

**Rapport de consultation sur le rapport intermédiaire (fin de phase
1) de Schwartz and Co relatif à l'étude de l'efficacité des
gestionnaires de réseau de distribution d'électricité et de gaz
naturel en Région wallonne pour la période réglementaire 2024-2028
(lot 2)**

12 novembre 2020

Version 1.01

Préparé par : Schwartz and Co, la CWaPE



SOMMAIRE

1. CONTEXTE ET OBJECTIF	4
2. REPONSES AUX REMARQUES DE RESA, AIEG, AIESH ET REW	5
2.1. Processus de consultation transparent et concertée	5
2.1.1. Remarques de RESA, AIEG, AIESH et REW	5
2.1.2. Réponse aux remarques de RESA, AIEG, AIESH et REW.....	5
2.1.3. Précisions apportées dans le rapport final.....	5
2.1. Méthode Advanced COLS	6
2.1.1. Remarques de RESA, AIEG, AIESH et REW	6
2.1.2. Réponse aux remarques de RESA, AIEG, AIESH et REW.....	6
2.1.3. Précisions apportées dans le rapport final.....	9
2.2. Constitution de l'échantillon	9
2.2.1. Remarques de RESA, AIEG, AIESH et REW	9
2.2.2. Réponse aux remarques de RESA, AIEG, AIESH et REW.....	9
2.2.3. Précisions apportées dans le rapport final.....	10
2.3. Analyse académique du professeur Axel Gautier.....	10
2.3.1. Méthode Advanced COLS.....	10
2.3.2. Taille de l'échantillon.....	12
2.4. Analyse académique du professeur Damien Ernst	14
2.4.1. Ensemble de données.....	14
2.4.2. Variables explicatives.....	15
3. RAPPORT DE CONSULTATION SUR LES REMARQUES D'ORES (VIA UN RAPPORT DE L'ENTREPRISE OXERA)	16
3.1. Processus d'évaluation des coûts	16
3.1.1. Remarques d'Oxera	16
3.1.2. Réponse aux remarques d'Oxera	17
3.1.3. Précisions apportées dans le rapport final.....	17
3.2. Méthodologie d'évaluation des coûts.....	18
3.2.1. « Advanced COLS » comme approche de développement de modèle	18
3.2.2. Rejet de la méthode SFA.....	23
3.2.3. Rejet de la méthode DEA	25
3.2.4. Estimation du benchmark et des niveaux d'efficacité appropriés.....	28
4. RAPPORT DE CONSULTATION AUX REMARQUES FORMULEES PAR ORES ET RESA, AIEG, AIESH ET REW TRAITANT DE POINTS IDENTIQUES	31
4.1. Comparabilité des coûts dans le cadre d'un benchmark international.....	31
4.1.1. Remarques d'Oxera	31
4.1.2. Remarques de RESA, AIEG, AIESH et REW	32
4.1.3. Réponse aux remarques de RESA, AIEG, AIESH et REW et d'Oxera.....	34
4.1.1. Précisions apportées dans le rapport final.....	34



4.2. Comparabilité des CAPEX	35
4.2.1. Remarques d'Oxera	35
4.2.2. Remarques de RESA, AIEG, AIESH et REW	36
4.2.3. Réponse aux remarques de RESA, AIEG, AIESH et REW et d'Oxera	37
4.2.4. Précisions apportées dans le rapport final	39
4.3. Approche hybride	39
4.3.1. Remarques d'Oxera	39
4.3.1. Remarques de RESA, AIEG, AIESH et REW	40
4.3.2. Réponse aux remarques de RESA, AIEG, AIESH et REW et d'Oxera	40
4.3.3. Précisions apportées dans le rapport final	42
4.4. Choix des variables explicatives des coûts	42
4.4.1. Remarques d'Oxera	42
4.4.2. Remarques de RESA, AIEG, AIESH et REW	44
4.4.3. Réponse aux commentaires de RESA, AIEG, AIESH et REW et d'Oxera	45
4.4.4. Précisions apportées dans le rapport final	47
4.5. Utilisation de panels et vérification de la robustesse du modèle	47
4.5.1. Remarques d'Oxera	47
4.5.2. Remarques de RESA, AIEG, AIESH et REW	49
4.5.3. Réponse aux remarques d'Oxera et de RESA, AIEG, AIESH et REW	49
4.5.4. Précisions apportées dans le rapport final	50



1. Contexte et objectif

Ce document répond aux remarques formulées par les 5 GRD wallons et leurs consultants sur le rapport intermédiaire du cabinet Schwartz and Co relatif à l'étude de l'efficacité des gestionnaires de réseau de distribution d'électricité et de gaz naturel en Région wallonne pour la période réglementaire 2024-2028, qui propose une méthode de mesure de l'efficacité des GRD wallons sur la base d'une analyse du retour d'expérience international en la matière et de l'expérience du cabinet.

Faisant suite à cet échange contradictoire formel et considérant que d'une part, les remarques ont toutes fait l'objet d'une réponse circonstanciée et détaillée, et d'autre part, qu'il n'y a pas de points bloquants, le cabinet Schwartz and Co est à même d'intégrer les remarques des 5 GRD wallons et de leurs consultants et de finaliser le rapport intermédiaire susmentionné.

La CWaPE pourra de ce fait prendre sa décision finale quant au modèle de mesure de l'efficacité qu'elle souhaite appliquer et procéder au lancement de l'option du lot 2, à savoir le calcul d'un facteur d'efficacité individuel pour chaque gestionnaire de réseau de distribution d'électricité et de gaz naturel en Région wallonne.

Les réponses ont été rédigées par :

- M. Régis Bourbonnais, Senior Advisor du cabinet Schwartz and Co
Docteur en économie mathématique et économétries
Maitre de conférences à l'Université de Paris-Dauphine
<https://regisbourbonnais.dauphine.fr/>
Auteurs des livres :
 - « Économétrie », Dunod, Collection Module économique, 10ème éd., 2018 (11^{ème} édition prévue en mars 2021).
 - « Analyse des séries temporelles : applications à l'économie et à la gestion », Dunod, 4ème éd., nouvelle édition Mars 2016.
 - « The Econometrics of energy systems », Palgrave, 2007 en collaboration avec Jan Horst Keppler et Jacques Girod sous la direction de Jean-Marie Chevalier.
- M. Hervé Schwartz, Président fondateur du cabinet Schwartz and Co
Ingénieur diplômé de l'Ecole Centrale Paris
- M. Hugo Jennequin, Consultant du cabinet Schwartz and Co
Ingénieur diplômé de l'Institut Polytechnique de Grenoble
- Olivier De Breuck, Conseiller à la CWaPE
Direction socio-économique et tarifaire



2. Réponses aux remarques de RESA, AIEG, AIESH et REW

2.1. Processus de consultation transparent et concertée

2.1.1. Remarques de RESA, AIEG, AIESH et REW

Le processus proposé à ce stade pour mesurer l'efficacité des GRDs wallons manque de transparence, ce qui inévitablement amènera des questions sur la validité du résultat. En particulier, tant la constitution et la standardisation de la base de coût que la constitution de l'échantillon doivent être beaucoup mieux documentés. Il n'y a quasi pas d'informations sur ces deux aspects essentiels qui doivent garantir que la comparaison qui sera réalisée ne sera pas biaisée. Ces questions sont d'autant plus importantes que l'échantillon qui sera constitué sera de taille très réduite. Le choix des points de comparaison est à ce titre essentiel et doit être mieux détaillé.

Comme mentionné dans le cahier des charges, la mission du lot 2 comprend aussi une « implication régulière des GRD actifs en Région wallonne » et à notre sens, celle-ci ne peut pas se limiter à une concertation unique par rapport aux résultats intermédiaires. Si la CWaPE décide d'exercer l'option du lot 2, il conviendrait que le rapport intermédiaire indique un calendrier précis des prochaines étapes de consultation et de concertation.

2.1.2. Réponse aux remarques de RESA, AIEG, AIESH et REW

Les commentaires attendus portent sur le rapport intermédiaire. La constitution et la standardisation de la base de coût font l'objet de l'option du lot 2 et n'ont pas encore démarré, de même que la constitution de l'échantillon : il est par conséquent exact de considérer que ces éléments ne puissent être entièrement clairs. Par ailleurs, ils ne font pas partie du scope du rapport intermédiaire soumis à la présente consultation.

Le processus ne peut être caractérisé par un manque de transparence à partir du moment où l'option du lot 2 n'a pas encore débuté.

En ce qui concerne l'implication des GRD dans l'option du lot 2, ces derniers seront impliqués dans le cadre de besoin d'informations complémentaires éventuels. Dans une perspective d'efficacité et de bonne gestion, l'option du lot 2 sera réalisée en une fois, sans étapes intermédiaires, et le rapport avec les résultats finaux sera communiqué aux GRD.

2.1.3. Précisions apportées dans le rapport final

Ces remarques n'appellent pas de compléments dans le rapport final.



2.2. Méthode Advanced COLS

2.2.1. Remarques de RESA, AIEG, AIESH et REW

Schwartz and Co exclut la méthode SFA pourtant souvent appliquée par les régulateurs et qui présente l'avantage de distinguer dans le calcul de l'inefficience la part d'inefficience de la part du bruit statistique.

Nous ignorons précisément dans quels pays européens ce modèle « advanced COLS » a été appliqué, en dehors de la France. Le tableau 3 du rapport intermédiaire (p. 12°) devrait être complété pour couvrir l'ensemble des pays européens.

Nous estimons que le modèle sélectionné ne peut pas l'être seulement sur base de tests statistiques. Le modèle doit non seulement être significatif et robuste d'un point de vue statistique, mais il doit également être significatif et compréhensible au niveau économique. Il doit représenter la technologie et l'environnement des GRD et être compris par ceux-ci. Il nous semble hasardeux de choisir un modèle exclusivement sur une base statistique sans s'interroger sur le fondement économique de celui-ci.

De façon fondamentale, nous considérons que le choix du modèle de benchmarking est indissociable du choix de l'échantillon. On peut d'ailleurs se demander si le choix de méthode de benchmarking proposé par le consultant n'est pas une conséquence directe du choix ou plutôt de la contrainte d'un échantillon limité.

2.2.2. Réponse aux remarques de RESA, AIEG, AIESH et REW

Méthode SFA

Nous avons analysé les méthodes utilisées par les régulateurs d'énergie dans un échantillon de 10 pays européens et la méthode SFA n'est utilisée quand dans 2 cas : le Danemark et l'Allemagne à travers une méthode hybride (voir détails au paragraphe 4.3.2). La méthode SFA n'est donc pas une méthode largement déployée par les régulateurs en Europe, et les autres méthodes paramétriques (COLS, MOLS, advanced COLS) sont beaucoup plus employées. Par ailleurs, comme présenté dans le rapport et comme étayé en réponse aux remarques d'ORES (voir détails au paragraphe 3.2.2.2), la méthode SFA nécessite un échantillon de données de taille importante (de l'ordre d'une centaine d'observations) pour pouvoir produire des résultats satisfaisants. RESA, AIEG, AIESH et REW semblent d'ailleurs d'accord avec cette affirmation (« *Si les méthodes basées sur un OLS permettent d'avoir une estimation de l'efficacité individuelle sur base d'un échantillon très réduit, ce n'est pas le cas avec la méthode SFA qui nécessite plus d'observations* », page 7 du retour de RESA, AIEG, AIESH et REW). Comme indiqué dans notre rapport intermédiaire, un benchmark international étant indispensable pour mesurer l'efficacité des GRD wallons, étant donné le nombre trop faible de GRD en Wallonie et dans les autres régions de Belgique pour assurer la robustesse statistique de la modélisation, la mise en œuvre de la méthode SFA n'est tout simplement pas faisable, puisque



cela nécessiterait la collecte et le retraitement des données de plus de 90 GRD étrangers pour assurer que la comparabilité des bases de coûts. Ceci n'est cependant pas l'unique problème de la méthode SFA. Comme indiqué dans notre rapport intermédiaire : « même si la méthode SFA distingue dans le calcul de l'inefficience, une part d'inefficience et une part de bruit statistique, il est nécessaire de présupposer la distribution statistique modélisant la part d'inefficience technique et la part de bruit et de disposer d'un large échantillon de GRD, ce qui est un choix arbitraire ». Le bénéfice supposé de la méthode SFA en matière de mesure d'efficacité par rapport aux méthodes advanced COLS et MOLS reste donc à prouver.

Méthodologie de sélection des variables dans la méthode Advanced COLS

Nous saluons le fait que RESA, AIEG, AIESH et REW partagent le fait que la vérification de la significativité et de la robustesse statistique du modèle soit indispensable, contrairement à ORES par l'intermédiaire de son consultant OXERA, qui dénigre l'importance des tests statistiques d'un modèle économétrique, à l'opposé des pratiques de base en matière d'économétrie (voir notre réponse à OXERA au paragraphe 3.2.1.2). L'importance de la vérification de la significativité et de la robustesse statistique du modèle exclut ainsi de facto la méthode DEA, comme nous l'avons décrit dans notre rapport intermédiaire.

Nous partageons l'affirmation de RESA, AIEG, AIESH et REW selon laquelle « le modèle doit non seulement être significatif et robuste d'un point de vue statistique, mais il doit également être significatif et compréhensible au niveau économique ». Nous avons en particulier indiqué dans notre rapport intermédiaire :

- dans le dernier paragraphe de la section 3.2.3, au sujet des modèles économétriques conçus à partir des approches « stagewise regression » et « all possible models » : « ces méthodes de détermination peuvent aboutir à des variables explicatives qui sont les plus pertinentes statistiquement, mais qui sont difficiles à appréhender d'un point de vue métier pour les GRD. Dans ce cas, le choix des variables explicatives peut être ajusté en pleine connaissance d'une perte de qualité statistique du modèle, en choisissant un modèle ayant une qualité statistique légèrement moindre que le modèle optimal, mais plus facilement appréhendable d'un point de vue métier ».
- dans la section 5.2 : « le jeu de variable exact sera fixé au moment de l'implémentation de la méthode, sur la base de notre expérience et du retour d'expérience international. Les variables retenues doivent être suffisamment corrélées avec la base de coût et faire également du sens d'un point de vue métier. »

Dans la méthode advanced COLS, le choix des variables explicatives n'est pas mécanique. Pour rappel, la méthode advanced COLS sélectionne les variables explicatives et la fonction de coûts la plus pertinente sur la base de deux méthodes :

- la méthode « stagewise regression » : la méthode de régression par étape (« stagewise regression ») est un processus de régression multilinéaire visant à sélectionner le meilleur



ensemble de variables explicatives, minimisant l'intercorrélation entre les variables explicatives en étudiant les résidus de régression ;

- la méthode « all possible models » : la méthode « all possible models » consiste à calculer et tester tous les modèles possibles à partir du jeu de variables explicatives disponibles en linéaire, en logarithmique et en ratio.

Le meilleur modèle économétrique ainsi identifié (par les critères d'Akaike et Schwarz) est ensuite validé par une méthodologie à deux étapes :

- **Etape 1 - validité statistique** : une série de tests statistiques est réalisée dans le but de valider statistiquement la pertinence et la robustesse du modèle. Les principaux tests utilisés sont décrits dans notre rapport intermédiaire et des compléments apportés aux paragraphes 3.2.1.2 et 4.5.3 du présent rapport.
- **Etape 2 – validité économique et opérationnelle** : la validité économique et opérationnelle du modèle économétrique est vérifiée en contrôlant que les variables explicatives du modèle sont cohérentes avec les pratiques métiers des GRD et que les signes attendus des coefficients sont conformes à l'intuition économique.

