

Contexte

Afin de disposer d'un indice lié au concept de défavorisation tel que défini par Townsend (Townsend et al, 1987), les unités INSERM U558 de Toulouse et U1086 de Caen ont construit, en collaboration avec le Townsend Institute, University of Bristol, un indice basé sur le concept de pauvreté relative. Il est basé sur la sélection de variables agrégées censitaires les plus associées à un indicateur individuel de la défavorisation objective et subjective. Sa construction repose sur une enquête européenne standardisée annuelle dénommée EU-SILC. Celle-ci dispose d'un questionnaire spécifique à chaque pays et d'un questionnaire commun à tous les pays. Cet indice, dénommé the European Deprivation Index, est destiné à être répliquable dans le temps et dans l'espace français et européen. Son score est une combinaison pondérée de variables écologiques.

Dans sa version publiée, l'EDI était construit sur le questionnaire spécifique à la France (Pornet et al., JECH 2012). Pour répondre à l'objectif de comparabilité entre pays européens, il a été recalculé (selon la même méthodologie) sur la base du questionnaire commun à tous les pays européens. La plateforme fournit désormais cette version de l'indice.

Sa construction repose sur différentes étapes, ci-dessous représentées (figure).

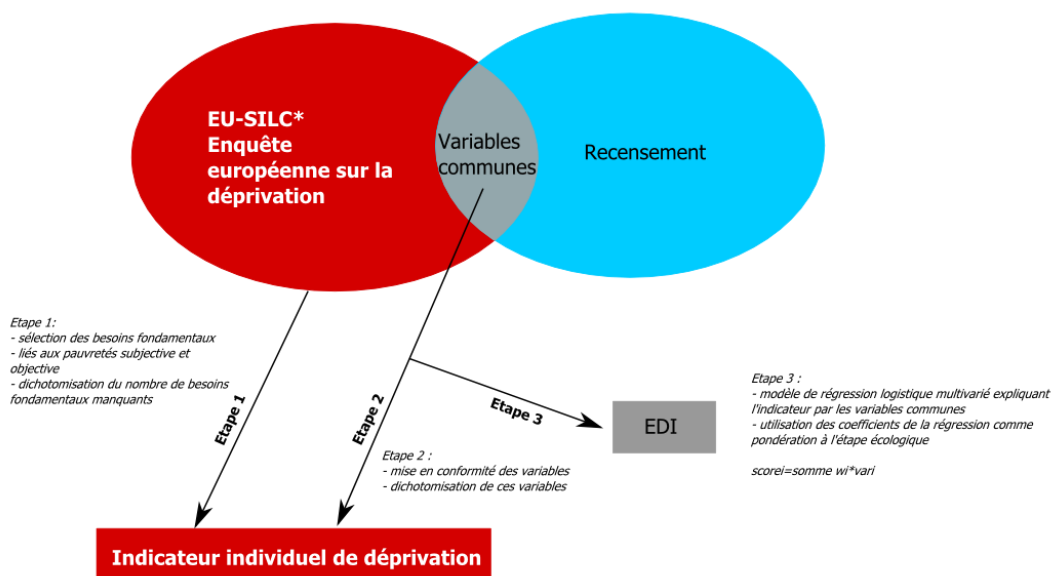


Figure : Schéma méthodologique de la construction de l'EDI (Pornet et al, 2012)

Etape 1 : Construction d'un indicateur individuel de défavorisation

1. Sélection des besoins fondamentaux

Les besoins fondamentaux sont sélectionnés dans l'enquête EU-SIL parmi les questions :

	Intitule	Answer
1	Using your own means to cover a necessary yet unplanned expense of 800 euros	Y/N
2	Taking a week's annual holiday away from home	Y/N
3	Eating a meal containing some meat or some fish or the vegetarian equivalent once every two days	Y/N
4	Keeping your house sufficiently warm	Y/N
5	Having a car	Y/N-cannot afford/ N-other reason
6	Having a computer	Y/N-cannot afford/ N-other reason
7	Having a washing machine	Y/N-cannot afford/ N-other reason
8	Having a colour TV	Y/N-cannot afford/ N-other reason
9	Having a phone (including mobile phone)	Y/N-cannot afford/ N-other reason

Pour les variables **1-4**, réponses oui/non, elles impliquent directement une capacité ou une habilité à pouvoir se permettre quelque chose. Moins de 50% des ménages doivent posséder ce bien pour qu'il soit considéré comme un besoin fondamental potentiel.

