

# Reconstituer la facture d'énergie totale des ménages à partir de deux sources distinctes (ENL et ENTD) :

le pseudo-appariement dans le modèle de microsimulation *Prometheus*

---

M. Clément, C. Peroumal, A. Rolland

Ministère de la Transition Écologique et Solidaire

CGDD\Seeidd

**Journées de Méthodologie Statistique**

***13 juin 2018 – session 6***

***Projections et microsimulation***

Photo : A. Bouissou/Terra



MINISTÈRE  
DE LA TRANSITION  
ÉCOLOGIQUE  
ET SOLIDAIRE

# PLAN

- I – Présentation du modèle de microsimulation *Prometheus*
  - A) Objectifs et utilisations du modèle
  - B) Structure du programme de microsimulation
  
- II – Le pseudo-appariement dans le modèle
  - A) Pourquoi faire ?

*Les étapes :*
  - B) Avant : sélection des variables d'appariement
  - C) Pseudo-appariement
  - D) Après : contrôle qualité de l'appariement

# Pourquoi construire *Prometheus* ?

*PR*Ogramme de *Microsimulation des Énergies du Transport et de l'Habitat pour ÉvalUations Sociales*

## ■ Né en 2013-2014, contexte :

- Volonté d'utiliser la **fiscalité énergétique** pour changer les comportements des ménages vers des pratiques plus durables
- Dans le même temps ... Hausse tendancielle des prix des énergies

## ■ Besoin pour le **Seeidd** de se doter de capacités d'analyse des **impacts sociaux** de la fiscalité énergétique...

- **Évaluer** l'impact de mesures fiscales (ex-ante et ex-post)
  - => Mieux **calibrer** les dispositifs
- En tenant compte de la **diversité des situations** (localisation, revenus, etc.)
  - => Connaître les disparités entre ménages, pour mieux les **compenser**

= Encourager **l'acceptabilité** de ces mesures et **promouvoir** une fiscalité énergétique favorisant la transition énergétique

# Que permet de faire PROMETHEUS ?

- **Prometheus permet d'estimer :**

- les **consommations** individuelles des ménages en **énergies (logement et carburants)** et les **factures** correspondantes
- Ses atouts : un modèle basé sur des **données « réelles »** (ENL de l'Insee et ENTD du SDES), **au niveau du ménage**