Si le modèle (variables explicatives et fonction de coûts) n'est pas ainsi validé, il n'est pas retenu et le meilleur modèle économétrique suivant est testé. L'itération s'arrête lorsque le modèle a satisfait aux deux étapes de validation.

Chronologie entre le choix du « modèle de benchmarking » et le choix de l'échantillon

Nous comprenons que par « modèle de benchmarking », RESA, AIEG, AIESH et REW font référence à la méthode de mesure de l'efficacité qui est en cours de définition dans le cadre de la présente étude, et pas au modèle économétrique qui est calculé sur la base d'un échantillon de GRD dans le cadre de la mise en œuvre de la méthode.

Ceci étant précisé, nous réfutons l'affirmation que le choix de la méthode de mesure de l'efficacité est indissociable du choix de l'échantillon. En effet, quelle que soit la méthode de mesure, les données utilisées sont les mêmes : il s'agit de la base de coût de chacun des GRD qui sont comparés et des variables explicatives des coûts, ces 2 types de données étant par définition indépendants de la méthode de mesure (des variables explicatives qui différencieraient selon les méthodes traduiraient un problème de fond de l'une ou l'autre des méthodes). La taille de l'échantillon qui peut être utilisé a un impact sur le choix de la méthode de mesure qui peut être utilisé, mais le choix de l'échantillon n'en a pas (s'il en avait cela traduirait un problème de fond de l'une ou l'autre des méthodes). La méthode doit être choisie pour ses qualités propres, et en particulier pour sa capacité à mesurer de manière effective et contrôlable en termes de qualité l'efficacité des GRD, et non pour les résultats plus ou moins favorables qu'elle donnerait pour tel ou tel GRD.

Par ailleurs, nous rappelons que les GRD ont été consultés par la CWaPE pendant le processus de rédaction du cahier des charges de l'appel d'offres de la mission remportée par Schwartz and Co.



Le cahier des charges décompose la mission en deux parties successives : la première partie, en cours, est consacrée au choix d'un modèle de mesure d'efficacité, tandis que la seconde, optionnelle, vise à mettre en œuvre ce calcul d'efficacité sur la base de données collectées au cours de cette seconde partie. Les GRD n'ont pas remis en cause cette méthodologie lors de leurs commentaires sur le cahier des charges.

2.2.3. Précisions apportées dans le rapport final

Le rapport final sera complété sur la base des réponses du paragraphe précédent.

2.3. Constitution de l'échantillon

2.3.1. Remarques de RESA, AIEG, AIESH et REW

L'étude ne détaille pas la manière dont l'échantillon sera constitué. Or, il s'agit d'un élément fondamental. La question du choix des GRDs internationaux qui seront intégrés dans l'étude de benchmarking reste vague.

Pourquoi avoir présélectionné l'Autriche, l'Allemagne, le Royaume Uni, les Pays-Bas et la France ? En fonction des pays et des GRD sélectionnés, la frontière d'efficacité sera très différente.

On peut préjuger que les GRD anglais sont efficaces, dans la mesure où ceux-ci ont peu investi dans leur réseau, avec pour conséquence une faible qualité de leur réseau. Compte tenu que le « UK » ne fera plus partie de l'Europe après le Brexit, on peut se poser la question de la pertinence d'intégrer les GRD anglais dans le benchmarking.

On peut aussi aisément préjuger que les GRD néerlandais sont efficaces, compte tenu de la densité de population des Pays-Bas et de sa géologie. Sur quelle base se fera le choix final ? Utilisera-t-on un tirage aléatoire ? Comment garantir l'établissement d'un échantillon non biaisé ? Ceci est particulièrement important au vu de l'hétérogénéité des GRDs wallons surtout en électricité.

Le choix des points de comparaison, dans un échantillon de 20 GRDs, est un des éléments clés du modèle et ce point n'est pas du tout abordé. Sans informations supplémentaires sur la manière dont l'échantillon est constitué, les résultats ne pourront pas être considérés comme robustes.

2.3.2. Réponse aux remarques de RESA, AIEG, AIESH et REW

Sélection des GRD constituant l'échantillon

Toute mesure d'efficacité d'un GRD est une mesure relative aux autres GRD composant l'échantillon. Nous notons tout d'abord que RESA, AIEG, AIESH et REW semblent vouloir éviter d'être comparé à certains GRD qui seraient plus efficaces qu'eux. En ce sens, si les GRD wallons sont comparés uniquement à des GRD préjugés « non efficaces », on peut s'interroger sur la



pertinence de conduire un exercice de mesure d'efficacité. Dans le cas où des mesures d'efficacité, dont les résultats sont publics, ont déjà été effectuées par le passé pour des GRD étrangers, il conviendra de sélectionner les GRD à intégrer dans le benchmark parmi les plus efficaces.

Ensuite, les préjugés exprimés par les GRD wallons sur l'efficacité des GRD anglais et néerlandais nous paraissent prématurés et hasardeux, car notre expérience des benchmark internationaux nous a montré que les intuitions basées sur l'un ou l'autre élément considéré séparément comme un indice d'efficacité ou d'inefficacité potentielle se révèlent souvent être trompeuses, l'efficacité prenant en compte un grand nombre d'effets conjugués propre à chaque GRD.

Prise en compte de la densité de population et de la géologie dans la méthode de mesure d'efficacité

Nous avons proposé de retenir parmi le jeu de variables explicatives à considérer dans le cadre de la méthode advanced COLS pour générer le meilleur modèle économétrique : la longueur du réseau, le nombre de clients ainsi que le ratio nombre de client / km de réseau ; le retour d'expérience des benchmarks montre en effet que la longueur du réseau et le nombre de clients sont parmi les variables explicatives les plus corrélées avec les coûts des GRD. Le ratio nombre de clients par kilomètre traduit directement l'effet de la densité de population dans la mesure d'efficacité. En effet, dans un habitant urbain, le nombre de clients par km de réseau est plus élevé que dans un habitat rural. Ainsi, un GRD rural ne sera pas pénalisé par rapport à un GRD urbain.

Concernant la géologie, son impact est majoritairement négligé par les régulateurs des pays que nous avons analysés en détail (Allemagne, Autriche, France et Finlande). Pour ces exemples, seul le régulateur allemand, et uniquement pour les GRD gaz prend en compte une variable explicative traduisant la géologie (longueur du réseau pondérée par la part de sol de type 4, 5 et 6 dans la surface totale couverte par le réseau, page 53 du rapport intermédiaire de Schwartz and Co).

2.3.3. Précisions apportées dans le rapport final

Ces remarques n'appellent pas de modifications du rapport intermédiaire.

2.4. Analyse académique du professeur Axel Gautier

2.4.1. Méthode Advanced COLS

2.4.1.1. Remarques du professeur Axel Gautier

Un des avantages mis en avant pour la méthode Advanced COLS par rapport aux autres méthodes paramétriques est l'absence de présupposé dans la définition de la fonction de coût. Ceci appelle plusieurs observations :

- L'absence d'une forme fonctionnelle sous-jacente n'est pas nécessairement un avantage. Une spécification log-linéaire correspond à une fonction de production sous-jacente de



type Cobb-Douglas. En théorie de la production, cette fonction est communément admise, tout comme l'est la fonction translog, cette dernière étant par ailleurs une forme fonctionnelle très flexible. Une fonction de coût non spécifiée à priori et déterminée par itération, n'a pas nécessairement un fondement théorique et de ce fait la relation entre inputs et outputs peut être plus difficile à comprendre et à appréhender pour les GRDs, notamment la nature des *rendements d'échelle*. Le modèle retenu dans le cadre français (page 66) et qui contient à la fois des variables en niveau et en ratio et des valeurs linéaires et en log n'est pas nécessairement simple à comprendre d'un point de vue économique.

- Avec un nombre limité de variables indépendantes du fait d'un échantillon réduit, le nombre de combinaisons possibles est limité.
- Que se passe-t-il si des variables sont significatives statistiquement, mais n'ont pas le signe attendu ? Que se passe-t-il si des variables sont économiquement significatives, mais pas statistiquement significatives dans le modèle ? Les explications fournies page 23 concernant ce dernier point ne sont pas claires.

2.4.1.2. Réponse aux remarques du professeur Axel Gautier

Choix de la forme fonctionnelle

Nous ne pouvons pas partager l'observation du professeur Axel Gautier qui mentionne que « l'absence d'une forme fonctionnelle sous-jacente n'est pas nécessairement un avantage ». En effet, l'absence de présupposition d'une forme fonctionnelle est très largement reconnue dans la littérature scientifique comme l'avantage principal des méthodes non-paramétriques, et c'est justement également le cas de la méthode paramétrique advanced COLS.

Par ailleurs, concernant les méthodes paramétriques, le professeur Axel Gautier mentionne que les formes fonctionnelles de type Cobb-Douglas et Translog sont communément admises en économétrie. Néanmoins, il n'explique pas s'il est plus pertinent de retenir l'une ou l'autre de ces formes fonctionnelles, ni à quelles conditions. La méthode Advanced COLS permet de retenir la forme fonctionnelle la plus pertinente statistiquement.

Les réponses de Schwartz and Co concernant la validité économique du modèle sélectionné sont présentées au paragraphe 2.2.2.

Nombre de combinaison de variables

Pour rappel, la sélection des variables explicatives et de la fonction de coût est effectuée dans la méthode Advanced COLS par la méthode de régression par étape (« stagewise regression ») et par la méthode « all possible models ».

En particulier, la méthode « all possible models », qui est considérée dans la littérature économétrique comme la méthode optimale, consiste à calculer et tester tous les modèles possibles



à partir des variables explicatives disponibles sous différentes formes (en logarithme, en linéaire et pour certaines d'entre elles en ratios). Le nombre de modèles testés est égal à $2^k - 1$ modèles, avec k le nombre de variables explicatives testées (chaque forme de variable comptant pour une variable). Avec 11 variables explicatives testées par exemple, le nombre de modèles testés est de 2047, ce qui ne constitue pas un nombre de combinaisons limité.

Validité du modèle

Les réponses à ces remarques sont présentées dans le paragraphe 2.2.2 en réponse aux remarques de RESA, AIEG, AIESH et REW.

2.4.1.3. Précisions apportées dans le rapport final

Ces remarques n'appellent pas de modifications du rapport intermédiaire.

2.4.2. Taille de l'échantillon

2.4.2.1. Remarques du professeur Axel Gautier

Le benchmark sera réalisé sur base d'un échantillon restreint de 15 à 20 GRDs pour l'électricité et le gaz, sans que l'on sache si l'utilisation d'un échantillon réduit est un choix délibéré ou le résultat d'une contrainte (non accès aux données). La taille réduite de l'échantillon pose plusieurs problèmes et questions :

- En économétrie, la qualité et la robustesse des estimations augmentent avec la taille de l'échantillon. L'utilisation d'un échantillon de très petite taille du point de vue économétrique donne des résultats qualitativement limités et peu robustes.
- Un échantillon de taille réduite limite le choix des méthodes d'estimation. Si les méthodes basées sur un OLS permettent d'avoir une estimation de l'efficacité individuelle sur base d'un échantillon très réduit, ce n'est pas le cas avec la méthode SFA qui nécessite plus d'observations, et dans une moindre mesure avec la méthode DEA. Cependant la qualité de l'estimation avec un échantillon réduit est limitée et le fait de pouvoir estimer une efficacité individuelle n'est pas en soi un gage de qualité. On peut se demander si le choix de méthode de benchmarking n'est pas une conséquence directe du choix (?) d'un échantillon limité.
- Le nombre de variables explicatives (outputs et variables de contrôle environnementales) utilisées dans le modèle dépend du nombre d'observations (l'étude mentionne la règle suivante : nombre d'observations \geq variables indépendantes + 11). La taille réduite de l'échantillon limite fortement le nombre de variables indépendantes qui peuvent être intégrées dans le modèle. L'étude mentionne 4 variables output et une variable de contrôle (coût du travail).

2.4.2.2. Réponse aux remarques du professeur Axel Gautier



La taille minimale requise de l'échantillon indiquée dans notre rapport (nombre d'observations \geq variables indépendantes +11) correspond à une règle empirique constituant une pratique reconnue en économétrie. Sans remettre en cause cette règle, le professeur Axel Gautier s'exprime de manière assez ambiguë sur la taille d'échantillon que nous proposons, puisqu'il qualifie notre échantillon d'abord de « taille réduite », puis indique que « l'utilisation d'un échantillon de « très petite taille » du point de vue économétrique donne des résultats qualitativement limités et peu robustes ». Comme nous le démontrons ci-après, un échantillon de taille réduite n'est pas un échantillon de très petite taille au sens où ce dernier porte atteinte à la qualité et la robustesse des résultats obtenus.

Il est inexact de dire qu'un échantillon de 15 à 20 GRD résulterait forcément en un modèle économétrique peu robuste. Il existe de nombreux exemples de modèles économétriques de qualité développés à partir d'échantillon réduits, mais respectant la règle susmentionnée, en particulier dans le monde de la régulation de l'énergie, par exemple :

- Le régulateur de l'énergie irlandais (CER) a mis en œuvre en 2015 la méthode COLS à travers un benchmark international de 15 GRD électricité (l'unique GRD électricité irlandais + 14 GRD électricité de Grande-Bretagne) pour préparer la période tarifaire 2016-2020¹.
- Le régulateur de l'énergie anglais (OFGEM) a mis en œuvre à de nombreuses reprises la méthode COLS à travers des benchmark portant sur les 14 GRD électricité de Grande-Bretagne.
- Le cabinet Schwartz and Co a conduit pour le régulateur français (CRE) une mesure d'efficacité des OPEX du plus grand GRD français sur la base d'un benchmark international basé sur la méthode Advanced COLS avec 28 points d'observation, qui a permis d'aboutir à un modèle économétrique de très bonne qualité à 4 variables explicatives².

Par ailleurs, ce sont les tests statistiques de validité du modèle qui constituent le juge de paix en matière de taille d'échantillon requis : s'il apparaissait lors de l'implémentation qu'aucun modèle de qualité n'était obtenu avec le nombre de GRD que nous prévoyons, alors nous augmenterions le nombre d'observations. Nous ne considérons donc pas du tout les 15 à 20 GRD comme des valeurs indépassables au moment de la mise en œuvre de la méthode retenue.

2.4.2.3. Précisions apportées dans le rapport final

Des compléments seront apportés au rapport final sur la base de nos réponses.

¹ Consultancy Support for Electricity Transmission and Distribution Revenue Controls (2016-2020), Commission for Energy Regulation, Jacobs, June 2015

² Rapport public de l'étude réalisée par Schwartz and Co pour la Commission de Régulation de l'Énergie (CRE) en France, publiée par la CRE le 18 novembre 2015, dans le cadre de la consultation publique de la CRE sur le prochain tarif d'utilisation des réseaux publics de distribution de gaz naturel de GRDF (dit « ATRD5 »)



2.5. Analyse académique du professeur Damien Ernst

2.5.1. Ensemble de données

2.5.1.1. Remarques du professeur Damien Ernst

Je note que Schwartz and Co propose comme méthode de mesure d'efficacité des GRDs une méthode de type Cols. Ces méthodes nécessitent de réaliser une régression linéaire à partir d'un ensemble d'apprentissage permettant de déterminer un modèle prenant en entrée des variables explicatives des coûts d'un GRD (nombre de clients servis, etc.) et en sortie ses coûts (totaux). Ces approches n'ont de sens que si le modèle appris est effectivement capable de bien représenter les coûts de GRDs non considérés pour l'apprentissage. Pour être plus précis, définissons un ensemble test de GRDs composé de GRDs dont les données n'ont pas été utilisées pour réaliser la régression linéaire. Le modèle est appris de manière correcte si la moyenne des erreurs sur l'ensemble de test est faible. L'erreur pour un élément de l'ensemble de test est définie comme étant la valeur absolue de la différence entre ses coûts totaux et ceux prédits par le modèle. Il est bien connu que l'origine des erreurs sur l'ensemble de test est double : la première vient du biais dans les modèles, que l'on ne discutera pas ici bien que particulièrement présent dans les techniques d'apprentissage supervisé basées sur la régression linéaire. La seconde vient de la variance du modèle, qui peut être définie comme étant sa sensibilité à des modifications de l'ensemble d'apprentissage. Un modèle appris ne peut dès lors être bon que si des modifications mineures dans l'ensemble d'apprentissage ne conduisent qu'à des variations faibles de ce dernier. Cependant, vu la taille extrêmement faible de l'échantillon de GRDs utilisés pour la comparaison - on parle ici de 15 à 20 GRDs - il est clair que la variance de votre modèle va être bien trop élevée, ce qui invalide complètement votre approche.

