

Les indicateurs de score du niveau socio-économique

Ces indicateurs ont émergé parce que les données socio-économiques individuelles sont non seulement difficilement accessibles, mais aussi « sensibles ». Le RGPD (Règlement Général sur la Protection des Données) encadre le traitement des données personnelles sur le territoire de l'Union européenne. Ainsi, les données sont collectées pour un but bien déterminé et légitime et ne sont pas traitées ultérieurement de façon incompatible avec cet objectif initial. Il devient donc très difficile d'utiliser des données socio-économiques individuelles. Pour contourner ce problème, il est d'usage de travailler sur des données agrégées (au niveau d'une commune ou d'un IRIS) afin de préserver le secret statistique et le principe d'anonymisation. Ainsi, plutôt que d'appréhender directement les caractéristiques socio-économiques d'un individu, on mesure l'environnement dans lequel l'individu ou les populations évoluent.

L'objectif des indicateurs de score est de synthétiser une réalité complexe à travers un chiffre unique. Sachant que l'environnement est un concept multidimensionnel, il n'y a pas de consensus sur les variables qu'on devrait utiliser (revenu, emploi, éducation, famille, logement etc ...). C'est pourquoi, plusieurs auteurs ont contribué au fil du temps en proposant des indices qui synthétiseraient le gradient social.

Nous avons principalement recensé dans la littérature 4 indicateurs qui feront l'objet de notre analyse : l'indice de Townsend, le « French Deprivation Index », l'European Deprivation Index, et l'indice de Lalloué.

L'indice de Townsend

Townsend est le premier à avoir proposé un indicateur synthétique du désavantage social dans le but de déterminer un « état de désavantage observable et mesurable par rapport à la communauté locale, la société ou la nation à laquelle appartient un individu, une famille ou un groupe » (Townsend, 1987). Ce score est exploitable à l'échelle infra-communale et il a été constitué à partir de 4 variables :

- ✓ Taux de ménages sans voitures
- ✓ Taux de ménages non-propriétaires
- ✓ Taux de ménages habitant un logement avec plus d'une personne par pièce
- ✓ Taux de chômage dans la population active

Townsend a utilisé la méthode du Z score : après avoir été passées à l'échelle logarithmique, les variables sont centrées et réduites et le score final correspond à la somme de ces 4 variables transformées pour chaque unité géographique.

Ce score étant pionnier, il a été adopté par les chercheurs notamment dans leurs études sur le bien-être social (Lahelma & Valkonen, 1990 ; Smith & Susan, 1989) ainsi que dans les études dans le domaine médical sur le taux de mortalité (Mays & Chinn, 1989). Contrairement à d'autres indices qui ont suivi, l'indice de Townsend n'inclut pas la composante de revenu. Cela s'explique par le fait que dans les années 1980, cette donnée était indisponible au Royaume-Uni.

L'une des critiques de cet indice est qu'il traduit une vision « urbaine » de la défavorisation. Depuis, les scores sont en majorité construits à partir d'analyse factorielle (ACP) et ils incluent le revenu médian.

French Deprivation Index (FDep)

Contrairement aux pays anglo-saxons, l'étude de disparités de santé est une thématique plus récente en France qui n'apparaît qu'à partir des années 2000. Auparavant, les études se focalisaient sur l'aspect individuel compte tenu du fait que les données étaient essentiellement issues d'enquêtes. Le premier indicateur proposé en France est celui créé par Rey et al (2009), le French Deprivation Index (FDep). Il s'agit d'un indice de défavorisation sociale constitué à partir de 4 variables issues du recensement de la population effectuée par l'Insee et des revenus fiscaux de la population (Insee) :

- ✓ Taux de chômage dans la population active de 15 à 64 ans
- ✓ Taux d'ouvriers dans la population active de 15 à 64 ans
- ✓ Taux de bacheliers dans la population non scolarisée des plus de 15 ans
- ✓ Revenu médian déclaré par Unité de Consommation (revenu moyen par UC en 1990)

