



## **L'état des statistiques Belgique : constats et recommandations**

*par Philippe DEFEYT - août 2013*

*« We don't know everything, but that doesn't mean we know nothing. »*

The Economist<sup>1</sup>

*« Je suis loin sans doute de prétendre que quelques tableaux numériques isolés peuvent suffire pour déterminer complètement tous les éléments si compliqués de nos sociétés modernes. Il faudrait, pour remonter des effets aux causes, ou pour conclure de ce qui est à ce qui sera, avoir égard à un ensemble de circonstances qu'il n'est point donné à l'homme de pouvoir embrasser : de là, la nécessité de négliger toujours, dans toute espèce d'appréciation, un certain nombre de circonstances dont il aurait fallu tenir compte. De là aussi l'absurdité des résultats auxquels conduit souvent cette énumération incomplète, ou le trop d'importance qu'on attache à un élément qui ne devrait être considéré que comme secondaire. La mauvaise foi pourra même porter à ne choisir dans une série de résultats, que ceux qui sont favorables au principe qu'on voudrait faire prévaloir, en passant sous silence ceux qui lui seraient contraires : et c'est ainsi, comme on l'a fort bien observé, que tout pourrait se prouver par les nombres de la statistique. Mais de ce que l'observation est difficile et de ce qu'il existe des ignorants ou des hommes de mauvaise foi, faut-il en conclure qu'on doit rejeter la statistique ? Non sans doute (...) »*

Adolphe QUETELET<sup>2</sup>

La présente note de l'Institut pour un Développement Durable trouve son origine dans un projet d'analyse – étudier l'évolution du marché du travail belge depuis le début de la crise (2008) sur base de l'Enquête sur les forces de travail<sup>3</sup> – et la préparation d'un exposé sur la pauvreté sur base des résultats de l'enquête SILC (*Statistics on Income and Living Conditions*).

Les travaux préparatoires liés à ces deux démarches m'ont progressivement amené à m'interroger sur la qualité de données très utilisées dans le débat socio-économique. Je prolonge cette analyse par quelques considérations plus générales sur l'état des statistiques en Belgique et par quelques recommandations.

Commençons par regarder d'un peu plus près le marché du travail.

<sup>1</sup> The Economist, July 6<sup>th</sup> 2013

<sup>2</sup> A. QUETELET, « Recherches statistiques sur le Royaume des Pays-Bas », Mémoire lu à la séance de l'Académie du 6 décembre 1828, Bruxelles, M. Hayez, Imprimeur de l'Académie royale, 1829, p. III

<sup>3</sup> Voir Philippe DEFEYT, « 2008-2012 : un marché du travail prévisible et étonnant à la fois », Institut pour un Développement Durable, juillet 2013

## Le marché du travail

Les résultats de l'Enquête sur les forces du travail (EFT) ont pris progressivement une place essentielle dans le débat socio-économique et politique en matière d'emploi et de chômage. Le lecteur peu familiarisé avec cette enquête en trouvera une présentation dans l'encadré suivant.

### Présentation de l'EFT

« L'enquête par sondage sur les Forces de Travail (EFT), organisée en Belgique par la Direction générale Statistiques et Information économique du SPF Economie, est une enquête socio-économique auprès des ménages dont le but principal est de classer la population en âge de travailler (15 ans et plus) en trois groupes (personnes occupées, chômeurs et inactifs), et de fournir des données descriptives et explicatives sur chacune de ces catégories. Cette enquête s'inscrit dans le cadre des enquêtes communautaires par sondage sur la population active coordonnées par l'Office statistique des Communautés européennes EUROSTAT en collaboration avec les Instituts nationaux de statistique et réalisées sous la responsabilité de ces derniers. Il s'agit d'obtenir des informations comparables au niveau européen notamment pour ce qui concerne les taux d'emploi et de chômage au sens du Bureau international du Travail (BIT), mais également de collecter et de diffuser des résultats non disponibles par ailleurs (mobilité des travailleurs, motifs du temps partiel, niveau de formation de la population, profession...). »

Pour la suite de l'analyse il importe d'avoir en tête deux définitions :

« Chômeurs (BIT) : Conformément aux normes établies par le Bureau international du travail, les chômeurs comprennent les personnes âgées de 15 ans et plus qui étaient :

(a) sans travail pendant la semaine de référence

(b) disponibles pour travailler, c'est-à-dire pour commencer une activité en tant que salarié ou non dans un délai de deux semaines suivant la semaine de référence

(c) à la recherche active d'un travail, c'est-à-dire qui avaient entrepris des démarches spécifiques en vue de trouver un emploi (salarié ou non) pendant une période de quatre semaines se terminant à la fin de la semaine de référence, ou qui avaient trouvé un travail à commencer plus tard, c'est-à-dire endéans une période maximale de trois mois.

(...)

Population ayant un emploi : Les personnes ayant un emploi sont toutes celles âgées de 15 ans et plus qui, au cours de la semaine de référence, ont effectué un travail d'une durée d'une heure au moins contre rémunération ou en vue d'un bénéfice, ou celles qui ne travaillaient pas mais occupaient un emploi dont elles étaient temporairement absentes. Les aidants familiaux sont également inclus. »

Source : SPF Economie

Les définitions de l'EFT n'étant pas semblables à celles d'autres approches statistiques, il est normal qu'il y ait des différences. C'est ce que confirme le tableau en haut de la page suivante pour les grands agrégats du marché du travail.

Des différences sont explicables. En matière de chômage on sait que les statistiques administratives surévaluent le nombre de personnes activement à la recherche d'un emploi ou vraiment disponibles sur le marché du travail. On peut même comprendre que les écarts varient dans le temps. Parce que, par exemple, l'attitude des personnes en matière de recherche d'un emploi peut varier en fonction de la conjoncture.

Mais les différences apparaissent trop grandes (par exemple en matière d'emploi) et/ou trop fluctuantes (par exemple en matière de chômage) pour trouver des explications convaincantes à celles-ci.