- **Prometheus permet de simuler :**

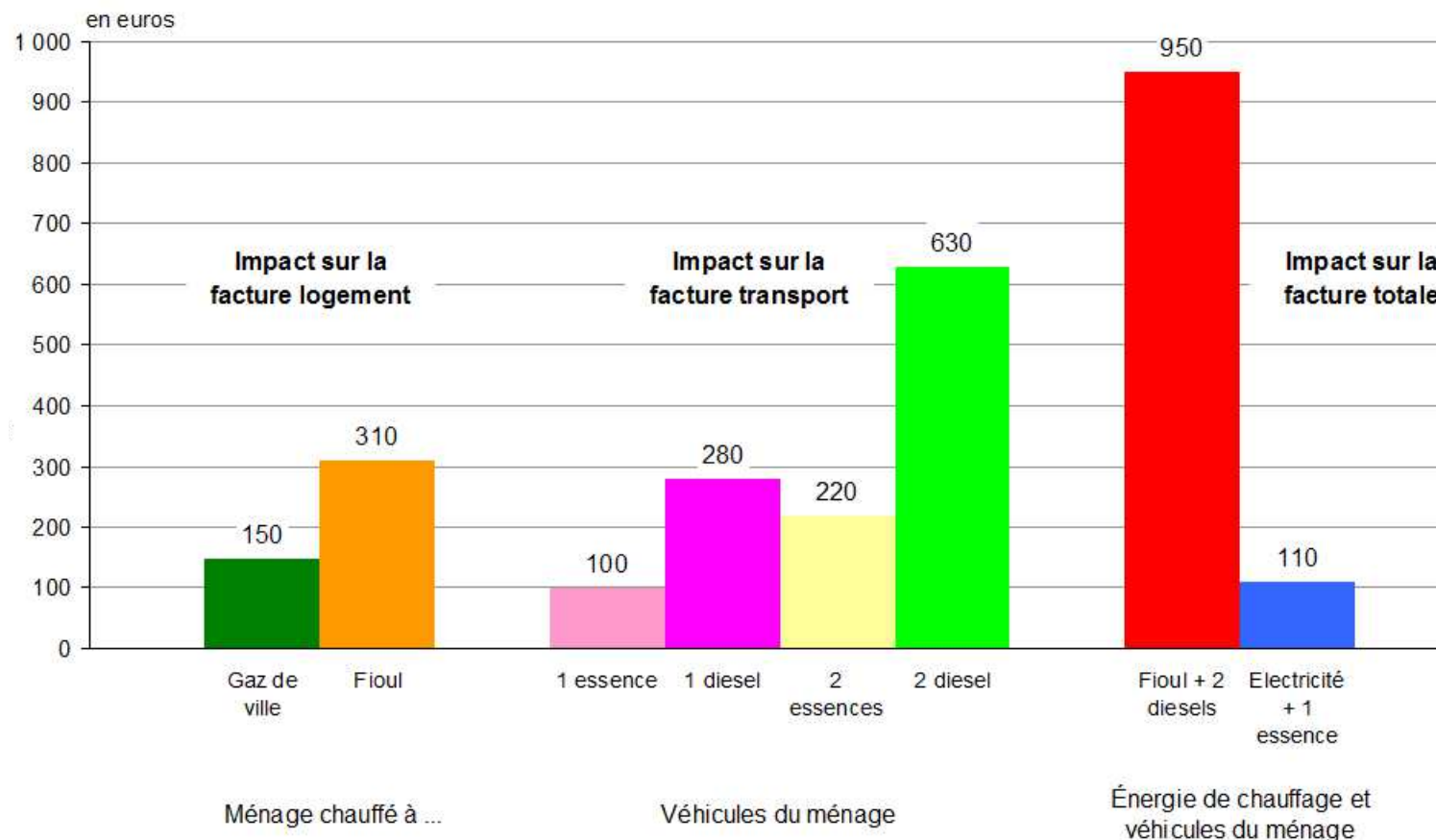
- une **taxe** (modification, création...) : la composante carbone,
- Une **mesure compensatoire** : le chèque énergie,
- une modification de prix, etc.

- **Utilisations :**

- **« Réponse à la demande »** : Comité pour la fiscalité énergétique (CFE) puis Comité de l'économie verte (CEV), Stratégie nationale bas carbone (SNBC), ONPE (précarité énergétique), etc.
- **Études** : exemple publication *Insee Références* : « La facture énergétique du logement a baissé depuis 2013 malgré la hausse de la fiscalité »

# Exemple 1 : évaluation de l'impact de la nouvelle trajectoire carbone

- Impact en 2022 relativement à 2017 de l'accélération de la trajectoire carbone et du rattrapage gazole-essence (PLF 2018)