Pour vous en convaincre, je vous conseille de réaliser l'expérience suivante. Si vous disposez des données liées à n GRDs, recalculez votre modèle n fois en utilisant à chaque fois en entrée un sous-ensemble différent de taille $n-1$ de votre ensemble d'apprentissage. Évaluez alors pour chacun des modèles les prédictions associées à un ensemble de points sélectionnés au hasard dans votre espace de variables d'entrée. Vous allez sans doute voir que d'un modèle à l'autre, lorsque n est de l'ordre de 10 à 20, les résultats varient significativement alors que vous n'avez que très légèrement modifié la base de données. Ceci est un signe du fait que votre modèle n'est absolument pas bon, ce qui hypothèque toute base scientifique à l'approche que vous avez développée.

2.5.1.2. Réponse aux remarques du professeur Damien Ernst

Les remarques du professeur Damien Ernst ne tiennent pas compte des pratiques économétriques communément admises. En effet :

- La notion de modèle « appris » : en économétrie, à la différence des algorithmes de Machine Learning en IA, cette notion n'existe pas. Les paramètres sont estimés à partir d'un échantillon, le modèle est ensuite passé au crible des tests statistiques. S'il satisfait l'ensemble des critères statistiques et économiques, il est validé.
- La variance du modèle : nous connaissons la variance des coefficients de régression, la variance résiduelle (estimation de la variance de l'erreur), la variance de l'erreur de prévision,



mais la variance du modèle est une notion qui nous est inconnue (le modèle n'est pas une variable aléatoire !). Nous imaginons que l'auteur veut parler de la non-stabilité des coefficients du modèle en fonction de la taille de l'échantillon. La réponse à cette objection est traitée au point au paragraphe 4.5.3. Nous vérifierons en effet la robustesse de notre modèle vis-à-vis de chaque point d'observation avec une analyse de l'effet de levier et des points d'influence, ainsi qu'une analyse des valeurs anormales.

Nous notons d'ailleurs que le professeur Damien Ernst est spécialisé dans le domaine de l'intelligence artificielle et de l'apprentissage par renforcement appliqué aux problématiques techniques des réseaux d'électricité³.

Concernant les objections sur la taille de l'échantillon, voir nos réponses au paragraphe 2.3.2.2.

2.5.1.3. Précisions apportées dans le rapport final

La description de la méthodologie d'identification et de gestion des valeurs anormales, de l'effet de levier et des points d'influence sera complétée dans le rapport final.

2.5.2. Variables explicatives

2.5.2.1. Remarques du professeur Damien Ernst

J'ai parcouru la liste des variables explicatives que Schwartz compte utiliser dans son approche de benchmarking. Pour que cette approche puisse avoir du sens, il est très important que les variables explicatives soient informatives et définies de manière très claire. Cela n'est malheureusement pas le cas pour de nombreuses d'entre elles, comme par exemple la pointe de charge. Je suis par exemple incapable de savoir à quelle valeur le consultant fait exactement référence lorsqu'il s'agit d'un GRD et je crains que les différentes interprétations qui puissent être faites de ces variables influencent très fortement les conclusions que vous pourriez tirer de vos études. De plus, à cet état de fait s'ajoute que Schwartz and Co propose dans sa méthodologie de transformer certaines variables explicatives en utilisant des opérateurs non linéaires avant de les utiliser comme entrée de votre méthode de benchmarking. J'attire votre attention sur le fait qu'en procédant de la sorte, l'approche deviendra dépendante des unités dans lesquelles les grandeurs sont exprimées (par exemple des mètres ou des kilomètres pour les longueurs de ligne), ce qui me semble très peu justifiable d'un point de vue scientifique.

2.5.2.2. Réponse aux remarques du professeur Damien Ernst

La pointe de charge est une notion très courante dans le secteur des réseaux d'électricité et de gaz. Nous partageons bien sûr la nécessité de définir précisément toutes les variables explicatives à collecter pour éviter les mauvaises interprétations et les erreurs de données qui pourraient en résulter. Dans le cas de la pointe de charge, il s'agit de la puissance maximale appelée par l'ensemble du réseau du GRD pour l'année de mesure de l'efficacité considérée.

³ <http://blogs.ulg.ac.be/damien-ernst/damien-ernst/>



L'utilisation d'opérateurs non linéaires dans les modèles économétriques est une pratique courante et ne soulève pas de problème spécifique. A titre d'exemple, les régulateurs suivants, qui utilisent des méthodes paramétriques ont sélectionné les formes fonctionnelles suivantes :

- Allemagne : fonction Translog pour les GRD gaz avec une méthode SFA ;
- Autriche : fonction Cobb-Douglas pour les GRD gaz et les GRD électricité avec une méthode MOLS.

Enfin un des intérêts bien connu des modèles économétriques est justement qu'ils permettent de traiter des variables explicatives exprimées dans des unités hétérogènes.

2.5.2.3. Précisions apportées dans le rapport final

Néant.

3. Rapport de consultation sur les remarques d'Ores (via un rapport de l'entreprise Oxera)

3.1. Processus d'évaluation des coûts

3.1.1. Remarques d'Oxera

D'après l'expérience que nous avons acquise en matière de benchmarking de l'efficacité dans plusieurs pays et secteurs européens et de benchmarking couvrant plusieurs juridictions, le processus d'évaluation des coûts implique généralement plusieurs consultations sur la méthode d'estimation, le développement du modèle, la construction des coûts et des drivers de coûts, et l'application des paramètres d'efficacité dans la formule de détermination des revenus.

Par exemple, dans le processus d'évaluation des coûts, Ofgem (le régulateur britannique de l'énergie) a organisé 14 réunions avec les parties prenantes du secteur avant de rédiger ses projets de décisions pour les réseaux de distribution de gaz, et a déjà organisé 10 réunions similaires avec les parties prenantes de la distribution d'électricité malgré le fait que le projet de décision ne sera pas publié avant l'été de 2022. Ofwat (le régulateur de l'eau pour l'Angleterre et le Pays de Galles) a mené des consultations similaires et a permis aux entreprises de soumettre leurs propres modèles d'évaluation des coûts lors d'une consultation en mars 2017, celle-ci a donné au régulateur un aperçu de ce que le secteur considère comme étant des facteurs importants de dépenses. Ofwat a publié sa décision finale en décembre 2019.

De toute évidence, le processus d'évaluation des coûts peut prendre plusieurs mois, voire plusieurs années, car les données et la méthodologie sont discutées avec les parties prenantes. Ce processus peut prendre encore plus de temps dans un contexte de benchmarking international, où les entreprises rapportent leurs données selon différentes directives de reporting et un travail important est nécessaire pour homogénéiser l'ensemble de données.



Lors du processus de consultation, il est important que les GRD aient accès aux mêmes données et codes de modélisation utilisés par le régulateur et son consultant. Cela donne aux GRD plus de détails sur le développement du modèle et l'approche d'estimation, et leur permet de faire des requêtes et des recommandations spécifiques sur des questions empiriques (telles que les erreurs de données).

Recommandation d'Oxera :

- Schwartz and Co devrait organiser une consultation du secteur tout au long du processus d'élaboration du modèle pour s'assurer que son ensemble final de modèles est solide d'un point de vue opérationnel, économique et statistique
- Tous les codes de modélisation, les fichiers d'analyse et les données brutes qui alimentent l'analyse intermédiaire de Schwartz and Co devraient être partagés avec les parties prenantes avant la publication du rapport final.
- Schwartz and Co devrait accorder suffisamment de temps aux GRD pour identifier les erreurs éventuelles dans l'analyse, apporter des éclaircissements et faire des commentaires sur les modèles.
- Schwartz and Co devrait citer toutes les références appuyant ses décisions de modélisation. De même, ils devraient présenter des preuves empiriques et opérationnelles de toute hypothèse qui alimente leur modèle final.
- Il est bien reconnu que tous les modèles (c'est-à-dire les inputs, les outputs et leur relation sous-jacente) sont des représentations imparfaites d'une réalité complexe. Par conséquent, Schwartz and Co devrait valider son (ses) modèle (s) préféré(s) par rapport à des alternatives (cela inclut à la fois l'approche d'estimation et la sélection des drivers de coûts), en tenant compte de l'analyse de sensibilité et des considérations de « marges d'erreur » qui sont essentielles dans de tels exercices.
- Dans le rapport intermédiaire, Schwartz and Co fait un certain nombre d'affirmations qui doivent être étayées par une analyse empirique (ces affirmations sont énumérées dans la section 4). Les choix de modélisation et les décisions devraient être réévalués une fois qu'ils auront accès à l'ensemble de données et que les entreprises auront suffisamment de temps pour répondre à ces révisions.

3.1.2. Réponse aux remarques d'Oxera

Nous prenons note des suggestions formulées par la société de consultance Oxera concernant le processus d'évaluation des coûts. Nous tâcherons d'en tenir compte, dans la mesure du possible, compte tenu des exigences et du processus définis dans le cahier des charges, du modèle de mesure de l'efficacité qui sera sélectionné, de l'expertise et du savoir-faire de la société Schwartz and Co, ainsi que de la confidentialité de certaines données. Par conséquent et pour cette dernière raison, il n'est pas prévu de partager les fichiers et les données brutes mais de partager le rapport final qui détaillera le calcul ainsi que les hypothèses et choix faits pour parvenir à l'obtention des résultats finaux.

3.1.3. Précisions apportées dans le rapport final



Néant : le processus de consultation est hors du champ de l'étude.

3.2. Méthodologie d'évaluation des coûts

3.2.1. « Advanced COLS » comme approche de développement de modèle

3.2.1.1. Remarques d'Oxera

Page 3 du rapport d'Oxera :

Schwartz and Co a l'intention d'utiliser la méthode « *Advanced COLS* » comme unique approche d'estimation. *Advanced COLS* est une approche exclusivement statistique de développement de modèle (parfois appelée « analyse des drivers de coûts ») basée sur l'approche d'estimation des moindres carrés ordinaires dite « *Ordinary Least Squares* » ou OLS. Selon la manière dont elle est appliquée, il peut s'agir d'une procédure entièrement mécanique. **Une limitation spécifique de cette approche qui a été négligée par Schwartz and Co est que les critères de sélection du modèle sont basés sur des tests statistiques qui ne sont pas appropriés dans un contexte d'analyse de l'efficacité. En outre, l'approche ne tient pas compte explicitement de la validité opérationnelle ou économique des modèles concurrents.**

Paragraphe 4.1 du rapport d'Oxera :

Schwartz and Co recommande d'utiliser la méthode « *advanced COLS* » comme seule approche d'estimation. **Advanced COLS est une approche statistique standard et automatisée du développement de modèle**, où soit: (i) les drivers de coûts doivent être inclus ou exclus dans le modèle un par un, en fonction de leur corrélation avec les coûts, jusqu'à ce qu'aucun nouveau driver de coûts ne soit ajouté ou supprimé (cette approche est appelée approche par étapes, qui peut être prospective ou rétrospective); ou (ii) tous les modèles possibles sont testés et celui qui obtient le score le plus élevé selon certains critères de sélection de modèle est sélectionné comme modèle final (approche dite de « tous les modèles possibles »). Il est à noter que cette approche du développement de modèle n'est pas propre uniquement à l'approche d'estimation COLS et peut être appliquée à toute approche d'estimation paramétrique, mais Schwartz and Co ne le reconnaît pas dans son rapport. Par exemple, le régulateur allemand de l'énergie, Bundesnetzagentur, dans son analyse de benchmarking des GRD gaz récemment achevée, considère les méthodes OLS et SFA au stade du développement du modèle. Cela faisait suite au rapport d'expert de Kumbhakar, Parthasarathy et Thanassoulis (2017) qui mettait en évidence les limites de l'approche OLS dans un contexte d'efficacité.

En théorie, il existe plusieurs façons d'identifier les drivers de coûts pertinents. **L'approche de Schwartz and Co n'est pas claire quant à la manière de détecter les relations non linéaires entre les coûts et les drivers de coûts. Par exemple, Ofwat teste la densité de population et la densité de population au carré comme drivers de coûts dans son modèle économétrique.** Cependant, la densité est statistiquement non significative lorsqu'elle est incluse dans le modèle sans le terme au carré et n'aurait donc pas été détectée en utilisant l'approche par étapes (stagewise approach en anglais). L'approche dite de « tous les modèles possibles » pourrait permettre de détecter une telle relation, mais seulement si le terme au carré était également considéré comme un



inducteur de coûts en première instance. Schwarz and Co déclare également qu'il prendra en considération les spécifications à la fois linéaires et logarithmiques; cependant, il n'indique pas comment il déterminera qu'elle spécification est la plus appropriée.

Surtout, l'approche de Schwartz and Co est mécanique et purement statistique - il n'y a aucune considération de la cohérence opérationnelle et économique des résultats du modèle dans le processus de sélection du modèle. De plus, les tests statistiques utilisés pour informer le processus de sélection du modèle reposent sur des hypothèses qui pourraient ne pas être valides sur l'échantillon et doivent être testées empiriquement.

Par exemple, le critère d'information Akaike (Akaike Information Criterion en anglais ou AIC qui est la mesure de la qualité du modèle utilisée dans l'approche «tous les modèles possibles») n'est une mesure valide de la qualité du modèle que lorsque l'échantillon est important. De plus, l'estimateur OLS utilisé par Schwartz and Co pour le processus de développement du modèle suppose que le terme d'erreur est symétrique et normalement distribué.

Cependant, si Schwartz and Co soupçonne que certains GRD fonctionnent de manière inefficace, le terme d'erreur sera biaisé, comme le montre l'encadré 4.1. Toute inférence statistique que Schwartz and Co propose d'utiliser pour sélectionner les drivers de coûts (et ils ont l'intention de se fonder uniquement ou principalement sur l'inférence statistique) ne sera donc pas concluante. Comme indiqué, certains régulateurs, tels que la Bundesnetzagentur, procèdent à un développement de modèle en utilisant des méthodes SFA (qui modélisent explicitement le terme d'erreur biaisé) ainsi que des méthodes OLS.

Recommandation d'Oxera :

- Schwartz and Co devrait tenir compte de l'intuitivité économique et opérationnelle de chaque modèle tout au long du processus d'élaboration du modèle.
- Schwartz and Co devrait utiliser un ensemble plus large de tests statistiques et d'outils de diagnostic pour soutenir son processus de développement de modèle. Cela lui permettrait de détecter d'importantes relations non linéaires entre les coûts et les facteurs de coûts et de différencier les modèles concurrentiels.
- De même, Schwartz and Co devrait examiner de manière appropriée si la forme fonctionnelle devrait être en niveaux ou en logarithmes ou si elle n'est pas concluante entre les deux. Bien qu'il existe des tests statistiques permettant d'étayer une approche plutôt qu'une autre, ces tests ne sont parfois pas concluants et la forme fonctionnelle devrait être fondée sur la cohérence opérationnelle et économique.
- Toutes les hypothèses qui sous-tendent les tests statistiques de Schwartz and Co (comme le t-test de Student) doivent être testées empiriquement, et les résultats de ces tests devraient être présentés dans le rapport de Schwartz and Co. Si les résultats des tests statistiques ne soutiennent pas les hypothèses de modélisation de Schwartz and Co ou que ces hypothèses ne sont pas valables pour les tests considérés comme concluants, ils devraient être consultés et validés par des moyens robustes.