Le FDep est issu de la 1^{re} composante principale (67-70% de la variance totale expliquée selon la période) d'une analyse en composante principale pondérée sur la population de ces quatre variables agrégées. Pour qu'il soit comparable inter-temporellement, le score est centré et réduit. Enfin, la méthodologie est appliquée à l'échelle de l'IRIS et de la commune, tandis que pour les échelles plus agrégées comme le département ou la région, la moyenne pondérée sur la population du score à l'échelle communale est utilisée. Le lien ci-contre donne accès à une base de données de FDep ([Indice de défavorisation sociale \(FDep\) par IRIS — Opendatasoft](#)).

Depuis son instauration, le FDep est amplement utilisé dans les revues médicales, car il sert de proxy de l'environnement dans lequel l'individu évolue et permet donc d'en voir l'impact sur le

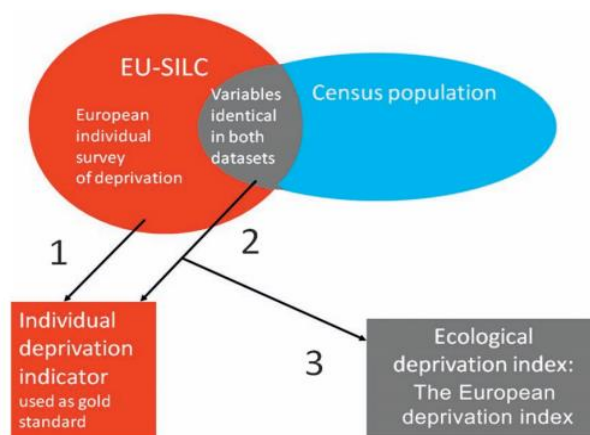
taux de mortalité, taux de dépistage etc... Concernant le cancer du sein, ce score a été utilisé notamment dans l'étude de Deborde et. al (2018) qui vise à quantifier la relation entre la défavorisation sociale du quartier de résidence de la femme, et sa participation au dépistage organisé.

European Deprivation Index (EDI)

Le EDI proposé par Pornet et al. (2012) est un indicateur du désavantage social. Contrairement au FDep qui a été utilisé dans le contexte national français, l'EDI est spécifique à chaque pays et il a pour ambition de comparer le niveau de désavantage social entre les Etats membres de l'UE. Cet indicateur combine à la fois des données individuelles ainsi que des données agrégées. Les données individuelles sont issues de l'enquête européenne EU – SILC (European Union – Statistics on Income and Living Conditions) effectuée en 2006. Cette population est constituée de 10 036 ménages, soit 24 940 individus. De l'autre côté, les données agrégées sont issues du recensement de la population de 1999 réalisé par l'Insee.

La construction de l'indice écologique de désavantage social s'effectue en trois temps. On sélectionne chaque variable individuelle associée au désavantage social individuel dont on estime le poids (effets du modèle multivarié dont l'estimation est donnée au tableau 1). Ces poids issus du modèle individuel sont ensuite introduits dans le modèle et multiplient les mesures agrégées (IRIS) correspondantes. Enfin, le tout est sommé pour constituer le score final. Comme l'indice précédent, celui-ci est aussi disponible à l'échelle de l'IRIS.

Figure 1: Etapes de construction du EDI



Source : Pornet et.al (2012)

Au niveau de la France métropolitaine, le modèle final de la régression logistique multivariée est présenté au tableau 1 (la seconde colonne donne la pondération associée à chaque variable).