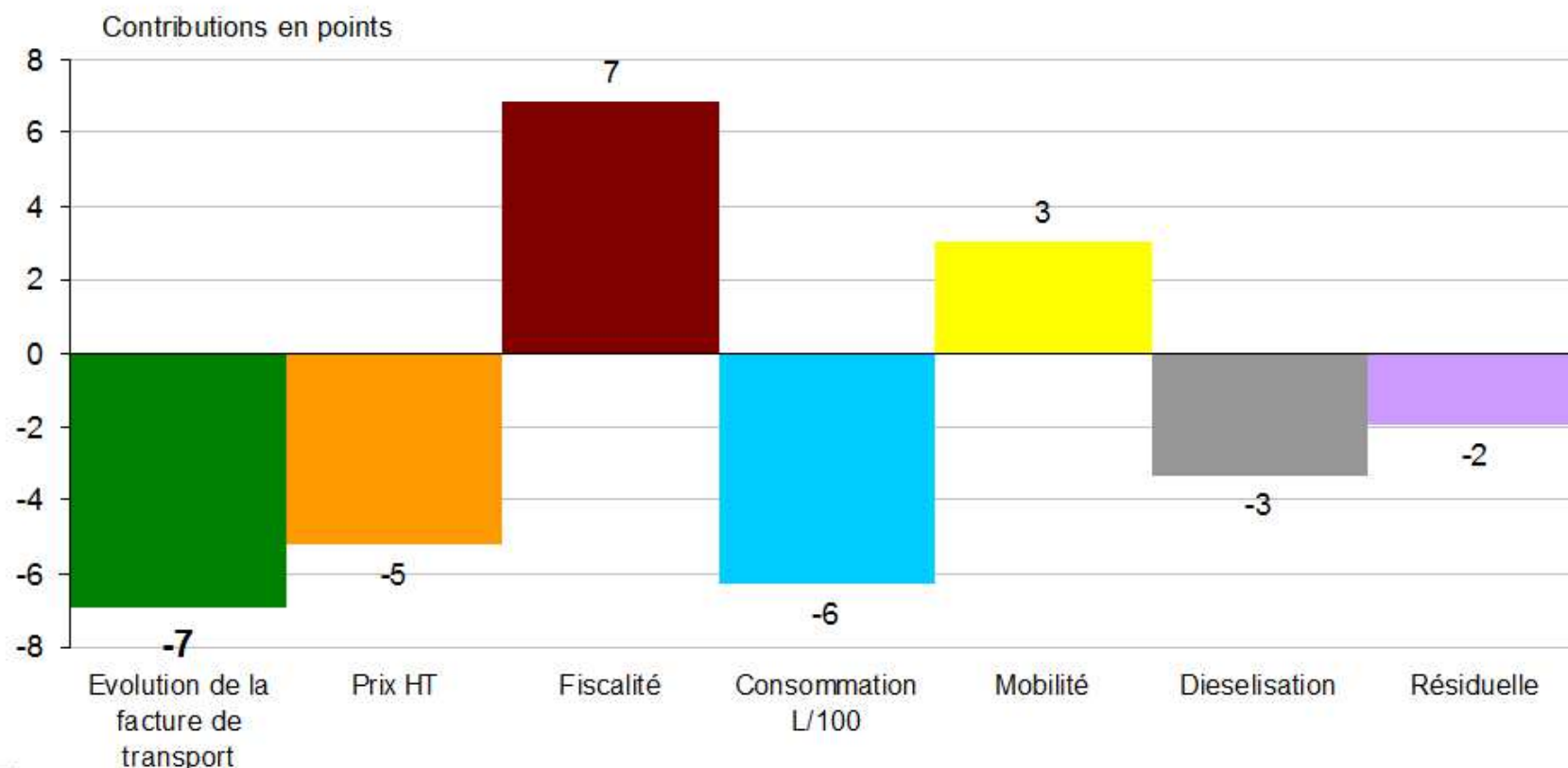
Source : estimations CGDD mars 2018 – modèle Prometheus

Champ : France métropolitaine, dépenses d'énergie dans les résidences principales, dépenses de carburant des ménages et des entrepreneurs individuels.

Note : surcoûts entre 2017 et 2022, exprimés en euros yc TVA, toutes choses égales par ailleurs (hors ajustements des consommations à la baisse) : consommations d'énergies (corrigées de la météo pour le chauffage), parcs et mobilités 2016 ; prix 2017.

## Exemple 2 : évolution de la facture de carburant depuis 10 ans, quels facteurs ?

- La facture de carburant des ménages a reculé de 7 % en moyenne entre 2006 et 2016



Source : estimations CGDD mars 2018 – modèle Prometheus.