3.2.1.2. Réponse aux remarques d'Oxera

Soi-disant non pertinence des tests statistiques dans le contexte de mesure d'efficience

Oxera nous indique tout d'abord à la page 3 de son rapport : « Une limitation spécifique de cette approche qui a été négligée par Schwartz and Co est que les critères de sélection du modèle sont basés sur des tests statistiques qui ne sont pas appropriés dans un contexte d'analyse de l'efficience. »

Nous ne partageons pas l'affirmation d'Oxera qui remet en question de la validité des tests statistiques d'un modèle économétrique. Oxera semblerait donc suggérer la possibilité d'utilisation d'un modèle économétrique sans validation de ce modèle par des tests statistiques, ce qui va à l'encontre des pratiques incontournables en économétrie. Les tests statistiques sont la force de la construction des modèles économétriques et il serait aberrant de considérer un modèle économétrique sans la réalisation de tests statistiques de validation.

Par ailleurs, nous n'avons pas trouvé dans le domaine public la référence fournie par Oxera pour étayer ses propos : « discussion dans Kumbhakar, S., Parthasarathy, S. and Thanassoulis, E. (2018), "Validity of Bundesnetzagentur's cost driver analysis and second-stage analysis in its efficiency benchmarking approach", février » et juger de la pertinence des arguments. Par ailleurs, nous notons que Kumbhakar, S, et Thanassoulis, E sont Associate de l'entreprise Oxera et que Parthasarathy, S est partner de l'entreprise Oxera. Nous notons que Parthasarathy, S est présenté comme un expert de la méthode SFA, qui est une méthode paramétrique. Nous nous étonnons donc d'autant plus qu'un expert des méthodes paramétriques puisse affirmer que les tests statistiques ne sont pas appropriés dans un contexte d'analyse de l'efficience. Quel que soit le contexte d'utilisation d'un modèle économétrique, les tests statistiques sont indispensables et constituent un prérequis.

Les affirmations d'Oxera sont par ailleurs contradictoires, puisque que d'un côté Oxera indique que les tests statistiques ne sont pas appropriés dans un contexte d'analyse d'efficience, d'autre part Oxera recommandent à Schwartz and Co d'utiliser un ensemble plus large de tests statistiques.

Sélection des variables explicatives et de la fonction de coûts et caractère soi-disant mécanique de la méthode Advanced COLS

Oxera nous indique également à la page 3 de son rapport : « en outre, l'approche ne tient pas compte explicitement de la validité opérationnelle ou économique des modèles concurrents ».

Cette affirmation est inexacte. En effet, nous indiquons dans notre rapport intermédiaire :

- dans le dernier paragraphe de la section 3.2.3, au sujet des modèles économétriques conçus à partir des approches « stagewise regression » et « all possible models » : « ces méthodes de détermination peuvent aboutir à des variables explicatives qui sont les plus pertinentes statistiquement, mais qui sont difficiles à appréhender d'un point de vue métier pour les



GRD. Dans ce cas, le choix des variables explicatives peut être ajusté en pleine connaissance d'une perte de qualité statistique du modèle, en choisissant un modèle ayant une qualité statistique légèrement moindre que le modèle optimal, mais plus facilement appréhendable d'un point de vue métier ».

- dans la section 5.2 : « le jeu de variable exact sera fixé au moment de l'implémentation de la méthode, sur la base de notre expérience et du retour d'expérience international. Les variables retenues doivent être suffisamment corrélées avec la base de coût et faire également du sens d'un point de vue métier. »

Nous partageons l'importance de la vérification de la validité opérationnelle et économique des modèles conçus avec une méthode économétrique. Cette préoccupation doit être mise en œuvre dès la présélection du jeu de variables explicatives sur la base desquelles la méthode advanced COLS va permettre de calculer les différents modèles économétriques possibles : comme indiqué, ces variables doivent avoir un sens métier et constituer des drivers des coûts des GRD.

Nous tenons par ailleurs à apporter les précisions suivantes :

- Dans la méthode advanced COLS, le choix des variables explicatives n'est pas mécanique. Pour rappel, la méthode advanced COLS sélectionne les variables explicatives et la fonction de coûts la plus pertinente sur la base de deux méthodes :
 - la méthode « stagewise regression » : la méthode de régression par étape (« stagewise regression ») est un processus de régression multilinéaire visant à sélectionner le meilleur ensemble de variables explicatives, minimisant l'intercorrélacion entre les variables explicatives en étudiant les résidus de régression ;
 - la méthode « all possible models » : la méthode « all possible models » consiste à calculer et tester tous les modèles possibles à partir du jeu de variables explicatives disponibles en linéaire, en logarithmique et en ratio (ceci permet justement de prendre en compte les éventuelles relations non linéaires entre les variables explicatives et les coûts).
- Le meilleur modèle économétrique ainsi identifié (selon les critères d'Akaike et Schwarz) est ensuite validé par une méthodologie à deux étapes :
 - **Etape 1 - validité statistique** : une série de tests statistiques est réalisée dans le but de valider statistiquement la pertinence et la robustesse du modèle. Les principaux tests utilisés sont décrits dans notre rapport intermédiaire et des compléments apportés aux paragraphes 3.2.1.2 et 4.5.3 du présent rapport.
 - **Etape 2 - validité économique et opérationnelle** : la validité économique et opérationnelle du modèle économétrique est vérifiée en contrôlant que les variables explicatives du modèle sont cohérentes avec les pratiques métiers des GRD et que les signes attendus des coefficients sont conformes à l'intuition économique.



- Si le modèle (variables explicatives et fonction de coûts) n'est pas ainsi validé, il n'est pas retenu et le meilleur modèle économétrique suivant est testé. L'itération s'arrête lorsque le modèle a satisfait aux deux étapes de validation.

Détection des relations non linéaires entre coûts et drivers de coûts

Le test de RAMSEY (test RESET) évoqué par Oxera permet de détecter des relations non linéaires entre les variables explicatives et les coûts des GRD. Ce test ne présente pas d'intérêt particulier avec la méthode advanced COLS car ce type de relations est testé directement dans le cadre de l'approche « all-possible models ».

Par ailleurs, les modèles de mesure d'efficacité utilisés dans le domaine de l'électricité et du gaz n'ont à notre connaissance jamais pris en compte la densité de population au carré. Pour les exemples de régulateurs européens que nous avons analysés en détail (Allemagne, Autriche, Finlande et France), seul le régulateur autrichien utilise des densités, qui sont des densités de points de connexion et non des densités de population. Ces densités de points de connexion ont une relation logarithmique ou linéaire avec les coûts⁴ (la fonction de production utilisée pour la méthode MOLS est une fonction log-linéaire). Par ailleurs, Oxera mentionne l'étude Ofwat qui teste la densité de population et la densité de population au carré comme drivers de coûts dans son modèle économétrique. Cet exemple concerne les réseaux d'eau qui ont des propriétés et des spécifications très différentes des réseaux d'énergie (électricité et gaz). Il n'est ainsi pas pertinent de comparer le choix des variables explicatives utilisées dans l'étude de l'Ofwat au choix de variables explicatives utilisées pour les réseaux d'électricité et de gaz.

Pertinence du critère d'information Akaike

Le critère d'information d'Akaike (AIC) est tout à fait pertinent, quelle que soit la taille de l'échantillon. Dans le cas d'échantillons de petite taille (typiquement pour $n/k < 40$, avec n le nombre d'observations et k le nombre de variables explicatives), il est utilisé dans sa forme dite corrigée (AICc), tel que préconisé par Burnham et Anderson⁵ (AICc tend vers l'AIC lorsque le nombre d'observation devient grand).

Terme d'erreur OLS symétrique et normalement distribué, et question de l'inférence statistique en présence d'inefficience

Oxera nous indique que « l'estimateur OLS utilisé par Schwartz and Co pour le processus de développement du modèle suppose que le terme d'erreur est symétrique et normalement

⁴ Electricity Distribution System Operators, 1 January 2019 - 31 December 2023 Regulatory Regime for the Fourth Regulatory Period, December 2018

Regulierungssystematik für die dritte Regulierungsperiode der Gasverteilernetzbetreiber
1. Jänner 2018 - 31. Dezember 2022

⁵ Burnham, Kenneth P. and David R. Anderson . 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach. 2d ed. New York: Springer-Verlag.



distribué ». Tout d’abord, cette formulation est inexacte d’un point de vue économétrique à deux niveaux : le terme d’erreur est un inconnu en économétrie et seul le résidu de la modélisation est un paramètre connu ; ensuite toute loi normale est symétrique.

Nous comprenons donc que ce qu’Oxera cherche à dire est que le résidu de la modélisation OLS suit une loi normale. Dans la méthode advanced COLS, comme dans toute méthode économétrique s’appuyant sur OLS, on fait l’hypothèse que le résidu de la régression suit une loi normale, et on vérifie également que c’est bien le cas, en appliquant par exemple le test de Jarque-Bera. Comme nous l’avons également indiqué dans notre rapport, nous vérifions également la présence éventuelle d’hétéroscédasticité par le test de White. Un modèle qui présente une hétéroscédasticité est par définition un modèle qui ne respecte pas l’hypothèse de normalité. Dans ce cas nous pouvons corriger le modèle et ainsi retrouver une loi normale.

3.2.1.3. Précisions apportées dans le rapport final

Le rapport final prendra en compte les éléments de réponses formulés par Schwartz and Co.

3.2.2. Rejet de la méthode SFA

3.2.2.1. Remarques d’Oxera

Nous notons que la méthode SFA a été mise en œuvre sur des bases de données plus petites que celle proposée par Schwartz and Co, y compris dans des exercices de benchmarking international⁶. De plus, la taille appropriée d’une base de données « nécessaire » pour toute méthode (y compris celle proposée par Schwartz and Co) est toujours une question empirique qui ne peut être décidée en l’absence des données et du modèle.

Schwartz and Co soutient que la méthode SFA nécessite des hypothèses « arbitraires » concernant la distribution de l’inefficience dans l’échantillon. ORES dans ses questions écrites, a fait valoir que la méthode de Schwartz and Co exigeait également une décision arbitraire concernant le benchmark d’efficience approprié, mais Schwartz and Co ne l’a pas reconnu dans sa réponse. En outre, il est possible d’évaluer plusieurs distributions de l’inefficience différentes et, soit, de sélectionner la distribution préférée en se basant sur des diagnostics statistiques, soit, de faire une triangulation entre les modèles basés sur différentes hypothèses. En ce sens l’hypothèse de distribution n’a pas besoin d’être une décision arbitraire et peut être éclairée par des preuves empiriques.

Recommandation d’Oxera :

- Schwartz and Co devrait tester les modèles SFA sur son ensemble de données et ne pas les exclure ex ante comme méthode. Si Schwartz and Co considère que la base de données est trop petite pour estimer les modèles SFA, il peut alors envisager de collecter davantage de données (par le biais de GRD supplémentaires ou d’une période supplémentaire, ou les deux).

⁶ Voir The Office of Rail and Road (2013), “PR13 Efficiency Benchmarkings of Network Rail using LICB”, août, p. 6; The Office of Rail and Road (2018), “PR18 Econometric top-down benchmarking of Network Rail A report”, juillet, p. 43; Sumicsid (2016), “Project E2GAS Benchmarking European Gas Transmission System Operators”, juin, pp. 44–45.



- Comme (presque) toutes les approches d'estimation et les tests statistiques donnent de meilleurs résultats sur des échantillons plus importants, Schwartz and Co devrait s'assurer que sa base de données est suffisamment importante pour permettre une estimation robuste de son modèle.

3.2.2.2. Réponse aux remarques d'Oxera

Rejet de l'utilisation d'une méthode SFA dans le cas d'échantillon de petite taille

Oxera indique que « la méthode SFA a été mise en œuvre sur des bases de données plus petites que celle proposée par Schwartz and Co, y compris dans des exercices de benchmarking international ». Les différents exemples présentés par Oxera montrent néanmoins que l'utilisation d'une méthode SFA sur un « petit échantillon » de données n'est pas pertinente et viennent ainsi contredire l'argumentaire d'Oxera :

- The Office of Rail and Road (2013), “PR13 Efficiency Benchmarkings of Network Rail using LICB“, août, p. 6 : l'échantillon analysé comporte 14 points et 21 modèles sont testés, dont deux méthodes SFA (PSFA et Pitt and Lee). A la suite des différentes analyses, les modèles SFA n'ont pas été retenus pour la mesure d'efficacité.

Figure 1. Modèles d'analyse testés dans le rapport « PR13 Efficiency Benchmarkings of Network Rail using LICB »

		Decomposition of unexplained variation			Components of inefficiency				Estimation technique
		Noise	Unobserved Heterogeneity	Inefficiency	Time Invariant	Time-varying			
						Independent	Trend		
						Linear	Quadratic		
COLS	COLS			✓		✓			OLS
	PSFA	✓		✓		✓			MLE
Fixed Effects COLS	COLSFE		✓	✓		✓			OLS
Fixed Effects as Inefficiency	FEI	✓		✓	✓				OLS
Pitt and Lee	PL	✓		✓	✓				MLE
Pit and Lee + Mundlak	PLMUN	✓	✓	✓	✓				MLE
Cuesta Quadratic	CUESTAQ	✓		✓				✓	MLE
Cuesta Linear + Quadratic Network Rail	CUESTAN	✓		✓				✓	MLE
Cuesta Linear	CUESTAL	✓		✓		✓			MLE
Cuesta Quadratic + Mundlak	CUESTAMUN	✓	✓	✓	✓			✓	MLE
CSS Fixed Effects Quadratic	CSSFQ	✓		✓	✓			✓	OLS
CSS Fixed Effects Linear	CSSFEL	✓		✓		✓			OLS
CSS Random Effects GLS	CSSRE	✓		✓				✓	GLS
CSS Random Effects OLS	CSSREO	✓		✓				✓	OLS
CSS Random Effects + Mundlak	CSSREMUN	✓	✓	✓				✓	OLS
Persistent/ Residual	PERRES	✓		✓	✓	✓			MLE
Persistent/ Residual + Mundlak	PERRESMun	✓	✓	✓	✓	✓			MLE
“True” Random Effects	TRERE	✓	✓	✓		✓			MLE
“True” Random Effects + Mundlak	TRREMUN	✓	✓	✓		✓			MLE
4 Components	COMP	✓	✓	✓	✓	✓			MLE
4 Components Mundlak	COMPmun	✓	✓	✓	✓	✓			MLE

Source : The Office of Rail and Road (2013), “PR13 Efficiency Benchmarkings of Network Rail using LICB“, août, p. 38

- The Office of Rail and Road (2018), “PR18 Econometric top-down benchmarking of Network Rail A report“, juillet, p. 43 : pour l'analyse des routes, l'échantillon analysé comporte 8 points de mesure par période temporelle et 5 périodes temporelles. La méthodologie a été conduite avec une méthode COLS et une méthode SFA. Néanmoins, après diverses analyses, la méthode SFA a été rejetée et le modèle de mesure d'efficacité



retenu est uniquement une méthode COLS : « *Given the small size of our dataset as well as difficulties in getting reliable data on some potential cost drivers, we preferred to base our final analysis on the results from the simplest but most widely used model specification i.e. the COLS model* ».

