

Mortalité et inégalités de revenu en France : une analyse multi-niveaux

Florence Jusot*

Septembre 2004

Proposition de communication pour le XXV Congrès International de la Population

Résumé

Les sociétés égalitaires sont-elles en meilleure santé ? Plusieurs études récentes ont mis en évidence une corrélation au niveau agrégé entre inégalités de revenu et espérance de vie. Cet article explore sur données de mortalité individuelles françaises les différentes hypothèses pouvant expliquer cette corrélation. Il s'agit de déterminer si cette corrélation est un artefact statistique dû la concavité de la relation entre santé et revenu (Gravelle, 1996) ou alors si les inégalités de revenu constituent l'un des déterminants de la santé individuelle (Wilkinson, 1996).

Pour pallier l'absence d'information sur le revenu dans les données démographiques disponibles, cette analyse a été fondée sur une enquête cas-témoins, construite à partir de deux bases de données fiscales : l'enquête Patrimoine au Décès en 1988 et l'enquête sur Revenus Fiscaux des ménages de l'année 1990.

Le risque de décès est fortement corrélé au niveau de revenu. Cette relation existe tout au long de la distribution des revenus, indépendamment de l'effet des catégories socioprofessionnelles. Les résultats indiquent une surmortalité liée à la pauvreté mais aussi un effet moins attendu, une sous-mortalité liée aux plus hauts revenus, remettant en cause l'hypothèse de concavité de la relation santé-revenu. Une analyse multi-niveaux permet de montrer que le niveau des inégalités de revenu de la région d'habitation constitue en lui-même un déterminant de la mortalité, après contrôle par des caractéristiques régionales, telles que l'offre de soins et le taux de chômage.

1 Introduction

Les sociétés égalitaires sont-elles en meilleure santé ? Dès 1975, Preston a mis en évidence à partir de comparaisons internationales que l'espérance de vie n'augmentait pas continûment avec le revenu par tête. Alors que dans les pays en développement, la richesse moyenne explique largement l'état de santé moyen des populations, cette relation disparaît parmi les pays riches. En revanche la mortalité moyenne est partout expliquée par le niveau des inégalités (Rodgers, 1979). En 1992, Wilkinson a réouvert le débat sur le lien entre inégalités et santé en mettant en évidence une corrélation de -0.81 entre l'espérance de vie de 11 pays industrialisés et l'indice de Gini des revenus et en émettant l'hypothèse que les inégalités étaient un déterminant essentiel de l'état de santé. Si cette proposition a suscité un vif intérêt, l'approche méthodologique fondée sur des comparaisons internationales a été remise en cause. Ce type d'étude ne correspond en aucun cas à une expérience contrôlée, puisque les politiques de santé mises en place et le

*IRDES, 10 rue Vauvenargues, 75018 Paris, email : jusot@irdes.fr

contexte socio-culturel différent beaucoup d'un pays à l'autre. De plus, la comparabilité même des statistiques a été controversée. En effet la mesure des revenus est contingente au protocole d'enquête et influence largement l'ampleur des inégalités constatée (Deaton, 2001). Pour pallier ces problèmes méthodologiques, plusieurs études ont testé la corrélation entre inégalités et santé moyenne sur données américaines. Kaplan et al. (1996) et Kawachi et al. (1997 a) trouvent une relation négative entre le niveau d'inégalités de chaque état et plusieurs indicateurs d'état de santé¹. Cette relation semble par ailleurs robuste au choix de l'indicateur d'inégalités (Kawachi et al., 1997 a)². Cette relation est par ailleurs confirmée en France, où l'espérance de vie de chaque région décroît avec le niveau des inégalités intrarégionales (Jusot, 2003). Ces résultats forts, réalisés sur données agrégées, ne sont pourtant pas directement interprétables car plusieurs hypothèses relatives aux déterminants de la santé individuelle peuvent impliquer une corrélation entre la santé moyenne d'une population et les inégalités (Deaton, 2001, Wagstaff et Van Doorslaer, 2000).

La première hypothèse retenue par Rodgers (1979) et connue aujourd'hui sous le terme d'artefact statistique (Gravelle, 1996) est l'existence d'une relation concave entre revenu et santé. Si la santé est une fonction croissante du revenu mais à rendements décroissants, toute augmentation des inégalités de revenu se traduira par une diminution de l'état de santé moyen de la population, puisque l'amélioration de l'état de santé d'une personne riche bénéficiant d'un transfert de revenu sera inférieure à la détérioration de l'état de santé due la perte de revenu d'une personne plus pauvre. De nombreux travaux, réalisés exclusivement sur données américaines, et donc dans un cadre d'accès aux soins très spécifique, ont montré l'existence de cette relation concave entre revenu et santé (e.g. Mellor et Milyo, 1998, Smith et Kington, 1997). Cependant, la convexité de la relation estimée entre mortalité et revenu semble insuffisante pour expliquer l'ampleur de la corrélation sur données agrégées entre inégalités et mortalité (Wolfon et al., 1999).

Selon la deuxième hypothèse formulée par Wilkinson (1996) et Kawachi (2000), les inégalités de revenu ont un effet direct sur la santé individuelle. Cette externalité négative serait induite par un niveau plus faible de capital social et une moindre cohésion sociale (Kawachi et Berkman, 2000) ou encore par un effet des comparaisons sociales sur la santé (Wilkinson, 1997, Kawachi, 2000). Les indicateurs de capital social utilisés par Putnam (1995)³ sont ainsi de bons prédicteurs des différences d'espérance de vie entre états américains (Kawachi et al., 1997 b).

Enfin, cette corrélation peut refléter, non pas un effet direct des inégalités, mais l'impact de caractéristiques géographiques, elles-mêmes corrélées aux inégalités, telles que les décisions politiques locales (Deaton, 2001, Kawachi, 2000). Les différences de revenus induisent des conflits d'intérêts, qui peuvent se traduire notamment par des politiques sociales moins généreuses. Ainsi, les états américains où la distribution des revenus est la plus inégalitaire ont à la fois un taux de mortalité plus élevé, un taux de chômage plus élevé, des taux de criminalité et d'incarcération plus élevés, une proportion de personnes non assurées pour les soins de santé plus élevée et un plus faible niveau d'éducation (Kaplan et al., 1996, Kawachi et al., 1999). Cependant, selon Legrand (1987), la corrélation entre espérance de vie et inégalités de revenu est robuste à l'introduction de la part des dépenses de santé financées par le secteur public parmi 32 pays industrialisés.

¹Les indicateurs de santé utilisés sont le taux de mortalité pour toutes les causes de décès, le taux de mortalité infantile, le taux de mortalité par cancers, par maladies coronariennes, le taux d'homicides et l'invalidité.

²Les auteurs trouvent un effet significatif des inégalités sur le taux de décès moyen de chaque état américain quelque soit l'indicateur retenu : indice de Gini, de Theil, d'Atkinson, ratio interdécile, part des revenus détenue par 50% de la distribution, 60% et 70%.

³Les indicateurs retenus sont le nombre d'associations et des indices de confiance entre citoyens, tels que la confiance que l'on a dans les autres ou le sentiment que les autres ne s'intéressent qu'à eux.

Nous proposons ici d'évaluer ces différentes hypothèses sur la base de données individuelles françaises de mortalité. Il s'agit de déterminer par un double niveau d'analyse si les différences de mortalité sont expliquées uniquement par des caractéristiques individuelles, et notamment par le niveau de revenu et la profession, ou si elles sont également dues aux caractéristiques de l'environnement social dans le lequel évolue l'individu, telles que le niveau des inégalités de revenu. Ainsi, cette étude vise à apporter des éléments de réponse à quelques-unes des nombreuses questions restant encore en suspens dans le domaine de recherche consacré aux inégalités de santé (Kawachi et al., 2002). L'étude de la forme du lien santé-revenu peut permettre de déterminer si les inégalités sociales de santé sont expliquées par un effet de la pauvreté, ou s'il existe un gradient socioéconomique de santé, c'est-à-dire, des différences d'état de santé entre chacun des niveaux de la hiérarchie économique et sociale. L'étude de la forme de ce lien et l'utilisation de deux indicateurs de statut économique, le revenu et la profession, permet ensuite d'apporter des informations sur les mécanismes générant ces inégalités sociales de santé. Enfin, cette étude nous permet de déterminer si les inégalités de santé entre territoires sont expliquées par des effets de composition ou si elles relèvent de dimensions spécifiques attachées à la structure sociale de l'environnement.

Par ailleurs, la réalisation de cette étude sur données françaises présente un double intérêt.

En premier lieu, le lien entre mortalité et revenu n'a jusqu'à présent pas pu être établi en France, en raison de l'absence d'information sur le revenu dans les données traditionnellement utilisées pour étudier la mortalité. Les travaux portant sur les différences de mortalité selon la profession et catégorie sociale (PCS) (Mesrine, 1999) ne permettent pas d'étudier la forme de la relation entre revenu et mortalité sur l'ensemble de la distribution des revenus. Or, même si les indicateurs de revenu et de catégorie sociale sont corrélés, la variance des revenus au sein de chaque groupe est deux fois plus élevée que la variance intergroupe (Jusot, 2003). En particulier, la distribution des salaires parmi les plus diplômés est beaucoup plus dispersée que dans l'ensemble de la population (Nauze-Fichet et Tomasini, 2002).

Plus généralement, la France est un cadre d'étude particulièrement adapté pour évaluer l'effet des inégalités sur la santé. Pour obtenir une variabilité des inégalités de revenu en coupe transversale, nous calculons les indicateurs à l'échelle régionale. L'homogénéité des politiques sociales et des politiques de santé sur l'ensemble du territoire permet donc de limiter les biais induits par des caractéristiques régionales omises. En effet, seules les études empiriques fondées sur données individuelles américaines ont jusqu'à présent pu mettre en évidence un effet des inégalités sur la santé (Mackenbach, 2002, Subramaniam et al., 2003 a). Or, comme le souligne Mellor et Milyo (1998, 1999, 2001), la corrélation entre inégalités et santé peut être largement attribuable aux différences de politiques sociales entre Etats et plus particulièrement aux conditions d'attribution de Medicaid.

Pour pallier l'absence d'information sur le revenu dans les données habituellement utilisées pour étudier la mortalité, cette étude s'appuie sur deux bases de données fiscales, l'enquête Patrimoine au décès en 1988 et l'enquête sur les Revenus Fiscaux de l'année 1990. L'analyse repose alors sur la comparaison des caractéristiques des personnes décédées en 1988 (enquête Patrimoine au décès) aux caractéristiques des personnes enquêtées en 1990 et donc survivantes en 1988. Nous mettons tout d'abord en évidence une forte corrélation entre revenu et mortalité, corrigée de la PCS. La relation trouvée ne correspond pas à l'hypothèse de concavité. En dehors d'une surmortalité liée à la pauvreté, nous mettons en évidence un effet moins attendu, un effet protecteur des plus hauts revenus. Par ailleurs, le niveau d'inégalités intrarégionales explique significativement la mortalité, après contrôle par l'offre de soins.

2 La construction d'une enquête cas-témoins

L'absence d'étude portant sur le lien entre mortalité et revenu en France s'explique essentiellement par le manque de statistiques disponibles sur le revenu. La plupart des enquêtes comportant des informations sur le revenu sont des enquêtes annuelles, qui par construction ne portent que sur la population vivante (Budget des Familles ou Revenus Fiscaux). Les enquêtes longitudinales devraient permettre de repérer les personnes décédées au cours du suivi. Cependant, les données de panel disponibles en France ne permettent pas d'étudier la mortalité (DADS, ESPS). La mort étant heureusement un événement assez rare et principalement concentré aux âges élevés, il faudrait des échantillons initiaux très importants, suivis pendant une très longue période ou comportant initialement une proportion importante de personnes âgées, afin d'obtenir un nombre statistiquement suffisant de défunts. En outre les décès ne sont pas toujours clairement spécifiés parmi les autres causes d'attrition. Les études de mortalité s'appuient donc sur des sources démographiques spécifiques. L'Echantillon Démographique Permanent permet de suivre individuellement une large cohorte. Par ailleurs, les taux de mortalité en population générale, publiés chaque année par l'INSEE, sont calculés à partir du nombre de décès enregistrés, du nombre de naissances enregistrées et de l'effectif de la population connu au moment du recensement. Cependant ces sources ne comportent pas d'information sur le revenu.

Nous utilisons ici les données de la seule enquête représentative comportant des informations sur le revenu de personnes décédées, l'enquête "Patrimoine au décès". Cependant cette enquête ne comportant que des informations sur des défunts, l'analyse est fondée sur une enquête cas-témoins, construite à partir des données de cette enquête et de l'enquête sur les "Revenus Fiscaux de 1990".

2.1 L'enquête Patrimoine au décès : un échantillon sélectionné

L'enquête "Patrimoine au décès en 1988" (PAD) a été réalisée conjointement par l'Institut National de la Statistique et des Etudes Economiques (INSEE) et la Direction Générale des Impôts (DGI) (Laferrère et Monteil, 1995). Elle porte sur un échantillon de 4570 personnes, âgées de 20 ans et plus, domiciliées en France métropolitaine, hors de la Corse, et décédées en 1988.

Cet échantillon est issu de l'Echantillon Démographique Permanent (EDP), géré par l'INSEE. L'EDP est un échantillon d'individus tirés aléatoirement, selon leur jour de naissance, pour lequel des informations sont collectées à partir des fichiers d'état civil et des recensements depuis 1968. Cet échantillon contient donc pour un échantillon représentatif de 1% de la population des données d'état civil : la date de naissance, le statut matrimonial, le nombre d'enfants, la date du décès et le lieu du décès ainsi que la profession enregistrée au moment du décès⁴. Pour compléter les informations connues dans l'EDP, des informations fiscales relatives au revenu et au patrimoine laissé au décès⁵ ont été collectées par l'administration fiscale, à partir des déclarations de revenu et les éléments de calculs des droits de succession pour les personnes de l'EDP décédées en 1988. Par contre, peu d'informations recueillies au recensement n'ont

⁴Aucune information n'est disponible sur les causes de décès car celles-ci ne sont pas mentionnées sur les actes de décès. Identifiées lors du constat du décès, elles sont inscrites sur des certificats de décès, et constituent la base des fichiers épidémiologiques gérés par l'INSERM.

⁵Cette enquête recense les différents éléments constituant le patrimoine laissé au décès et les donations antérieures, les caractéristiques juridiques de la succession (déclaration de succession, présence d'un testament, régime matrimonial, donation au dernier vivant), ainsi que les caractéristiques des héritiers et bénéficiaires des donations antérieures.

été reportées dans l'enquête PAD. En particulier, le niveau de diplôme et le type d'activité au moment des recensements (inactifs, retraités, chômeurs, actifs occupés) ne sont pas connus.

Cependant, bien que l'on dispose d'informations sur l'âge au décès et le revenu déclaré par le foyer fiscal du défunt de l'une des deux années précédant le décès, pour un échantillon à la fois représentatif et de taille suffisante, cette base ne nous permet que d'étudier de manière précise le lien entre ces deux variables⁶. En l'absence d'effet de génération sur la mortalité et sur la distribution des caractéristiques individuelles (profession, état matrimonial), les déterminants du risque de décès pourraient être analysés par des modèles de durées (Cox, 1984). Ces méthodes permettent d'étudier les paramètres déterminant la survenue d'un changement d'état (ici le décès) à partir d'un échantillon d'individus suivi depuis une même date initiale. En recherche épidémiologique, il est fréquent que les individus ne puissent être suivis à partir de la même date. Par exemple, dans le cadre d'études portant sur des pathologies très rares, les contraintes d'échantillonnage conduisent à étudier la survie de personnes entrées en traitement à des dates différentes. Lorsque les dates de début d'observation sont suffisamment homogènes, les durées sont analysées en supposant une date origine commune (Cox, 1984). Les dates origine ne peuvent cependant pas ici être considérées comme homogènes, puisque les dates de naissance couvrent environ un siècle. Ces modèles supposent également que le phénomène expliquant le changement d'état est lui-même stable sur l'ensemble de la période étudiée. Or, il est évident que le taux de mortalité a largement évolué sur longue période. En outre la distribution même des caractéristiques (profession, état matrimonial) est marquée par de forts effets d'âge et de génération. Ainsi, dans un monde non stationnaire, l'utilisation de ces modèles conduit davantage à mettre en évidence les caractéristiques initiales des différentes générations, conditionnellement au décès en 1988, que les déterminants de la durée de vie. Pour s'en convaincre, il suffit de comparer l'âge moyen au décès dans l'enquête PAD des cadres et des agriculteurs. Alors que l'espérance de vie des cadres est supérieure à celle des agriculteurs (Mesrine, 1999), l'âge moyen au décès des agriculteurs est ici de 77 ans et celle des cadres de 71,5 ans. Ce "paradoxe" s'explique par le fait que les agriculteurs étaient plus nombreux dans les générations anciennes, alors que la proportion de cadres dans la population active a augmenté tout au long du XX^e siècle.

Puisque les déterminants de la mortalité ne peuvent être rigoureusement déduits de la distribution des âges au décès, il devient évident que l'analyse des déterminants du risque de décès ne peut être fondée sur le seul échantillon de l'enquête PAD, dans la mesure où il est parfaitement sélectionné selon le critère que nous souhaitons étudier⁷ : comme tous les individus de l'échantillon sont décédés en 1988, il manque un dénominateur, c'est-à-dire le nombre de vivants au début de l'année 1988, pour calculer la probabilité de décès durant l'année 1988. Néanmoins, cette enquête permet de connaître la distribution des caractéristiques parmi une population de défunts. Nous proposons donc de construire une enquête cas-témoins afin de déduire, de l'observation de la distribution des caractéristiques conditionnelle au décès une information sur la probabilité de décès, conditionnelle aux caractéristiques.