**T1. Approche globale du marché du travail – X 1.000 et écarts en %  
Comparaison Bureau fédéral du Plan (BfP) – Enquête sur les forces de travail (EFT)**

	2008	2009	2010	2011	2012
<b>Emploi total</b>					
BfP	4.539	4.531	4.563	4.622	4.629
EFT	4.421	4.421	4.489	4.509	4.524
Ecart	-118	-110	-74	-113	-105
En % BfP	-2,6%	-2,4%	-1,6%	-2,4%	-2,3%
<b>Chômage</b>					
BfP	500	551	565	545	559
EFT	380	380	406	347	369
Ecart	-120	-171	-159	-198	-190
En % BfP	-24,1%	-31,1%	-28,1%	-36,4%	-34,0%
<b>Population active</b>					
BfP	5.039	5.082	5.127	5.167	5.188
EFT	4.800	4.800	4.895	4.856	4.893
Ecart	-238	-282	-233	-311	-295
En % BfP	-4,7%	-5,5%	-4,5%	-6,0%	-5,7%

**Sources : BfP et EFT – Calculs : IDD**

Pour soutenir cette affirmation comparons d'abord les statistiques des personnes ayant un emploi. On observe à la lecture du tableau suivant (voir page suivante) :

- une plus ou moins bonne concordance entre les deux sources statistiques utilisées ici en ce qui concerne les salariés, surtout en fin de période
- un écart énorme - environ 200.000 emplois ! - pour ce qui est de l'emploi salarié privé et donc, forcément un écart opposé mais du même ordre de grandeur pour l'emploi public ; certes, il s'agit ici de définitions différentes (par exemple le Bureau fédéral du Plan met les entreprises publiques dans le secteur « privé ») ; dans son acception la plus large (ONSS secteur public + ONSS-APL) l'emploi public concerne fin 2011 environ 1.100.000 travailleurs, mais y compris tous les enseignants qui travaillent dans l'enseignement "privé" (au moins 200.000) ; en fonction de la question posée dans le questionnaire EFT, voir ci-après, on a du mal à comprendre que plus de un million de salariés déclarent travailler dans le public<sup>4</sup> ;

**EFT – Activité principale – Question 5**

B1. Activité principale

5. Au cours de la semaine de référence, quelle catégorie caractérise le mieux votre statut professionnel ?

(Instruction pour l'enquêteur : une seule réponse possible)

- Secteur privé – ouvrier(ère)
- Secteur privé – employé(e)
- Secteur public – fonctionnaire statutaire
- Secteur public – contractuel(le)
- Indépendant(e) sans personnel
- Indépendant(e) avec personnel
- Aidant(e) non rémunéré(e)

- une différence de presque 100.000 unités pour l'emploi indépendant en fin de période

<sup>4</sup> La question du nombre de travailleurs "publics" (abusivement et improprement souvent assimilés aux "fonctionnaires") est d'autant plus importante que ce nombre donne lieu à de nombreuses lectures politico-idéologiques.

d'observation

- des fluctuations importantes, difficilement explicables, des écarts observés entre ces deux sources.

On notera, d'une manière générale, que les catégories de la question 5 et l'obligation de ne cocher qu'une seule case ne correspondent plus aux réalités vécues par de nombreux travailleurs et à la fluidité des catégories socioprofessionnelles.

**T2. Personnes actives ayant un emploi – X 1.000 et écarts en %  
Comparaison Bureau fédéral du Plan – Enquête sur les forces de travail**

	2008	2009	2010	2011	2012
<b>Total salariés</b>					
BfP	3.823	3.811	3.837	3.886	3.888
EFT	3.766	3.766	3.844	3.864	3.875
Ecart	-57	-44	8	-22	-13
En % BfP	-1,5%	-1,2%	0,2%	-0,6%	-0,3%
<b>Dont salariés du secteur privé</b>					
BfP	3.002	2.977	2.996	3.045	3.051
EFT	2.754	2.754	2.823	2.845	2.856
Ecart	-249	-223	-172	-200	-195
En % BfP	-8,3%	-7,5%	-5,8%	-6,6%	-6,4%
<b>Dont salariés du secteur public</b>					
BfP	821	834	841	841	838
EFT	1.013	1.013	1.021	1.019	1.019
Ecart	192	179	180	178	181
En % BfP	23,4%	21,5%	21,4%	21,1%	21,6%
<b>Indépendants et aidants</b>					
BfP	716	720	726	736	741
EFT	654	654	645	645	649
Ecart	-61	-66	-81	-91	-92
En % BfP	-8,6%	-9,2%	-11,2%	-12,3%	-12,4%
<b>Emploi total</b>					
BfP	4.539	4.531	4.563	4.622	4.629
EFT	4.421	4.421	4.489	4.509	4.524
Ecart	-118	-110	-74	-113	-105
En % BfP	-2,6%	-2,4%	-1,6%	-2,4%	-2,3%

**Sources : BfP et EFT – Calculs : IDD**

Le tableau du haut de la page suivante compare – par catégorie d'âge et en % – l'écart entre le nombre de chômeurs déterminé sur base administrative (tous ceux qui sont considérés comme demandeurs d'emploi inoccupés – DEI) et celui qui découle de l'EFT.

A nouveau il est normal qu'il y ait un écart, même important, et que cet écart fluctue en fonction de la conjoncture. Mais à nouveau les fluctuations des écarts sont trop importantes ou trop différentes d'une catégorie d'âge à l'autre pour pouvoir être expliquées.

**Note méthodologique :** La dernière colonne du tableau ci-dessous devrait correspondre parfaitement à la 8<sup>ème</sup> ligne de données du premier tableau (Chômage – en % du BfP). Il y a néanmoins de minimes différences dans la mesure où le Bureau fédéral du Plan apporte quelques corrections aux données des DEI de Belgostat, qui lui même les collecte auprès des offices régionaux de l'emploi.