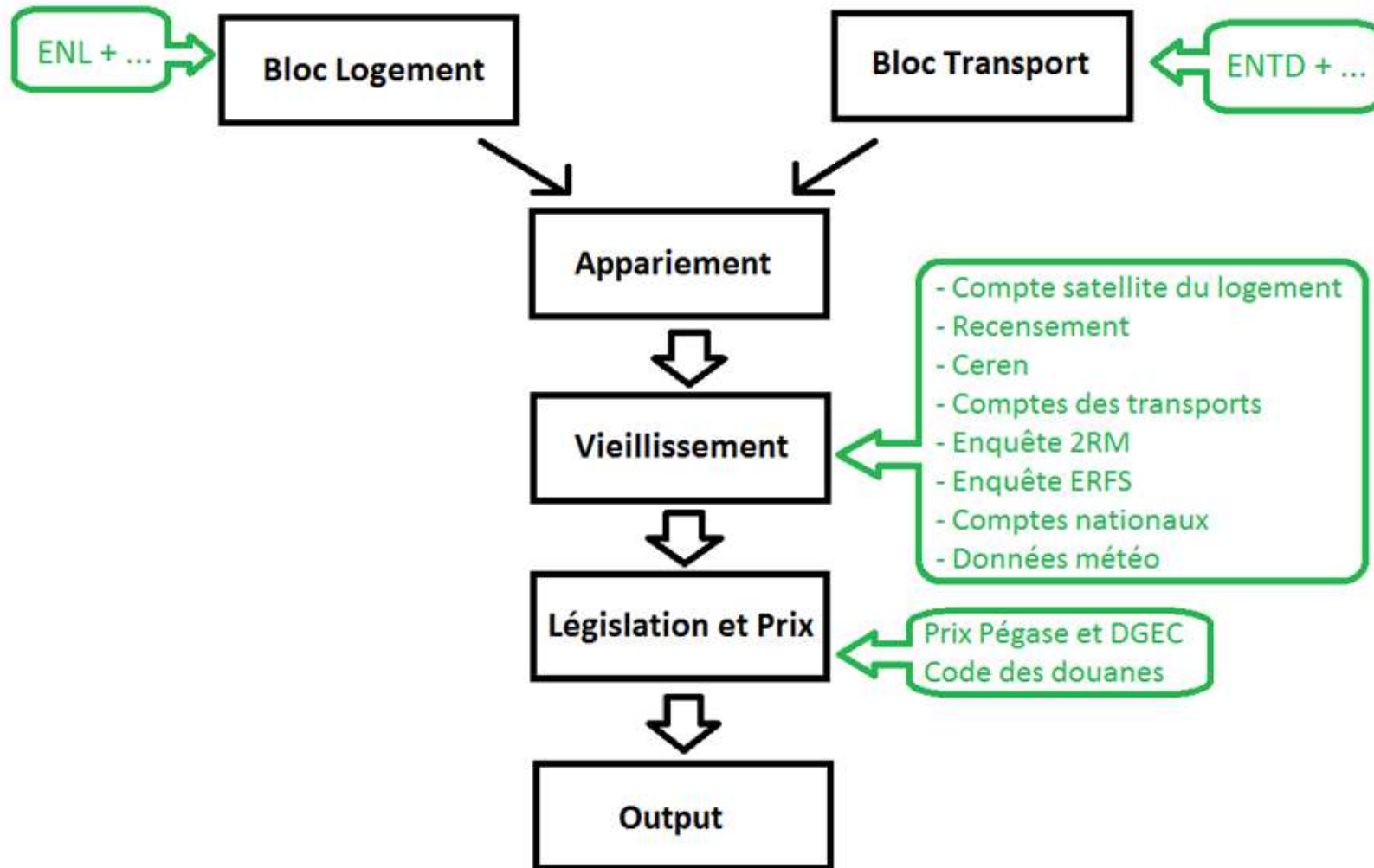
Champ : France métropolitaine, dépenses d'énergie dans les résidences principales, dépenses de carburant des ménages et des entrepreneurs individuels.