⁶En utilisant une méthode d'estimation indirecte de la mortalité, qualifiée de *variable r procedure* (Preston, 1999), il est possible de mettre en évidence des différences de fonction de survie selon le niveau des opportunités de revenu permanent sous l'hypothèse de stabilité du différentiel moyen des taux de mortalité sur longue période (Jusot, 2003).

⁷Manski (1995) qualifie ce problème de "response-based sampling". En économie, il est qualifié de "choice-based sampling" et en épidémiologie de "case control sampling".

2.2 Le principe des enquêtes cas-témoins

Les enquêtes cas-témoins sont fréquemment utilisées en épidémiologie pour mesurer l'association entre un facteur de risque et une maladie (Schlesselman, 1982, Bouyer et al., 1995). Elles permettent de mettre en évidence cette association à partir d'un échantillon représentatif de personnes ayant connu l'événement auquel on s'intéresse, appelé groupe de cas, et d'un échantillon représentatif de personnes n'ayant pas connu l'événement, appelé groupe témoin, sans pour autant connaître la véritable occurrence de l'événement considéré. Dans ce cadre, l'échantillon de l'enquête PAD correspond au groupe de "cas", puisqu'ils sont tous décédés. Pour constituer un groupe "témoin", nous utilisons les données de l'enquête sur les Revenus Fiscaux des ménages en 1990.

La méthodologie des enquêtes cas-témoins repose sur le principe suivant. Supposons que l'on cherche à connaître l'association entre un facteur de risque, codé par une variable dichotomique X prenant les valeurs 0 et 1, et la survenue d'un événement, ici le décès, codé par la variable dichotomique D prenant les valeurs 0 et 1. Le principe de ces enquêtes repose sur la comparaison, entre le groupe de cas et le groupe de témoins, de la fréquence relative d'exposition au facteur de risque, c'est-à-dire, de la fréquence de la caractéristique $X = 1$, relative à la fréquence de la caractéristique de la catégorie de référence $X = 0$. Le rapport de ces fréquences correspond à l'odds ratio, noté $O.R$, qui apporte une information sur le risque relatif de décès des personnes exposées au facteur de risque, par rapport au risque de décès des personnes non exposées, noté $R.R$.

Dans le groupe de cas, la fréquence relative d'exposition correspond à :

$$\frac{P(X = 1 | D = 1)}{P(X = 0 | D = 1)} = \frac{P(D = 1 | X = 1) \times P(X = 1)}{P(D = 1 | X = 0) \times P(X = 0)}$$

De même, dans le groupe de témoins :

$$\frac{P(X = 1 | D = 0)}{P(X = 0 | D = 0)} = \frac{P(D = 0 | X = 1) \times P(X = 1)}{P(D = 0 | X = 0) \times P(X = 0)}$$

Le rapport de ces fréquences relatives est égal à l'odds ratio, c'est-à-dire au ratio des rapports de chances entre les individus exposés à la caractéristique $X = 1$ et les individus exposés à la caractéristique $X = 0$:

$$\frac{\frac{P(X = 1 | D = 1)}{P(X = 0 | D = 1)}}{\frac{P(X = 1 | D = 0)}{P(X = 0 | D = 0)}} = \frac{\frac{P(D = 1 | X = 1)}{P(D = 0 | X = 1)}}{\frac{P(D = 1 | X = 0)}{P(D = 0 | X = 0)}} = O.R$$

L'odds ratio apporte alors une information sur le risque relatif de décès par rapport à la catégorie de référence, noté $R.R$:

$$R.R = \frac{P(D = 1 | X = 1)}{P(D = 1 | X = 0)}$$

On peut en effet montrer que (Manski, 1995) :

$$\begin{cases} O.R < 1 & \Rightarrow & O.R \leq R.R < 1 \\ O.R = 1 & \Rightarrow & R.R = 1 \\ O.R > 1 & \Rightarrow & 1 < R.R \leq O.R \end{cases}$$

Si l'odds ratio est égal à 1, la variable X n'a donc aucune influence sur le risque de décès, si l'odds ratio est inférieur à 1, la caractéristique $X = 1$ est associée à un risque plus faible de décès, et enfin, si l'odds ratio est supérieur à 1, elle est associée à un risque de décès plus élevé. En outre, lorsque la probabilité de survenue de l'événement $D = 1$ est faible, la valeur du risque relatif tend vers la valeur de l'odds ratio (Cornfield, 1951).

En effet :

$$R.R = \frac{P(D = 1 | X = 1)}{P(D = 1 | X = 0)} = \frac{\frac{P(X=1|D=1) \times P(D=1)}{P(X=1|D=0) \times P(D=0) + P(X=1|D=1) \times P(D=1)}}{\frac{P(X=0|D=1) \times P(D=1)}{P(X=0|D=0) \times P(D=0) + P(X=0|D=1) \times P(D=1)}}$$

D'où :

$$\lim_{P(D=1) \rightarrow 0} R.R = \frac{\frac{P(X=1|D=1) \times P(D=1)}{P(X=1|D=0) \times P(D=0)}}{\frac{P(X=0|D=1) \times P(D=1)}{P(X=0|D=0) \times P(D=0)}} = \frac{P(X=1|D=1)}{P(X=0|D=1)} = \frac{P(X=1|D=0)}{P(X=0|D=0)} = O.R$$

La probabilité de décès des hommes âgés de plus de 20 ans en 1988 n'étant que 1.4%, nous pourrions donc interpréter les odds ratios estimés comme des risques relatifs de décès par rapport à la catégorie de référence.

2.3 L'enquête Revenus Fiscaux de 1990 : un groupe témoin pertinent

Pour constituer un groupe "témoin", c'est à dire un échantillon de survivants, nous utilisons les données de l'enquête sur les Revenus Fiscaux des ménages en 1990 car cette enquête est la source la plus directement comparable à l'enquête Patrimoine au décès.

Réalisée également par l'INSEE et la DGI, elle s'inscrit dans une série d'enquêtes débutant en 1956 (Campagne et al, 1996). Elle porte sur un échantillon aléatoire de ménages ordinaires (excluant donc les ménages collectifs) tiré pour les deux tiers (20267) dans le recensement de la population de 1990 et pour le tiers (7360) dans l'enquête Actifs Financiers de 1991-1992. Sur la base des déclarations de revenus, l'administration fiscale a reconstitué le revenu économique de ces ménages. En dehors des informations socio-démographiques contenues dans les fichiers sources, on connaît donc pour 79 686 individus le revenu déclaré par son foyer fiscal et par agrégation le revenu déclaré du ménage.

Notre méthode revient à supposer que ces deux échantillons sont issus d'une même population initialement vivante au début de l'année 1988, et que les déterminants de la probabilité de décès en 1988 sont estimés en comparant les caractéristiques d'un échantillon représentatif des personnes décédées à la fin de l'année 1988, l'échantillon de l'enquête PAD, aux caractéristiques d'un échantillon représentatif des personnes survivantes à la fin de l'année 1988, l'échantillon de l'enquête Revenus Fiscaux. Cependant, l'échantillon de l'enquête Revenus Fiscaux n'est pas exactement représentatif des survivants de l'année 1988, puisqu'il ne comporte que des personnes vivantes en 1990, et donc ayant survécu à la fin de l'année 1989. Néanmoins, le biais induit par l'absence des défunts de 1989 sera très limité, puisque la distribution des caractéristiques des survivants de la fin de l'année 1989 sera peu différente de la distribution des caractéristiques des survivants de la fin de l'année 1988 car la probabilité de décès en population générale est très faible.

2.4 Les contraintes de concaténation

Comme notre étude repose sur la comparaison de deux populations, issues de deux enquêtes différentes, la robustesse des résultats repose sur la représentativité des enquêtes, et sur la

comparabilité de leur contenu. La concaténation de ces deux bases nous a donc contraint à limiter l'information utilisée.

La représentativité des enquêtes est d'une part garantie par leur mode de tirage aléatoire et d'autre part par la prise en compte des pondérations respectives de chacune des enquêtes dans l'analyse économétrique. L'homogénéité de ces sources provient essentiellement du mode de recueil du revenu, c'est-à-dire la saisie rétrospective des revenus déclarés à l'administration fiscale. Les deux enquêtes permettent de connaître pour chaque individu, le revenu déclaré⁸ et la composition de son foyer fiscal. Par contre, les deux enquêtes n'ont pas été réalisées au même niveau d'agrégation. Alors que l'enquête Revenus Fiscaux a été réalisée à l'échelle des ménages, l'enquête PAD ne donne aucune information sur le ménage de chaque défunt.

Le revenu fiscal, utilisé dans cette étude, diffère donc des variables habituellement retenues. En premier lieu, il n'est pas possible d'individualiser les revenus dans l'enquête PAD. En second lieu, la notion de foyer fiscal est strictement juridique et ne recouvre pas exactement la notion de ménage. Pour déterminer le revenu disponible de l'agent, nous avons utilisé l'échelle d'équivalence estimée par l'INSEE à partir de l'enquête Budget des Familles de 1989 (INSEE, 1995). Cette échelle attribue un poids de 1 pour la première personne du foyer fiscal et de 0.35 aux personnes suivantes. Cette échelle ne tient pas compte de l'âge, puisque l'âge des autres personnes du foyer fiscal n'est pas indiqué dans l'enquête PAD. Enfin, le revenu fiscal ne comprend que les revenus déclarés et en particulier ne comprend pas les revenus de transferts (allocations, minimum vieillesse et revenu minimum d'insertion à partir de 1990). Afin d'éviter les biais d'estimation induits par la déclaration des revenus, nous avons exclu de l'analyse les agriculteurs et les indépendants (y compris les professions libérales). En effet, les bénéficiaires agricoles et les bénéficiaires industriels et commerciaux ne donnent pas toujours lieu à une déclaration exacte et les forfaits calculés par le fisc sont peu représentatifs du revenu disponible de ces ménages (Campagne et al, 1996).

Nous n'avons retenu que les individus chefs de ménage dans l'enquête Revenus Fiscaux. Cette dernière restriction s'explique par la sensibilité de la catégorie sociale au codage. La typologie utilisée dans cette enquête diffère de la typologie utilisée au recensement, qui permet en particulier de reclasser les retraités dans leur catégorie d'origine. La PCS au sens du recensement n'est renseignée que pour les chefs de ménage, le classement des autres individus ayant fait l'objet de redressement. Or, le changement de nomenclature est loin d'être sans effet : il conduit à des discordances dans un tiers des cas (Campagne et al, 1996). Nous avons donc choisi de privilégier l'homogénéité de codification à la taille de l'échantillon. Cependant, pour éviter les biais, nous avons restreint l'échantillon à la population masculine, puisque les hommes sont le plus souvent chef de ménage. En outre, nous avons exclu les personnes inactives (hors retraités reclassés) et les retraités dont la profession d'origine était inconnue. Finalement, notre étude porte sur un échantillon de 13399 personnes, dont 1438 personnes décédées et 11961 survivantes.

3 La méthodologie empirique

Nous proposons une analyse *ceteris paribus* de la mortalité reposant sur l'estimation des déterminants de la probabilité de décès en 1988 par un modèle LOGIT. Ce modèle revient à supposer que le capital santé individuel inobservable H_i est expliqué par des caractéristiques observables X_i et un résidu u_i , distribué selon une loi logistique. Conformément aux modèles de

⁸Les revenus des défunts de 1988 sont ajustés par le taux d'inflation calculé par l'INSEE sur l'ensemble des ménages urbains.

capital santé⁹ (Grossman 1972, Erlich et Chuma, 1990), la mort survient lorsque le niveau de capital santé devient inférieur à un seuil minimal H_{\min} .

$$H_i = \alpha + X_i\beta + u_i \quad (1)$$

$$P(H_i < H_{\min}) = F(H_{\min} - (\alpha + X_i\beta)) \quad (2)$$

Comme la méthodologie cas-témoins ne requiert pas de connaître le nombre exact de défunts et de survivants, les effectifs de défunts et de survivants n'ont pas été pondérés afin d'obtenir une probabilité de décès réaliste. En effet, la taille respective de chaque sous-population influence la valeur de la constante α du modèle mais ne biaise pas la valeur estimée des paramètres β , et donc des odds-ratios qui sont égaux à l'exponentiel de β , puisque celle-ci repose sur la comparaison de la distribution des caractéristiques des défunts à la distribution des caractéristiques des survivants. Par ailleurs, pour une taille globale d'échantillon (cas+témoins) fixée, la puissance maximale est obtenue pour un nombre de cas égal au nombre de témoins (Schlesselman, 1982, Bouyer et al., 1995).

Afin de tester la représentativité de la base de données créée, le modèle 0 propose d'estimer l'effet conditionnel sur la probabilité de décès de 4 dimensions traditionnellement utilisées : l'âge, introduit par une fonction linéaire par morceaux, la catégorie sociale (PCS à un chiffre), le statut matrimonial et l'environnement géographique. À ce niveau de l'analyse, l'environnement géographique est apprécié par l'introduction de variables indicatrices des Zones d'Équipement et d'Aménagement du Territoire (ZEAT). Les résultats présentés en annexe correspondent aux estimations réalisées sur l'ensemble de l'échantillon, le traitement séparé de la population âgée de moins de 65 ans et de plus de 65 donnant des résultats très similaires¹⁰.

3.1 Une relation duale entre santé et revenu

Les modèles de capital santé (Grossman 1972, Erlich et Chuma, 1990) proposent un cadre d'analyse théorique du lien existant entre santé et revenu. Ces modèles supposent que l'agent dispose d'un capital santé initial (sa dotation génétique) qui est soumis à une dépréciation à chaque période (les effets biologiques de l'âge). Le décès survient lorsque le niveau de capital santé de l'agent devient inférieur à un niveau minimal. À chaque période, l'agent fait un arbitrage entre la consommation de biens de consommations (qui lui apporte de l'utilité) et l'investissement dans sa santé, qui est réalisé grâce à la consommation de biens médicaux ou d'autres activités bénéfiques à la santé (sport, alimentation équilibrée). L'état de santé est donc théoriquement croissant du revenu disponible, puisque ce dernier détermine les ressources dont dispose l'individu pour acquérir différents biens, y compris des biens médicaux.

On constate ainsi de fortes inégalités sociales de consommations de soins. Des différences de structure de consommation ont été largement mises en évidence (par exemple Bocognano, 1999) : les plus pauvres consomment moins de soins de spécialistes, de soins dentaires et d'optiques et davantage de soins hospitaliers et infirmiers. Cette sous-consommation de soins ambulatoires résulte principalement d'un recours aux soins moins fréquent et plus tardif au système de soins, et non d'une dépense plus faible une fois l'épisode de soins engagé (Breuil-Genier et al. 1999). Ces différences semblent en partie refléter un impact du revenu sur la décision d'investissement

⁹La détermination endogène de la durée de vie dans le modèle fondateur de Grossman (1972) a été largement controversée. Le modèle de demande de longévité d'Erlich et Chuma (1990) propose spécifiquement un arbitrage entre longévité et qualité de vie (Pour un exposé plus complet sur ce débat, voir Grossman, 2000).

¹⁰Pour les résultats par groupe d'âge, voir Jusot (2003).

en capital santé, puisque, par exemple, 14% des personnes interrogées lors de l'enquête Santé Protection Sociale de 1998 déclaraient avoir renoncé à des soins pour raisons financières, cette proportion atteignant 24% parmi les personnes disposant d'un revenu inférieur à 3 000 FF par unité de consommation (Bocognano, 1999). Par ailleurs, le montant total des dépenses de santé augmente légèrement avec le revenu (Grignon et Polton, 2000).

Dans une deuxième étape, nous testons donc l'effet du revenu par unité de consommation sur le risque de décès (modèle 1), en contrôlant par l'âge, le statut matrimonial et la zone d'habitation. Dans le modèle 2 (et suivants), nous introduisons les PCS. L'utilisation de ces deux indicateurs de statut socioéconomique nous permet alors de séparer les effets des conditions de travail et de l'éducation, approximés par les PCS, de l'effet du revenu disponible, puisque l'analyse de la distribution des revenus au sein de chaque PCS montre que ces variables sont imparfaitement corrélées (tableau 1).

L'introduction directe du revenu dans l'analyse économétrique pose cependant un problème d'endogénéité car la relation entre santé et revenu est en fait duale. Outre l'effet protecteur du revenu disponible sur la santé par l'accès aux soins, l'état de santé a un effet inverse sur les revenus du travail. Dans les modèles de capital santé (Grossman, 1972, Erlich et Chuma, 1990), l'état de santé détermine le temps de travail de l'agent car son niveau de capital santé définit le temps en bonne santé disponible pour le travail, la consommation de biens ou l'investissement dans le capital santé. Dans le cadre d'un arbitrage travail-loisir, l'offre de travail est en partie déterminée par l'état de santé, puisqu'un état de santé dégradé rend le travail plus pénible, voire impossible (Couffignal et al., 2002a, b). Les modèles de salaires d'efficience (Leibenstein, 1957) proposent par ailleurs un effet de la santé sur la productivité et donc sur le salaire.