### **T3. Chômeurs-demandeurs d'emploi – écarts en %** **Comparaison Enquête sur les forces de travail - Belgostat**

	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60 et +	Total
<b>2.008</b>	-24,2%	-5,7%	-13,4%	-11,4%	-24,2%	-26,6%	-34,8%	-54,8%	-46,1%	10,6%	-24,8%
<b>2.009</b>	-33,7%	-16,5%	-22,5%	-19,7%	-29,7%	-31,2%	-38,0%	-55,6%	-56,7%	3,5%	-31,5%
<b>2.010</b>	-20,8%	-17,2%	-13,5%	-25,1%	-22,4%	-28,4%	-31,1%	-44,5%	-64,8%	-3,5%	-28,4%
<b>2.011</b>	-26,1%	-27,1%	-23,6%	-28,0%	-37,4%	-30,5%	-39,4%	-50,7%	-68,8%	-36,9%	-36,6%
<b>2.012</b>	-33,4%	-26,0%	-17,6%	-24,7%	-30,5%	-31,6%	-36,0%	-52,3%	-67,2%	-16,7%	-34,1%

**Sources : Belgostat et EFT – Calculs : IDD**

On peut aussi, ce sera la dernière comparaison concernant le marché du travail, comparer la structure par âge de l'emploi salarié, en prenant comme référence les données de la Banque carrefour de la sécurité sociale (BCSS).

#### **Notes méthodologiques :**

- *La comparaison la plus récente est seulement possible pour 2010, la BCSS n'ayant pas – à la date de fin juin 2013 – actualisé sa banque de données pour 2011 .*
- *Attention : au vu des incertitudes (presque 9.000 salariés dont l'âge est inconnu en décembre 2009 pour la BCSS et on ne connaît l'âge des travailleurs qui travaillent à l'étranger que globalement dans l'EFT), la structure par âge constitue une approximation dans les deux sources.*

### **T4. L'emploi salarié par âge – X 1.000 et écarts en %**

	De 15 à 24 ans	De 25 à 49 ans	De 50 ans à plus	Total
BCSS	324	2.432	886	3.642
EFT	309	2.557	868	3.735
Ecart	-15	126	-18	93
En % BCSS	-4,6%	5,2%	-2,0%	2,6%

**Sources : BCSS et EFT – Calculs et estimations : IDD**

Ces observations amènent tout naturellement à s'interroger sur la confiance que l'on peut accorder à, par exemple, les analyses concernant les marchés du travail des jeunes et des travailleurs âgés. Peut-on, par exemple, partager les analyses récentes plutôt optimistes sur l'évolution du chômage des jeunes ? Peut-on faire une étude un peu sérieuse sur le phénomène des NEETs (les jeunes qui ne sont ni à l'emploi, ni en stage et ni aux études ou en formation) ? Que valent, autre exemple, les analyses sur l'évolution de la population active des plus de 55 ans faites sur base de l'EFT ?

Quel crédit attribuer au niveau du taux de chômage tel que calculé par Eurostat et à ses fluctuations, surtout quand ces observations concernent des sous-ensembles de la population totale, dès lors qu'il "manque" environ 100.000 personnes à l'emploi et que l'estimation du nombre de chômeurs est à ce point fluctuante par rapport à d'autres définitions/approches statistiques et conceptuelles du chômage ?

## **Les revenus**

Deux sources essentielles – l'enquête SILC et l'enquête sur le budget des ménages – sont supposées nous permettre de mieux connaître les revenus des ménages (privés), leur décomposition et leur répartition en fonction de diverses caractéristiques (taille du ménage, décile de revenus, âge de la personne de référence dans le ménage, etc.). Les deux encadrés suivants

proposent une courte présentation de ces deux sources.

#### Présentation de l'enquête SILC

« L'objet de cette statistique communautaire est d'établir un cadre commun pour la production systématique de statistiques communautaires sur le revenu et les conditions de vie (statistiques EU-SILC), englobant des données transversales et longitudinales comparables et actuelles sur le revenu ainsi que sur le nombre de pauvres et d'exclus et sur la composition de ce groupe social au niveau national et au niveau de l'Union européenne. »

Un concept est ici central :

« Risque de pauvreté : Le seuil de pauvreté équivaut à 60% de la médiane, à l'échelle individuelle, du revenu disponible. Les personnes qui ont un revenu inférieur à ce seuil sont considérées comme pauvres. »

Source : SPF Economie

#### Présentation de l'enquête sur le budget des ménages (EBM – HBS en anglais)

##### But et description sommaire

EU-HBS (European Union – Household Budget Survey) est une enquête sur les dépenses et les revenus des ménages belges ; c'est un outil important pour dresser, tant au niveau belge qu'europpéen, un état des habitudes de consommation de la population sur une année.

L'objectif de cette enquête est de mettre en place un cadre global pour la production d'informations statistiques 'communautaires' sur la consommation des ménages à partir de données transversales sur les postes budgétaires (montant, composition...) aux échelons national et européen.

L'enquête est aussi la source principale pour composer le panier de l'indice des prix à la consommation. Le choix des produits témoins et leur pondération dans le panier sont revus en profondeur tous les huit ans. La dernière révision a eu lieu en 2006 (année de base 2004) et la prochaine aura lieu en 2014 (année de base 2012). Le panier des produits est également réactualisé tous les deux ans à l'intérieur des groupes de produits (dernière actualisation en 2010).

##### Population étudiée

Les ménages privés qui résident en Belgique.

Source : SPF Economie

Sur base de ces sources j'ai comparé le revenu disponible global des ménages découlant de ces deux sources avec le revenu disponible total estimé par l'Institut des comptes nationaux (voir tableau en haut de la page suivante).

Faut-il vraiment s'étonner de constater que le revenu disponible global des ménages est sous-estimé par l'enquête SILC et l'EBM. On sait en effet, par exemple, que les revenus de la propriété sont systématiquement sous-déclarés par les ménages, pour diverses raisons (intérêts capitalisés, peur du fisc, volonté de discrétion sur le patrimoine).

##### **Note méthodologique :**

- ***Pour permettre une comparaison à périmètre semblable, j'ai ajouté les revenus imputés des logements des propriétaires aux données de SILC, cette enquête excluant ces revenus.***
- ***J'ai fait l'hypothèse que le revenu moyen des personnes habitant en ménage collectif était semblable à celui du restant de la population (elles représentent un peu plus 1% de la population totale).***

Une (grande) partie de la différence dans l'estimation du revenu disponible total des ménages – mais pas tout – s'explique par les écarts en ce qui concerne les revenus mobiliers. Par exemple, l'EBM évalue le montant des revenus mobiliers pour 2010 à 145 €/personne (!) alors que l'estimation macroéconomique pour les seuls intérêts et dividendes est de 2.016 €/personne !