- Sumicsid (2016), “Project E2GAS Benchmarking European Gas Transmission System Operators“, juin, pp. 32 : L'échantillon analysé comporte 22 points de mesure. La mesure d'efficacité a été réalisée avec une méthode DEA. La méthode SFA a été utilisée comme une simple méthode de validation des résultats compte tenu de la faible taille de l'échantillon : « *In a study of European gas TSOs, the number of observations is too small for a fullscale application of SFA as main instrument. We have therefore used DEA as our base estimation approach. As part of the robust check, we have additionally estimated the same model using SFA. Part of the motivation for this is also to discipline the modelling effort.* ». Nous relevons également l'approche incohérente de cette étude qui mentionne que l'utilisation d'une méthode SFA n'est pas pertinente, car la base de données est trop petite, mais qui conduit néanmoins un calcul avec la méthode SFA comme simple méthode de validation des résultats.

Voir également notre réponse sur la méthode SFA au paragraphe 2.2.2.

3.2.2.3. Précisions apportées dans le rapport final

Le rapport final prendra en compte les éléments de réponses formulés par Schwartz and Co.

3.2.3. Rejet de la méthode DEA

3.2.3.1. Remarques d'Oxera

La méthode DEA a été utilisée par un certain nombre de régulateurs et d'universitaires pour évaluer l'efficacité des entreprises et des prestataires de services publics (par exemple dans les domaines de la santé et de l'éducation). La méthode DEA utilise des modèles de programmation linéaire pour estimer le niveau minimum de TOTEX nécessaire à un GRD pour atteindre ses objectifs (c'est-à-dire les facteurs de coûts pris en compte dans le modèle), compte tenu des données relatives aux coûts et aux outputs de tous les GRD. En tant que méthode non paramétrique, elle présente l'avantage unique par rapport aux autres méthodes examinées par Schwartz and Co, de ne pas exiger d'hypothèses ex ante strictes concernant la forme fonctionnelle de la fonction de coûts (c'est-à-dire la relation entre les coûts et les drivers de coûts). Il suffit d'identifier les intrants (inputs) (par exemple TOTEX ou OPEX et CAPEX séparément) et les sortants (outputs) pertinents, et l'hypothèse de rendements d'échelle, qui peuvent être validées statistiquement et / ou opérationnellement.

Selon Schwartz and Co, il est essentiel que la « qualité statistique » du modèle de benchmarking soit non seulement assurée, mais aussi quantifiable. Par conséquent, les méthodes non paramétriques sont écartées. Cet argument est réitéré dans la réponse de Schwartz and Co aux questions de l'ORES.



Nous considérons que l'argument selon lequel les méthodes non paramétriques puissent être écartées, car il serait impossible d'assurer et de quantifier la « qualité du modèle » est inapproprié pour deux raisons.

Premièrement, les considérations statistiques (telles que la validation des hypothèses) sont réalisables et prises en compte dans les applications de la méthode DEA dans un contexte réglementaire, et nous renvoyons Schwartz and Co aux travaux fondateurs de Banker (1993, 1996) et de Simar et Wilson (1998, 2000) pour les inciter à reconsidérer leur points de vue à ce sujet. Deuxièmement, étant donné les limites des tests statistiques dans un contexte de mesure d'efficacité et dans les bases de données de petite et moyenne taille avec lesquelles les régulateurs ont tendance à travailler, nous ne comprenons pas la raison pour laquelle Schwarz and Co a accordé un poids aussi important à l'analyse statistique lors de la détermination des approches d'estimation les plus pertinentes.

Nous soulignons que la qualité statistique n'est qu'un des éléments à prendre en compte lors de l'élaboration d'un modèle. En outre, un large éventail de techniques a été mis au point pour évaluer la qualité des modèles utilisant des méthodes non paramétriques (par exemple, la méthode DEA). Ainsi, la méthode DEA (et d'autres méthodes non paramétriques) ne peuvent être ignorées sur cette base. En effet, c'est la seule méthode de la liste des méthodes examinées par Schwartz and Co qui peut fournir une vue alternative aux méthodes paramétriques et elle a été considérée explicitement comme la méthode de base ou comme une méthode de validation dans presque tous les pays où Schwartz and Co a l'intention d'échantillonner des données.

Le développement du modèle peut être effectué en utilisant des méthodes d'efficacité paramétriques telles que la méthode SFA et la spécification du modèle identifié peut être soumise à une estimation de l'efficacité en utilisant la DEA. Des tests de robustesse supplémentaires peuvent être entrepris dans l'approche DEA en considérant plusieurs spécifications de modèles en termes de drivers de coûts pour identifier les GRD qui sont particulièrement sensibles à certains drivers de coûts (ou certains substituts de drivers de coûts) même si ces drivers de coûts ne sont pas significatifs au niveau du secteur (c'est-à-dire non identifiés à l'aide des approches OLS).

Comme indiqué, la méthode DEA a été largement utilisée dans la régulation à travers l'Europe, y compris dans les exercices de benchmarking internationaux. Par exemple, la méthode DEA est utilisée pour définir des objectifs d'efficacité pour les réseaux d'énergie en Allemagne et en Autriche, et nous apprenons que le régulateur néerlandais (l'Autorité pour les consommateurs et les marchés, ACM) a l'intention d'utiliser les résultats de la méthode DEA pour fixer des objectifs d'efficacité aux GRT néerlandais de gaz et d'électricité (comme dans la précédente période réglementaire). Nous ne voyons pas d'élément exceptionnel dans l'étude de Schwartz and Co qui justifierait d'exclure ex ante l'utilisation de la méthode DEA.

Recommandation d'Oxera :

- Schwartz and Co devrait considérer les résultats de la modélisation DEA comme faisant partie de la base de preuves pour informer les niveaux d'efficacité estimés des GRD



- Si la méthode DEA doit être entièrement exclue de la base de preuves, Schwartz and Co devrait fournir une justification solide de cette décision

Les deux approches (SFA et DEA) ont des précédents réglementaire importants, y compris dans les juridictions analysées dans le rapport de Schwartz and Co. Par contre, la méthode advanced COLS proposée par Schwartz and Co combine les inconvénients de ces deux méthodes en ce qu'elle est déterministe et présuppose des hypothèses paramétriques fortes.

3.2.3.2. Réponse aux remarques d'Oxera

Absence de tests statistiques pertinents permettant de mesurer la qualité du modèle DEA

Nous partageons le fait qu'un certain nombre de tests peuvent être effectués sur une méthode DEA, dont certains ont été présentés par Schwartz and Co dans son rapport intermédiaire au paragraphe 3.3.1.6 (« Pour tester la robustesse d'un modèle DEA, des analyses de sensibilités sur les résultats peuvent être conduites, sans garantie sur la valeur statistique de cette approche. Une première manière de tester la sensibilité des résultats d'un modèle DEA consiste à ajouter ou à retirer des organisations dans le modèle DEA. Une autre manière de tester la sensibilité des résultats d'un modèle DEA consiste à modifier les valeurs des outputs et des inputs. Cette approche permet de déterminer la variation maximale, dans les variables d'un GRD donné, afin que celui-ci conserve sa mesure d'efficacité »).

Cependant, nous maintenons fermement qu'il n'existe pas de test statistique permettant d'évaluer la validité de la mesure d'efficacité obtenue par une méthode DEA.

Nous réfutons les arguments apportés par Oxera qui vise à démontrer que la validité statistique d'un modèle DEA peut être établie. En effet, comme présenté par Oxera, les travaux de Banker sont utilisés pour tester les hypothèses dans le cadre de la détermination des économies d'échelle et de l'identification des outliers (valeurs aberrantes). Ces tests ne permettent en aucun cas d'évaluer statistiquement la qualité du modèle DEA.

L'absence de tests statistique permettant d'évaluer la validité de la mesure d'efficacité obtenue par une méthode DEA est d'ailleurs largement étayée dans la littérature :

- « *The user of data envelopment analysis (DEA) has little available guidance on model quality. The technique offers none of the misspecification tests or goodness of fit statistics developed for parametric statistical methods.*», On the quality of the data envelopment analysis model, Pedraja-Chaparro, Salinas-Jiménez et Smith, 1999
- « *Since DEA is a non-parametric approach the impact the respective input factors on the efficiency can not be determined. Furthermore, DEA does not regard possible noise in the data and outliers can have a large effect on the outcomes* » (Efficiency Analysis of German Electricity Distribution Utilities – Non-Parametric and a Parametric Test, Christian von Hirschhausen and Andreas Kappeler, 2006, p7-8)



- « *There are no estimates or significance tests of the parameters of the production function, a potentially serious problem if results are sensitive to the specification of inputs and outputs* » (Data envelopment analysis and its application to the measurement of efficiency in higher education, Jill Johnes , 2005, p5)
- « *Resultant weights associated with the input variables have no economic interpretation. They simply define the relative contribution of reference points on the frontier to the estimation of efficient or capacity output for the point under examination* » (Measuring and assessing capacity in fisheries, Appendix D, S. Pascoa, J.E. Kirkley, D.Gréboval, C.J. Morrison-Paul, 2003)

De plus, les régulateurs utilisant la méthode DEA que nous avons interrogés au cours de l'étude nous ont confirmé au cours d'entretiens qu'aucun test statistique ne permettait de mesurer la qualité d'un modèle DEA.

Enfin, il est intéressant de noter que le directeur général du GRD Réseaux d'Énergies de Wavre avait indiqué lors de réunion de lancement de la mission avec les GRD du 24 avril 2020 (compte-rendu transmis aux GRD le 30 avril 2020) avoir eu une mauvaise expérience avec le benchmarking basé sur la méthode DEA mis en œuvre à l'époque, en 2007-par la CREG. Aucun autre GRD n'avait alors contredit ce constat.

3.2.3.3. Précisions apportées dans le rapport final

Le rapport final prendra en compte les éléments de réponses formulés par Schwartz and Co.

3.2.4. Estimation du benchmark et des niveaux d'efficacité appropriés

3.2.4.1. Remarques d'Oxera

Les approches OLS, telles qu'envisagées par Schwartz and Co, ne peuvent pas faire la distinction entre les erreurs statistiques, l'inefficacité et l'hétérogénéité des GRD (c'est-à-dire les caractéristiques propres aux gestionnaires de réseaux) qui ne sont pas prises en compte dans la spécification du modèle. Ainsi, en raison de l'incertitude liée à l'estimation des coûts efficaces à l'aide des approches OLS, si une correction est nécessaire, les régulateurs ne déplacent généralement pas la courbe de régression vers l'entreprise la plus efficace, mais vers un autre point de référence (benchmark).

Ceci est destiné à tenir compte de l'impact du bruit statistique sur les scores d'efficacité estimés des entreprises. En effet, si les données sont de qualité relativement faible ou si le modèle n'a pas pris en compte de manière adéquate, toutes les caractéristiques opérationnelles pertinentes (ce qui est probable dans un exercice de benchmarking international moins développé), le score d'efficacité estimé d'une entreprise peut être entièrement tributaire du bruit statistique. C'est pourquoi l'efficacité estimée est parfois appelée « écart de coûts » ou « écart résiduel », plutôt que « inefficacité ».



Il est important de noter que la sélection d'un point de référence (benchmark) n'est pas une simple question de discrétion réglementaire - elle doit être éclairée par la confiance du régulateur dans le(s) modèle(s) qui, à son tour, doit être éclairée par des mesures (quantifiables) de la qualité du modèle et de l'incertitude statistique. Nous soulignons que les régulateurs et les tribunaux ont estimé que certaines corrections apportées à la courbe de coûts moyens sont trop stricts compte tenu de la qualité des modèles économétriques, et ont recommandé le maintien du point de référence (benchmark) moyen.

En outre, il a été démontré que de tels ajustements ad hoc peuvent sur- ou sous-compenser des entreprises particulières, même si, par coïncidence, l'ajustement peut être globalement correct dans l'ensemble du secteur.

Recommandation d'Oxera :

- Schwartz and Co devrait envisager des mesures statistiques de l'incertitude/précision et une validation robuste des résultats afin d'éclairer le point de référence (benchmark) approprié pour les GRD wallons concernés. Les méthodes statistiques pourraient inclure une analyse de sensibilité : (i) des simulations de Monte Carlo ; (ii) une analyse d'intervalles de confiance ; et (iii) la méthode SFA (si elle n'est pas déjà considérée comme l'approche d'estimation primaire).

3.2.4.2. Réponse aux remarques d'Oxera

Décalage de la frontière d'efficacité

Oxera évoque le fait que la distance à la frontière d'efficacité positionnée dans le cadre d'une méthode déterministe de type COLS (positionnement sur le GRD le plus efficace) ne peut être attribuée intégralement à de l'inefficacité, et intègre également du bruit de sources multiples.

Nous partageons ce point et l'avons largement documenté ce point dans notre rapport intermédiaire.

Compte tenu de cette incertitude, nous recommandons, au paragraphe 5.1 de notre rapport intermédiaire, de ne pas utiliser une méthode COLS (dont la courbe de régression est décalée sur le GRD le plus efficace), mais que le modèle optimal obtenu par la méthode advanced COLS soit décalé, soit au premier décile, soit au second décile, soit au premier quartile en fonction de l'objectif que la CWaPE souhaite fixer.

Cette approche nous semble préférable à un décalage sur base de l'erreur standard de la régression OLS, qui demeure malgré tout envisageable. Ce choix peut être réalisé comme un objectif de régulation plus ou moins fort sur les GRD wallons et ainsi être décidé après l'obtention des résultats du modèle. Nous réfutons que le choix du décalage de la frontière d'efficacité devrait être étayé



par des mesures statistiques. Cette pratique n'est d'ailleurs pas observée chez le régulateur autrichien qui utilise une méthode MOLS⁷.

Nous maintenons également que l'utilisation d'une méthode SFA pour tenter de décomposer cette distance entre une part de bruit et une part d'inefficience n'est pas faisable dans le cas présent et ne présente aucune garantie d'un résultat meilleur qu'avec la méthode que nous préconisons (voir nos réponses sur la méthode SFA aux paragraphes 2.2.2 et 3.2.2.2).

Analyse de sensibilité

Comme indiqué en réponse aux remarques des GRD concernant la base de coût au paragraphe 4.1.3, nous prévoyons de mener des analyses de sensibilités sur les données de la base de coûts. Ces analyses de sensibilité sur la base de coûts seront traduites en analyses de sensibilité sur les mesures d'efficacité des différents GRD et permettront d'adopter une approche conservatrice vis-à-vis des GRD wallons en cas d'incertitudes.

Point de référence moyen

Oxera affirme que « les régulateurs et les tribunaux ont estimé que certaines corrections apportées à la courbe de coûts moyens sont trop stricts comptes tenus de la qualité des modèles économétriques, et ont recommandé le maintien du point de référence (benchmark) moyen ». Ces propos ne sont néanmoins étayés que par un seul exemple dans le domaine de l'eau. Concernant les régulateurs européens dans le domaine de l'énergie, pour les 10 régulateurs que nous avons analysés (voir détail au paragraphe 4.3.2), aucun régulateur n'a utilisé la courbe des coûts moyens comme frontière d'efficacité.

3.2.4.3. Précisions apportées dans le rapport final

Des compléments seront apportés au rapport final sur la base des réponses de Schwartz and Co.