En l'absence d'information sur le taux de salaire horaire et le temps de travail, l'état de santé H_{it} et le revenu Y_{it} sont déterminés dans un modèle structurel, comprenant deux équations simultanées¹¹ :

$$\begin{cases} H_{it} = a_0 + a_1 Y_{it} + a_2 X_{it}^h + a_3 Z_{it}^h + u_{it} \\ Y_{it} = b_0 + b_1 H_{it} + b_2 X_{it}^y + b_3 Z_{it}^y + v_{it} \end{cases} \quad (3)$$

où les caractéristiques personnelles de l'individu, comme son capital humain, et les caractéristiques de l'emploi, comme la pénibilité, peuvent affecter la santé (respectivement X_{it}^h et Z_{it}^h) et le niveau de salaire (X_{it}^y et Z_{it}^y).

L'estimation directe de a_1 dans la première équation conduit donc à surestimer l'effet du revenu sur la santé et le biais de simultanéité est une fonction croissante de l'effet de la santé sur le statut économique. La séparation rigoureuse de l'effet direct du revenu sur la santé de l'effet inverse, qualifié de "healthy worker effect", n'est en fait possible que si l'on montre un effet sur l'état de santé d'un instrument du revenu, strictement exogène à l'état de santé. Or, comme le soulignent Adams et al. (2003), un tel instrument est très difficile à trouver¹². Une solution est d'étudier l'évolution de l'état de santé de sous-populations dont le revenu est par construction indépendant de l'état de santé¹³. Une autre stratégie consiste à étudier l'évolution

¹¹Pour un modèle structurel plus complet, voir Haveman et al. (1994).

¹²On peut toutefois citer l'attribution aléatoire de taux de couverture d'assurance maladie différents lors de l'expérience de la Rand (Newhouse, 1993), ou encore l'expérience naturelle constituée par une modification du montant des pensions en 1977 aux Etats-Unis, n'ayant touché que les générations postérieures à 1917 (Snyder et Evans, 2002).

¹³Les populations d'enfants offrent un cadre d'analyse intéressant si l'on suppose que leur offre de travail est nulle et que l'état de santé des enfants n'a pas d'impact sur la décision d'offre de travail des parents. (Case et al., 2002). L'utilisation des revenus du conjoint (Smith et Kington, 1997) peut permettre de mettre en évidence un effet direct des ressources sur la santé, si l'on admet toutefois que l'offre de travail de l'un des conjoints est

de l'état de santé sur une période en contrôlant par l'état de santé initial (Adams, 2003, Duleep, 1986, McDonough, 1997).

Dans l'enquête utilisée, nous ne disposons pas d'information sur l'état de santé antérieur au décès. Nous utilisons alors la nature des revenus perçus par l'individu, et non leur montant, comme proxy de l'état de santé initial (Smith et Kington, 1997). En particulier la perception d'une pension d'invalidité est conditionnée par un mauvais état de santé. Dans les enquêtes utilisées ici, les pensions d'invalidité sont confondues avec les autres pensions et rentes et notamment les retraites et nous ne pouvons pas repérer les retraités dans l'enquête PAD. Cependant, selon l'enquête ESTEV, les troubles physiques augmentent la probabilité de départ anticipé à la retraite (Saurel-Cubizolles et al., 2001). Nous supposons alors que la part des revenus provenant de pensions et de rentes avant 60 ans permet d'approximer un mauvais état de santé initial, et donc de limiter le biais de simultanéité (modèle 3 et suivants), puisque l'impact de la santé sur la productivité ne semble pas clairement validé empiriquement¹⁴. L'analyse de la mortalité parmi les plus de 65 ans ne peut pas être contrôlée du "healthy worker effect". Cependant, l'indépendance du montant des retraites à l'évolution récente de l'état de santé permet de réduire les biais d'estimation (Adams, 2003, Smith et Kington, 1997). Toutefois, nous ne pouvons pas exclure l'hypothèse d'effets de long terme de l'état de santé sur les revenus d'activité antérieurs et donc sur le montant actuel des retraites.

3.2 Pauvreté absolue versus revenu absolu

Dans une troisième étape, nous étudions la forme de la relation existant entre santé et revenu, puisque l'hypothèse de l'artefact statistique de Gravelle repose sur la concavité de cette relation. Il existe en fait deux hypothèses concurrentes dans la littérature relatives à la forme du lien existant entre le revenu et la santé.

Selon la première hypothèse, qualifiée d'hypothèse de pauvreté absolue, les différences de mortalité se limitent à une surmortalité des plus pauvres, imputable à des conditions de travail pénibles, à des logements précaires et à de mauvaises conditions d'accès aux soins. Selon cette hypothèse, l'état de santé serait croissant avec le revenu mais ce, uniquement jusqu'à un seuil de pauvreté (Wagstaff et al., 2000). Au delà de ce seuil, l'état de santé ne serait pas amélioré par des ressources supplémentaires.

La création de la Couverture Maladie Universelle en France répond à cette logique¹⁵. Les

indépendante de l'état de santé de l'autre. D'autres travaux étudient l'effet du revenu sur l'apparition de nouvelles pathologies dans des populations âgées, dont le montant des retraites ne dépend pas de l'évolution récente de la santé (Smith et Kington, 1997, Adams et al., 2003).

¹⁴On peut en effet penser que le biais de simultanéité sera limité dans cet échantillon, puisque, pour construire la base de données, nous avons dû exclure les inactifs (hors retraités). Or l'effet de la santé sur l'inactivité est largement validé (Grandfils et LeFur, 1993, Mesrine 2000), alors que l'effet de la santé sur les salaires est moins clairement validé. L'effet de la santé sur la productivité paraît en fait limité à la population pauvre en France. Selon, l'enquête sur les bénéficiaires du RMI, le taux de croissance du salaire mensuel, corrigé du nombre d'heures travaillées, diminue avec les problèmes de santé (Rioux, 2001). En revanche, selon l'enquête sur les Conditions de vie, le handicap ou les accidents de santé graves n'ont pas d'incidence sur l'investissement en capital humain. La baisse temporaire du revenu s'explique uniquement par la perte d'expérience due à l'éloignement du marché du travail (Lechene et Magnac, 1998). De même selon l'enquête Health and Retirement Survey, l'effet des chocs de santé non anticipés sur le revenu s'explique par la réduction du nombre d'heures travaillées et non par une réduction du taux de salaire horaire (Smith, 1999).

¹⁵Lors des travaux de la commission portant sur la création de la CMU, la ministre de l'Emploi et de la Solidarité Martine Aubry précisait que «l'espérance de vie varie considérablement en fonction des critères socioprofessionnels. La réduction de l'inégalité devant l'accès aux soins constitue un des axes majeurs de la politique de lutte contre les exclusions menée par le gouvernement.» (audition du 10 mars 1999).

difficultés d'accès aux soins des plus pauvres pouvaient en effet s'expliquer par le manque de couverture du risque maladie, même en présence d'une couverture obligatoire de soins (Dourgnon et al., 2001). La Sécurité sociale ne couvre que 75% des dépenses de santé et l'absence de souscription à une assurance complémentaire était largement corrélée à la pauvreté. En 1998 par exemple, seuls 52% des personnes ayant un revenu inférieur à 2000FF par UC par mois étaient couvertes par une assurance complémentaire (Bocagnano, 1999). En dépit des difficultés de mise en évidence empirique de l'impact des consommations de soins sur l'état de santé en population générale (Couffinal et al 2002 a, b), l'absence de couverture complémentaire pouvait expliquer une dégradation de l'état de santé des plus pauvres, induite par un renoncement aux soins ou un recours plus tardif au système de soins. Par ailleurs l'expérience de la Rand a permis de montrer que la gratuité des soins n'induisait une véritable amélioration de l'état de santé que pour les populations pauvres et les populations à risques, comme les enfants ou les hypertensifs (Newhouse, 1993).

L'hypothèse de pauvreté absolue a cependant été largement remise en cause en Grande-Bretagne, notamment par le rapport Black (Townsend, 1982, 1988) qui a mis en évidence que, en dépit de l'instauration d'un système de santé accessible à tous (National Health System) et de l'assez large disparition de la pauvreté absolue, les différences de risque de mortalité des hommes âgés de 15 à 64 ans entre catégories socioprofessionnelles s'étaient accrues entre 1931 et 1971. La pertinence de l'hypothèse de la pauvreté absolue a été partiellement réhabilitée au regard de l'importance des conditions de vie dans l'enfance sur l'état de santé à l'âge adulte et de la forte reproduction sociale. Si les classes sociales inférieures ont connu des conditions de pauvreté absolue dans leur enfance, leur faible état de santé à l'âge adulte pourrait provenir de ces conditions de vies précaires antérieures. Cette hypothèse qualifiée de "early life hypothesis" repose sur l'importance des effets de très long terme sur la santé (Wadsworth, 1999, Case et al, 2002). Par exemple, Barker (1997) montre que les risques de maladie cardiovasculaire à l'âge de 50 ou 60 ans sont en partie déterminés in utero et Ravelli (1998) montre, à partir d'une expérience naturelle constituée par le cas de la Hollande entre 1944 et 1945, que le diabète à l'âge adulte pouvait être expliqué par l'exposition de la mère à une famine durant la période de grossesse.

Selon la seconde hypothèse, qualifiée d'hypothèse de revenu absolu, les inégalités sociales de santé ne se réduisent pas à une opposition entre pauvres et non-pauvres ou manuels et non-manuels. Le revenu disponible aurait un effet continu sur la santé qui serait à l'origine d'un gradient économique de santé. En France, l'existence du gradient social de mortalité a été mise en évidence selon les PCS, les employés vivant en moyenne plus longtemps que les ouvriers, les cadres plus longtemps que les professions intermédiaires (Mesrine, 1999). Selon Deaton (2001) et Wagstaff et al.(2000), l'amélioration de l'état de santé induite par un supplément de ressources se réduirait avec l'augmentation du niveau de revenu, conformément à l'hypothèse de rendements décroissants de la fonction de production de capital santé dans le modèle d'Erlich et Chuma (1990). Plusieurs travaux, réalisés exclusivement sur données américaines ont montré ainsi l'existence de cette relation continue et concave entre revenu et santé (Mellor et Milyo (1999), Smith et Kington (1997), Wolfson (1999)). Cependant, ces travaux ne sont pas contrôlés par la profession et surtout ont porté une faible attention sur les plus hauts revenus. Le revenu est le plus souvent introduit dans les estimations sous des formes conduisant à conclure à un relation concave entre revenu et santé : sous forme logarithmique (Wolfson, 1999) ou quadratique (Mellor et Milyo, 1999) (par ailleurs très restrictive). Lorsque le revenu est spécifié en tranche, la tranche de revenu la plus élevée ne permet souvent pas d'étudier l'effet des très hauts revenus. Par exemple, Smith et Kington (1997) étudient la concavité de la relation existant entre revenu

et limitations fonctionnelles selon les tertiles.

Pour séparer l'hypothèse de pauvreté absolue de l'hypothèse d'un effet du revenu absolu, nous introduisons le revenu par unité de consommation dans l'analyse selon plusieurs spécifications. Pour tester l'hypothèse de concavité, nous utilisons une forme logarithmique (modèles 1, 2 et 3) et une forme quadratique (modèle 4). Nous utilisons ensuite une forme polynomiale de degré 3 (modèle 5) afin d'étudier plus spécifiquement l'impact des plus hauts revenus sur la santé. Pour étudier la relation existant entre mortalité et revenu tout au long de la distribution, le revenu est introduit en quintiles, sans contrôle par la profession dans le modèle 6, puis en introduisant les PCS dans le modèle 7. Les quintiles de revenu sont calculés en tenant compte d'une part des pondérations respectives de chaque enquête et d'autre part en repondérant l'échantillon total pour le rendre représentatif de la population vivante au premier janvier 1988, c'est-à-dire, pour que les proportions de défunts et de survivants permettent de retrouver le taux de décès de la population masculine âgée de plus de 20 ans en 1988, soit 1.415%.

3.3 L'effet direct des inégalités

Dans une dernière étape, nous testons l'hypothèse d'un effet des inégalités de revenu sur la santé individuelle, en analysant l'impact du niveau des inégalités de chaque région en 1990, évalué à partir de l'enquête Revenus Fiscaux, sur le risque de décès en 1988 (modèles 8 à 11). Cette analyse est menée en prenant en compte les différentes critiques formulées dans la littérature, relatives à la mesure des inégalités d'une part et à la méthodologie d'estimation utilisée d'autre part (Kawachi, 2000, Subramaniam et al., 2003a).

Selon Kawachi (2000), le choix des indicateurs d'inégalités paraît souvent arbitraire et, en tout cas, n'est pas vraiment discuté. Pour améliorer la robustesse des résultats, plusieurs indicateurs d'inégalités sont ici utilisés : l'indice de Gini, la mesure d'entropie de Theil, l'indice d'Atkinson (en faisant varier la norme d'aversion aux inégalités) et la variance des logarithmes des revenus¹⁶. Par ailleurs, Kawachi (2000) a critiqué les études ne prenant pas en compte les taxes et les revenus de transferts, ainsi que la taille des ménages dans le calcul des inégalités. Ici, le niveau d'inégalités intrarégionales est calculé à partir des revenus disponibles par unité de consommation des ménages de l'enquête sur les Revenus Fiscaux de 1990 (tableau 11). Le revenu disponible des ménages correspond au revenu fiscal après impôts et revenus de transfert, calculés par l'INSEE d'après les barèmes de l'année. Les revenus disponibles ont été ajustés de la taille des ménages en utilisant l'échelle d'équivalence estimée par l'INSEE (1995) que nous avons auparavant utilisée pour ajuster le revenu fiscal au niveau individuel.

Enfin, l'analyse de l'impact des inégalités de revenu sur la santé individuelle, relevant de l'analyse contextuelle, requiert une méthodologie particulière, multi-niveaux (Diez-Roux 2000, Chaix et Chauvin, 2002, Subramaniam et al., 2003a). En effet, selon Subramaniam et al. (2003a), la non-utilisation de ce type de modélisation est à l'origine des résultats contradictoires trouvés à partir des mêmes données aux Etats-Unis¹⁷. Cette approche permet de prendre en compte

¹⁶Seules les estimations des coefficients des variables régionales sont reproduites ici pour l'ensemble des indicateurs d'inégalités utilisés. Nous proposons les estimations complètes pour l'indicateur d'entropie de Theil, les autres résultats étant disponibles auprès de l'auteur.

¹⁷A partir du Current Population Survey, Mellor et Milyo (2003) ne trouvent pas d'effet du niveau des inégalités de l'état de résidence sur l'état de santé déclaré, en utilisant une procédure d'estimation classique, et en introduisant des variables indicatrices correspondant aux 9 divisions géographiques du recensement. Sur les mêmes données, Subramaniam et al. (2003a) montrent, en revanche, que l'état de santé déclaré décroît avec le niveau des inégalités de l'état de résidence à l'aide des plusieurs analyses multi-niveaux. Cette seconde étude montre ainsi, à l'aide d'un modèle à deux niveaux (individus, Etat), que l'effet des inégalités de revenu de l'Etat

la structure hiérarchique des données en diagnostiquant l'existence d'un impact de l'environnement géographique sur le risque de décès, en traitant les biais induits par l'autocorrélation des résidus au sein de groupes plus homogènes, ici les régions, et en permettant d'expliquer cette autocorrélation grâce à des variables explicatives contextuelles.

Précédemment, nous avons supposé que le capital santé individuel inobservable H_i était expliqué par des caractéristiques individuelles observables X_i et un résidu u_i , les résidus u_i étant supposés indépendants et distribués selon une loi logistique :

$$H_i = \alpha + X_i\beta + u_i \quad (4)$$

A présent, nous considérons que les individus d'une même région peuvent être soumis à un risque de mortalité spécifique. Nous estimons donc l'impact des caractéristiques individuelles sur le risque de décès, sans introduire de variables indicatrices régionales, mais en estimant un effet aléatoire régional. En estimant les résidus simultanément aux deux niveaux d'analyse, l'analyse multi-niveaux permet de séparer la constante du modèle précédent en une constante fixe et un résidu aléatoire attaché à chaque région (Rasbash et al., 2002). Cela revient à supposer que l'état de santé H_{ij} d'un individu i appartenant à la région j est expliqué à la fois par ses caractéristiques individuelles X_{ij} , un résidu individuel e_{ij} , une constante α_0 et un effet aléatoire de groupe u_{0j} :

$$H_{ij} = \alpha_0 + u_{0j} + X_{ij}\beta + e_{ij} \quad (5)$$

Les résidus e_{ij} individuels sont supposés indépendants des résidus des groupes u_{0j} , et les résidus de chaque type sont supposés indépendants entre eux. Dans le cas d'une variable continue, les deux types de résidus sont supposés distribués selon une loi normale :

$$\begin{aligned} e_{ij} &\rightsquigarrow N(0, \sigma_e) \\ u_{0j} &\rightsquigarrow N(0, \sigma_u) \end{aligned}$$

Dans le cas d'une variable expliquée dichotomique, la probabilité que le niveau de capital santé H_{ij} de l'individu i , appartenant au groupe j , soit inférieur au seuil H_{\min} , s'exprime comme la somme de la probabilité expliquée au niveau individuel π_{ij} , et d'un résidu individuel e_{ij} distribué selon une loi logistique :

$$P(H_{ij} < H_{\min}) = \pi_{ij} + e_{ij} \quad (6)$$

La probabilité π_{ij} est expliquée par les caractéristiques individuelles de l'individu, X_{ij} et le terme aléatoire représentant la région u_{0j} , normalement distribué : $\text{logit}(\pi_{ij}) = \alpha_0 + u_{0j} + X_{ij}\beta$ avec $u_{0j} \rightsquigarrow N(0, \sigma_u)$.