# PLAN

- I – Présentation du modèle de microsimulation *Prometheus*
  - A) Objectifs et utilisations du modèle
  - B) Structure du programme de microsimulation
  
- II – Le pseudo-appariement dans le modèle
  - A) Pourquoi faire ?  
*Les étapes :*
  - B) Avant : sélection des variables d'appariement
  - C) Pseudo-appariement
  - D) Après : contrôle qualité de l'appariement

I– B) Structure du programme de micro-simulation

# Étapes de construction du modèle et sources utilisées





# PLAN

- I – Présentation du modèle de microsimulation *Prometheus*
  - A) Objectifs et utilisations du modèle
  - B) Structure du programme de microsimulation
  
- II – Le pseudo-appariement dans le modèle
  - A) Pourquoi faire ?
    - Les étapes :*
    - B) Avant : sélection des variables d'appariement
    - C) Pseudo-appariement
    - D) Après : contrôle qualité de l'appariement

## II– Le pseudo-appariement

# Pourquoi faire ?

- **Input : 2 bases**

- « **Logement** », base ENL :
  - consommations d'énergies du logement pour chaque ménage ENL en 2013
- « **Transport** », base ENTND :
  - consommations de carburants pour chaque ménage ENTND en 2008

- **Objectif : pouvoir simuler les dépenses totales (logement et carburant) au niveau d'un même ménage**

- **Appariement :**

- ENL choisie comme source « base » du modèle
  - Plus récente, plus robuste, plus grosse
- A chaque ménage de l'ENL on apparie un ménage comparable de l'ENTND :
  - On attribue au ménage ENL « receveur » les consommations, les véhicules et la mobilité du ménage ENTND « donneur » apparié.
  - Chaque ménage ENL a ainsi in fine des consommations de carburants « 2008 »

## II– Le pseudo-appariement

# Les étapes de l'appariement

### Les étapes :

#### A) Sélection des variables d'appariement :

1. Quelles sont les informations communes aux deux enquêtes ?
2. On recode et on compare les distributions de ces variables
3. On sélectionne et hiérarchise les variables pertinentes en termes d'explication de la facture de carburants

#### B) Pseudo-appariement :

1. Tirage aléatoire simple par strate avec condition de nombre de donneurs minimal

#### C) Contrôle de la qualité des résultats obtenus :

1. Distribution du nombre de variables de stratification utilisées pour l'appariement
2. Le nombre de donneurs distincts utilisés est-il de bonne qualité?
3. Comparaison des caractéristiques mobilité / véhicules avant et après appariement



## II– Le pseudo-appariement

# Les étapes de l'appariement

### Les étapes :

#### A) Sélection des variables d'appariement :

1. Quelles sont les informations communes aux deux enquêtes ?
2. On recode et on compare les distributions de ces variables
3. On sélectionne et hiérarchise les variables pertinentes en termes d'explication de la facture de carburants

#### B) Pseudo-appariement :

1. Tirage aléatoire simple par strate avec condition de nombre de donneurs minimal

#### C) Contrôle de la qualité des résultats obtenus :

1. Distribution du nombre de variables de stratification utilisées pour l'appariement
2. Le nombre de donneurs distincts utilisés est-il de bonne qualité?
3. Comparaison des caractéristiques mobilité / véhicules avant et après appariement



# Quelles sont les informations communes aux deux enquêtes ?

- Les informations **directement** liées à la **mobilité** du ménage :
  - Le nombre de **véhicules** du ménage
  - Le temps de trajet domicile-travail de la PR et de son conjoint (**mobilité contrainte**)
- Des caractéristiques socio-démographiques du ménage **possiblement liées** à la mobilité des ménages :
  - **Taille** du ménage
  - **Revenu** par UC
  - **Âge** de la Personne de Référence (PR)
  - **Sexe** de la PR
  - Situation d'**activité** de la PR
  - **Niveau d'éducation** de la PR
  - Taille de l'**unité urbaine**
  - Possession d'une **résidence secondaire**
  - Etc.