**T5. Revenu disponible global des ménages – en milliards € et écarts en %  
Comparaison SILC – EBM – Comptabilité nationale**

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
CN	149,7	157,2	158,6	160,2	163,6	168,2	177,5	186,6	196,5	200,2	201,1
<b>Enquête sur le budget des ménages (EBM)</b>											
EBM	137,4	134,6	140,7	142,5	148,0	155,6	162,6	165,5	173,3	188,3	188,3
Ecart	-12,2	-22,6	-17,9	-17,7	-15,6	-12,6	-14,9	-21,1	-23,1	-11,9	-12,9
En % CN	-8,9%	-16,8%	-12,7%	-12,4%	-10,6%	-8,1%	-9,2%	-12,7%	-13,4%	-6,3%	-6,8%
<b>Enquête sur les revenus et les conditions de vie (SILC)</b>											
SILC	-	-	-	-	148,6	155,9	157,1	165,3	174,9	178,2	184,0
Ecart	-	-	-	-	-15,0	-12,3	-20,4	-21,3	-21,5	-22,0	-17,1
En % CN	-	-	-	-	-10,1%	-7,9%	-13,0%	-12,9%	-12,3%	-12,3%	-9,3%

**Sources : BfP, EBM, ICN et SILC – Calculs et estimations : IDD**

Ces deux sources de données (EBM et SILC) sont essentielles dans le débat socio-économique belge.

L'EBM sert, dans son volet dépenses, à la construction de l'indice des prix à la consommation, notamment aux pondérations des différents postes de consommation. Certes, il n'est pas impossible que les réponses en matière de dépenses soient de meilleure qualité que celles relatives aux revenus, mais peut-on vraiment y croire ? Rappelons aussi que les résultats de cette enquête sont aussi essentiels aux comptes nationaux pour l'estimation macroéconomique des dépenses des ménages et leur répartition entre postes.

Les données SILC servent au calcul du seuil de risque de pauvreté monétaire et donc du taux de pauvreté. Il n'y a pas de lien automatique entre le revenu moyen et le revenu médian. Un revenu moyen peut être sous-estimé par rapport à la réalité sans que cela n'influe sur le revenu médian (qui sert de référence pour le calcul du seuil de pauvreté). Il suffit, par exemple, que les revenus sous-déclarés soient concentrés dans les déciles les plus élevés. Mais on peut aussi faire l'hypothèse que les revenus mobiliers sous-déclarés dans l'enquête SILC soient en partie logés chez des personnes âgées dont la pension publique est modeste.

Il est donc difficile de savoir si les doutes nés à l'examen des résultats de l'enquête SILC permettent de mettre en cause le seuil de pauvreté et/ou le taux de pauvreté. Il faudrait pour cela entreprendre des travaux qui dépassent l'ambition de cette note. Cependant, mon intuition – raisonnée – est que le seuil de pauvreté est quelque peu plus élevé que celui estimé (1.000 €/mois pour une personne isolée) et le taux de pauvreté plus élevé pour les moins de 65 ans et plus faible pour les plus de 65 ans. Mon intuition est aussi que les coefficients de Gini et les écarts 1er quintile-5ème quintile calculés sur base des données SILC et utilisés notamment par l'OCDE sous-estiment les écarts dans la distribution des revenus<sup>5</sup>.

### **Quelques données socio-démographiques**

Les enquêtes SILC et EBM produisent aussi des résultats sur la taille moyenne des ménages et le nombre d'unités de consommation. Cette dernière information est particulièrement cruciale pour le calcul du seuil et du taux de pauvreté puisque la norme européenne définit le seuil de pauvreté comme étant égal à 60% du revenu médian par unité de consommation (UC).

<sup>5</sup> Voir, par exemple : Philippe DEFÉYT, « Distribution et redistribution des revenus : Indicateurs pour la Belgique – 1985-2008 », Institut pour un Développement Durable, février 2009 (voir : <http://www.iddweb.eu/docs/EBM.pdf>)



### L'unité de consommation (UC)

« L'unité de consommation modifiée est une échelle d'équivalence qui est appliquée pour adapter les dépenses de consommation en fonction de la taille et de la composition du ménage. Un coefficient de 1 est attribué au premier adulte, de 0,5 aux autres personnes de plus de 13 ans et de 0,3 aux enfants de 13 ans ou moins (échelle modifiée de l'OCDE). »

Source : [http://statbel.fgov.be/fr/statistiques/collecte\\_donnees/enquetes/budget\\_des\\_menages/](http://statbel.fgov.be/fr/statistiques/collecte_donnees/enquetes/budget_des_menages/)

Concrètement, une famille de, par exemple, deux adultes avec 2 jeunes enfants comptera 4 personnes mais seulement 2,1 unités de consommation. En moyenne, chaque membre du ménage vaut donc 0,525 UC. Autre exemple : les membres d'un couple vaudront chacun en moyenne 0,75 UC.

**NB** : Une véritable réflexion devrait s'engager sur le calcul des unités de consommation. Celles utilisées depuis des lustres me semblent dépassées !

Le tableau suivant compare – pour le nombre de ménages privés, la taille moyenne des ménages et le nombre d'unités de comparaison – les résultats de ces deux enquêtes avec les données calculées à partir des données du SPF Economie.

### T6. Nombre de ménages privés, taille moyenne des ménages et nombre d'unités de consommation – Comparaison de trois sources : Statistiques démographiques, EBM et SILC