Cette procédure permet d'obtenir des coefficients non biaisés par l'autocorrélation éventuelle des résidus individuels au sein de chaque région pour les variables explicatives individuelles. Ce

est plus fort lorsque l'on prend en compte l'autocorrélation des résidus individuels au niveau de chaque état par un effet aléatoire régional, en dehors de tout contrôle par les divisions. En outre, cette étude reporte un effet significatif des inégalités sur la santé lorsque les résidus sont estimés sur trois niveaux (individu, Etat, division). Enfin, Subramaniam et al. (2003a) ne trouvent plus d'impact des inégalités lorsque les variables indicatrices des divisions sont introduites dans l'estimation, même en tenant compte d'un effet aléatoire au niveau de l'Etat. Selon Subramaniam et al., la non-significativité du niveau d'inégalités intra-Etat sur la santé déclarée montrée par Mellor et Milyo s'explique donc par l'utilisation d'effet fixe pour les divisions, puisque cette procédure revient à analyser séparément chaque zone constituée alors d'un nombre insuffisant d'Etats. En outre, ils soulignent que ce contrôle n'est pas explicitement justifié par un mécanisme supposé intervenir au niveau de la division, mais plutôt par la volonté de corriger une auto-corrélation des résidus individuels au niveau de chaque état, qui ne pouvait être traitée par l'introduction d'une variables indicatrice pour chaque état.

type de modélisation constitue une alternative intéressante aux autres types de traitement de données hiérarchisées. En estimant un seul paramètre pour l'ensemble des groupes, la variance σ_u des résidus u_{0j} , et non $n - 1$ paramètres dans le cas d'introduction d'un effet fixe pour chacun des n groupes, cette méthode permet de traiter de manière plus souple la hiérarchisation des données. Elle permet donc de prendre en compte les effets de groupe lorsque le nombre élevé de groupes, ou la faiblesse des effectifs au sein de chaque groupe, interdisent l'introduction d'effet fixe de groupe par des variables indicatrices, ou encore une procédure d'estimation groupe par groupe. En outre, cette procédure permet de diagnostiquer la présence d'un impact du contexte sur la variable expliquée. Si la variance intergroupe, c'est-à-dire la variance des résidus u_{0j} , est significativement différente de zéro, cela signifie que la région d'habitation a un impact sur la santé individuelle. Il s'agit donc, dans une deuxième étape, de comprendre les mécanismes par lesquels l'environnement joue sur la santé. Cette deuxième étape consiste alors à expliquer la variance intergroupe par des facteurs contextuels, c'est-à-dire par des variables qui caractérisent les régions.

Nous introduisons alors dans le modèle des caractéristiques régionales représentant les mécanismes supposés expliquer les différences géographiques de mortalité. Cela revient à supposer que l'état de santé H_{ij} d'un individu i appartenant au groupe j est expliqué à la fois par ses caractéristiques individuelles X_{ij} , un résidu individuel e_{ij} , une constante α_0 , des variables contextuelles C_j mesurées au niveau du chaque groupe j , et un effet de groupe aléatoire u_{0j} :

$$H_{ij} = \alpha_0 + C_j\gamma + u_{0j} + X_{ij}\beta + e_{ij} \quad (7)$$

Cette modélisation permet de tester l'impact des variables régionales sur l'état de santé individuelle ($\gamma \neq 0$), ici le niveau intra-régional des inégalités, et d'évaluer la contribution des inégalités de revenu à la variance intergroupe u_{0j} et donc aux différences géographiques de mortalité, tout en conservant un terme aléatoire pour chaque région qui permet ainsi d'obtenir une estimation non biaisée des coefficients des variables contextuelles par l'omission d'autres variables contextuelles pertinentes.

Néanmoins, nous avons retenu plusieurs caractéristiques régionales pouvant constituer des facteurs de confusion c'est-à-dire les caractéristiques régionales logiquement indépendantes du niveau des inégalités, mais qui sont en réalité corrélées à ces dernières : le taux de chômage (source INSEE) et l'offre de soins. L'offre régionale de soins est représentée par la densité de médecins, salariés et libéraux, du secteur ambulatoire (généralistes et spécialistes) pour 100 000 habitants en 1988 (Base Eco-Santé, CREDES). L'utilisation de cette variable de contrôle répond à un double objectif. D'une part, cette variable peut avoir un impact sur la santé, puisqu'une faible densité médicale peut être un facteur de diminution du recours aux soins, via une augmentation du coût des soins. Par ailleurs, la question de la répartition géographique des producteurs de soins est aujourd'hui une question centrale dans les débats politiques. Il semble donc intéressant d'évaluer l'impact sur la santé de cette répartition. Cependant, la décision d'installation des médecins sur le territoire ne se fait sans doute pas au hasard. En particulier, on peut envisager que les médecins s'installent dans les régions où la clientèle est, soit plus nombreuse, en raison par exemple d'une forte proportion de personnes âgées, soit plus riche. Nous utilisons le taux de chômage régional (source INSEE) comme indicateur du niveau économique de la région. L'utilisation d'un contrôle du niveau économique de la région par le revenu moyen n'est pas retenue ici, cette variable renvoyant davantage à la compréhension des mécanismes sous-jacents expliquant l'impact des inégalités, et notamment à l'effet des comparaisons sociales. Or comme le souligne Subramaniam et al. (2003a), l'introduction d'un nombre trop important de variables

de contrôle, et notamment de variables médiatrices de l'effet des inégalités, est à l'origine d'un surajustement du modèle, qui expliquerait la disparition de l'effet des inégalités.

4 les Résultats

L'analyse des déterminants traditionnels de la mortalité (modèle 0, tableau 4 annexe B) donne des résultats cohérents avec les analyses démographiques précédentes (e.g. Vallin, Meslé et Valkonen, 2001). Le risque de décès est croissant de l'âge. Nous retrouvons le gradient social de mortalité selon la hiérarchie des PCS, bien que la probabilité de décès des ouvriers ne soit pas significativement différente de celle des employés¹⁸. Le statut matrimonial est également corrélé au risque de décès. Les personnes mariées semblent protégées par rapport aux célibataires, aux divorcés et aux veufs. Enfin, l'existence de disparités géographiques de mortalité (Salem, Rican et Jouglà, 2000) reste confirmée dans le cadre d'une analyse multivariée puisque nos résultats montrent une surmortalité dans le Nord-Pas-de-Calais.

4.1 L'étude du lien santé-revenu

L'analyse descriptive semble montrer un lien entre mortalité et revenu, puisque les défunts ont des revenus par unité de consommation moins élevés que les survivants (annexe A). Ainsi, le revenu moyen des défunts est significativement inférieur à celui des survivants (84000 contre 97000) et la distribution des revenus des survivants domine en tout point la distribution des revenus des défunts, les seuils de chaque quintile étant supérieurs (tableau 2). En outre ces résultats ne sont pas entièrement expliqués par des effets de carrière affectant le revenu, puisque l'analyse séparée de la population des moins de 65 ans et des 65 ans et plus donnent des résultats équivalents. Lorsque les quintiles sont définis sur l'ensemble de la population vivante au 1er janvier 1988 (tableau 3), les défunts sont plus concentrés dans le bas de distribution que les survivants. En fait, la part de personnes décédées dans chaque quintile diminue tout au long de la distribution des revenus et ce, quelle que soit la population considérée (population totale, moins de 65 ans, 65 ans et plus) alors que la proportion des survivants croît avec le niveau de revenu¹⁹. Même si ces résultats sont à considérer avec précaution puisque le revenu est partiellement corrélé à l'âge, ces éléments semblent toutefois indiquer que les défunts ne sont pas "tirés aléatoirement" dans la population mais que la mort touche plus fréquemment le bas de la distribution des revenus.

4.1.1 Une forte corrélation entre mortalité et revenu

L'effet positif du revenu sur la santé est confirmé par l'analyse économétrique (annexe B). Le revenu fiscal par unité de consommation réduit significativement le risque de décès (modèle 1, tableau 4). Le revenu apparaît avoir un effet distinct de la profession sur le risque de mortalité puisque les deux indicateurs de statut socioéconomique sont significatifs (modèle 2, tableau 5),

¹⁸L'absence de différence de risque de décès entre ces deux catégories est cohérente avec les résultats des analyses démographiques précédentes. Elle s'explique tout d'abord par le fait que nous n'avons pas pu distinguer les ouvriers qualifiés des ouvriers non qualifiés. En outre, selon Mesrine (1999), ce sont les personnels de services directs aux particuliers, inclus dans la PCS des employés, et non les ouvriers, qui ont l'espérance de vie à 35 ans la plus faible.

¹⁹La distribution des défunts et des survivants au sein des différents quintiles de revenu a été calculée à partir des effectifs non pondérés par les pondérations respectives de chaque enquête.

l'introduction du logarithme du revenu par unité de consommation ne réduisant que faiblement le caractère explicatif des PCS.

Les pensions et rentes semblent être une bonne proxy de l'état de santé puisque leur coefficient est significativement positif et que leur introduction dans le modèle réduit la valeur du coefficient du logarithme du revenu (modèle 3, tableau 5). Même si l'interprétation précise en termes de causalité reste encore délicate, le revenu semble avoir un effet causal sur la mortalité puisque qu'il reste largement significatif.

Le statut matrimonial explique encore significativement la probabilité de décès. Ce résultat contredit l'hypothèse souvent avancée selon laquelle la corrélation entre mariage et longévité serait en partie due à une corrélation entre mariage et statut économique (Vallin, Meslé et Valkonen, 2001). Notre analyse ne permet pas, en revanche, de séparer les autres hypothèses expliquant ce lien. La surmortalité des célibataires peut provenir soit d'un effet de sélection par la santé sur le marché du mariage, soit d'un effet néfaste du célibat ou encore d'un effet protecteur du mariage. La surmortalité des veufs, qui ne peut pas s'expliquer par un effet de sélection, peut en revanche refléter soit un effet protecteur du mariage, soit un effet néfaste du veuvage.

La surmortalité du Nord reste significative après contrôle par le revenu. Ce résultat contredit l'hypothèse avancée par Salem, Rican et Jouglà (2000), selon laquelle cette surmortalité serait largement expliquée par un effet de composition induit par la forte proportion d'emplois manuels et les difficultés économiques de la région. Par ailleurs, on retrouve l'effet protecteur de la zone Sud-Ouest documenté par la littérature démographique après contrôle par le revenu. Ces résultats suggèrent donc d'explorer des caractéristiques régionales spécifiques, comme le niveau des inégalités ou encore l'offre de soins.

4.1.2 Elasticité de la probabilité de décès au revenu

A partir des estimations de l'impact du logarithme du revenu sur la probabilité de décès, nous avons calculé la valeur de l'élasticité de la probabilité de décès au revenu (annexe C). Cet indicateur est plus facilement interprétable que la valeur des coefficients estimés par la régression logistique. En outre, une étude récente, réalisée à partir des fichiers administratifs de l'Echantillon Interrégime, a permis de calculer, pour quelques générations de retraités, l'élasticité de la probabilité de survie à 4 ans au montant des droits à la retraite en France entre 1997 et 2001 (Bommier, et al., 2003). Même si les deux échantillons ne sont pas directement comparables, la comparaison de nos résultats aux résultats de cette enquête nous permet de valider la procédure d'enquête cas-témoin utilisée ici.

Toutefois, le calcul des élasticités n'est pas immédiat en raison de la nature peu habituelle des données utilisées ici. En effet, l'élasticité de la probabilité de décès correspond à la variation relative de la probabilité de décès p induite par une augmentation relative du revenu y .

En considérant un individu i représentatif, choisi comme disposant d'un revenu y_i , égal à la moyenne des revenus de l'échantillon, elle correspond au produit de l'effet marginal du revenu et du rapport entre le revenu moyen et la probabilité de décès associée p_i :

$$\eta = \frac{\partial p}{\partial y} \cdot \frac{y_i}{p_i}$$

Dans le cas particulier du modèle LOGIT, l'effet marginal du logarithme du revenu sur le décès peut être déduit simplement de la probabilité de décès de l'individu représentatif prédite p_i et

du coefficient β estimé correspondant au logarithme du revenu :

$$\frac{\partial p}{\partial \ln y} = p_i(1 - p_i)\beta$$

L'élasticité-revenu de la probabilité de décès correspond alors à :

$$\eta = (1 - p_i)\beta \quad (8)$$

La valeur calculée de l'élasticité est donc dépendante de la probabilité de décès prédite pour l'individu représentatif. Or dans notre cas, la part relative des défunts et des survivants ne permet pas de calculer une probabilité de décès réaliste. Comme les défunts sont environ 10 fois trop représentés dans notre enquête cas-témoin, ce calcul conduit à sous-estimer d'environ 10 % l'élasticité calculée. Deux possibilités de correction sont envisagées.

La première solution consiste à repondérer les observations pour rendre l'échantillon représentatif du taux de décès en 1988. Cette procédure ne modifie que légèrement la valeur des coefficients calculés pour les paramètres, mais modifie leur significativité, puisque cette pondération diminue la quantité d'information apportée par les défunts. Comme nous l'avons souligné précédemment, l'utilisation de cette pondération réduit la puissance statistique et élargit donc les intervalles de confiance des coefficients.

La deuxième solution consiste à approximer la valeur de l'élasticité à partir de la valeur de l'odds ratio, qui, dans le cas d'événements rares, est proche de celle du risque relatif.

On peut en effet calculer une estimation η^e de l'élasticité η pour un petit accroissement Δy du revenu autour de la valeur y_i :

$$\eta^e = \frac{P(y_i + \Delta y) - P(y_i)}{P(y_i)} \times \frac{\Delta y}{y_i}$$

Pour cet accroissement Δy , l'odds ratio $O.R_i$ s'écrit :

$$O.R_i = \frac{P(y_i + \Delta y)}{P(y_i)} \times \frac{1 - P(y_i)}{1 - P(y_i + \Delta y)}$$

On a donc :

$$\eta^e = \frac{1 - P(y_i)}{1 - P(y_i + \Delta y)} \times (O.R_i - 1) \times \frac{\Delta y}{y_i}$$

Comme la probabilité de décès est très proche de 0, on a donc :

$$\eta^e \simeq (O.R_i - 1) \times \frac{\Delta y}{y_i}$$

Ces élasticités ont été calculées séparément pour l'échantillon total, l'échantillon restreint aux moins de 65 ans et l'échantillon restreint aux 65 ans et plus, pour un individu disposant du revenu moyen de l'échantillon, employé, marié, résidant dans le Bassin Parisien et âgé de 65 ans pour l'échantillon total, 55 ans parmi les moins de 65 ans et 75 ans parmi les plus de 65 ans. Les deux méthodes proposées ont été appliquées à l'échantillon brut et à l'échantillon repondéré par le taux de décès en 1988. Pour estimer l'élasticité à partir de la valeur de l'odds ratio, nous avons calculé l'odds ratio $O.R_i$, correspondant, pour cet individu représentatif, à un accroissement du revenu annuel par unité de consommation de 12000FF²⁰.

²⁰En effet, l'odds ratio associé à chaque variable dans les résultats de la régression, noté $O.R$, est égal à l'exponentiel du coefficient β . Dans le cas particulier de la spécification du revenu en logarithme, l'odds ratio correspond à la modification du rapport de chance induite par une augmentation d'environ 272% du revenu et ne permet donc pas d'approcher la valeur de la dérivée de la fonction de probabilité.

Les deux méthodes de calcul des élasticités donnent des résultats assez proches. Sans contrôle par la profession (modèle 1, tableau 6), l'élasticité de la probabilité de décès au revenu s'établit à environ -0.4 pour la population totale, -0.5 pour les moins de 65 ans et -0.3 pour les plus de 65 ans. Le risque de décès est donc fortement associé revenu, quel que soit le groupe d'âge, même si cette corrélation décroît avec l'âge. La décroissance de l'impact du revenu avec l'âge peut sans doute s'expliquer par un effet de sélection par la santé réduisant la part de la population à risques aux âges élevés. Ce résultat contredit partiellement les résultats de McDonough et al. (1997) qui trouvent une relation très fortement réduite au-delà de 65 ans. On peut alors envisager que le plus grand différentiel de couverture du risque maladie des plus pauvres entre ces deux groupes d'âges aux Etats Unis peut expliquer une décroissance avec plus marquée de l'effet du revenu sur le risque de décès. En effet, l'accès généralisé à la couverture maladie Medicare à partir de 65 ans a sans doute un effet plus fort sur la réduction des inégalités sociales de santé aux Etats Unis, où les plus pauvres n'ont pas toujours accès à Medicaid avant 65 ans. Ces élasticités peuvent apparaître assez élevées. Cependant, ces valeurs restent encore inférieures aux élasticités estimées par Bommier et al. (2003), qui s'échelonnent entre -0.62 pour les hommes âgés de 67 ans à -0.14 pour les hommes âgés de 91 ans.

L'introduction des PCS dans le modèle 2 réduit d'environ 0.1 la valeur des élasticités (modèle 2 et tableau 7). L'élasticité de la probabilité de décès est alors d'environ -0.3 pour la population totale, -0.4 pour les moins de 65 ans et enfin de -0.2 pour les plus de 65 ans. En revanche, l'introduction des pensions et des rentes (modèle 3 et tableau 8) modifie très peu la valeur des élasticités, la prise en compte de cette variable ne réduisant que d'environ 0.02 la valeur absolue des élasticités calculées.