⁷ Electricity Distribution System Operators - 1 January 2019 - 31 December 2023 - Regulatory Regime for the Fourth Regulatory Period, E-Control



4. Rapport de consultation aux remarques formulées par ORES et RESA, AIEG, AIESH et REW traitant de points identiques

4.1. Comparabilité des coûts dans le cadre d'un benchmark international

4.1.1. Remarques d'Oxera

La Wallonie compte actuellement deux GRD gaz et cinq GRD électricité. Cela rend l'analyse comparative au sein de la région wallonne difficile, même si elle peut encore être faisable. Par exemple, Ofgem compare les trois GRT d'électricité britanniques les uns aux autres. Les régulateurs ont également envisagé des méthodes alternatives telles que le benchmarking à un niveau plus granulaire et au fil du temps, et des approches bottom-up telles que l'analyse d'experts, la modélisation de référence et le processus de benchmarking lorsque les comparateurs nationaux sont limités. Un benchmarking international a également été envisagé à cet égard.

Schwartz and Co affirme utiliser des données d'Autriche, de France, d'Allemagne, des Pays-Bas, du Royaume-Uni et / ou d'autres régions belges pour estimer son modèle. Bien que la nécessité d'améliorer l'ensemble de données nationales soit connue, l'introduction de données provenant d'autres juridictions crée des complications supplémentaires. Par exemple, nous notons ce qui suit :

- Les GRD peuvent effectuer différentes tâches dans différentes juridictions qui ont une incidence sur la relation entre les coûts et les drivers de coûts. Par exemple Schwartz and Co a reconnu que la plupart des GRD européens n'ont pas à servir des clients protégés (une exigence légale de la régulation wallonne) et a proposé d'ajuster les données de dépenses des GRD wallons en conséquence.
- De même, certains GRD peuvent posséder ou appartenir aux GRT connectés à leur réseau (ou faire partie d'un groupe de propriété plus large) et bénéficier d'économies d'échelle et de gamme, tandis que d'autres ne le pourraient pas.
- Les coûts « contrôlables » peuvent avoir des définitions différentes selon les régions. Par exemple, en Allemagne, les coûts de main d'œuvre non salariaux (tels que les pensions et les cotisations à la sécurité sociale) sont considérés comme non contrôlables.
- Les paramètres qui alimentent la construction des dépenses en capital (CAPEX), tels que le taux de dépréciation, la réévaluation des actifs, le taux d'amortissement et la durée de vie utile des actifs, peuvent différer d'une juridiction à l'autre.
- Les GRD n'opèrent pas tous sur le même marché et devront donc probablement faire face à des prix des intrants (input) différents qui doivent être pris en compte dans la modélisation. Schwartz and Co a reconnu la nécessité de tenir compte des différences de coûts de main-d'œuvre entre les GRD, mais il ne traite pas des différences dans les prix des autres facteurs de production.
- Les directives en matière de reporting sur les coûts et les drivers de coûts peuvent différer d'une juridiction à l'autre, et il est essentiel que les données qui alimentent le modèle de



benchmarking soient présentées de manière cohérente et que des vérifications croisées robustes soient prises en compte.

Cette hétérogénéité dans l'échantillon de GRD doit être prise en compte dans le processus de benchmarking, par le biais d'exclusions et de normalisations de données (par exemple l'ajustement des données sur les coûts) et la procédure de développement du modèle (par exemple, en incluant des drivers de coûts appropriés pour capturer cette hétérogénéité et isoler les anomalies). Non seulement il s'agit d'une exigence générale pour développer un modèle robuste, mais c'est aussi une exigence du droit wallon que les comparaisons de coûts soient faites à l'aide de «données homogènes, transparentes et fiables».

En effet, Schwartz and Co a reconnu certaines des limites du benchmarking international dans son rapport intermédiaire, et a fourni une description de générale de la manière dont il entend les traiter. Cependant, en réponse aux questions concernant des ajustements spécifiques des données internationales, Schwartz and Co a déclaré qu'il « effectuera les retraitements nécessaires pour la mise en œuvre du modèle » sans fournir de détails supplémentaires. Nous comprenons que tous les ajustements apportés aux données seront décrits et consultés avant le rapport final. À cet égard, comme l'exercice de benchmarking international des GRD est moins avancé que d'autres initiatives similaires, il peut être prudent de ne pas attribuer automatiquement l'écart estimé entre les dépenses d'un GRD et celles prévues par le modèle à l'inefficacité, mais plutôt de l'appeler « écart de coûts » ou « écart résiduel » pour un examen plus approfondi. Nous nous attendons à ce que la cohérence des directives en matière de reporting et la qualité globale des données s'améliorent avec le temps.

Recommandation d'Oxera :

- Schwartz and Co devrait prendre en compte toutes les complications d'un benchmarking international et documenter pleinement les étapes à suivre pour garantir que les GRD soient évalués sur une base comparable.
- Comme Schwartz and Co l'a reconnu, certains des ajustements des données nécessitent « des estimations ou approximations qui peuvent biaiser les résultats ». Par conséquent, il est essentiel que Schwartz and Co teste la sensibilité de son analyse aux hypothèses qui alimentent la construction des données et évalue une «marge d'erreurs» appropriée pour en tenir compte.
- La taille de l'échantillon proposée est relativement petite (voir la section 4.3), et une erreur de données dans n'importe quel GRD peut avoir un impact significatif sur le modèle global. Cette question est particulièrement pertinente, car COLS est une méthode déterministe d'estimation de l'efficacité. Par conséquent, Schwartz and Co devrait s'assurer que les résultats ne sont pas sensibles aux petits changements des données et du modèle.

4.1.2. Remarques de RESA, AIEG, AIESH et REW

Comme le stipule l'art. 4 §2 15° du décret tarifaire du 19 janvier 2017 : « Toute comparaison avec d'autres gestionnaires de réseau est réalisée entre des sociétés ayant des activités similaires et opérant dans des circonstances analogues ». Il convient dès lors de standardiser la base de coût pour tenir compte e.a. :



- d'un périmètre d'activité potentiellement différent ;
- d'un périmètre technique potentiellement différent ;
- d'obligations spécifiques imposées par le régulateur et/ou les autorités ;
- d'une allocation des coûts entre les différentes activités pour les GRD qui en exercent plusieurs.

A ce stade, nous n'avons cependant pas de vue sur la méthode qui sera utilisée par le consultant pour effectuer ces retraitements ou ajustements. Il s'agira d'un exercice difficile, voire impossible dans la mesure où comme l'indique le consultant « L'accès aux données de coût des GRD étrangers est complexifié et limité à une faible granularité (sauf si un régulateur accepte de partager ses données) ». Nous souhaiterions connaître les régulateurs disposés à partager ses données.

Toute comparaison devra également s'assurer que le périmètre technique des GRD est le même. La limite technique entre Transport et Distribution varie considérablement en Europe. Certains GRD opèrent des réseaux haute tension (ex. GRD anglais et hollandais) tandis que d'autres se limitent aux réseaux basse et moyenne tension. Les premiers seront favorisés par rapport aux derniers dans un exercice de benchmarking. A ce stade, nous ignorons comment Schwartz tiendra compte de cette contrainte. Le rapport intermédiaire est muet sur ce point.

Dans le périmètre d'activités des GRD belges, une mission importante impactant considérablement les coûts contrôlables, concerne l'exécution des missions de services publics ou obligations de services publics (OSP) imposées par les autorités (régionales). Compte tenu que les GRD étrangers, ont peu ou pas d'OSP, il convient – comme le propose le consultant à juste titre - d'exclure les OSP des Opex contrôlables. Il peut également y avoir des obligations spécifiques imposées par le régulateur en termes par exemple d'enfouissement des lignes électriques. Il conviendra également d'en tenir compte.

En Europe, certains GRD sont également actifs dans des activités concurrentielles (sous certaines conditions) et/ou exercent d'autres activités que la gestion des réseaux de distribution d'électricité et de gaz (par ex. des réseaux de chaleur). C'est le cas en particulier des Stadtwerke en Allemagne et en Autriche.

Enfin, se pose également la question de la répartition des OPEX entre OPEX contrôlables et OPEX non contrôlables. Une même dépense peut être considérée par un régulateur comme une dépense non contrôlable et par un autre régulateur comme une dépense contrôlable. Comparer des bases de coûts contrôlables avec une définition des coûts contrôlables différente d'un régulateur à l'autre induira à nouveau un biais. Nous ignorons comment le consultant compte traiter concrètement ce problème.

En conclusion, la standardisation des dépenses nécessite un relevé clair des activités des différents GRD dans les différents pays considérés et d'une méthode d'allocation des coûts claire et



transparente. L'étude ne détaille pas la procédure qui sera utilisée pour standardiser la partie OPEX de la base de coût.

4.1.3. Réponse aux remarques de RESA, AIEG, AIESH et REW et d'Oxera

Les GRD évoquent dans leurs commentaires une préoccupation que nous partageons et que nous avons déjà évoquée au paragraphe 5.2 de notre rapport intermédiaire, comme suit :

«

- Le périmètre d'activité des GRD peut varier entre la Belgique et les pays étrangers intégrés au benchmark (cf. l'exemple des OSP cité précédemment). Il est ainsi nécessaire d'effectuer des retraitements dans les bases de coûts des différents GRD afin de les comparer à périmètre constant. Il existe ainsi un risque que les retraitements à effectuer nécessitent des estimations ou approximations qui peuvent biaiser les résultats. Il convient donc d'être très vigilant dans la mise en œuvre de ces retraitements pour éviter de tels biais. Concernant la partie CAPEX de la base de coût, il convient également d'éliminer les différences induites par les financements externes des actifs.
- Les contraintes exogènes portant sur les GRD peuvent différer d'un pays à l'autre avec des conséquences sur les coûts à la hausse ou à la baisse. Dans un benchmark international il convient donc d'identifier autant que possible ce type de contraintes et d'en tenir compte soit dans la constitution de la base de coût (par retraitement) soit de manière préférentielle directement dans les variables explicatives. Un exemple concret est le coût moyen du travail qui diffère d'un pays à l'autre. Ce point se traite très bien à travers la méthode advanced COLS en intégrant aux variables explicatives le coût moyen du travail. »

Comme présenté dans le rapport intermédiaire (au paragraphe 5.2), nous veillerons à ce que les bases de coûts, qui sont composées d'OPEX et d'amortissements couvrent le même périmètre de coûts et soient retraitées des effets exogènes pouvant biaiser la comparaison, dans les limites de ce qui est matériellement faisable avec les données recueillies. Pour les coûts dont les retraitements présenteraient trop d'incertitudes, nous conduirons des analyses de sensibilité et adopterons une approche conservatrice vis-à-vis des GRD wallons, qui aura pour conséquence de maximiser l'efficacité mesurée des GRD wallons plutôt que de la minimiser. Les retraitements effectués ainsi que les analyses de sensibilité seront documentés.

4.1.1. Précisions apportées dans le rapport final

Le rapport final sera complété avec les éléments de méthodologie supplémentaires que nous avons détaillés dans nos réponses.



4.2. Comparabilité des CAPEX

4.2.1. Remarques d'Oxera

Les CAPEX sont généralement « atypiques » et irréguliers d'une entreprise à l'autre et dans le temps ; un GRD peut investir dans un segment important du réseau au cours d'une année donnée, ce qui sera enregistré comme une forte augmentation des dépenses. Pour éviter que les estimations d'efficience ne soient dictées par la position d'un GRD dans le cycle d'investissement (ou son « âge »), les régulateurs envisagent souvent des méthodes de « lissage » ou de normalisation des CAPEX. Schwartz and Co discute de deux de ces méthodes.

- CAPEX comptable. Ceci est basé sur l'approche du « coût historique amorti » (CHA) pour la mesure du CAPEX, où la mesure du CAPEX diminue au cours de la durée de vie d'un actif.
- CAPEX standardisé. Ceci est basé sur l'approche de l'annuité pour la mesure du CAPEX, où la mesure du CAPEX est constante sur la durée de vie de l'actif.

Schwartz and Co affirme que le CAPEX standardisé n'est pas approprié, car il favorise les GRD avec des investissements récents et ne correspond pas à la « réalité des coûts », tandis que l'utilisation du CAPEX comptable permet aux GRD de choisir la meilleure stratégie pour combiner OPEX et CAPEX. Cependant, Schwartz and Co ne cite aucune source pour étayer cela et n'explique pas en quoi cette conclusion est appropriée. En effet, ORES a demandé des sources et des précisions supplémentaires sur l'argument de Schwartz and Co, mais Schwartz and Co a déclaré que l'argument « ne demande à priori pas de documentation complémentaire ».

Il est à noter que la conclusion de Schwartz and Co est unilatérale, étant donné que ce sujet fait toujours partie du débat sur la régulation. En effet, certains des inconvénients mentionnés pour l'approche de l'annuité sont présentés comme des avantages dans d'autres exercices de benchmarking, y compris d'autres exercices de benchmarking international. Par exemple, dans TCB18, Sumicsid déclare que : « [...] une annuité réelle doit être utilisée, car l'application d'amortissements nominaux (même standardisés) introduirait immédiatement un biais en faveur des investissements tardifs. »

Cela remet en question la validité de la conclusion de Schwartz and Co.

Compte tenu des avantages et des inconvénients des deux mesures et du précédent réglementaire selon lequel les deux sont généralement prises en compte dans le benchmarking, il n'est pas évident que l'approche coût historique amorti (CHA) soit clairement la méthode la plus appropriée dans ce contexte.

Recommandation d'Oxera :

- Schwartz and Co devrait prendre en compte et corriger les différences de niveaux de prix entre les GRD sur les CAPEX.
- Schwartz and Co devrait envisager d'utiliser de multiples approches à la mesure des CAPEX dans le processus de benchmarking (y compris l'approche CHA et annuité),



conformément au précédent réglementaire. Si Schwartz and Co estime qu'une approche est meilleure que l'autre, il devrait clairement expliquer pourquoi c'est le cas, fournir des preuves empiriques et citer ses sources, le cas échéant.

- Les paramètres qui alimentent l'estimation des CAPEX (tels que la durée de vie utile des actifs, le rendement du capital et la classification des actifs) doivent être suffisamment motivés et appliqués de manière cohérente à travers les GRD. Cela est particulièrement important dans un contexte de benchmarking international, où ces paramètres sont susceptibles de varier d'une juridiction à l'autre.

4.2.2. Remarques de RESA, AIEG, AIESH et REW

Le consultant propose d'appliquer sa méthode « advanced COLS », à une base de coût constituée des TOTEX contrôlables hors OSP, avec prise en compte des CAPEX sur base comptable.

Le consultant souligne à juste titre que « La base de coûts doit être strictement identique entre les différents GRD analysés afin de garantir la pertinence des résultats ». Cette exigence que nous partageons entièrement doit non seulement être énoncée mais également appliquée. Un élément essentiel repris dans le cahier des charges de la CWaPE concernant l'option du Lot 2 est que la « réalisation de l'option inclut la standardisation des données utilisées dans les modèles ». Le consultant semble avoir omis cette exigence, d'ailleurs non reprise à la page 7 de son rapport lorsqu'il énumère les objectifs du Lot 2 (voir aussi annexe 1).

La question de la standardisation concerne l'ensemble de la base de coût et il est important d'avoir une base de coût standardisée. La mesure d'inefficience est calculée comme la différence entre la meilleure pratique estimée 'TOTEX best practice' et le TOTEX réel. Si cette dernière mesure est biaisée, à la hausse ou à la baisse, la mesure d'inefficience est automatiquement biaisée.

Il faut donc absolument garantir que la base de coût sera comparable d'un GRD à l'autre.

L'utilisation des CAPEX sur base comptable plutôt qu'une base standardisée est problématique à plusieurs points de vue.