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
	<b>Nombre de ménages privés (X 1.000)</b>										
Stat. Démographiques (1)	4.258	4.298	4.340	4.382	4.421	4.461	4.503	4.546	4.588	4.628	4.676
EBM (2)	4.157	4.155	4.079	4.140	4.158	4.246	4.369	4.440	4.493	4.539	4.593
Différence (2)-(1)	101	143	261	243	263	215	134	106	96	90	82
SILC (3)	-	-	-	4.364	4.447	4.488	4.517	4.563	4.613	4.654	4.705
Différence (3)-(1)	-	-	-	-18	26	28	15	17	25	26	29
	<b>Taille moyenne des ménages privés</b>										
Stat. Démographiques (1)	2,38	2,36	2,35	2,34	2,33	2,32	2,32	2,31	2,31	2,31	2,30
EBM (2)	2,44	2,45	2,50	2,48	2,48	2,44	2,39	2,36	2,36	2,35	2,34
Différence (2)-(1)	0,06	0,08	0,15	0,14	0,15	0,12	0,07	0,06	0,05	0,05	0,04
SILC (3)	-	-	-	2,35	2,32	2,31	2,31	2,30	2,29	2,29	2,29
Différence (3)-(1)	-	-	-	0,01	-0,01	-0,01	-0,01	-0,01	-0,01	-0,01	-0,01
	<b>Nombre d'UC pour 100 personnes</b>										
Stat. Démographiques (1)	67,69	67,83	67,96	68,10	68,22	68,32	68,40	68,47	68,49	68,50	68,51
EBM (2)	66,61	66,61	66,17	66,24	66,11	66,68	67,35	67,44	67,52	67,79	67,82
Différence (2)-(1)	-1,07	-1,22	-1,79	-1,86	-2,11	-1,64	-1,05	-1,02	-0,97	-0,71	-0,70
SILC (3)	-	-	-	66,52	67,45	68,73	67,92	68,08	68,22	67,62	68,39
Différence (3)-(1)	-	-	-	-1,58	-0,78	0,41	-0,47	-0,39	-0,27	-0,88	-0,13

**Source : EBM, SILC et SPF Economie – Calculs et estimations : IDD**

On commencera par rappeler que les données administratives, supposées être exhaustives et correctes, ne sont évidemment pas exemptes de "faiblesses". C'est ainsi que l'on sait qu'il y a des domiciliations fictives qui, toutes choses égales par ailleurs, augmentent le nombre de ménages privés, diminuent la taille moyenne de ceux-ci et augmentent le nombre d'unités de consommation. Il y a aussi des personnes âgées non domiciliées dans leur maison de repos ce qui conduit à augmenter le nombre de ménages privés (en tout cas quand la personne était isolée avant d'entrer en maison de repos), à surestimer la taille moyenne des ménages privés et à augmenter le nombre d'unités de consommation.



Ceci précisé, le tableau ci-dessus fait apparaître des écarts plus ou moins importants, plus ou moins significatifs ; par exemple

- un écart de 263.000 ménages privés en 2004 et de 0,15 pour la taille moyenne des ménages suivant que l'on prend en considération le registre de la population ou les données de l'EBM ; on peut se poser des questions donc sur la qualité des informations qui ont servi à revoir l'indice des prix en 2006 puisque la base de l'index est 2004 !
- un écart de 1,3% en 2009 dans l'estimation du nombre d'unités de consommation suivant que l'on prend en considération le registre de la population ou les données SILC ; impossible de dire si cet écart, relativement modeste au demeurant, est ou non significatif pour le calcul du seuil de pauvreté et/ou du taux de pauvreté ; mais s'ajoutant à d'autres faiblesses relevées ci-dessus, on commence à douter... ; rappelons à cet égard que les revenus augmentent lentement dans le bas de l'échelle des revenus et que donc une très faible variation du seuil de pauvreté peut très vite augmenter ou diminuer le nombre de pauvres de plusieurs dizaines (centaines?) de milliers de personnes.

## Et alors ?

Les statistiques belges ont connu leur heure de gloire grâce à l'action – souvent visionnaire – d'Adolphe Quetelet. Il est notamment à l'initiative du premier recensement décennal (1846) à portée scientifique. On trouvera une présentation de la vie et de l'action d'Adolphe Quetelet sur le site BESTOR<sup>6</sup>. Cette notice renvoie aussi à des reproductions de certains de ses livres, dont le très connu « Recherches statistiques sur le Royaume des Pays-Bas »<sup>7</sup>.

Mais je ne peux m'empêcher de penser que le système statistique belge n'est plus tout à fait à la hauteur des enjeux sociétaux, même si d'importants efforts ont été faits ces dernières années en termes de production et d'accessibilité statistiques. .

Certes, appréhender la complexité du réel est difficile, très difficile. Encore plus dans un monde aux changements rapides et où les classifications deviennent floues et les limites poreuses. Un(e) jeune peut passer chaque semaine quelques jours chez lui, d'autres chez son ami(e) et d'autres encore chez ses parents. Où réside-t-il ? Comment appréhender sa consommation ? Des personnes peuvent avoir jusqu'à 3 statuts socioprofessionnels différents. C'est ainsi que certains bénéficiaires du revenu d'intégration peuvent, simultanément, être répertoriés comme travailleurs (à temps partiel ou intérimaire) et chômeurs (allocation de garantie de revenus) ! On voit le risque d'utiliser plusieurs sources statistiques différentes si on ne tient pas compte des doubles ou triples comptages. Et peu d'utilisateurs de données socio-économiques ont vraiment conscience de ces recouvrements.

L'internationalisation de l'économie et de la société rendent également cette appréhension plus compliquée. Il faut par exemple bien appréhender les emplois occupés par des travailleurs étrangers qui demeurent en Belgique pour une période limitée pour mieux comprendre l'évolution du marché du travail "belge".

La "privatisation" de données stratégiques constitue un autre obstacle pour la recherche comme pour l'action publique. Améliorer le calcul de l'évolution des prix passe par l'accès à des données collectées par des entreprises privées dans les chaînes de distribution. Toutes les données relatives aux emballages pour boissons ne sont pas disponibles. Voilà deux enjeux, parmi beaucoup d'autres, où l'élaboration et le suivi d'une politique publique dépendent de sources

<sup>6</sup> Lien : [http://wiki.arts.kuleuven.be/wiki/index.php/Quetelet,\\_Lambert-Adolphe-Jacques\\_\(1796-1874\)](http://wiki.arts.kuleuven.be/wiki/index.php/Quetelet,_Lambert-Adolphe-Jacques_(1796-1874))

<sup>7</sup> Lien : [http://gdz.sub.uni-goettingen.de/dms/load/img/?PPN=PPN129323640\\_0005&DMDID=dmdlog17](http://gdz.sub.uni-goettingen.de/dms/load/img/?PPN=PPN129323640_0005&DMDID=dmdlog17)

privées plus ou moins accessibles. Mais la recherche aussi s'en trouve handicapée. Comment, par exemple, décrire l'évolution de la générosité des belges si les statistiques relatives aux collectes dans les églises, parmi d'autres, ne sont pas rendues publiques. Il n'y a pas que les groupes internationaux qui gardent par devers eux des informations stratégiques.