4.1.3 Un impact de la pauvreté et de la richesse

La comparaison des modèles 3 à 7 permet enfin d'analyser plus particulièrement la forme de la relation entre mortalité et revenu (annexe D).

Les résultats font apparaître une corrélation entre pauvreté et santé. La spécification en quintiles (modèles 6 et 7, tableau 10) met en évidence un risque spécifique lié à la pauvreté, qui peut sans doute s'expliquer par une différence d'accès aux soins, indépendamment des conditions de travail et des habitudes culturelles, approximées par les PCS. Le risque de mortalité du premier quintile est significativement plus élevé que le risque de décès de tous les autres niveaux de revenus, et ce dans chaque groupe d'âges. La valeur des odds ratios associés aux quintiles de revenu permet en outre de mieux percevoir l'ampleur des inégalités sociales de santé en France. Sans contrôle par la profession, le risque de décès du premier quintile est environ 2.5 fois plus élevé que le risque de décès du cinquième quintile (modèle 6, tableau 10). Après ajustement par la PCS, le risque relatif du premier quintile reste encore 2 fois plus élevé (modèle 7, tableau 10). Cependant, nos résultats invalident l'hypothèse de pauvreté absolue, puisque l'effet du revenu ne se limite pas au bas de la distribution. Conformément à l'analyse de McDonough et al. (1997), la probabilité de décès décroît tout au long de la distribution des revenus (modèles 6 et 7, tableau 10).

L'hypothèse de rendements décroissants de la fonction de production de capital santé, avancée par Erlich et Chuma (1990) et qui serait à l'origine de l'artefact statistique de Gravelle (1996), n'est pas non plus vérifiée. La spécification sous forme logarithmique (modèles 2 et 3, tableau 5) et la spécification sous forme quadratique (modèle 4, tableau 9) pourraient suggérer une relation concave entre le revenu et la santé, puisque les coefficients du revenu et du revenu au carré sont

de signes opposés²¹. Cependant, les spécifications sous forme polynomiale de degrés 3 (modèle 5, tableau 9) et surtout sous forme de quintiles (modèles 6 et 7, tableau 10) invalident cette hypothèse. La forme cubique suggère que la relation entre santé et revenu est en fait de forme logistique. Les résultats du modèle 6 confirment que l’effet marginal du revenu ne se réduit qu’au milieu de la distribution puisque les risques de mortalité des quintiles 2 et 3 ne sont pas significativement différents. Au delà, nous mettons en évidence une sous-mortalité associée aux revenus les plus élevés (quintiles 4 et 5) dans tous les groupes d’âges étudiés, le risque de décès associé au quatrième quintile étant même significativement supérieur au risque de décès du cinquième quintile, en l’absence de contrôle par la profession (modèle 6, tableau 10). Cet effet plus inattendu de la richesse n’a jusqu’à présent pas été mis en évidence dans la littérature. Il ne semble pas être dû à un biais d’échantillonnage entre les deux enquêtes, puisque les résultats par quintiles restent robustes à l’exclusion des outliers²². La question n’est alors plus seulement de comprendre pourquoi la pauvreté tue mais pourquoi la richesse protège.

4.2 Les inégalités, source d’externalités en santé

La première étape de l’analyse d’un effet du contexte sur le risque individuel de décès consiste à évaluer la variabilité interrégionale de mortalité à l’aide d’une analyse multi-niveaux (annexe E).

Nous avons tout d’abord estimé un modèle “vide”, c’est-à-dire en excluant toutes les variables explicatives. Ce premier modèle estime donc simplement la valeur de la constante et la variance des résidus régionaux. Il permet de constater que, bien que d’une valeur faible, la variance régionale de la probabilité de décès est significativement différente de zéro (tableau 12). Ensuite, nous avons inclus les différentes caractéristiques individuelles utilisées précédemment mais en excluant les variables indicatrices des zones géographiques (modèle 8, tableau 13). L’ajustement par les variables individuelles réduit la valeur estimée de la variance des résidus régionaux u_{0j} . Cependant, elle reste tout de même significativement différente de zéro. Cela signifie qu’il subsiste des différences de mortalité entre régions après correction des effets de composition, et donc que la recherche de facteurs contextuels est pertinente. L’introduction d’un terme aléatoire au niveau régional modifie peu le pouvoir explicatif des variables individuelles. Nous avons comparé les coefficients des paramètres, leur écart-type, les statistiques de Student et les odds ratios associés obtenus avec l’analyse multi-niveaux (logiciel MLwiN) aux valeurs obtenues en estimant le modèle par un logit classique (logiciel SAS). La prise en compte de l’autocorrélation des résidus au sein des régions renforce légèrement la significativité des déterminants individuels de l’état de santé, puisque les statistiques de Student associées à chaque variable sont légèrement plus fortes. En particulier, le risque du décès du quatrième quintile devient significativement supérieur à celui du cinquième quintile, même après ajustement par les professions.

L’analyse de l’impact des différentes variables contextuelles pouvant expliquer la variance inter-régions du risque de décès est réalisée en deux étapes. Une première étape, plus exploratoire, consiste uniquement en une régression logistique classique destinée à sélectionner les variables contextuelles pertinentes. Dans une seconde étape, la robustesse de l’impact des variables sélectionnées est testée par une analyse multi-niveaux.

²¹Dans un modèle LOGIT, il est normal que les coefficients des variables continues soient proches de zéro et que leurs odds ratios soient très proches de 1. L’interprétation est donc fondée sur le signe des coefficients ainsi que sur leur niveau de significativité.

²²L’estimation sur un échantillon réduit aux 98 premiers centiles de la distribution des revenus par UC dans l’enquête sur les Revenus Fiscaux donne des résultats équivalents.

4.2.1 Analyse exploratoire

Dans le modèle 9 (tableau 14), le niveau des inégalités, le taux de chômage et l'offre de soins, sont introduits séparément. Les résultats ne font alors apparaître aucun impact significatif sur le risque de décès de ces dimensions lorsqu'elles sont introduites seules dans le modèle, après contrôle par les déterminants individuels de l'état de santé.

En revanche, l'introduction simultanée de ces facteurs dans le modèle 10 (tableaux 14, 15) permet de mettre en évidence deux effets attendus au regard de la littérature : le risque de décès est significativement croissant avec le niveau des inégalités intra-régionales et significativement décroissant avec la densité médicale.

L'offre de soins apparaît donc être un facteur de confusion important de l'impact des inégalités de revenu sur la santé. Cet effet est expliqué par le fait que les régions les plus inégalitaires ont aussi une densité de médecins plus élevée (Ile-de-France, PACA). Après contrôle par l'offre de soins, il existe un effet significatif des inégalités sur le risque de décès qui est robuste au choix de l'indicateur (tableau 15). Cependant, le degré de significativité est contingent à ce dernier. En particulier, l'indice de Gini est assez peu corrélé avec le risque de décès.

Nos résultats suggèrent un effet positif de l'offre de soins ambulatoires sur la santé individuelle. En particulier, la sous-mortalité de la région Midi-Pyrénées et de la zone Méditerranée, documentée par la littérature démographique, semble largement être expliquée par la forte densité de médecins, bien que l'interprétation causale soit délicate. La répartition géographique des médecins sur le territoire ne serait donc pas neutre, y compris dans les pays développés, où les plus faibles densités sont souvent jugées satisfaisantes.

Par contre, le taux de chômage n'est pas corrélé avec le risque de décès. Même si le taux de chômage de la région n'a pas d'impact contextuel, on pouvait attendre un impact significatif de cette variable, en tant que proxy du statut d'activité individuel, puisque les données utilisées dans cette étude ne nous ont pas permis de tester l'impact²³ du statut de chômeur sur le risque de décès, pourtant largement documenté (Mesrine, 2000). La relation trouvée jusqu'à présent entre probabilité de décès et statut socioéconomique (revenu et profession) pouvait en partie cacher la relation existante entre chômage et la santé, en raison, d'une part, de la décroissance du risque de chômage avec la hiérarchie professionnelle, et d'autre part, de la perte de revenu induite par le chômage à profession donnée. La non-significativité des taux de chômage régionaux, conjuguée au maintien de l'impact du revenu et des catégories sociales sur le risque de décès, confortent donc la robustesse des résultats trouvés jusqu'ici.

4.2.2 Analyse multi-niveaux

L'analyse multi-niveaux permet de confirmer l'absence d'effet des inégalités, de l'offre de soins ou encore du taux de chômage, lorsque l'on introduit dans le modèle une seule variable contextuelle (modèle 9 bis, tableaux 16 et 17). En conséquence, la variance des résidus régionaux est toujours significativement différente de zéro (tableau 17). Le modèle 10 bis offre, en revanche, des résultats assez différents de ceux du modèle 10 (tableaux 16 et 18). L'analyse multi-niveaux réduit fortement la significativité du niveau des inégalités après contrôle par l'offre de soins et le niveau de chômage. Nous retrouvons toutefois un effet délétère des inégalités lorsque celles-ci sont mesurées par la variance des logarithmes et par l'indicateur d'Atkinson (norme -0.5). Par contre, l'effet de l'offre de soins est robuste à l'introduction des effets aléatoires régionaux.

²³Il existe en fait une relation duale entre chômage et santé. La dégradation de l'état de santé augmente le risque de chômage (Sauré-Cubizolles et al., 2001). Par ailleurs, le chômage a un effet pathogène (Bungener et Pierret, 1994).

Pour comprendre les différences induites par les méthodes d'estimation, l'impact du niveau des inégalités est testé dans le modèle 11, après contrôle par l'offre de soins et les caractéristiques individuelles, en spécifiant le niveau des inégalités, non plus comme une variable continue, mais en 3 groupes de 7 régions correspondant à des niveaux croissants d'inégalités (inégalités faibles, moyennes et fortes). Cette spécification permet non seulement de vérifier que l'impact trouvé des inégalités ne résulte pas de l'effet spécifique d'une région particulière, telle que l'Ile-de-France ou de la région P.A.C.A., où les inégalités de revenu sont très élevées, mais également de tester, sans imposer une forme linéaire, l'hypothèse de croissance du risque de décès avec le niveau des inégalités (Kennedy et al., 1998). L'estimation est réalisée à l'aide d'une régression logistique classique et par une analyse multi-niveaux.

Le modèle 11 montre alors l'existence d'une corrélation entre inégalités de revenu et mortalité après contrôle par l'offre de soins, quelle que soit la méthode d'estimation utilisée (tableaux 19 et 20). Pour tous les indicateurs, le risque de décès dans les régions les plus égalitaires est significativement inférieur au risque de décès des autres régions. En revanche, le risque de décès des régions les plus inégalitaires n'est pas significativement supérieur à celui des régions ayant un degré moyen d'inégalités. Lorsque le niveau d'inégalités est mesuré par l'indicateur d'Atkinson (normes 0.5, 0.25 et -0.5), l'odds ratio associé aux régions les plus inégalitaires apparaît même inférieur à celui associé aux régions où les inégalités ont été jugées moyennes, mais les différences ne sont pas significatives. Cette non-linéarité de l'effet des inégalités peut alors expliquer la quasi absence d'impact des inégalités lorsque le niveau des inégalités est introduit par une variable continue.

L'introduction du niveau des inégalités en tranche et de l'offre de soins réduit fortement la valeur de la variance interrégionale résiduelle (tableau 21). Toutefois, la contribution des inégalités à l'explication des différences régionales de décès diffère selon l'indicateur d'inégalités utilisé. L'effet des inégalités est plus marqué lorsque le niveau des inégalités est évalué par l'indice de Theil et l'indice d'Atkinson (normes 0.5, 0.25 et -0.5). Dans ces cas, la variance interrégionale ne peut plus être considérée comme significativement différente de zéro. Sans pouvoir conclure à l'existence d'un impact causal des inégalités et de l'offre de soins sur l'état de santé, ces résultats indiquent que ces deux variables sont corrélées au risque individuel de décès et que les corrélations ont le signe attendu au regard de la littérature. Par ailleurs, les inégalités et l'offre de soins semblent expliquer une part importante des différences régionales de mortalité en France.

5 Discussion

A l'issue de cette recherche, il semble donc que l'hypothèse de Wilkinson (1996), selon laquelle "la santé est fortement affectée par l'étendue des différences économiques et sociales au sein d'une population" est validée dans le cas de la France. Ce résultat peut paraître surprenant au regard de la littérature. En 2002, Mackenbach écrivait que "l'hypothèse d'inégalités de revenu" était un bon exemple de démarche scientifique. Pour lui, cette hypothèse devait être réfutée, car l'effet des inégalités sur la santé n'avait pu être montré que sur données américaines. En 2003, Subramaniam et al. (2003a) soulignaient que, avant d'être réfutée, cette hypothèse devait être testée dans d'autres pays, et en particulier des pays plus inégalitaires. Ainsi, un effet significatif des inégalités sur la santé a été mis en évidence au Chili (Subramaniam et al., 2003b). Dans notre étude, l'effet des inégalités de revenu n'a pu être mis en évidence qu'après contrôle par l'offre régionale de soins. Ce résultat semble montrer que le manque de prise en compte de facteurs

de confusion peut expliquer l'absence de validation empirique de cette hypothèse en dehors des Etats-Unis. En particulier, l'offre de soins n'a pas, à notre connaissance, été utilisée dans les études précédentes.

Ces résultats ne permettent pas, en revanche, de comprendre les mécanismes par lesquels les inégalités ont un effet sur la santé. Deux mécanismes sont fréquemment évoqués pour expliquer cet effet. D'une part, les inégalités sont à l'origine d'une externalité négative induite par un niveau plus faible de capital social, et par la moindre cohésion sociale des sociétés inégalitaires (Kawachi et Berkman, 2000). D'autre part, les inégalités de revenu ont un impact sur la santé via le stress psychosocial induit par les comparaisons sociales (Wilkinson, 1997, Kawachi, 2000). Ainsi, les plus pauvres souffriraient plus encore de leur faible niveau de vie en présence de personnes disposant d'un niveau de vie plus élevé.

Ces deux hypothèses sont difficiles à discriminer sans indicateur spécifique. Toutefois, pour tester plus précisément l'hypothèse de capital social, nous avons testé l'impact du niveau régional de criminalité sur le risque de décès, puisque la criminalité serait un symptôme de moindre capital social selon Kawachi et al. (1999). La criminalité est estimée par le nombre de crimes et délits constatés par les services de la police nationale et de la gendarmerie, pour 1 000 habitants, en 1988 (Ministère de l'Intérieur, 1990). Les résultats des modèles 12 et 13 (tableau 22) ne mettent en évidence aucun impact significatif de la criminalité, avec ou sans contrôle par l'offre de soins, que le modèle soit estimé par une régression logistique classique ou par une analyse multi-niveaux (Jusot, 2003). L'existence de phénomène de comparaisons sociales paraît donc plus satisfaisante pour expliquer l'effet négatif des inégalités car cette hypothèse est en fait très cohérente avec l'ensemble nos résultats.

Cette étude met en évidence une forte corrélation entre le risque de décès et le niveau de revenu, indépendamment des effets d'éducation et de conditions de travail, traduits par les catégories socioprofessionnelles. Cette relation existe tout au long de la distribution des revenus. En particulier, nous avons mis en évidence une sous-mortalité associée aux plus hauts revenus. Selon les modèles de capital santé, l'effet du revenu sur la santé passe par une plus grande consommation de soins. Les analyses empiriques de la consommation de soins, réalisées dans des pays où pour l'essentiel l'accès aux soins est garanti à tous, ne semblent pas indiquer de corrélation très forte entre le montant total des soins consommés et le revenu. De plus l'effet de la consommation de soins sur l'état de santé est difficile à mettre en évidence, sauf pour les bas revenus ou pour les personnes en très mauvaise santé. La relation trouvée ici entre revenu et santé paraît donc difficilement attribuable en totalité à une sous-consommation de soins puisque d'une part, les pathologies les plus lourdes sont systématiquement couvertes par la sécurité sociale et que d'autre part les problèmes d'accès à une couverture complémentaire concernent surtout le bas de la distribution (Dourgnon et al, 2001).

Si les difficultés d'accès aux soins peuvent permettre de comprendre pourquoi la pauvreté tue, il nous semble que les différences de conditions matérielles semblent plus difficilement expliquer pourquoi la richesse protège. Ce phénomène peut alors, en partie, s'interpréter comme un impact de la position sociale sur le risque de décès puisque le revenu est le reflet du rang de l'individu dans la société, au-delà du montant de ses ressources disponibles (Nauze-Fichet et Tomasini, 2002). Par ailleurs, nous avons mis en évidence un impact des différentes catégories sociales sur le risque de décès et non simplement une surmortalité des ouvriers. Ce résultat est alors cohérent avec les effets pathogènes du stress induit par la hiérarchie et le manque de contrôle de l'environnement, documentés par la littérature épidémiologique (Marmot, 2000). En outre, une analyse menée séparément selon la catégorie sociale a permis de montrer que le risque de décès est décroissant avec le revenu au sein de chaque groupe (Jusot, 2003). Si l'on

considère que la PCS constitue un groupe de référence pertinent pour l'individu, ce résultat peut s'expliquer par la frustration induite par la comparaison de son revenu à celui de ses pairs. Enfin, les résultats du modèle 14 (tableau 23) montrent que l'état de santé des personnes disposant d'un revenu inférieur au revenu médian est davantage affecté par le niveau des inégalités, et par l'offre de soins, que celui des personnes les plus riches. Or ce résultat est cohérent avec l'hypothèse de déprivation relative, selon laquelle l'état de santé d'un individu est affecté par la somme des revenus supérieurs au sien, puisqu'à revenu moyen constant, une augmentation des inégalités détériore l'état de santé des plus pauvres, l'écart entre leur revenu et le revenu du plus riche étant plus important.