Tableau 1 : Poids des variables liées au désavantage social

Final model of multivariate logistic regression selecting components of EDI, which were associated with the individual deprivation indicator, French EU–SILC 2006* (N=19 253)†

	β	95% CI	p Value
Overcrowding	0.11	0.02 to 0.20	0.0171
No access to a system of central or electric heating	0.34	0.25 to 0.43	<0.0001
Non-owner	0.55	0.50 to 0.60	<0.0001
Unemployment	0.47	0.38 to 0.55	<0.0001
Foreign nationality	0.23	0.13 to 0.32	<0.0001
No access to a car	0.52	0.46 to 0.58	<0.0001
Unskilled worker—farm worker	0.37	0.27 to 0.46	<0.0001
Household with more than six persons	0.45	0.34 to 0.56	<0.0001
Low level of education	0.19	0.13 to 0.25	<0.0001
Single-parent household	0.41	0.33 to 0.48	<0.0001

*Data weighted on non-response and adjusted on sample design of the French EU–SILC survey 2006.

†Missing data: n=989.

EDI, European deprivation index; EU–SILC, European Union—Statistics on Income and Living Conditions.

Ouédraogo et.al (2014) ont utilisé l’EDI dans leur étude des déterminants du dépistage organisé. L’EDI a été divisé en terciles, où le premier tercile représente les zones les plus favorisées et le troisième les plus défavorisées. Avec la méthode du logit multivarié, ils concluent que l’EDI est significatif, c’est-à-dire que le taux de dépistage organisé est statistiquement différent en fonction du niveau de défavorisation de l’environnement dans lequel la femme réside.

L’indice de défavorisation contextuelle

Lalloué et.al (2012) ont proposé un indice défavorisation contextuelle, créé à partir de données de recensement de la population et des revenus fiscaux de l’Insee à l’échelle de l’IRIS pour 3 grandes agglomérations françaises, Lille, Lyon et Marseille. Ici, un des objectifs est de proposer un indice spécifique à chaque ville et de le comparer à l’indice global rassemblant les 3 agglomérations.

L’indice est construit en deux temps. Tout d’abord, 48 variables sont extraites du fichier de recensement. Celles-ci couvrent les grands déterminants du gradient social, à savoir l’emploi, la famille et le ménage, le niveau d’études, le logement, le revenu et le statut de l’immigration. Les auteurs ont basé la sélection des variables sur l’utilisation d’un critère statistique plutôt qu’en se référant à la littérature. Ensuite, la méthode de l’analyse en composante principale (ACP) a été utilisée afin d’exclure les variables redondantes et de garder les variables les plus corrélées à la première composante principale (voir ci-dessous).

Enfin, au lieu d’utiliser une approche par quintile, ils ont procédé à une classification hiérarchique ascendante (CAH) qui consiste à agréger successivement les observations afin de former les groupes les plus similaires. Ils parviennent ainsi à trois classes de défavorisation contextuelle, ce qui s’avère plus efficace statistiquement qu’une simple division en quintile.

Variables utilisées pour l'indice de défavorisation contextuelle de Lalloué et.al (2012) :

Emploi

- ✓ Taux d'emplois précaires dans la population active
- ✓ Taux d'emplois stables dans la population active
- ✓ Taux de travailleurs indépendants dans la population active

Logement

- ✓ Taux de ménages sans voiture
- ✓ Taux de ménages avec deux voitures ou plus
- ✓ Taux de résidences principales non occupées par leur propriétaire
- ✓ Taux de résidences principales avec plus d'une personne par pièce
- ✓ Nombre moyen de personnes par pièce

Famille et ménage

- ✓ Taux de familles monoparentales

Revenu

- ✓ Revenu médian par unité de consommation

Niveau d'études

- ✓ Taux de diplômes professionnels (BEP, CAP, etc) dans la population âgée de 15 ans ou plus
- ✓ Taux de bacheliers (technique ou général) dans la population âgée de 15 ans ou plus
- ✓ Taux de personnes sans diplôme dans la population âgée de 15 ans ou plus

Statut en matière d'immigration

- ✓ Taux de personnes de nationalité étrangère (non française)
- ✓ Taux d'immigrants étrangers depuis le dernier recensement sur la population totale