## II – A) Sélection et hiérarchisation des variables d'appariement

# Recodage et comparaison des distributions

### ■ Exemple de l'équipement en véhicules

- Questionnaires Recensement (variable appariée ENL) et ENT D

**13** De combien de voitures les habitants de ce logement disposent-ils ?  
*Ne comptez pas les voitures ou les fourgonnettes à usage exclusivement professionnel.*

• Aucune .....  0

• 1 .....  1

• 2 .....  2

• 3 ou plus .....  3

**JNBVEH**

**De combien de voitures particulières votre ménage dispose-t-il au total ?**

Qu'elles soient au nom d'un des membres du ménage, d'un proche ou au nom d'une société, y-compris de leasing. Donnez le nombre précis.

**JNBCCVUL**

**De combien de camping-cars, fourgonnettes, ou véhicules utilitaires légers votre ménage dispose-t-il au total ?**

Compter les véhicules dont la conduite exige un permis B, pas les voiturettes, quads, camions ou tracteurs. Donnez le nombre précis.

- Distributions :

Nb voitures et VUL	Part des ménages (%)	
	ENL(RP 2011)	ENTD
0	17,8	19,1
1	47,3	44,7
2	29,8	29,6
3 ou plus	5,1	6,6

- Champ et années légèrement différents

## II – A) Sélection et hiérarchisation des variables d'appariement

# Recodage et comparaison des distributions

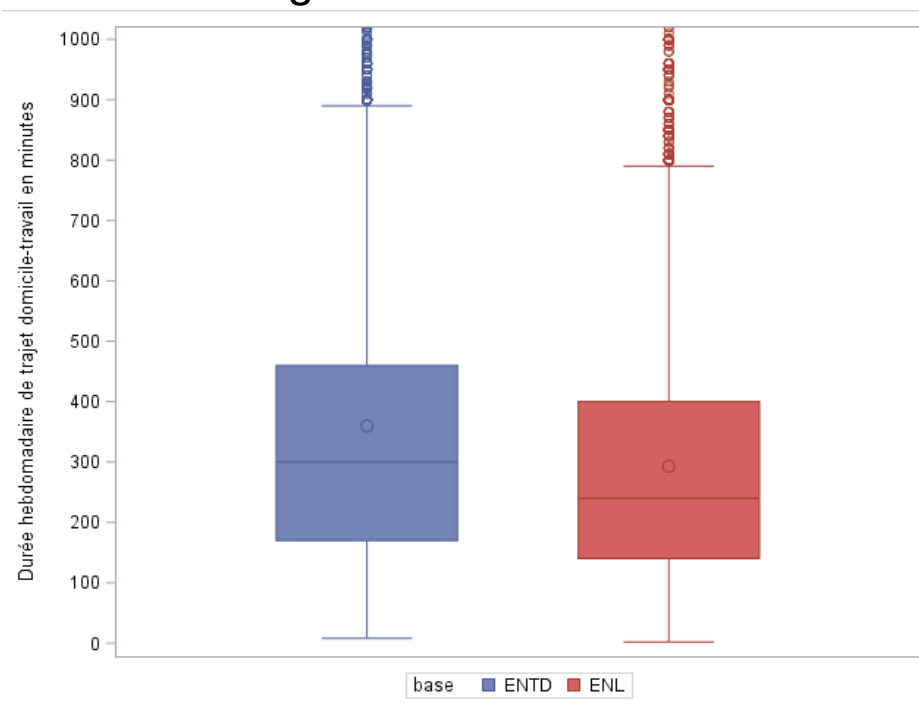
### ■ Exemple de la mobilité contrainte

- Dans l'ENL : temps de trajets hebdomadaires domicile-travail PR et conjoint
- Dans l'ENTD : distances domicile trajet – études – garderie de tous les membres du ménage + temps de trajets

=> On reconstitue dans l'ENTD une variable homogène à celle de l'ENL :

- Les distributions sont différentes :

Tps de parcours hebdomadaire domicile-travail (en min)	ENTD	ENL
Q1	170	140
Médiane	300	240
Q3	460	400



=> On utilise donc la variable catégorielle donnant le décile d'appartenance.