L'ensemble de ces résultats appelle donc de futures recherches sur l'effet des interactions sociales sur la santé²⁴.

6 Bibliographie

Adams P. Hurd M.D., McFadden D., Merrill A., Ribeiro T. (2003), "Healthy, Wealthy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status", *Journal of Econometrics*, 112, 1 : 3-56.

Barker D.J.P. (1997), "Maternal Nutrition, Fetal Nutrition and Disease in Later Life", *Nutrition*, 13, 9 : 807-813.

Bocognano A. et al. (1999), "Santé, Soins et Protection Sociale en 1998", *rapport CREDES, série Résultats*, 1282.

Bommier A., Magnac T., Rapoport B., Roger M., (2003), "Droits à la retraite et mortalité différentielle", *Document de travail LEA*, 0303.

Bouyer J. et al. (1995), *Epidémiologie : Principes et méthodes quantitatives*, Les Editions INSERM.

Breuil-Genier P., Grandfils N., Raynaud D. (1999), "Revenus, assurance et santé : le problème de l'accès aux soins des plus démunis", *Les cahiers du Gratice* 15 : 243-76.

Bungener M., Pierret J. (1994), "De l'influence du chômage sur l'état de Santé", in *Trajectoires sociales et Inégalités*, eds Bouchayer F. et Verger D, Recherche sur les conditions de vie, ERES : 43-61.

Campagne N., Contencin D., Roineau C. (1996), "Les Revenus Fiscaux des Ménages en 1990", *Insee Résultats*, 453, *Emploi-Revenus*, 103.

Case A., Lubotsky D., Paxson C. (2002), "Economic Status and Health in Childhood : The Origins of the Gradient", *The American Economic Review*, 92, 5 : 1308-34.

Chaix B., Chauvin P. (2002), "L'apport des modèles multiniveau dans l'analyse contextuelle en épidémiologie sociale : une revue de la littérature", *Revue d'Epidémiologie et de Santé Publique*, 50 : 489-499.

Cornfield J. (1951), "A Method of Estimating Comparative Rates from Clinical Data : Application to Cancer of the Lung, Breast, and Cervix", *Journal of the National Cancer Institute*, 11 : 1269-1275.

Couffinal A., Dourgnon P., Geoffard P.Y., Grignon M., Jusot F., Naudin F. (2002a), "Rôle de la Couverture Maladie dans l'Insertion sur le Marché du Travail",
http://www.credes.fr/En_ligne/WorkingPaper/pdf/couvmaltrav.pdf.

²⁴L'introduction de questions portant sur les déterminants sociaux de la santé dans l'enquête Santé Protection Sociale de l'année 2004 permettra sans doute de déterminer l'impact respectif sur la santé du capital social et des effets de comparaisons sociales.

- Couffinal A., Dourgnon P., Geoffard P.Y., Grignon M., Jusot F., Naudin F. (2002b), "Comment évaluer l'impact de la complémentaire CMU sur l'emploi?", *Question d'Economie de la Santé, Synthèse*, 59, CREDES.
- Cox D.R., Oakes D. (1984), *Analysis of Survival Data*, Chapman & Hall.
- Deaton A. (2001), "Health, Inequality and Economic Development", *NBER WP* 8318.
- Diez-Roux A.V. (2000), "Multilevel Analysis in Public Health Research", *Annual Review of Public Health*, 21 : 171-92.
- Dourgnon P., Grignon M., Jusot F.(2001), "L'assurance maladie réduit-elle les inégalités sociales de santé?", *Question d'Economie de la Santé, Synthèse*, 43, CREDES.
- Duleep H.O. (1986), "Measuring the Effect of Income on Adult Mortality Using Longitudinal Administrative Record Data", *Journal of Human Resources*, 21, 2 : 238-251.
- Erlich I., Chuma H. (1990), "A model of the demand for longevity and the value of the life extension", *Journal of Political Economy*, 98 : 761-782.
- Grandfils N., Le Fur P. (1993), "Les inactifs pour raison de santé", *rapport CREDES*, n°962.
- Gravelle H. (1996), "How much of the relation between population mortality and unequal distribution of income is a statistical artefact", *British Medical Journal*, 316 : 382-385.
- Grignon M., Polton D., (2000), "Inégalités d'accès et de recours aux soins", in *Mesurer les inégalités*, DRESS : 188-200.
- Grossman M. (1972), "On the concept of health capital and the demand for health", *Journal of Political Economy*, 80 : 223-255.
- Grossman M. (2000), "The Human Capital Model", in *Handbooks of Health Economics*, eds Culyer A.J., Newhouse J.P. : 348-408.
- Haveman R. et al. (1994), "Market work, wages and men's health", *Journal of Health Economics*, 13, 2 : 163-182.
- INSEE (1995), "Revenus et Patrimoine des ménages, édition 1995", *Synthèses*, 1.
- Jusot F. (2003), *Revenu et Mortalité : Analyse économique des inégalités sociales de santé en France*, Thèse de doctorat de l'EHESS, soutenue le 12 décembre 2003.
- Kaplan G.A., Pamuk E.R., Lynch J.W., Cohen R.D., Balfour J.L. (1996), "Inequality in income and mortality in the United States : analysis of mortality and potential pathways", *British Medical Journal*, 312 : 999-1003.
- Kawachi I. (2000), "Income Inequality and Health", in *Social Epidemiology*, eds Berkman L.F. et Kawachi I., Oxford University Press.
- Kawachi I., Berkman L. (2000), "Social Cohesion, Social Capital, and Health", in *Social Epidemiology*, eds Berkman L.F. et Kawachi I., Oxford University Press.
- Kawachi I., Kennedy B.P. (1997 a), "The Relationship of Income Inequality to Mortality : Does the Choice of Indicator Matter?", *Social Science and Medicine*, 45, 7 : 1121-27.
- Kawachi I., Kennedy B.P., Lochner K., Prothrow-Stith D. (1997 b), "Social Capital, Income Inequality, and Mortality", *American Journal of Public Health*, 87, 9 : 1491-98.
- Kawachi I., Kennedy B.P., Wilkinson R.G. (1999), "Crime : social disorganization and relative deprivation", *Social Science and Medicine*, 48, 6 : 719-731.
- Kawachi I., Subramanian S.V., Almeida-Fiho N.(2002), "A glossary for health inequalities", *Journal of Epidemiology and Community Health*, 56 : 647-652.
- Kennedy B.P., Kawachi I. Glass R., Prothrow-Stith D. (1998), "Income distribution, socioeconomic status, and self rated health in the United States : multilevel analysis", *British Medical Journal*, 317 : 917-21.
- Laferrère A., Monteil P. (1995), "Le Patrimoine au décès en 1988", *INSEE Résultats, Consommations - Modes de Vie*, 390, 71.

- Lechene V., Magnac T. (1994), "Analyse des déterminants des salaires", in *Trajectoires sociales et Inégalités*, eds Bouchayer F., Verger D, Recherche sur les conditions de vie, ERES : 221-243.
- Legrand J.L. (1987), "Inequalities in Health. Some International Comparisons", *European Economic Review*, 31 : 182-191.
- Leibenstein H. (1957), "The Theory of Underemployment in Backward Economies", *Journal of Political Economy*, 65, 2 : 91-103.
- Mackenbach J.P. (2002), "Income Inequality and Population Health : Evidence favouring a negative correlation between income inequality and life expectancy has disappeared", *British Medical Journal*, 2002, 324 :1-2.
- Manski C.F. (1995), "Response-Based Sampling", in *Identification Problems in the Social Sciences*, Cambridge : Harvard University Press.
- Marmot M. (2000), "Multilevel Approches to Understanding Social Determinants", in *Social Epidemiology*, eds Berkman L., Kawachi I., Oxford University Press.
- McDonough P. Duncan G.J., Williams D., House J.(1997), "Income Dynamics et Adult Mortality in United States, 1972 through 1989", *American Journal of Public Health*, 87, 9 : 1476-83.
- Mellor J., Milyo J. (1998), "Income Inequality and Health Status in the United States : Evidence from the Current Population Survey", *Tufts University, Economics Department, WP* 98-15.
- Mellor J., Milyo J. (1999), "Income Inequality and Health Status in the United States : Evidence from the Current Population Survey", *Princeton, NJ : Robert Wood Johnson Found.*
- Mellor J., Milyo J. (2001), "Re-examining the evidence of an ecological association between income inequality and health", *Journal of Health Politics, Policy and Law*, 26, 3 : 487-522.
- Mellor J., Milyo J. (2003), "Is Exposure to Income Inequality a Public Health Concern? Lagged Effects of Income Inequality on Individual and Population Health", *Health Services Research*, 38, 1 : 137-151.
- Mesrine A. (1999), "Les différences de mortalité par milieu social restent fortes", *La Société Française, Données Sociales* : 228-35 .
- Mesrine A. (2000), "la surmortalité des chômeurs : un effet catalyseur du chômage", *Economie et Statistique*, 334, 2 : 33-48.
- Ministère de l'Intérieur (1990), *Aspects de la criminalité et de la délinquance constatés en France en 1989 - par les services de police et de gendarmerie*, La Documentation Française.
- Nauze-Fichet E., Tomasini M. (2002), "Diplôme et Insertion sur le marché du travail : approches socio-professionnelle et salariale du déclassement", *Economie et Statistique*, 354 : 21-43.
- Newhouse J.P. (1993), *Free for all? Lessons from the Rand experiment*, Harvard University Press.
- Preston S.H. (1975), "The changing relation between mortality and level of economic development", *Population Studies*, 29 : 231-48.
- Preston S.H., Elo I.T., Stewart Q. (1999), "Effects of Age Misreporting on Mortality Estimates at older Ages", *Population Studies*, 53 : 165-177.
- Putnam R.D. (1995), "Bowling alone : America 's declining social capital", *Journal of Democracy*, 6 : 65-78.
- Rasbash J. et al.(2002), "A user's guide to MLwiN", *Center for Multilevel Modelling, Institute of Education*, University of London.
- Ravelli A. et al.(1998), "Glucose tolerance in adults after prenatal exposure to famine", *Lancet*, 351 : 173-76.

- Rodgers G.B. (1979), "Income and inequality as determinants of mortality : an international cross-section analysis", *Population Studies*, 33 : 343-51.
- Salem G., Rican S., Jouglu E. (2000), *Atlas de la santé en France, volume 1, les causes de décès*, John Libbey Eurotext.
- Saurel-Cubizolles M.J. et al. (2001), "Etat de santé perçu et perte d'emploi ", in *Travail-Santé-Vieillesse : Relation et Evolution*, ed Cassou B., Edition Octares.
- Schlesselman J.J.(1982), *Case-control studies. Design, control, analysis*, Oxford University Press.
- Smith J.P., Kington R. (1997), "Demographic and Economic Correlates of Health in Old age", *Demography*, 34, 1 : 159-170.
- Snyder S., Evans W. (2002), "The Impact of Income on Mortality : Evidence from the Social Security Notch", *NBER WP* 9197.
- Subramaniam S.V., Blakely T., Kawachi I. (2003 a), "Income Inequality as a Public Health Concern : Where do we stand? ", *Health Services Research*, 38, 1 : 153-167.
- Subramaniam S.V., Delgado L., Jadue L., Vega J., Kawachi I. (2003 b). "Income inequality and health : multilevel analysis of Chilean communities ", *Journal of Epidemiology and Community Health*, forthcoming.
- Townsend P., Davidson (1982), *Inequalities in Health : The Black Rapport*, London : Penguin.
- Townsend P., Davidson, Whitehead M. (1988), *Inequalities in Health : The Health Divide*, London : Penguin.
- Vallin J., Meslé F., Valkonen T. (2001), *Tendances en matière de mortalité et mortalité différentielle*, Editions du Conseil de l'Europe.
- Wadsworth M. (1999), "Early Life", in *Social Determinants of Health*, eds Marmot M., Wilkinson R.G., Oxford University Press : 44-63.
- Wagstaff A., van Doorslaer E. (2000), "Income Inequality and Health : What Does the Literature Tell Us? ", *Annual Review of Public Health*, 21 : 543-67.
- Wilkinson R.G. (1992), "Income distribution and life expectancy", *British Medical Journal*, 304 : 165-68.
- Wilkinson R.G. (1996), *Unhealthy Societies : the Afflictions of Inequality*, Routledge, London.
- Wilkinson R.G. (1997), "Income, Inequality and Social Cohesion ", *American Journal of Public Health*, 87, 9 : 953-957.
- Wolfson M., Kaplan G., Lynch J., Ross N., Backlund E. (1999), "Relation between income inequality and mortality : empirical demonstration", *British Medical Journal*, 319 : 953-957.

A Analyse descriptive

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Total
Ouvriers	1671 (32.9%)	1395 (27.5%)	1117 (22.0%)	683 (13.4%)	214 (4.2%)	5080
Employés	336 (17.0%)	465 (23.5%)	485 (24.5%)	448 (22.6%)	247 (12.5%)	1981
Intermédiaires	304 (9.1%)	484 (14.5%)	675 (20.3%)	959 (28.8%)	908 (27.3%)	3330
Cadres	100 (3.3%)	115 (3.8%)	239 (7.9%)	657 (21.8%)	1897 (63.1%)	3008
Total	2411	2459	2516	2747	3266	13399

Tableau 1 : Effectif non pondéré de chaque PCS par quintile de revenu par UC

Revenu par UC	Population totale		Moins de 65 ans		65 ans et plus	
	Survivants	Défunts	Survivants	Défunts	Survivants	Défunts
moyenne	97198	83842	96032	77013	103019	87017
écart-type	69737	56186	67888	52885	78079	57406
effectif	11961	1438	10282	447	1679	991
Décile 1 (seuil max)	37893	37727	36471	27321	46514	41484
Décile 2	51471	48771	49934	40913	58129	52039
Décile 3	62413	56627	61438	51167	67993	59353
Décile 4	72721	64235	72061	57250	75775	67364
Médiane	83027	73062	82724	65216	84727	76187
Décile 6	94416	81588	93890	74998	96437	83804
Décile 7	109357	92116	109233	85500	109891	94440
Décile 8	129231	109852	128962	105242	130367	111411
Décile 9	166923	134864	166168	132104	170721	137163
Décile 10	1727560	816855	1727560	612009	1176817	816855

Tableau 2 : Distribution des revenus par UC selon le décès

Revenu par UC	Population totale		Moins de 65 ans		65 ans et plus	
	Défunts	Survivants	Défunts	Survivants	Défunts	Survivants
Quintile 1	21.8%	17.5%	27.7%	17.3%	25.9%	17.6%
Quintile 2	25.9%	17.4%	27.7%	17.6%	21.7%	17.9%
Quintile 3	21.7%	18.4%	18.3%	18.4%	21.2%	18.7%
Quintile 4	16.4%	21.0%	13.0%	21.0%	16.6%	21.1%
Quintile 5	14.2%	25.6%	13.2%	25.8%	14.5%	24.7%
Ensemble	1438	11961	447	10282	991	1679

Tableau 3 : Proportion de défunts et de survivants (hors pondération) par quintile de revenu (calculés séparément selon le groupe d'âge)

B Effet du revenu sur la mortalité

Population Totale		Modèle 0		Modèle 1		
Explicatives	Coef ²⁵	O.R	IC 95%	Coef	O.R	IC 95%
Log revenu				-0.466***	0.628	(0.569 - 0.692)
Ouvrier	1.035***	2.814	(2.177 - 3.637)			
Employé	1.069***	2.911	(2.208 - 3.838)			
Intermédiaire	0.776***	2.173	(1.641 - 2.878)			
Cadre	ref	1				
moins de 40	0.046**	1.047	(1.008 - 1.087)	0.042**	1.043	(1.004 - 1.083)
40-54	0.112***	1.119	(1.091 - 1.147)	0.118***	1.125	(1.098 - 1.154)
55-64	0.076***	1.079	(1.050 - 1.109)	0.081***	1.085	(1.055 - 1.115)
65-74	0.057***	1.059	(1.031 - 1.087)	0.059***	1.060	(1.033 - 1.088)
75-84	0.123***	1.13	(1.094 - 1.168)	0.114***	1.121	(1.085 - 1.158)
85 et plus	0.074**	1.077	(1.003 - 1.156)	0.074**	1.077	(1.003 - 1.156)
Marié	ref	1		ref	1	
Célibataire	0.483***	1.62	(1.288 - 2.038)	0.438***	1.533	(1.217 - 1.933)
Veuf	0.702***	2.018	(1.636 - 2.491)	0.731***	2.077	(1.685 - 2.562)
Divorcé	0.402***	1.495	(1.120 - 1.996)	0.358**	1.430	(1.071 - 1.911)
Ile de France	-0.056	0.945	(0.766 - 1.167)	-0.036	0.965	(0.781 - 1.191)
B Parisien	ref	1		ref	1	
Nord	0.299**	1.349	(1.040 - 1.749)	0.263**	1.301	(1.004 - 1.685)
Est	0.148	1.16	(0.914 - 1.471)	0.133	1.142	(0.900 - 1.449)
Ouest	0.121	1.129	(0.903 - 1.410)	0.100	1.105	(0.885 - 1.381)
Sud-Ouest	-0.203	0.816	(0.635 - 1.050)	-0.248*	0.780	(0.606 - 1.004)
Centre Est	-0.012	0.988	(0.782 - 1.249)	-0.018	0.982	(0.776 - 1.241)
Méditerranée	-0.137	0.872	(0.689 - 1.104)	-0.189	0.827	(0.653 - 1.047)
Constante	-5.0086			1.0451		
N	13399			13399		
-2 log L (const)	9134.858			9134.858		
-2 log L	6635.167			6635.212		

Tableau 4 : Probabilité de décès en 1988 (modèle LOGIT)
(hommes, âgés de plus de 20 ans)

²⁵Seuils de significativité : * 10%, ** 5%, *** 1%.