Pour les variables **5-9**, au moins 50% des ménages doivent posséder le bien pour qu'il puisse être retenu comme pouvant être un besoin fondamental potentiel. Parmi ces besoins pré-sélectionnés, ceux non possédés par moins de 50 % des ménages par incapacité financière (et non par choix) étaient considérés comme des besoins fondamentaux potentiels.

2. Sélection des besoins liés aux pauvretés objective et subjective

2.1. Construction de la pauvreté objective

La pauvreté objective est directement disponible dans l'enquête EU-SILC à partir de l'item : "poverty indicator" dont la valeur dépend du niveau de vie. Un ménage est considéré comme "pauvre" si son niveau de vie (i.e. le revenu par unité de consommation) est inférieur à 60% du niveau de vie médian.

2.2. Construction de la pauvreté subjective

Obtenue à partir de la question "ability to make ends meet" :

- 1: with great difficulty
- 2: with difficulty
- 3: with some difficulty
- 4: fairly easily
- 5: easily
- 6: very easily

La variable a été dichotomisée (0, 1) selon sa corrélation avec la pauvreté objective.

Modèle de régression logistique :

pauvreté objective (pauvre=1/pas pauvre=0) = pauvreté subjective

Tous les modèles ont été testés (modalité 1 vs 2, 3, 4, 5, 6; modalité 1, 2 vs 3, 4, 5, 6). Le modèle avec la plus grande valeur au test de Wald a été retenu.

2.3. Sélection des besoins liés aux deux pauvretés

Régressions logistiques univariées:

- pauvreté objective (pauvre=1/pas pauvre=0)= chaque besoin fondamental potentiel (réf=pas pauvre)
- pauvreté subjective (pauvre=1/pas pauvre=0)= chaque besoin fondamental potentiel (réf=pas pauvre)

Régressions logistiques multivariées :

- pauvreté objective (pauvre=1/pas pauvre=0)= tous les besoins fondamentaux associés aux deux pauvretés (réf=pas pauvre)
- pauvreté subjective (pauvre=1/pas pauvre=0)=tous les besoins fondamentaux associés aux deux pauvretés (réf=pas pauvre)

➔ Les besoins fondamentaux sont ceux pour lesquels la p-value est significative au seuil de 5% pour les deux modèles : nous obtenons ainsi n besoins manquants.

PS: Les modèles multivariés sont réalisés sans méthode de sélection, les variables sont exclues en même temps.

3. Construction d'un indicateur binaire de pauvreté individuel (IDI)

Dichotomisation du nombre de besoins fondamentaux manquants par incapacité financière pour créer l'indicateur individuel à l'aide de régressions logistiques :

- pauvreté objective (pauvre=1/pas pauvre=0)= n besoins manquants versus n-1
- pauvreté subjective (pauvre=1/pas pauvre=0)= n besoins manquants versus n-1
- ⇒ Meilleur modèle pour les deux pauvretés, ou en cas de discordance, par celui obtenu pour la pauvreté objective → n besoins manquants.

Etape 2 : sélection des variables communes expliquant l'indicateur de pauvreté individuel

1. Sélection des variables disponibles à l'échelle individuelle et à l'échelle agrégée

Les variables doivent être identiquement formulées.

2. Dichotomisation des variables plurimodales selon leur relation avec l'indicateur individuel

Modèle de régression logistique :

IDI = combinaison de chaque modalité de variables selon leur type.

Le recodage dépend du type de la variable:

- qualitative : modalité n versus les autres
- quantitative ou qualitative ordinale : variable \geq valeur versus $<$ valeur

Etape 3 : construction du score à l'échelle agrégée

1. Sélection des variables associées à l'indicateur individuel

- Régression logistique univariée

IDI(pauvre/pas pauvre)=chaque variable disponible aux deux échelles

- Régression logistique multivariée

IDI(pauvre/pas pauvre)=les variables précédemment sélectionnées

PS : Aucune méthode de sélection n'est appliquée, les variables non significatives sont retirées "en même temps".

2. Pondération des variables pour l'indice écologique

Utilisation des coefficients de la régression logistique précédente comme pondération lors de la construction du score à l'échelle écologique (données du recensement).

$$Score_i = \sum_1^n \beta_i variable_i$$

i=unité géographique

n=nombre de variables dans le modèle

Publications

- Pornet et al, JECH2012
- Guillaume et al, JECH2016