- L'utilisation d'une valeur comptable pour le CAPEX plutôt qu'une valeur standardisée, implique que le coût du capital n'est pas directement comparable entre les GRD en raison, de règles d'amortissement potentiellement différentes d'un GRD ou d'un pays à l'autre, d'un financement du capital différent d'un GRD à l'autre, et (surtout) de la non prise en compte de l'évolution des prix, le capital étant comptabilisé à la date d'achat, sans tenir compte de l'inflation. De ce fait, un GRD dont le capital est relativement plus ancien aura un CAPEX comptable moins élevé qu'un GRD ayant un capital plus récent. Avec comme conséquence un biais d'efficience dans le benchmarking en faveur des GRD dont le capital est plus ancien. Ceci étant explicitement reconnu dans l'étude du consultant (page 60, §2). La non-standardisation du coût du capital introduit toute une série de biais dans la mesure d'efficience, en fonction de l'âge du capital, de son mode de financement et de la comptabilisation de celui-ci, et des règles d'amortissement.
- Le problème est particulièrement important du fait de l'importance relative des CAPEX dans les TOTEX.



- Tous les pays mentionnés dans l'étude, à l'exception de la France qui a réalisé le benchmark sur base des OPEX uniquement, utilisent un CAPEX standardisé, l'Autriche et la Finlande utilisant uniquement une valeur standardisée.

Nous plaillons en faveur d'une mesure standardisée plutôt qu'une mesure comptable pour le coût du capital. L'utilisation d'une valeur non-standardisée est source de biais dans le calcul des mesures d'efficacité et fragilise la démarche. Nous ne voyons également pas pourquoi il conviendrait d'éliminer les différences induites par les financements externes des actifs, comme semble le suggérer Schwartz ?

4.2.3. Réponse aux remarques de RESA, AIEG, AIESH et REW et d'Oxera

Nous rappelons tout d'abord que la part « CAPEX » de la base de coût qui est comparée à travers l'exercice de benchmark est constituée des charges nettes liées aux immobilisations, au sens de la méthodologie tarifaire 2019-2023 de la CWaPE, c'est-à-dire les charges d'amortissement, les charges de désaffectation (y inclut les charges de désaffectation des plus-values historiques et iRAB), les moins-values et plus-values sur la réalisation d'actifs régulés. Le coût du capital, sous forme de marge équitable en Wallonie ou de coût moyen pondéré du capital (WACC) dans les autres pays, n'est donc pas intégré dans les charges nettes liées aux immobilisations et n'a pas lieu d'être pris en compte dans le benchmarking. Nous rappelons également que nous avons utilisé le terme de CAPEX standardisés au sens des réglementations allemandes et autrichiennes : il s'agit donc des amortissements calculés sur la base de la valeur de remplacement pour chaque actif immobilisé traduisant une valeur de marché des actifs. A ce titre, nous notons que RESA, AIEG, AIESH et REW confondent les notions en essayant de faire correspondre la notion de CAPEX standardisés telle que nous la présentons dans notre rapport intermédiaire avec celle d'harmonisation des coûts évoqués dans le cahier des charges de l'option du lot 2, qui signifie une mise en cohérence des bases de coûts pour les rendre comparables.

Nous partageons l'affirmation de RESA, AIEG, AIESH et REW selon laquelle « il faut donc absolument garantir que la base de coût sera comparable d'un GRD à l'autre », mais ceci ne signifie pas qu'il convient de gommer les différences d'efficacité entre les GRD, en particulier au niveau des CAPEX, sous prétexte d'harmonisation.

RESA, AIEG, AIESH et REW affirme sur la base des inputs du professeur Axel Gauthier que l'utilisation des CAPEX au sens comptable ne permet pas la comparaison entre GRD surtout en raison « de la non prise en compte de l'évolution des prix, le capital étant comptabilisé à la date d'achat, sans tenir compte de l'inflation. De ce fait, un GRD dont le capital est relativement plus ancien aura un CAPEX comptable moins élevé qu'un GRD ayant un capital plus récent. Avec comme conséquence un biais d'efficacité dans le benchmarking en faveur des GRD dont le capital est plus ancien. »

Cette affirmation est basée sur une analyse tronquée, puisqu'elle ne prend pas en compte l'effet de l'âge du réseau sur les OPEX, or nous comparons une base TOTEX et non une simple base CAPEX. Un GRD avec un réseau plus âgé aura également des OPEX plus élevés pour la



maintenance et l'exploitation de son réseau qu'un GRD avec un réseau plus jeune, ce qui joue donc en sens inverse de l'effet sur les CAPEX. L'utilisation des CAPEX standardisés, si elle corrige l'effet âge au niveau des amortissements, qui peut effectivement être considéré comme exogène à l'efficacité CAPEX identique, introduit un biais que l'on ne peut pas corriger, en effaçant la différence de valeur des amortissements entre GRD sans possibilité d'ajuster les OPEX (il n'est pas possible de retraiter les OPEX d'un effet âge, et aucun régulateur s'appuyant sur une base de CAPEX standardisée ne le fait, ce qui est une faiblesse évidente de la méthode s'appuyant sur des CAPEX standardisés). L'approche sur la base de CAPEX comptable permet donc de mieux respecter le cahier des charges de la CWaPE que l'approche basée sur des CAPEX standardisés, puisqu'elle ne défavorise pas forcément un GRD avec un réseau plus récent qu'un autre, selon l'équilibre CAPEX-OPEX des GRD, tandis que l'approche sur la base de CAPEX standardisés favorise obligatoirement les GRD avec un réseau récent, étant donnée la non-possibilité de corriger les OPEX d'un effet âge.

De plus, contrairement à ce qu'indiquent RESA, AIEG, AIESH et REW, le régulateur autrichien n'utilise pas uniquement une base CAPEX standardisée pour les GRD électricité mais également une base CAPEX comptable.

Par ailleurs, l'utilisation des CAPEX comptables que nous préconisons, ne signifie pas que des retraitements de ces CAPEX ne sont pas requis pour assurer une comparaison objective entre les GRD, en conformité avec le cahier des charges du lot 2 de l'option, et comme nous l'indiquons déjà dans notre rapport intermédiaire. Assurer une comparaison objective signifie que les retraitements doivent permettre d'éliminer des différences de coûts qui ne proviennent pas de différences d'efficacité entre GRD mais de phénomènes exogènes. Concernant la base de coûts CAPEX, le recours à un benchmark international nécessite de considérer 3 effets afin d'en évaluer l'impact en matière de comparaison :

- La manière dont les financements de tiers (provenant des utilisateurs de réseaux sous forme de tarifs de type non périodique pour reprendre la terminologie en Wallonie ou d'autres tiers sous forme de subside) sont pris en compte dans le calcul des amortissements.
- Les durées d'amortissement des actifs.
- Les types de charges pris en compte (au-delà des charges d'amortissement il s'agit de vérifier la prise en compte des charges de désaffectation, les moins-values et plus-values sur la réalisation d'actifs régulés)

Pour le financement de tiers, nous identifierons dans les différents pays intégrés au benchmark sur quelle base d'investissement les amortissements des actifs immobilisés sont calculés. En effet selon les pays, les actifs peuvent être amortis sur la base des investissements bruts (c'est-à-dire sans déduction des financements des tiers ou de subsides, qui sont alors comptabilisés comme produits extratarifaires, venant donc baisser la valeur des charges nettes d'exploitation), ou sur la base des investissements nets (c'est-à-dire que les financements des tiers, subsides inclus, sont retranchés de la valeur brute des investissements, comme c'est le cas en Wallonie) ou sur une base intermédiaire (en France par exemple, pour certains GRD gaz, les amortissements sont calculés sur la base des investissements nets des subsides, mais pas des financements des tiers, qui sont comptabilisés



comme des produits extratarifaires, venant donc baisser la valeur des charges nettes d'exploitation). Nous évaluerons l'impact de ces différences entre pays au niveau des TOTEX et appliquerons le retraitement qu'il convient pour assurer la comparaison sur une base aussi neutre et objective que possible. Si un retraitement n'était pas réalisable, notamment en raison de la granularité des données disponibles, nous réaliserons une analyse de sensibilité afin de retenir une mesure d'efficacité ne défavorisant pas les GRD wallons.

Concernant les différences de durée d'amortissement entre pays, celles-ci peuvent en effet introduire des biais de comparaison au niveau des amortissements. Dans le cadre d'un benchmark international, il n'est pas faisable de retraiter intégralement les amortissements de chaque GRD pour les aligner sur les durées d'amortissement de la Région wallonne, mais il est en revanche possible d'estimer l'impact de ces différences notamment à travers des analyses de sensibilité, que nous mènerons afin de retenir une mesure d'efficacité ne défavorisant pas les GRD wallons.

4.2.4. Précisions apportées dans le rapport final

Le rapport final sera complété avec les éléments de réponses apportés par Schwartz and Co.

4.3. Approche hybride

4.3.1. Remarques d'Oxera

Comme tous les modèles économiques dits « top-down » sont des simplifications d'opérations très complexes (plus encore dans un contexte de benchmarking international), les régulateurs utilisent généralement des informations provenant de plusieurs modèles et méthodes, appliqués à différents niveaux d'agrégation, pour s'assurer que les résultats ne sont pas déterminés par des hypothèses impliquant des facteurs spécifiques, des méthodes et des niveaux d'agrégation particuliers (Schwartz and Co appellent cela l'approche «hybride»). Même en présence d'un grand nombre de comparateurs et lorsque les modèles sont correctement spécifiés et que toutes les hypothèses sous-jacentes sont satisfaites, il y aura toujours une incertitude autour des prévisions qui en résultent.

En d'autres termes, les modèles top-down ne peuvent pas fournir d'estimations ponctuelles précises des coûts efficaces (même si l'incertitude peut être évaluée empiriquement sous certaines hypothèses). Cependant, Schwartz and Co semble avoir négligé ces limites bien connues et entend plutôt à n'utiliser qu'un seul modèle et une seule méthode pour estimer les scores d'efficacité des GRD. Cette approche suppose qu'un seul modèle utilisant l'estimateur Advanced COLS peut fournir des informations utiles concernant l'efficacité des GRD, mais Schwartz and Co n'a pas fourni de preuves suffisantes pour étayer cette hypothèse forte. En effet, dans sa réponse aux questions d'ORES, Schwartz and Co va plus loin en déclarant qu'ils n'ont même pas l'intention d'utiliser des modèles alternatifs pour valider leur modèle final.

La proposition de Schwartz and Co est incompatible avec les meilleures pratiques réglementaires qui visent à prendre en compte plusieurs sources de preuves (y compris les propositions des



gestionnaires de réseau) pour améliorer la robustesse globale de l'évaluation. Par exemple, dans RIIO-GD1 et RIIO-ED1, Ofgem a utilisé un faisceau d'évidences multiples et a triangulé son point de vue avec celui des gestionnaires de réseau (dans un ratio de 75:25) pour refléter les limites de la modélisation et les imperfections des données. L'utilisation appropriée de plusieurs benchmarks / bases de données probantes peut réduire une partie de l'incertitude associée aux coûts de modélisation dans un benchmarking.

Compte tenu de l'incertitude considérable liée à l'utilisation d'un seul modèle pour estimer l'efficacité des GRD et compte tenu du précédent réglementaire convaincant pour l'utilisation de plusieurs modèles et méthodes pour fixer les paramètres réglementaires permettant d'informer les objectifs de réduction des coûts, nous recommandons à Schwartz and Co de réexaminer son évaluation de l'approche hybride.

Recommandation d'Oxera :

- Schwartz and Co ne devrait pas se limiter à un seul modèle pour évaluer l'efficacité des GRD et devrait plutôt envisager d'utiliser un certain nombre de modèles robustes mis en concurrence. Schwartz and Co (et la CWaPE) peuvent alors utiliser les informations de chaque modèle pour fixer des objectifs d'efficacité pour les GRD.
- Si Schwartz and Co décide d'utiliser uniquement l'approche Advanced COLS, Schwartz and Co devrait présenter des preuves suffisantes qu'aucune autre méthode d'estimation ne fournit d'informations utiles concernant l'efficacité des GRD
- Au minimum, Schwartz and Co devrait présenter des éléments probants provenant d'autres modèles et approches d'estimation (par exemple, la méthode DEA, la méthode SFA, autres ensembles de drivers de coûts) pour valider (ou mettre en doute) son analyse de base.

4.3.1. Remarques de RESA, AIEG, AIESH et REW

Les expériences des régulateurs allemands et autrichiens montrent que les résultats en matière d'efficacité peuvent être très différents d'une méthode à l'autre et que dès lors il est utile de combiner les méthodes. Ce n'est pas au motif que des méthodes de calcul d'efficacité pourraient conduire à des résultats très différents, voire contradictoires (le même GRD est efficace dans un modèle et inefficace dans un autre modèle) qu'il faut exclure d'office les modèles hybrides. Cela serait d'ailleurs contraire au principe de transparence.

Il importe que le modèle d'efficacité soit robuste et cela ne peut pas être établi exclusivement sur base de tests statistiques. La possibilité de calculer plusieurs modèles, en utilisant différentes méthodes (modèle hybride), en utilisant plusieurs échantillons de données, permet de comparer les résultats entre eux et, le cas échéant, de renforcer la confiance dans le modèle sélectionné.

4.3.2. Réponse aux remarques de RESA, AIEG, AIESH et REW et d'Oxera

Pratiques réglementaires communes chez les régulateurs européens



Les pratiques réglementaires observées dans différents pays européens montrent que l'utilisation d'une méthode hybride ne constitue pas une pratique dominante : sur les 10 régulateurs analysés, seulement 3 utilisent une méthode de mesure d'efficacité hybride (voir tableaux suivants).

Par ailleurs, Oxera présente de manière erronée dans son tableau 4.1 la méthode utilisée par l'OFGEM comme une méthode hybride. L'OFGEM utilise bien une seule et même méthode qui est la méthode paramétrique COLS.

Tableau 1. Comparaison des différentes méthodes de mesure d'efficacité par les régulateurs de 10 pays européens

Pays	Régulateur	GRD	Réalisation de la mesure	Période de régulation	Méthode utilisée
Allemagne ⁸	BNetzA	Electricité et gaz	Systématique	2018 à 2022 pour le gaz et 2019 à 2023 pour l'électricité	Méthode hybride (SFA + DEA)
Autriche ⁹	E-Control	Electricité et gaz	Systématique	2018 à 2022 pour le gaz et 2019 à 2023 pour l'électricité	Méthode hybride (MOLS + DEA)
France ¹⁰	CRE	Gaz	Ponctuelle	2016 à 2020	Advanced COLS
Finlande ¹¹	Energiavirasto	Electricité	Systématique	2016 à 2019 et 2020 à 2023	StoNED
Irlande ¹²	CER	Electricité	Systématique	2016 à 2020	COLS
Royaume-Uni ¹³	Ofgem	Electricité	Systématique	2015 à 2023	COLS (Pooled OLS)
Suède ¹⁴	Energimarknadsinspektionen	Electricité	Systématique	2016 à 2019	DEA
Norvège ¹⁵	NVE	Electricité	Systématique	2013 à 2018	DEA
Danemark ¹⁶	Forsyningstilsynet	Electricité	Systématique	2020 à 2022	Méthode hybride (SFA + DEA)

⁸ Effizienzvergleich verteilernetzbetreiber gas (3. Rp)

Effizienzvergleich Verteilernetzbetreiber Strom der dritten Regulierungsperiode (EVS3)

⁹ Electricity Distribution System Operators, 1 January 2019 - 31 December 2023 Regulatory Regime for the Fourth Regulatory Period, December 2018

Regulierungssystematik für die dritte Regulierungsperiode der Gasverteilernetzbetreiber

1. Jänner 2018 - 31. Dezember 2022

¹⁰ Rapport public de l'étude réalisée par Schwartz and Co pour la Commission de Régulation de l'Energie (CRE) en France, publiée par la CRE le 18 novembre 2015, dans le cadre de la consultation publique de la CRE sur le prochain tarif d'utilisation des réseaux publics de distribution de gaz naturel de GRDF (dit « ATRD5 »)

¹¹ Regulation methods in the fourth regulatory period of 1 January 2016 – 31 December 2019 and the fifth regulatory period of 1 January 2020 – 31 December 2023, Energiavirasto, November 2015

¹² Consultancy Support for Electricity Transmission and Distribution Revenue Controls (2016-2020), Commission for Energy Regulation, Jacobs, June 2015

¹³ RIIO-ED1: Final determinations for the slowtrack electricity distribution companies Business plan expenditure assessment, Ofgem, 2014

¹⁴ Benchmarkingrapport, Afsluttende rapport, Février 2017, Benchmarkingekspertgruppen

¹⁵ National Report 2019, VNE, 2020

¹⁶ Effektiviseringspotentialet i eldistributionssktoren, Forsyningstilsynet, Avril 2020



Portugal ¹⁷	ERSE	Electricité et gaz	Systématique	2020 à 2021 pour le gaz et 2018 à 2020 pour l'électricité	DEA
------------------------	------	--------------------	--------------	---	-----

Source : Schwartz and Co

Justification scientifique d'une méthode hybride

Les méthodes hybrides n'ont pas de fondement scientifique et sont le résultat d'une négociation entre les GRD et le régulateur à l'issue d'un rapport de force. L'utilisation de plusieurs méthodes de mesure d'efficacité permet aux GRD dans ces pays de maximiser leur chance d'obtenir une mesure d'efficacité favorable, sans considération objective du degré de validité de chacune des méthodes considérées. Or toutes ces méthodes hybrides associent une méthode paramétrique (MOLS ou SFA) à une méthode DEA, dont nous avons montré le problème fondamental d'absence de vérification de la qualité du modèle obtenu sur la base de critères statistiques objectifs. Nous avons également mené une analyse sur les données de mesure d'efficacité publiées par le régulateur allemand pour l'année 2019. Il apparaît que la méthode hybride utilisée par le régulateur allemand augmente sensiblement les résultats de mesure d'efficacité obtenus en moyenne par les GRD d'électricité et de gaz (à l'exception des résultats de mesure d'efficacité obtenus en moyenne par les GRD d'électricité en utilisant la méthode SFA).