Population Totale	Modèle 2			Modèle 3		
Explicatives	Coef	O.R	IC 95%	Coef	O.R	IC 95%
Log revenu	-0.366***	0.693	(0.620 - 0.775)	-0.347***	0.707	(0.632 - 0.792)
Pension (<60 ans)				0.011***	1.011	(1.007 - 1.015)
Ouvrier	0.719***	2.053	(1.561 - 2.700)	0.726***	2.067	(1.570 - 2.720)
Employé	0.858***	2.358	(1.775 - 3.133)	0.829***	2.29	(1.723 - 3.045)
Intermédiaire	0.631***	1.88	(1.414 - 2.500)	0.641***	1.899	(1.427 - 2.527)
Cadre	ref	1		ref	1	
moins de 40	0.045**	1.046	(1.007 - 1.087)	0.048**	1.049	(1.010 - 1.089)
40-54	0.115***	1.122	(1.095 - 1.151)	0.099***	1.104	(1.076 - 1.133)
55-64	0.079***	1.082	(1.053 - 1.113)	0.103***	1.108	(1.076 - 1.141)
65-74	0.057***	1.059	(1.031 - 1.087)	0.056***	1.058	(1.030 - 1.086)
75-84	0.118***	1.125	(1.089 - 1.162)	0.118***	1.125	(1.089 - 1.162)
85 et plus	0.076**	1.079	(1.004 - 1.159)	0.076**	1.079	(1.005 - 1.159)
Marié	ref	1		ref	1	
Célibataire	0.417***	1.517	(1.204 - 1.911)	0.395***	1.485	(1.178 - 1.871)
Veuf	0.709***	2.032	(1.645 - 2.509)	0.701***	2.017	(1.633 - 2.490)
Divorcé	0.349**	1.418	(1.060 - 1.897)	0.320**	1.377	(1.028 - 1.843)
Ile de France	0.008	1.009	(0.815 - 1.247)	0.021	1.021	(0.825 - 1.264)
B Parisien	ref	1		ref	1	
Nord	0.267**	1.306	(1.008 - 1.693)	0.242*	1.274	(0.982 - 1.653)
Est	0.138	1.148	(0.905 - 1.457)	0.131	1.14	(0.898 - 1.448)
Ouest	0.105	1.11	(0.888 - 1.388)	0.091	1.095	(0.875 - 1.369)
Sud-Ouest	-0.243*	0.784	(0.608 - 1.010)	-0.250*	0.779	(0.605 - 1.004)
Centre Est	-0.009	0.991	(0.783 - 1.253)	0.009	1.009	(0.797 - 1.277)
Méditerranée	-0.171	0.843	(0.665 - 1.068)	-0.182	0.834	(0.658 - 1.057)
Constante	-0.6957			-0.8894		
N	13399			13399		
-2 log L (const)	9134.858			9134.858		
-2 log L	6596.246			6565.691		

Tableau 5 : Probabilité de décès en 1988 (modèle LOGIT)
(hommes, âgés de plus de 20 ans)

C Elasticité de la probabilité de décès au revenu

Modèle 1	Population Totale		Moins de 65 ans		Plus de 65 ans	
Revenu	97000		96000		102000	
pondération	hors	avec	hors	avec	hors	avec
β	-0.466	-0.438	-0.546	-0.517	-0.360	-0.356
$O.R$	0.628	0.645	0.579	0.596	0.698	0.701
$I.C.95\%$	(0.569-0.692)	(0.516-0.806)	(0.511-0.657)	(0.442-0.804)	(0.595-0.818)	(0.523-0.938)
p_i	0.1597	0.0210	0.0691	0.0091	0.2243	0.0303
effet marg	-0.060	-0.009	-0.035	-0.005	-0.063	-0.010
$O.R_i$	0.947	0.950	0.938	0.941	0.961	0.964
η	-0.395	-0.429	-0.508	-0.512	-0.279	-0.345
η^e	-0.428	-0.403	-0.498	-0.473	-0.334	-0.330

Tableau 6 : Elasticité-revenu de la probabilité de décès en 1988 (sans contrôle par la profession)

Modèle 2	Population Totale		Moins de 65 ans		Plus de 65 ans	
Revenu	97000		96000		102000	
pondération	hors	avec	hors	avec	hors	avec
β	-0.366	-0.355	-0.446	-0.444	-0.239	-0.253
$O.R$	0.693	0.701	0.640	0.642	0.788	0.777
$I.C.95\%$	(0.620-0.775)	(0.542-0.908)	(0.556-0.738)	(0.453-0.908)	(0.657-0.945)	(0.556-1.086)
p_i	0.1482	0.0204	0.0570	0.0074	0.2450	0.0325
effet marg	-0.046	-0.007	-0.024	-0.003	-0.044	-0.008
$O.R_i$	0.958	0.959	0.949	0.949	0.974	0.972
η	-0.312	-0.347	-0.420	-0.441	-0.180	-0.244
η^e	-0.338	-0.328	-0.409	-0.407	-0.223	-0.236

Tableau 7 : Elasticité-revenu de la probabilité de décès en 1988, ajustée par les professions

Modèle 3	Population Totale		Moins de 65 ans	
Revenu	97000		96000	
pondération	hors	avec	hors	avec
β	-0.346	-0.343	-0.419	-0.430
$O.R$	0.707	0.710	0.657	0.650
$I.C.95\%$	(0.632-0.792)	(0.546-0.924)	(0.569-0.760)	(0.453-0.934)
p_i	0.1521	0.0210	0.0487	0.0062
effet marg	-0.045	-0.007	-0.019	-0.003
$O.R_i$	0.960	0.961	0.952	0.951
η	-0.294	-0.335	-0.399	-0.428
η^e	-0.320	-0.317	-0.386	-0.395

Tableau 8 : Elasticité-revenu de la probabilité de décès en 1988, ajustée par la profession et les pensions reçues

D Etude de la forme santé revenu

Population Totale	Modèle 4			Modèle 5		
Explicatives	Coef	O.R	IC 95%	Coef	O.R	IC 95%
revenu	-0.005***	0.995	(0.993 - 0.997)	-0.011***	0.989	(0.985 - 0.993)
revenu ²	4 E-6***	1.000	(1.000 - 1.000)	3 E-5***	1.000	(1.000 - 1.000)
revenu ³				-2 E-8**	1.000	(1.000 - 1.000)
Pension (<60 ans)	0.011***	1.011	(1.007 - 1.015)	0.011***	1.011	(1.007 - 1.014)
Ouvrier	0.637***	1.891	(1.417 - 2.522)	0.660***	1.935	(1.442 - 2.596)
Employé	0.736***	2.087	(1.555 - 2.802)	0.793***	2.21	(1.636 - 2.987)
Intermédiaire	0.567***	1.763	(1.317 - 2.360)	0.636***	1.889	(1.402 - 2.546)
Cadre	ref	1		ref	1	
moins de 40	0.049**	1.050	(1.011 - 1.091)	0.049**	1.050	(1.011 - 1.091)
40-54	0.099***	1.104	(1.076 - 1.133)	0.100***	1.105	(1.077 - 1.134)
55-64	0.100***	1.105	(1.073 - 1.138)	0.101***	1.106	(1.074 - 1.140)
65-74	0.056***	1.058	(1.030 - 1.086)	0.056***	1.057	(1.030 - 1.085)
75-84	0.117***	1.124	(1.088 - 1.162)	0.117***	1.125	(1.088 - 1.162)
85 et plus	0.079**	1.082	(1.007 - 1.163)	0.075**	1.078	(1.004 - 1.158)
Marié	ref	1		ref	1	
Célibataire	0.405***	1.500	(1.191 - 1.889)	0.388***	1.474	(1.169 - 1.857)
Veuf	0.702***	2.018	(1.634 - 2.492)	0.709***	2.031	(1.644 - 2.509)
Divorcé	0.325**	1.384	(1.034 - 1.853)	0.314**	1.370	(1.023 - 1.834)
Ile de France	0.039	1.040	(0.839 - 1.288)	0.036	1.037	(0.837 - 1.284)
B Parisien	ref	1		ref	1	
Nord	0.241*	1.273	(0.981 - 1.651)	0.237*	1.267	(0.977 - 1.644)
Est	0.127	1.136	(0.895 - 1.442)	0.129	1.138	(0.896 - 1.445)
Ouest	0.090	1.094	(0.875 - 1.368)	0.086	1.090	(0.872 - 1.363)
Sud-Ouest	-0.241*	0.786	(0.610 - 1.012)	-0.253*	0.777	(0.602 - 1.001)
Centre Est	0.009	1.009	(0.797 - 1.278)	0.009	1.009	(0.796 - 1.277)
Méditerranée	-0.179	0.836	(0.659 - 1.060)	-0.188	0.829	(0.653 - 1.052)
Constante	-4.2612			-4.032		
N	13399			13399		
-2 log L (const)	9134.858			9134.858		
-2 log L	6567.510			6555.328		

Tableau 9 : Probabilité de décès en 1988 (modèle LOGIT)
(hommes, âgés de plus de 20 ans)

Population Totale		Modèle 6		Modèle 7		
Explicatives	Coef	O.R	IC 95%	Coef	O.R	IC 95%
Quintile 1	0.916***	2.499	(1.995 - 3.130)	0.688***	1.989	(1.542 - 2.566)
Quintile 2	0.793***	2.211	(1.778 - 2.749)	0.557***	1.745	(1.368 - 2.226)
Quintile 3	0.622***	1.862	(1.495 - 2.319)	0.394***	1.482	(1.166 - 1.884)
Quintile 4	0.358***	1.431	(1.139 - 1.798)	0.175	1.191	(0.938 - 1.513)
Quintile 5	ref	1		ref	1	
Ouvrier				0.670***	1.955	(1.466 - 2.607)
Employé				0.797***	2.219	(1.651 - 2.981)
Intermédiaire				0.642***	1.901	(1.421 - 2.542)
Cadre						
Pension (<60 ans)	0.011***	1.011	(1.007 - 1.015)	0.011***	1.011	(1.007 - 1.014)
moins de 40	0.049**	1.05	(1.011 - 1.091)	0.050***	1.052	(1.012 - 1.092)
40-54	0.101***	1.106	(1.078 - 1.134)	0.099***	1.103	(1.075 - 1.132)
55-64	0.101***	1.106	(1.074 - 1.139)	0.101***	1.106	(1.073 - 1.139)
65-74	0.057***	1.058	(1.031 - 1.087)	0.055***	1.057	(1.029 - 1.085)
75-84	0.115***	1.122	(1.086 - 1.160)	0.118***	1.126	(1.089 - 1.163)
85 et plus	0.074**	1.077	(1.003 - 1.156)	0.074**	1.077	(1.003 - 1.157)
Marié	ref	1		ref	1	
Célibataire	0.420***	1.522	(1.209 - 1.918)	0.405***	1.499	(1.190 - 1.889)
Veuf	0.728***	2.071	(1.678 - 2.556)	0.710***	2.033	(1.646 - 2.512)
Divorcé	0.335**	1.398	(1.046 - 1.869)	0.326**	1.386	(1.035 - 1.855)
Ile de France	0.016	1.016	(0.821 - 1.257)	0.039	1.039	(0.839 - 1.288)
B Parisien	ref	1		ref	1	
Nord	0.229*	1.258	(0.969 - 1.631)	0.237*	1.267	(0.977 - 1.645)
Est	0.119	1.127	(0.887 - 1.430)	0.129	1.138	(0.896 - 1.445)
Ouest	0.080	1.083	(0.867 - 1.355)	0.084	1.088	(0.870 - 1.361)
Sud-Ouest	-0.250*	0.779	(0.605 - 1.003)	-0.249*	0.78	(0.605 - 1.005)
Centre Est	-0.004	0.996	(0.787 - 1.262)	0.006	1.006	(0.794 - 1.274)
Méditerranée	-0.193	0.824	(0.651 - 1.044)	-0.182	0.834	(0.657 - 1.058)
Constante	-4.702			-5.119		
N	13399			13399		
-2 log L (const)	9134.858			9134.858		
-2 log L	6592.141			6561.539		

Tableau 10 : Probabilité de décès en 1988 (modèle LOGIT)
(hommes, âgés de plus de 20 ans)

E Analyse contextuelle : l'effet des inégalités de revenu sur la santé

E.1 Analyse descriptive

Régions	Rev moyen	Gini	Theil	Varlog	Atk (0.5)	Atk (0.25)	Atk (-0.25)	Atk (-0.5)	Offre de soins	Tx de chô.m.
ZEAT Ile-de-France Ile-de-France	127 463	0.307	0.179	16.197	0.081	0.116	0.186	0.226	404.55	8.4
ZEAT Bassin Parisien										
Champagne-Ardennes	85 103	0.311	0.238	14.972	0.093	0.128	0.185	0.210	218.53	10.8
Picardie	88 673	0.269	0.129	14.665	0.061	0.090	0.146	0.174	205.18	10.9
Hte. Normandie	92 419	0.280	0.147	14.558	0.067	0.097	0.153	0.180	220.32	12.1
Centre	90 101	0.268	0.125	14.164	0.059	0.087	0.143	0.172	224.62	8.8
B. Normandie	86 400	0.271	0.138	14.046	0.063	0.093	0.149	0.178	218.86	10
Bourgogne	89 077	0.266	0.128	13.466	0.059	0.086	0.137	0.161	227.10	9.5
ZEAT Nord Nord-Pas-de-Calais	80 313	0.271	0.146	15.474	0.064	0.093	0.150	0.180	232.88	13.5
ZEAT Est										
Lorraine	85 094	0.270	0.144	15.220	0.064	0.093	0.149	0.178	242.87	10.2
Alsace	100 776	0.287	0.194	15.024	0.078	0.109	0.165	0.193	295.04	6.8
Franche-Comté	82 893	0.258	0.123	12.931	0.056	0.081	0.128	0.150	237.71	9.4
ZEAT Ouest										
Pays-de-Loire	84 857	0.258	0.118	13.324	0.056	0.082	0.133	0.159	226.46	10.7
Bretagne	84 939	0.270	0.144	14.158	0.064	0.094	0.152	0.185	242.46	10.1
Poitou-Charente	90 469	0.297	0.173	14.218	0.076	0.109	0.169	0.198	237.44	10.9
ZEAT Sud-Ouest										
Aquitaine	90 519	0.277	0.136	14.165	0.064	0.094	0.153	0.182	304.75	10.9
Midi-Pyrénées	94 043	0.289	0.147	14.410	0.069	0.101	0.163	0.194	315.27	9.5
Limousin	91 209	0.276	0.134	14.033	0.063	0.093	0.151	0.180	260.94	8.2
ZEAT Centre-Est										
Rhône-Alpes	97 509	0.274	0.137	13.785	0.063	0.092	0.148	0.177	260.4	8.2
Auvergne	86 551	0.304	0.228	14.377	0.088	0.121	0.178	0.204	247.9	9.9
ZEAT Méditerranée										
Languedoc-Roussillon	84 366	0.291	0.160	14.411	0.073	0.106	0.171	0.208	341.89	13.6
PACA	92 270	0.296	0.168	15.418	0.075	0.109	0.173	0.206	384.67	11.8

Table 11 : Caractéristiques Régionales

E.2 Des différences régionales de mortalité inexplicées

Variance inter-régionale	Modèle vide	Modèle individuel
Valeur estimée	0.019478	0.017145
Ecart-type	0.008690	0.007979
T Student	2.241427	2.148766

Tableau 12 : Une variance interrégionale significativement différente de zéro

Modèle 8	Régression Logistique				Analyse multi-niveaux			
	Coef	St dev	T	O.R	Coef	St dev	T	O.R
Quintile 1	0.665***	0.128	5.193	1.944	0.677***	0.125	5.412	1.969
Quintile 2	0.554***	0.122	4.554	1.740	0.558***	0.111	5.018	1.746
Quintile 3	0.393***	0.121	3.252	1.482	0.394***	0.117	3.361	1.483
Quintile 4	0.174	0.121	1.434	1.190	0.177**	0.086	2.060	1.194
Quintile 5	ref			1	ref			1
Ouvrier	0.687***	0.146	4.695	1.988	0.682***	0.145	4.715	1.978
Employé	0.789***	0.150	5.244	2.201	0.792***	0.136	5.841	2.209
Intermédiaire	0.645***	0.148	4.360	1.906	0.647***	0.137	4.738	1.909
Cadre	ref			1	ref			1
Pensions	0.011***	0.002	5.860	1.011	0.011***	0.002	5.943	1.011
moins de 40	0.049**	0.019	2.541	1.051	0.050***	0.018	2.805	1.051
40-54	0.098***	0.013	7.496	1.103	0.098***	0.013	7.840	1.103
55-64	0.099***	0.015	6.576	1.104	0.100***	0.013	7.669	1.105
65-74	0.053***	0.014	3.956	1.055	0.054***	0.014	3.855	1.056
75-84	0.120***	0.017	7.156	1.127	0.119***	0.019	6.203	1.126
85 et plus	0.069*	0.036	1.904	1.072	0.070	0.046	1.531	1.073
Marié	ref			1	ref			1
Célibataire	0.401***	0.118	3.410	1.494	0.401**	0.173	2.322	1.494
Veuf	0.718***	0.107	6.689	2.051	0.717***	0.137	5.233	2.048
Divorcé	0.307**	0.148	2.072	1.359	0.318*	0.170	1.867	1.374
Var. inter-régions					0.017**	0.008	2.149	
Constante	-5.098	(0.188)			-5.107	(0.156)		
Effectif (niveau 1)	13 399				13 399			
Effectif (niveau 2)					21			