Toute donnée nécessite une analyse serrée. Sur l'échantillon et sa représentativité. Sur son exhaustivité s'il s'agit de données administratives. Sur son interprétation. Sur ses limites. Même quand il s'agit de données solides, cette interprétation est indispensable. Comme une prise de distance critique. Cela est de moins en moins le cas dans les médias. Mais aussi dans certaines utilisations politiques de données politiquement utiles. Comment peut-on continuer à dire, par exemple, qu'il y a 170.000 travailleurs dans les titres-services (c'était encore le cas de certaines déclarations en mai 2013) alors que ces 170.000 constituent une estimation de la Ministre de l'emploi fédérale datant de mars 2012 et qu'elle porte sur le nombre total de personnes qui ont presté au moins un jour de travail dans le secteur. La réalité est autre : il y avait fin 2012 environ 120.000 personnes à l'emploi dans ce secteur.<sup>8</sup> Un autre exemple d'absence totale d'analyse sur la signification et l'interprétation des données se trouve dans une publication du SPF Economie sur l'évolution du coefficient de Gini calculé sur base des statistiques fiscales. Une partie de l'augmentation de ce coefficient est tout simplement expliquée par un changement dans la prise en compte des déclarations fiscales à revenu nul ou à faible revenu.<sup>9</sup>

Le refus de plus en plus fréquent de répondre aux enquêtes ne facilite pas les choses. « Depuis quand est-on obligé de répondre à des questions de vie privée ? » s'indignait une personne contactée par le SPF Economie pour répondre à l'Enquête sur les forces de travail.

Mais il faut aussi une réelle volonté politique pour améliorer la qualité et la quantité de données utiles à la société. Elle manque un peu aujourd'hui. Comment expliquer, par exemple, que le SPF Economie puisse publier la liste des prénoms les plus populaires (d'un intérêt limité à mes yeux) mais ne sache rien dire, ou presque, sur, par exemple, les expulsions domiciliataires ou sur le taux d'analphabétisme en Belgique (depuis des années on dit qu'il est de 10% sans aucune base scientifique un petit peu rigoureuse).

Les producteurs de statistiques publics doivent aussi garantir – ce qui est moins le cas qu'auparavant me semble-t-il – une grande transparence :

- sur les marges d'incertitude ; c'est ainsi par exemple que l'INS, quand cet organisme aujourd'hui fondu dans le SPF Economie était encore responsable des comptes nationaux, publiait les écarts d'estimation entre les trois approches des comptes nationaux (dépenses, revenus, valeur ajoutée) ; ce n'est plus le cas aujourd'hui des comptes nationaux publiés par l'ICN
- sur les méthodes de production (sources des données de base, corrections apportées, mise en perspective critique de sources différentes, etc.) des données ; difficile, par exemple, d'avoir une vision très claire de la manière dont les comptables nationaux estiment la masse des loyers<sup>10</sup>
- sur les corrections apportées aux données brutes
- et, pour les producteurs de "prévisions", sur les corrections apportées aux résultats bruts

<sup>8</sup> Le lecteur intéressé trouvera plus d'explications dans la note « L'emploi en titres-services : mise en perspective et mises au point », Philippe DEFEYT, mars 2013 à l'adresse : <http://www.iddweb.eu/docs/titreserv.pdf>

<sup>9</sup> Le lecteur intéressé trouvera plus d'explications dans la note « Evolution de l'inégalité de revenu avant et après impôts en Belgique : un commentaire critique des résultats publiés par l'INS », Philippe DEFEYT, mai 2010 à l'adresse : <http://www.iddweb.eu/docs/lneqRev.pdf>

<sup>10</sup> Voir : Philippe DEFEYT, « L'indice des prix et la comptabilité nationale sous-estiment la hausse des loyers », Institut pour un Développement Durable, août 2011 (<http://www.iddweb.eu/docs/EvoLoyers.pdf>)

des modèles.

D'une manière générale les intervalles de confiance devraient être plus systématiquement indiqués pour toutes les données et pas seulement globalement. Ils devraient en tout être facilement disponibles.

Les producteurs de statistiques devraient améliorer encore l'accessibilité des données existantes, même si des progrès ont été observés ces dernières années.

Accessibilité technique d'abord (possibilité ou pas de construire ses propres tableaux, techniques de téléchargement, pourcentage des données accessibles en ligne, rapidité de la mise à disposition, publication ou pas d'estimations provisoires, possibilité ou pas de remonter aux données de base...). Les écarts d'accessibilité sont ici énormes, par exemple, entre l'ONSS et l'ONSS-APL pour ce qui concerne le nombre de données sur l'emploi et leur accessibilité par internet (l'ONSS constitue ici un modèle pour d'autres producteurs de statistiques<sup>11</sup>) ou entre le SPF Economie et Eurostat pour le nombre d'années publiées (4 années seulement de données SILC sur la pauvreté sur le site du SPF Economie, toutes les années sur le site d'Eurostat) ou la facilité pour élaborer ses propres tableaux. Les tableaux relatifs aux statistiques de population élaborés par le SPF Economie constituent un contre-exemple : titres peu clairs, dispersion à plusieurs endroits du site, années manquantes...<sup>12</sup>

Accessibilité également en matière de présentation. La présentation des comptes nationaux a été complètement bouleversée par les lignes de conduites européennes, ce qu'on appelle communément les comptes SEC. Peut-être est-ce bien utile pour l'analyse macroéconomique poussée et pour les comparaisons au sein de l'Union. Mais l'utilisation de ces comptes pour l'analyse politique, syndicale, militante en a été compliquée, inutilement ! D'une manière générale les données transformées (pourcentage, taux de croissance...) peuvent intéresser le grand public tandis que les données de base intéresseront plus les chercheurs.

Plus que jamais, nos sociétés ont besoin de points de repère, de balises, pour se comprendre et orienter leur action, collective et individuelle.

Cela nécessite en Belgique en tout cas un sursaut, pour améliorer nos statistiques et leur utilisation.