## II – A) Sélection et hiérarchisation des variables d'appariement

# Hiérarchisation des variables

- Préalablement dans le modèle : modèle d'**Imputation du kilométrage annuel** du ménage sur les données ENTD (Bloc Transport)
- => On utilise les résultats de cette régression linéaire pour déterminer les **variables avec le plus fort pouvoir explicatif**.

Variable	Estimation des paramètres	Significativité	Référence
<b>Nb de voitures ou VUL</b>			1 essence
1 diesel	4400	***	
2 diesels	16800	***	
2 essence	7400	***	
1 diesel + 1 essence	12400	***	
<b>Nb de deux-roues</b>			0 deux-roues
1 deux-roues	3400	***	
2+ deux-roues	10600	***	
<b>Mobilité contrainte (km TEG)</b>	0,2	***	



## II – A) Sélection et hiérarchisation des variables d'appariement

# Hiérarchisation des variables

(suite)

Variable	Estimation des paramètres	Significativité	Valeur de réf
<b>log du Revenu disponible par UC</b>	1100	***	
<b>Âge de la Personne de référence</b>			< 50 ans
De 50 à 65 ans	-1100	***	
De 65 à 75 ans	-2700	***	
de plus de 75 ans	-4800	***	
<b>Taille unité urbaine</b>			< 5000 hab
5.000 à 50 000 habitants	-700	*	
50 000 à 200 000 habitants	-1700	***	
200 000 à 2 000 000 habitants	-1800	***	
Paris	-2500	***	

## II – A) Sélection et hiérarchisation des variables d'appariement

# Hiérarchisation des variables

- Étude des corrélations linéaires

Variable	Corrélation linéaire
Nb voitures et VUL	0,47
Durée Trajet domicile travail	0,33
Age de la PR	-0,24
Log Revenu/UC	0,20

## II– Le pseudo-appariement

# Les étapes de l'appariement

### Les étapes :

#### A) Sélection des variables d'appariement :

1. Quelles sont les informations communes aux deux enquêtes ?
2. On recode et on compare les distributions de ces variables
3. On sélectionne et hiérarchise les variables pertinentes en termes d'explication de la facture de carburants

#### B) Pseudo-appariement :

1. Tirage aléatoire simple par strate avec condition de nombre de donneurs minimal

#### C) Contrôle de la qualité des résultats obtenus :

1. Distribution du nombre de variables de stratification utilisées pour l'appariement
2. Le nombre de donneurs distincts utilisés est-il de bonne qualité?
3. Comparaison des caractéristiques mobilité / véhicules avant et après appariement

## II – B) Pseudo-appariement

# Nb de variables de stratification

- **Pseudo-appariement** à partir des caractéristiques communes entre l'ENL et l'ENTD corrélées à la mobilité (11 variables)

- Nb de voitures
- Âge de la Personne de Référence (PR)
- Taille de l'unité urbaine
- Déciles de durée des déplacements contraints
- Type d'habitat au voisinage du logement
- Quantiles de revenus par UC
- Résidence secondaire
- Sexe de la PR
- Taille du ménage
- Activité de la PR
- Niveau d'éducation de la PR

NbVarJointure	Pctage cumulé(%)
2	100
3	100
4	98
5	96
6	88
7	84
8	72
9	48
10	40
11	18

*Tirage aléatoire par strate avec nombre de donneurs par strate  $\geq 5$*

## II– Le pseudo-appariement

# Les étapes de l'appariement

### Les étapes :

#### A) Sélection des variables d'appariement :

1. Quelles sont les informations communes aux deux enquêtes ?
2. On recode et on compare les distributions de ces variables
3. On sélectionne et hiérarchise les variables pertinentes en termes d'explication de la facture de carburants

#### B) Pseudo-appariement :

1. Tirage aléatoire simple par strate avec condition de nombre de donneurs minimal

#### C) Contrôle de la qualité des résultats obtenus :

1. Distribution du nombre de variables de stratification utilisées pour l'appariement
2. Le nombre de donneurs distincts utilisés est-il de bonne qualité?
3. Comparaison des caractéristiques mobilité / véhicules avant et après appariement