Tableau 13 : Impact des caractéristiques individuelles sur le risque de décès (modèle 8)
Comparaison des résultats de la régression logistique et de l'analyse multi-niveaux

E.3 Analyse exploratoire

Population Totale	Modèle 9			Modèle 10		
Caractéristiques	Coef	O.R	I.C 95%	Coef	O.R	I.C 95%
Régions						
Theil	0.185	1.204	(0.956 - 1.516)	0.306**	1.358	(1.058 - 1.743)
Offre de soins				-0.001**	0.999	(0.998 - 1.000)
Chômage				-0.003	0.997	(0.961 - 1.035)
Individus						
Quintile 1	0.674***	1.961	(1.526 - 2.521)	0.653***	1.921	(1.493 - 2.473)
Quintile 2	0.566***	1.760	(1.386 - 2.235)	0.537***	1.712	(1.345 - 2.178)
Quintile 3	0.395***	1.485	(1.172 - 1.882)	0.372***	1.451	(1.143 - 1.841)
Quintile 4	0.179	1.195	(0.943 - 1.516)	0.162	1.176	(0.927 - 1.492)
Quintile 5	ref	1		ref	1	
Pensions	0.011***	1.011	(1.007 - 1.015)	0.011***	1.011	(1.007 - 1.015)
Ouvrier	0.700***	2.014	(1.511 - 2.685)	0.681***	1.976	(1.481 - 2.637)
Employé	0.796***	2.217	(1.650 - 2.978)	0.793***	2.210	(1.645 - 2.970)
Intermédiaire	0.658***	1.931	(1.444 - 2.582)	0.649***	1.913	(1.430 - 2.559)
Cadre	ref	1		ref	1	
moins de 40	0.049**	1.051	(1.012 - 1.091)	0.050***	1.051	(1.012 - 1.092)
40-54	0.098***	1.103	(1.075 - 1.132)	0.098***	1.103	(1.075 - 1.132)
55-64	0.100***	1.105	(1.073 - 1.138)	0.100***	1.105	(1.073 - 1.138)
65-74	0.053***	1.054	(1.027 - 1.083)	0.053***	1.054	(1.027 - 1.083)
75-84	0.119***	1.127	(1.090 - 1.164)	0.120***	1.127	(1.091 - 1.164)
85 et plus	0.069*	1.071	(0.997 - 1.150)	0.070*	1.072	(0.998 - 1.151)
Marié	ref	1		ref	1	
Célibataire	0.399***	1.490	(1.183 - 1.876)	0.411***	1.508	(1.197 - 1.900)
Veuf	0.720***	2.055	(1.665 - 2.537)	0.728***	2.072	(1.678 - 2.558)
Divorcé	0.300**	1.349	(1.009 - 1.804)	0.326**	1.385	(1.036 - 1.854)
Constante	-5.401			-5.198		
Effectif	13 399			13 399		
-2 Log L (const)	9134.858			9134.858		
-2 Log L	6575.098			6570.071		

Tableau 14 : L'impact des variables contextuelles sur le risque de décès
(régression logistique, mesure d'entropie de Theil)

Coefficient (Logit)	Inégalités de revenus	Offre de soins	Taux de Chômage
Gini	0.551*	-0.002**	-0.003
Theil	0.306**	-0.001**	-0.003
Variance des Logs	0.170***	-0.002***	-0.018
Atkinson (0.5)	0.970**	-0.001**	-0.003
Atkinson (0.25)	0.806**	-0.002**	-0.003
Atkinson (-0.25)	0.776**	-0.002***	-0.006
Atkinson (-0.5)	0.826***	-0.003***	-0.007

Tableau 15 : Impact des caractéristiques régionales (en continu) sur le risque de décès
(modèle 10, régression logistique)

E.4 Analyse multi-niveaux

Population totale	Modèle 9 bis				Modèle 10 bis			
Caractéristiques	Coef	St.dev	T	O.R	Coef	St.dev	T	O.R
Régions								
Theil	0.186	0.128	1.455	1.205	0.274	0.181	1.512	1.315
Offre de soins					-0.001*	0.001	-1.857	0.999
Chômage					-0.004	0.021	-0.187	0.996
Individus								
Quintile 1	0.679***	0.126	5.409	1.972	0.671***	0.127	5.303	1.957
Quintile 2	0.561***	0.111	5.033	1.752	0.550***	0.111	4.974	1.733
Quintile 3	0.392***	0.117	3.352	1.480	0.382***	0.117	3.280	1.466
Quintile 4	0.178**	0.085	2.080	1.194	0.171**	0.087	1.967	1.186
Quintile 5	ref			1	ref			1
Pensions	0.011***	0.002	5.944	1.011	0.011***	0.002	5.632	1.011
Ouvrier	0.690***	0.150	4.602	1.993	0.679***	0.149	4.560	1.972
Employé	0.797***	0.136	5.879	2.218	0.795***	0.137	5.805	2.215
Intermédiaire	0.654***	0.136	4.806	1.923	0.649***	0.139	4.668	1.913
Cadre	ref			1	ref			1
moins de 40	0.050***	0.018	2.803	1.051	0.050***	0.018	2.799	1.051
40-54	0.098***	0.013	7.786	1.103	0.098***	0.013	7.748	1.103
55-64	0.100***	0.013	7.708	1.105	0.100***	0.013	7.634	1.105
65-74	0.054***	0.014	3.830	1.055	0.054***	0.014	3.810	1.056
75-84	0.119***	0.019	6.138	1.127	0.119***	0.019	6.171	1.126
85 et plus	0.070	0.046	1.517	1.072	0.071	0.047	1.521	1.073
Marié	ref			1	ref			1
Célibataire	0.402**	0.174	2.310	1.494	0.407**	0.175	2.332	1.502
Veuf	0.719***	0.136	5.289	2.053	0.723***	0.137	5.295	2.060
Divorcé	0.316*	0.172	1.839	1.371	0.328*	0.174	1.882	1.388
Constante	-5.400***	0.263	-20.556		-5.130***	0.319	-16.091	
Var. inter-régions	0.016*	0.008	1.880		0.014*	0.007	1.875	

Tableau 16 : L'impact des variables contextuelles sur le risque de décès
(analyse multi-niveaux, mesure d'entropie de Theil)

Variable contextuelle		Variance des résidus		
Variable retenue	significativité	Estimation	Ecart-type	T Student
Gini	N.S	0.019**	0.008	2.284
Theil	N.S	0.016*	0.008	1.880
Variance de Logs	N.S	0.016*	0.008	1.927
Atkinson (0.5)	N.S	0.017**	0.008	2.060
Atkinson (0.25)	N.S	0.018**	0.008	2.133
Atkinson (-0.25)	N.S	0.018**	0.008	2.222
Atkinson (-0.5)	N.S	0.018**	0.008	2.238
Offre de soins	N.S	0.017**	0.007	2.366
Chômage	N.S	0.019**	0.008	2.383

Tableau 17 : Impact des inégalités (en continu)
(modèle 9 bis, analyse multi-niveaux)

Mesures des Inégalités	Variables contextuelles			Variance des résidus		
	Inég. (coef)	Offre de soins	Taux de Chômage	Estimation	Ecart-type	T student
Gini	0.473	-0.002*	N.S	0.016**	0.008	2.132
Theil	0.274	-0.001*	N.S	0.014*	0.007	1.875
Variance de Logs	0.170***	-0.002***	N.S	0	0	0
Atkinson (0.5)	0.856	-0.002*	N.S	0.014*	0.007	1.932
Atkinson (0.25)	0.708	-0.002*	N.S	0.014*	0.007	1.919
Atkinson (-0.25)	0.683	-0.002**	N.S	0.012*	0.007	1.794
Atkinson (-0.5)	0.749**	-0.003**	N.S	0.009	0.006	1.467

Tableau 18 : Impact des caractéristiques régionales (en continu)
sur le risque de décès (coefficient modèle 10 bis, analyse multi-niveaux)

Modèle 11	Régression Logistique				Analyse Multi-niveaux			
Caractéristiques	Coef	St dev	T	O.R	Coef	St dev	T	O.R
Régions								
Theil faible	ref			1	ref			1
Theil moyen	0.226**	0.091	2.496	1.254	0.222***	0.078	2.853	1.248
Theil élevé	0.361***	0.123	2.942	1.435	0.348***	0.120	2.895	1.416
Offre de soins	-0.002***	0.001	-2.736	0.998	-0.002**	0.001	-2.225	0.998
Individus								
Quintile 1	0.644***	0.129	5.011	1.904	0.664***	0.128	5.188	1.942
Quintile 2	0.521***	0.123	4.247	1.684	0.538***	0.113	4.775	1.712
Quintile 3	0.360***	0.122	2.966	1.434	0.374***	0.118	3.179	1.453
Quintile 4	0.159	0.121	1.309	1.172	0.168*	0.088	1.913	1.183
Quintile 5	ref			1	ref			1
Pensions	0.011***	0.002	5.806	1.011	0.107***	0.002	57.407	1.113
Ouvrier	0.684***	0.147	4.653	1.983	0.682***	0.150	4.551	1.978
Employé	0.793***	0.151	5.255	2.209	0.795***	0.138	5.756	2.214
Intermédiaire	0.648***	0.149	4.364	1.912	0.649***	0.140	4.627	1.913
Cadre	ref			1	ref			1
moins de 40	0.050***	0.019	2.591	1.051	0.052***	0.018	2.893	1.054
40-54	0.098***	0.013	7.504	1.103	0.098***	0.013	7.683	1.103
55-64	0.099***	0.015	6.583	1.104	0.100***	0.013	7.517	1.105
65-74	0.054***	0.014	3.985	1.055	0.054***	0.014	3.866	1.056
75-84	0.120***	0.017	7.156	1.127	0.119***	0.019	6.138	1.127
85 et plus	0.069*	0.036	1.893	1.071	0.070	0.046	1.533	1.073
Marié	ref			1	ref			1
Célibataire	0.423***	0.118	3.585	1.527	0.416**	0.176	2.366	1.517
Veuf	0.738***	0.108	6.854	2.092	0.731***	0.137	5.335	2.076
Divorcé	0.335**	0.149	2.255	1.398	0.333*	0.176	1.899	1.396
Constante	-4.769	0.262	-18.188		-4.744	0.289	-16.429	
variance interrégion					0.010*	0.005	1.856	
Effectif (niveau 1)	13 399				13 399			
Effectif (niveau 2)					21			
-2 Log L	6566.35							

Tableau 19 : L'impact des inégalités sur le risque de décès
(mesure d'entropie de Theil)

Odd Ratio (analyse multi-niveaux)	Inégalités faibles	Inégalités moyennes	Inégalités fortes
Gini	1	1.197*	1.264**
Theil	1	1.248***	1.416***
Variance de Logs	1	1.206*	1.221***
Atkinson (0.5)	1	1.280***	1.278**
Atkinson (0.25)	1	1.262***	1.260**
Atkinson (-0.25)	1	1.207**	1.329**
Atkinson (-0.5)	1	1.287***	1.186

Tableau 20 : Impact des inégalités (en tranche) sur le risque de décès
(modèle 11, analyse multi-niveaux)

Variance inter-régions	Valeur estimée	Ecart-type	T Student
Modèle individuel	0.0171**	0.0080	2.1488
Gini	0.0138**	0.0054	2.5603
Theil	0.0096*	0.0052	1.8555
Variance des Logs	0.0132**	0.0056	2.3723
Atkinson (0.5)	0.0092*	0.0055	1.6868
Atkinson (0.25)	0.0102*	0.0057	1.7837
Atkinson (-0.25)	0.0104**	0.0051	2.0340
Atkinson (-0.5)	0.0094	0.0062	1.5160

Tableau 21 : Contribution des inégalités de revenu (en tranche)
à l'explication des différences régionales de mortalité

E.5 Un effet du capital social ou un effet de comparaisons sociales

Population Totale	Modèle 12			Modèle 13		
Caractéristiques	Coef	O.R	I.C 95%	Coef	O.R	I.C 95%
Régions						
Criminalité	-0.002	0.998	(0.994 - 1.001)	9.4E-5	1.000	(0.992 - 1.008)
Offre de soins				-6.7E-4	0.999	(0.997 - 1.001)
Chômage				-0.003	0.997	(0.954 - 1.041)
Individus						
Quintile 1	0.653***	1.922	(1.494 - 2.472)	0.652***	1.918	(1.490 - 2.470)
Quintile 2	0.539***	1.714	(1.349 - 2.179)	0.536***	1.708	(1.342 - 2.174)
Quintile 3	0.384***	1.468	(1.157 - 1.861)	0.381***	1.463	(1.153 - 1.856)
Quintile 4	0.165	1.180	(0.930 - 1.496)	0.163	1.177	(0.928 - 1.494)
Quintile 5	ref	1		ref	1	
Pensions	0.011***	1.011	(1.007 - 1.015)	0.011***	1.011	(1.007 - 1.015)
Ouvrier	0.677***	1.968	(1.477 - 2.623)	0.672***	1.958	(1.468 - 2.610)
Employé	0.786***	2.194	(1.633 - 2.946)	0.785***	2.192	(1.632 - 2.943)
Intermédiaire	0.638***	1.894	(1.417 - 2.531)	0.636***	1.888	(1.412 - 2.524)
Cadre	ref	1		ref	1	
moins de 40	0.049**	1.051	(1.012 - 1.091)	0.050**	1.051	(1.012 - 1.091)
40-54	0.098***	1.103	(1.075 - 1.132)	0.098***	1.103	(1.075 - 1.132)
55-64	0.099***	1.104	(1.072 - 1.137)	0.099***	1.104	(1.072 - 1.137)
65-74	0.053***	1.055	(1.027 - 1.083)	0.054***	1.055	(1.027 - 1.083)
75-84	0.120***	1.127	(1.091 - 1.165)	0.120***	1.127	(1.091 - 1.165)
85 et plus	0.070*	1.072	(0.999 - 1.151)	0.070*	1.073	(0.999 - 1.152)
Marié	ref	1		ref	1	
Célibataire	0.405***	1.499	(1.190 - 1.888)	0.408***	1.504	(1.193 - 1.895)
Veuf	0.721***	2.056	(1.666 - 2.538)	0.722***	2.058	(1.667 - 2.540)
Divorcé	0.320**	1.378	(1.030 - 1.843)	0.323**	1.382	(1.033 - 1.849)
Constante	-4.963			-4.855		
Effectif	13 399			13 399		
-2 Log L (const)	9134.858			9134.858		
-2 Log L	6576.154			6575.714		

Tableau 22 : L'impact de la criminalité sur le risque de décès
(analyse multi-niveaux, mesure d'entropie de Theil)

Modèle 14	Revenu < Revenu Médian			Revenu > Revenu Médian		
Caractéristiques	Coef	O.R	I.C	Coef	O.R	I.C
Régions						
Theil faible	ref	1		ref	1	
Theil moyen	0.274**	1.315	(1.039 - 1.664)	0.138	1.148	(0.871 - 1.514)
Theil élevé	0.364**	1.440	(1.042 - 1.988)	0.330*	1.391	(0.959 - 2.019)
Offre de soins	-0.003***	0.997	(0.995 - 0.999)	-0.001	0.999	(0.997 - 1.001)
Individus						
Log revenu	-0.210**	0.811	(0.680 - 0.966)	-0.677***	0.508	(0.382 - 0.677)
Pensions	0.011***	1.011	(1.007 - 1.015)	0.006	1.006	(0.997 - 1.015)
moins de 40	0.016	1.016	(0.971 - 1.063)	0.155***	1.168	(1.064 - 1.281)
40-54	0.129***	1.137	(1.098 - 1.178)	0.059***	1.061	(1.021 - 1.102)
55-64	0.079***	1.082	(1.038 - 1.128)	0.125***	1.133	(1.085 - 1.183)
65-74	0.054***	1.056	(1.018 - 1.095)	0.059***	1.060	(1.021 - 1.102)
75-84	0.094***	1.099	(1.050 - 1.150)	0.141***	1.151	(1.098 - 1.207)
85 et plus	0.108**	1.114	(1.004 - 1.236)	0.034	1.034	(0.936 - 1.142)
Marié	ref	1		ref	1	
Célibataire	0.487***	1.627	(1.223 - 2.164)	0.219	1.245	(0.803 - 1.930)
Veuf	0.950***	2.587	(1.905 - 3.512)	0.559***	1.748	(1.300 - 2.351)
Divorcé	0.397**	1.487	(1.004 - 2.202)	0.212	1.236	(0.790 - 1.931)
Constante	-1.354			4.119		
Défunts	852			586		
Survivants	5267			6694		
-2 Log L (const)	4698.46			4234.703		
-2 Log L	3376.591			3099.967		

Tableau 23 : L'impact des inégalités sur le risque de décès
selon le groupe de revenu