0.271	-0.497
1975-83.gen 3	-0.853***	-0.861***	-1.815***	-2.018***
1983-87.gen 3	-0.958***	-0.960***	-0.618**	-0.828***
1987-89/91.gen 3	-1.439***	-1.485***	-0.883***	-1.218***
1991-92.gen 3	n.a.	n.a.	-2.488**	-2.904***
Ethnicité Dakar/Bamako				
Lebou/Malinké	0.135	0.128	0.048	-0.099
Peul	-0.269	-0.123	0.029	-0.167
Toucouleur/Soninké	-0.040	-0.029	-0.214	-0.168
Mande/Sonrai	0.217	0.248	-0.345	-0.531
Serrer/Minianka	-0.182	-0.374*	-0.030	0.123
Diola	0.211	-0.126	n.a.	n.a.
Divers nord	0.647	0.895	n.a.	n.a.
Divers sud	-0.085	-0.279	n.a.	n.a.
Autres	-0.745*	-0.423	-0.300	-0.359
Caste	**			
Griot	-0.943**	-0.703*	-0.278	0.018
Forgeron	-0.236	0.038	-0.211	-0.429
Cordonnier	-0.873	-0.792	-0.649	-0.462
Autres	0.488	0.595	0.064	0.222
Nb d'observations	910	910	747	747
Nombre d'événements	375	375	160	160

Niveaux de significativité: *** P<1%, ** P<5%, et * P<10%. Catégorie de référence: sans instruction, ne parle pas français, né en ville, y a résidé depuis l'âge de 15 ans, étudiant, en 1975-1983, né en 1945-54 (Dakar) ou 1947-56 (Bamako), d'ethnie Bambara (Bamako) ou Wolof (Dakar), et non casté.

Bocquier Philippe. (1997)

Biais de sélection et prise en compte des migrants dans
l'analyse des biographies

In : L'apport des collectes biographiques pour la connaissance
de la mobilité : contributions aux débats

Paris : INED, 13 p. multigr. L'Apport des Collectes

Biographiques pour la Connaissance de la Mobilité :

Rencontre Internationale, Paris (FRA), 1997/06/12-13