La première démarche doit être ici d'améliorer la culture statistique, dans tous les secteurs (administration, médias, partis politiques, enseignement...). Montrer toute l'utilité de disposer d'informations quantitatives mais aussi développer le sens critique dans la lecture et l'utilisation de celles-ci. Apprendre à mettre en perspective des sources différentes peut dire beaucoup sur des évolutions socio-économiques et sociétales, pour autant que le contenu, le champ de couverture et les limites de chaque source de données soient bien connus. Indiquer et apprendre à lire les intervalles de confiance, introduire de la rigueur dans l'utilisation des concepts et catégories (illustration : fonctionnaire et travailleur dans le secteur public ce n'est pas la même chose pas plus que nombre de travailleurs et nombre d'emplois), faire des comparaisons sur plusieurs années, etc., sont des démarches qui peuvent aider à mieux comprendre et mieux utiliser les données statistiques. Enfin, on peut aussi penser qu'une plus grande culture

---

<sup>11</sup> Notamment dans les risques que cet organisme accepte de prendre en publiant, certes avec de nombreux et bienvenus caveats, des estimations avancées en matière d'emploi.

<sup>12</sup> Voir : <http://statbel.fgov.be/fr/statistiques/chiffres/population/> et <http://statbel.fgov.be/fr/modules/publications/?f=cHVibGJlYXRpb25UeXBIPS9TdGF0aXN0aXF1ZXMgLSBQb3B1bGF0aW9u&selectedCat=Statistiques+-+Population>

statistique, plus largement partagée, permettrait de faire plus et mieux émerger les données existantes dans de nombreux lieux (par exemple : le suivi des personnes (re)prises à l'emploi dans un service d'insertion socioprofessionnelle de CPAS ou le nombre et contenu des jugements en matière d'expulsions dans un tribunal de paix), potentiellement utiles pour l'action des acteurs concernés et, plus largement, pour la recherche et la réflexion politique, mais qui restent en friche. Autre exemple : les statistiques de fréquentation des canaux et cours d'eau wallons par les péniches pourraient donner lieu à un indicateur conjoncturel intéressant dans la mesure où ce pourrait être un indicateur avant-coureur d'un retournement de conjoncture.

#### De la relativité des résultats statistiques...

Un des principaux obstacles à la bonne compréhension et à une utilisation pertinente des statistiques réside bien sûr dans l'absence d'une claire conscience du caractère conventionnel des statistiques, c'est-à-dire que les résultats (ex : le nombre d'emplois privés ou la consommation totale des ménages) dépendent de la convention adoptée.

L'actualité récente nous fournit deux beaux exemples :

1. Une modification dans la comptabilisation des immatriculations en Belgique – qui vise à éliminer plus rapidement les véhicules qui sont désimmatriculés – aboutit à des changements dans les immatriculations absolues, les évolutions à un an d'écart et les parts de marché de chaque marque.
2. Le PIB américain a "augmenté" en juillet 2013 d'environ 2% en valeur absolue à la faveur d'une nouvelle manière de comptabiliser les dépenses de recherche et développement et devrait encore augmenter d'un peu plus de 1% plus tard parce que le Bureau d'analyse économique américain comptabilisera désormais la production et l'achat des biens culturels ou de divertissements comme des investissements.

Si on veut améliorer l'utilité de certaines récoltes de statistiques (je pense ici en particulier à l'EFT et à l'enquête SILC, mais pas uniquement), il faudrait combiner beaucoup plus qu'aujourd'hui les données administratives (Banque carrefour, déclarations fiscales...) et les données résultant d'enquêtes. Une telle approche permettrait

- d'alléger et/ou de réorienter certaines enquêtes ; pourquoi, par exemple – dans l'EFT – demander à quelqu'un son statut socioprofessionnel il y a un an alors que l'information existe à la Banque carrefour (en principe en tout cas...) ; par contre lui demander pourquoi son statut ou son emploi a changé peut être très intéressant pour comprendre les réalités et stratégies individuelles sur le marché du travail
- de vérifier les informations données par le répondant et ainsi essayer de repérer et surtout essayer d'expliquer les écarts chiffrés ou les différences de représentation qui peuvent exister entre les catégories administratives et/ou d'enquêtes et la manière dont les personnes appréhendent leurs réalités
- et par là même de comprendre mieux les représentations individuelles et leur plus ou moins grand écart avec les catégories statistiques : comment par exemple se voit un(e) jeune qui n'est ni en emploi, ni en stage ni en formation-enseignement ?

Pour faire cela il faut lever certains obstacles mis à la connaissance des réalités socio-économiques par une vision par trop restrictive de la protection de la vie privée.

Il faut aussi, beaucoup plus qu'aujourd'hui, utiliser et comparer de manière critique plusieurs sources statistiques pour mieux comprendre ce qui se passe. Deux exemples :

- seule la comparaison régulière de plusieurs sources en matière d'emploi – organismes de sécurité sociale, Banque Carrefour et EFT – permet(trait) de bien/mieux appréhender les réalités macroéconomiques et sociologiques du marché du travail belge
- une comparaison du stock de logements recensés par le cadastre, des mises en chantier de logements et de l'évolution du nombre de ménages est nécessaire pour mieux comprendre

la dynamique de la construction résidentielle.

Il faut aussi un débat public sur les priorités des efforts statistiques. Comme souvent la demande excède l'offre possible. Mais il y a certainement des priorités. En tout cas on devrait pouvoir en discuter. Illustration : tout le monde prétend être préoccupé par le devenir des jeunes de moins de 25 ans et, en particulier, par le phénomène des NEETs. Mais, en même temps, la Communauté Wallonie-Bruxelles ne s'est jamais donnée les moyens pour, comme la Communauté flamande, disposer très rapidement des statistiques de fréquentation de l'enseignement supérieur, ce qui rend plus compliquée l'analyse du marché du travail des moins de 25 ans. Alors qu'en Flandre on dispose de telles informations en octobre-novembre de chaque année pour l'année en cours, il faut en Communauté Wallonie-Bruxelles attendre plusieurs années ! Je pense aussi, par exemple, qu'une enquête à intervalles réguliers (tous les 5 ans par exemple) sur l'analphabétisme permettrait de mieux (ré)orienter l'action associative et publique dans la lutte pour l'alphabétisation.