## II – C) Contrôle de la qualité des résultats de l'appariement

# Nombre de donneurs distincts

- **62 % des donneurs utilisés** (sur les 20 000 ménages) : **est ce trop peu ?**

Nb donneurs distincts	Nb receveurs	Ratio receveur/donneur
12437	27137	2,1

- Exercices de comparaison... avec le cas sans variable de stratification

Soit  $X_i$  la variable aléatoire prenant ses valeurs dans  $\{0, 1\}$ .  $X_i = 1$  si le donneur  $i$  est tiré au moins une fois pour être apparié à un ménage de l'ENL.

On a donc :

$$P(X_i = 0) = \left[ \frac{\text{card}(ENTD) - 1}{\text{card}(ENTD)} \right]^{\text{card}(ENL)}$$

$$P(X_i = 1) = 1 - P(X_i = 0)$$

$$E \left[ \sum_i^{\text{card}(ENTD)} X_i \right] = \sum_i^{\text{card}(ENTD)} P(X_i = 1)$$

=> Espérance de **74 %** des donneurs ENTD appariés

## II – C) Contrôle de la qualité des résultats de l'appariement

### Nombre de donneurs distincts

- ... Avec le cas théorique avec 1 variable de stratification : le nombre de voitures

Nb de voitures	Donneur	Receveurs	Nb moyen donneurs appariés	Part donneurs
0	3297	5111	2598	78,8%
1	7872	12790	6322	80,3%
2	7149	7897	4780	66,9%
3 ou plus	1860	1339	955	51,3%
Total	20178	27137	14654	72,6%

TABLE 1 – Espérance du nombre de donneurs appariés

- 72,6 %... => 62 % avec 11 variables semble cohérent

## II – C) Contrôle de la qualité des résultats de l'appariement

# Comparaison des caractéristiques avant et après appariement

	ENTD 2008 (avant appariement)	Après appariement (sur N=50 appariements)			Cible SDES 2008
		Moyenne	Ecart-type	Ecart	
Taux de <b>diesélisation</b> du parc	55,5 %	53,7 %	0,3	-3,3%	55,4 %
<b>Km moyen</b> annuel Véhicules <b>GO</b>	16 000	15 600	85	-2,3 %	15 500
<b>Km moyen</b> annuel Véhicules <b>ES</b>	9 600	9 300	65	-3,4 %	8 900
<b>conso unitaire</b> (L/100km) véhicules <b>GO</b>	6,53	6,55	0,01	+0,2 %	6,90
<b>conso unitaire</b> (L/100km) véhicules <b>ES</b>	7,18	7,14	0,01	-0,6 %	7,76





# Conclusion

- Le pseudo-appariement dans **PROMETHEUS**
  - C'est tout neuf !
    - Développé l'année passé,
    - Auparavant, seulement les dépenses d'énergies du logement dans le modèle.
  - Cela permet :
    - De prendre en compte la totalité de la facture énergétique pour un même ménage,
    - D'envisager des études sur la précarité énergétique (Logement) mais aussi sur la vulnérabilité énergétique globale (Logement ET Transport).
  - La suite ?
    - La prochaine enquête « mobilité des personnes » 2018-2019 du SDES, sur le terrain actuellement,
- Prometheus : **un outil à exploiter, et qui s'enrichit encore ...**
  - Étude de la facture de carburants sous l'angle de la **mobilité contrainte**,
  - Bloc « **comportements** » : fonctions de réponse des demandes en énergies des consommateurs,
    - Introduction d'élasticité-prix.

# FIN



*Liberté • Égalité • Fraternité*  
RÉPUBLIQUE FRANÇAISE

MINISTÈRE  
DE LA TRANSITION  
ÉCOLOGIQUE  
ET SOLIDAIRE

Ministère de la Transition écologique et solidaire

[www.ecologique-solidaire.gouv.fr](http://www.ecologique-solidaire.gouv.fr)