Parmi les questions sur les priorités : pourquoi abandonne-t-on la production de certaines statistiques pourtant bien utiles ? Je pense par exemple au chiffre d'affaires TVA par secteur, élément essentiel pour l'analyse conjoncturelle, qui n'a pas été publié pendant deux ans ou aux travaux menés en son temps par l'Institut économique agricole sur la récolte simultanée de statistiques en quantité et en valeur de la consommation alimentaire<sup>13</sup>.

Il faudrait aussi que l'apparition de nouveaux phénomènes socio-économiques et sociétaux (je pense par exemple aux groupements d'achats, aux donneries, au rôle croissant des grands-parents dans la prise en charge et l'éducation de leurs petits-enfants, à la cohabitation de jeunes adultes, etc.) puissent plus rapidement qu'aujourd'hui faire l'objet d'études préliminaires permettant, notamment, de débroussailler les évolutions socio-économiques et sociologiques et, le cas échéant et si nécessaire, de jeter les bases d'une production statistique plus régulière. Je pense à cet égard qu'on pourrait et devrait mieux mobiliser ici la masse des étudiants en fin de cycle qui doivent faire des mémoires ou autres travaux de fin d'étude pour contribuer à cet effort. Ne pas reproduire systématiquement tous les ans certaines enquêtes (a-t-on vraiment besoin de calculer le taux de pauvreté chaque année?) permettrait de libérer des moyens pour de telles "avancées". Notre système de production statistique devrait pouvoir rendre compte de l'extrême mobilité des évolutions socio-économiques et sociologiques.

D'une manière générale il faut pouvoir se donner les moyens de mettre en place des récoltes de statistiques et des analyses ponctuelles permettant de "cadrer" un peu mieux certains débats publics. Un bel exemple de ce genre de démarche et qui prouve sa pertinence est la récente étude de l'IBSR sur les causes des accidents des motocyclistes<sup>14</sup> dont les résultats tendent à relativiser très fort l'impact des dégradations du réseau routier souvent mises en évidence dans les manifestations des motards.

La question démocratique est ici essentielle : qui décide et où de l'évolution de la production statistique ? Une fois de plus l'Europe, au travers de son office statistique Eurostat, apparaît comme leader. Mais comment et où se décident ses priorités sinon dans des comités plus ou moins transparents et plus ou moins ouverts en matière de diversité culturelle et scientifique ? Au

---

<sup>13</sup> « Jusqu'en 1991, l'Institut économique agricole (IEA) a recueilli des données sur les achats ménagers, en quantité et en valeur, d'un panel de consommateurs regroupant environ 3.000 ménages représentatifs de la population belge en vue de prévoir la consommation des ménages, notamment en essayant d'établir les relations entre les quantités achetées, leur valeur, les habitudes d'achat et les facteurs revenus, saisonniers, régionaux et démographiques. » (<http://popups.ulg.ac.be/Base/document.php?id=1419>)

<sup>14</sup> Voir : <http://ibsr.be/frontend/files/userfiles/files/motac-rapport-fr-2013.pdf>

niveau fédéral le Conseil supérieur de Statistique<sup>15</sup> est une institution (trop) discrète dont on mesure mal sa capacité d'influence sur la production statistique et sa capacité d'innover en matière de statistiques. Les organismes régionaux qui peu ou prou donnent leur avis sur la production statistique sont plus discrets encore.

Les producteurs publics de statistiques doivent aussi améliorer l'accessibilité de leurs données et faciliter la vie de leurs utilisateurs. Faire sortir les données là où elles existent et les rendre facilement et rapidement accessibles constitue un enjeu important, y compris pour les entreprises privées. Dans cette perspective il est peut-être temps de lancer en Belgique et dans ses entités fédérées un débat solide sur ce que les anglo-saxons appellent l'*Open data* (voir l'annexe).

Enfin, les pouvoirs publics doivent aussi forcer les producteurs privés de statistiques à participer plus activement à cet effort pour une société plus transparente. Trop de données sont confinées dans les entreprises ; je pense par exemple à l'évolution détaillée des CA dans le secteur de la consommation courante des ménages – d'une importance cruciale pour caler au mieux les modifications à apporter à l'indice des prix à la consommation – ou aux données sur les différents types d'emballages de boissons ou autres produits de consommation.

---

<sup>15</sup> <http://statbel.fgov.be/fr/statistiques/organisation/css/>

### Open data : A new goldmine

*Making official data public could spur lots of innovation.*

AFTER a Soviet missile shot down a South Korean airliner that strayed into Russian airspace in 1983, President Ronald Reagan made America's military satellite-navigation system, GPS, available to the world. Entrepreneurs pounced. Car-navigation, precision farming and 3m American jobs now depend on GPS. Official weather data are also public and avidly used by everyone from insurers to ice-cream sellers.

But this is not enough. On May 9th Barack Obama ordered that all data created or collected by America's federal government must be made available free to the public, unless this would violate privacy, confidentiality or security. "Open and machine-readable", the president said, is "the new default for government information."

This is a big bang for big data, and will spur a frenzy of activity. Pollution numbers will affect property prices. Restaurant reviews will mention official sanitation ratings. Data from tollbooths could be used to determine prices for nearby billboards. Combining data from multiple sources will yield fresh insights. For example, correlating school data with transport information and tax returns may show that academic performance depends less on income than the amount of time parents spend with their brats.

Over the next few months federal agencies must make an inventory of their data and prioritise their release. They must also take steps not to release information that, though innocuous on its own, could be joined with other data to undermine privacy—a difficult hurdle.

Many countries have moved in the same direction. In Europe the information held by governments could be used to generate an estimated €140 billion (\$180 billion) a year. Only Britain has gone as far as America in making data available, however. For example, it requires the cost of all government transactions with citizens to be made public. Not all public bodies are keen on transparency. The Royal Mail refuses to publish its database of postal addresses because it makes money licensing it to businesses. On May 15th an independent review decried such practices, arguing that public-sector data belong to the public.

Rufus Pollock of the Open Knowledge Foundation, a think-tank, says most firms will eventually use at least some public-sector information in their business. But no one has a clue what breakthroughs open data will allow, just as Reagan never guessed that future drivers would obey robot voices telling them to turn left.

Source : The Economist, may 18<sup>th</sup> 2013