Type de mesure	GRD Electricité	GRD Gaz
Moyenne - DEA base comptable	86%	82%
Moyenne - SFA base comptable	93%	75%
Moyenne - DEA base standardisée	87%	85%
Moyenne - SFA base standardisée	94%	75%
Moyenne - Meilleur des quatre mesures	95%	96%

Source : Analyse Schwartz and Co sur la base des données publiées par la BNetzA
(ARegV31Tabelle_2020_geschwaerzt)

4.3.3. Précisions apportées dans le rapport final

Le rapport final sera complété avec les éléments de réponses apportés par Schwartz and Co.

4.4. Choix des variables explicatives des coûts

4.4.1. Remarques d'Oxera

Les modèles économiques top-down sont des représentations intrinsèquement simplistes d'une réalité complexe qui visent à saisir un large éventail de caractéristiques de fonctionnement en utilisant un nombre limité de facteurs de coûts et en simplifiant les hypothèses de modélisation. En examinant les modèles qui contrôlent les différents facteurs de coûts, il est essentiel d'examiner

¹⁷ Proveitos permitidos e ajustamentos para o ano gás 2020-2021 das empresas reguladas do setor do gás natural, ERSE, Juin 2020

Proveitos permitidos e ajustamentos para 2020 das empresas reguladas do setor elétrico, ERSE, Décembre 2019



si les entreprises jugées efficaces ne sont pas indûment récompensées ou pénalisées pour des caractéristiques opérationnelles particulières qui ne seraient pas prises en compte dans le modèle. Par conséquent, la meilleure pratique en matière de régulation consiste à envisager de multiples modèles qui utilisent différentes variables ou variables de substitution (proxies en anglais) pour saisir des caractéristiques opérationnelles spécifiques et à trianguler les résultats de celles-ci dans le but d'obtenir un résultat final non biaisé.

Sur la base des Réponse de Schwartz and Co aux questions de l'ORES, nous comprenons qu'ils sont réceptifs aux suggestions concernant le choix des drivers de coûts. À cet égard, nous notons que plusieurs drivers de coûts et substituts de drivers de coûts (proxies) pour les caractéristiques susmentionnées ne sont pas actuellement pris en considération, notamment les suivants :

- **Prix des intrants (input) non liés à la main-d'œuvre :** Les prix des intrants (inputs) associés à l'achat de service de maintenance externe, à la sous-traitance de personnel, aux services de consultants, aux fournitures de bureau et aux coûts des centres de contrôle peuvent varier d'une juridiction à l'autre, et il convient d'en tenir compte en utilisant un indice de niveau de prix approprié. Il convient de noter qu'il est couramment pratiqué dans le cadre de benchmarking international d'ajuster les données de coûts avec un indice de niveau des prix, au lieu de contrôler les indices de niveau des prix en tant que facteurs de coûts dans le modèle.
- **Exigences réglementaires :** Les GRD des différentes juridictions peuvent effectuer des tâches légèrement différentes en raison des différences dans les exigences réglementaires. Schwartz and Co a déjà identifié une de ces différences : les GRD wallons sont tenus de servir des clients protégés, ce qui est quelque peu unique à la région wallonne. Schwartz and Co a proposé de supprimer le coût associé à cette mesure. Il est essentiel que d'autres différences dans les exigences réglementaires soient étudiées et prises en compte, soit en ajustant les données sur les coûts, soit en saisissant leur impact par le biais des drivers de coûts.
- **Environnement d'exploitation :** Comme la principale tâche d'un GRD consiste à transporter l'énergie à travers une région, les caractéristiques de l'environnement de la région (telles que la densité de la population, les caractéristiques du sol, le climat et la topographie) peuvent être des facteurs de coûts significatifs. Les données relatives à ces caractéristiques peuvent être collectées à partir de sources accessibles au public.
- **Qualité de service :** Il est bien établi dans la théorie de la production que la qualité n'est pas gratuite. Les mesures de la qualité du service relative à la distribution de gaz et d'électricité comprennent les pertes de distribution, les interruptions de l'approvisionnement (leur fréquence et durée) et les mesures de l'état des actifs. Des exemples spécifiques de la manière dont la qualité du service peut être intégrée dans le processus d'évaluation des coûts sont présentés ci-dessous.

En plus de ce qui précède, certaines implémentations réglementaires ont considéré les mesures fondées sur les actifs comme des facteurs de coûts, car on estime qu'elles sont fortement corrélées aux résultats, tels que la capacité du réseau et la complexité topographique. Bien que ces mesures



présentent des limites, elles peuvent contribuer à représenter avec parcimonie un ensemble plus large de caractéristiques.

En outre, il faut faire preuve de prudence dans l'utilisation de variables telles que la pointe de charge et le volume d'énergie transportée, qui peuvent être très volatiles d'une année à l'autre et peuvent ne pas être bien corrélées avec les dépenses liées à la capacité du réseau. Nous constatons également que les niveaux de tension peuvent varier considérablement d'un pays européen à l'autre, ce qui peut avoir un impact significatif sur le coût « efficient » par unité d'énergie dans la distribution d'électricité.

Il est important de noter que l'ensemble des drivers de coûts proposés par Schwartz and Co ne tient pas compte de la qualité des services fournis et qu'il est donc incompatible avec les principes établis par la loi.

La prise en compte de la qualité de service dans le processus de benchmarking des coûts est devenue une question importante ces dernières années, car les régulateurs demandent aux entreprises de fournir davantage (par exemple par le biais d'objectifs environnementaux) à un meilleur rapport qualité-prix. Schwartz and Co dispose de plusieurs méthodes pour comptabiliser la qualité du service dans le processus d'évaluation des coûts :

- créer une valeur monétisée de la qualité du service basée sur les préférences des consommateurs et ajuster les données sur les dépenses des entreprises en conséquence
- ajouter des mesures de la qualité de service en tant qu'input dans un modèle de DEA (parallèlement aux dépenses) pour contrôler leur effet
- ajouter des mesures de la qualité du service en tant que facteurs de coûts distincts dans un modèle économétrique
- ajuster le benchmark d'efficacité de manière à ce qu'il ne soit pas influencé par des entreprises qui n'ont pas atteint les objectifs de qualité

Recommandation d'Oxera :

- Schwartz and Co devrait envisager de tenir compte explicitement de la qualité de service dans son exercice de benchmarking.
- Schwartz and Co devrait également examiner l'impact des conditions exogènes (par exemple, la densité de population, les prix des intrants autres que la main-d'œuvre, les caractéristiques du sol) sur les coûts efficaces des entreprises.
- Si les données relatives à certains facteurs de dépenses importants ne sont pas disponibles, Schwartz and Co devrait être conscient de l'impact que peut avoir l'omission d'une variable sur son modèle, mettre en doute son analyse en conséquence et proposer des moyens d'en tenir compte (par exemple en appliquant une marge d'erreur appropriée aux scores d'efficacité estimés).

4.4.2. Remarques de RESA, AIEG, AIESH et REW

Il existe des alternatives possibles comme variables explicatives. Il importe que les GRD soient impliqués dans le jeu des variables explicatives possibles et que ces variables (ex. : pointe de charge)



fassent l'objet d'une définition stricte afin d'assurer une interprétation et application uniforme entre GRD.

Comme susmentionné, il conviendra de tenir compte des facteurs environnementaux conformément à l'art. 4 §2 15° du décret tarifaire du 19 janvier 2017 qui stipule que « toute méthode de contrôle des coûts reposant sur des techniques de comparaison tient compte des différences objectives existant entre GRD et qui ne peuvent être éliminées à l'initiative de ces derniers ».

Se pose également la question de l'intégration de la qualité des réseaux. L'art. 4 §2 15° du décret tarifaire du 19 janvier 2017 stipule que « Toute décision utilisant des techniques de comparaison des coûts tient compte de la qualité des services rendus ».

La question qui se pose est comment le consultant va en tenir compte dans sa méthode de comparaison ? Le rapport intermédiaire est muet sur ce point.

4.4.3. Réponse aux commentaires de RESA, AIEG, AIESH et REW et d'Oxera

Prix des intrants (input) non liés à la main-d'œuvre

Les prix des intrants mentionnés par Oxera comme étant non liés à la main d'œuvre le sont en fait pour la plus grande partie. En effet les coûts des services de maintenance externe, de la sous-traitance de personnel, des services de consultants et des centres de contrôle sont directement liés aux coûts de main d'œuvre, dont la différence structurelle entre pays est bien prise en compte dans la méthode que nous proposons. Par ailleurs, le coût des fournitures de bureau, également mentionné par Oxera, représente une part marginale des coûts d'un GRD.

Exigences réglementaires

Si des exigences réglementaires différenciées entre pays autres que celles que nous avons déjà identifiées pouvait être génératrices de biais de comparaison, nous en tiendrons compte par des retraitements ou des analyses de sensibilité, selon les possibilités.

Environnement d'exploitation

La question de la prise en compte de la densité de population et des caractéristiques du sol sont traitées au paragraphe 2.3.2.

A notre connaissance, les effets climatiques sur les coûts de GRD n'ont jamais été pris en compte dans le cadre de benchmark internationaux par les régulateurs, en particulier en raison de la difficulté à isoler l'effet climatique sur les coûts, et nous jugeons la prise en compte de ces effets hasardeuse et difficilement faisable et dans le cadre de l'option du lot 2. Par ailleurs les pays



susceptibles d'être intégrés au benchmark sont tous situés à proximité de la Belgique en zone climatique tempérée, ce qui limite fortement l'intérêt de prendre en compte cet effet potentiel.

Qualité de services

Le CEER conduit de façon périodique une analyse de la qualité de service de fourniture d'électricité et de gaz au niveau européen. Les indicateurs les plus communément admis pour mesurer la qualité de service des GRD sont le SAIDI et le SAIFI :

- System Average Interruption Duration Index (SAIDI) ;
- System Average Interruption Frequency Index (SAIFI).

Néanmoins, ces indicateurs peuvent être calculés de façon différente selon les pays et ainsi ne pas être comparables entre eux : *« It is important to emphasise that indicators for CoS are currently not perfectly harmonised between different countries. Notably, the following definitions might differ depending on the country:*

- *The voltage levels EHV, HV, MV and LV;*
- *Exceptional events; and*
- *Indicators such as SAIDI, SAIFI, can be calculated with slightly different methods depending on the country (for example due to national legislation).»*,

(CEER Benchmarking Report 6.1 on the, Continuity of Electricity and Gas Supply, 2018, page 9).

Au périmètre de la Wallonie, ces indicateurs ne sont pas encore suffisamment homogènes entre les différents GRD. La CWaPE a réalisé courant 2018 et 2019 un audit spécifique sur les processus de collecte et de traitement des données rentrées dans le cadre du « rapport qualité » des gestionnaires de réseau. Il est indiqué dans ce rapport que : *« Une fois par an, la CWaPE reçoit un rapport qualité des gestionnaires de réseau de distribution (GRD). Les données communiquées dans ce rapport sont l'objet d'une analyse et sont ensuite discutées lors d'une réunion de travail CWaPE-GRD, dans le cadre de l'analyse des plans d'adaptation. Il ressort de ces échanges qu'en dépit d'une récente remise à plat des définitions utilisées dans des lignes directrices (gaz et électricité), l'on ne peut en tout temps garantir la parfaite exactitude des données communiquées par les GRD. En effet, la méthodologie de collecte et de traitement employée et/ou l'interprétation des données par les agents techniques ou administratifs peuvent potentiellement différer d'un GRD à l'autre ou au sein même d'un GRD. »* (page 5 du rapport CD-20d23-CWaPE-0072 - audit des rapports qualité des GRD, avril 2020).

Oxera affirme que : « l'ensemble des drivers de coûts proposés par Schwartz and Co ne tient pas compte de la qualité des services fournis et qu'il est donc incompatible avec les principes établis par la loi ». Cette affirmation est une interprétation, par Oxera, de la loi qui stipule que : « Toute décision utilisant des techniques de comparaison des coûts tient compte de la qualité des services rendus »¹⁸. En aucun cas la loi n'impose l'utilisation de paramètre de qualité de service comme des variables explicatives du modèle de mesure d'efficacité. Par ailleurs, la CWaPE propose dans les lignes directrices relatives aux indicateurs de performance des gestionnaires de réseau de gaz et

¹⁸ Moniteur Belge (2017), 'Décret relatif à la méthodologie tarifaire applicable aux gestionnaires de réseaux de distribution de gaz et d'électricité', Art. 4 § 2 15o.



d'électricité pour la période tarifaire 2024 à 2028¹⁹ de suivre et d'inciter les indicateurs SAIFI et SAIDI chez les GRD wallons. Dans ce contexte, la méthodologie tarifaire pour les GRD d'électricité et de gaz sur la période 2024 à 2028 tiendra compte de la qualité des services rendus et sera en cohérence avec la législation en vigueur.

En conclusion concernant la qualité de service, il serait hasardeux d'utiliser des indicateurs de qualité de service directement comme des variables explicatives du modèle :

- Ceux-ci ne sont à ce jour pas suffisamment comparables entre eux au niveau européen ni au niveau des GRD wallons.
- La réglementation Wallonne n'impose pas l'utilisation de paramètres de qualité de service directement comme variables explicatives du modèle de mesure d'efficacité.
- Ce n'est pas une pratique observée chez les régulateurs européens que nous avons analysés, à l'exception du régulateur finlandais qui intègre la variable explicative des coûts réglementaires de panne pour ses GRD d'électricité.

4.4.4. Précisions apportées dans le rapport final

Le rapport final sera complété avec les réponses apportées par Schwartz and Co.

4.5. Utilisation de panels et vérification de la robustesse du modèle

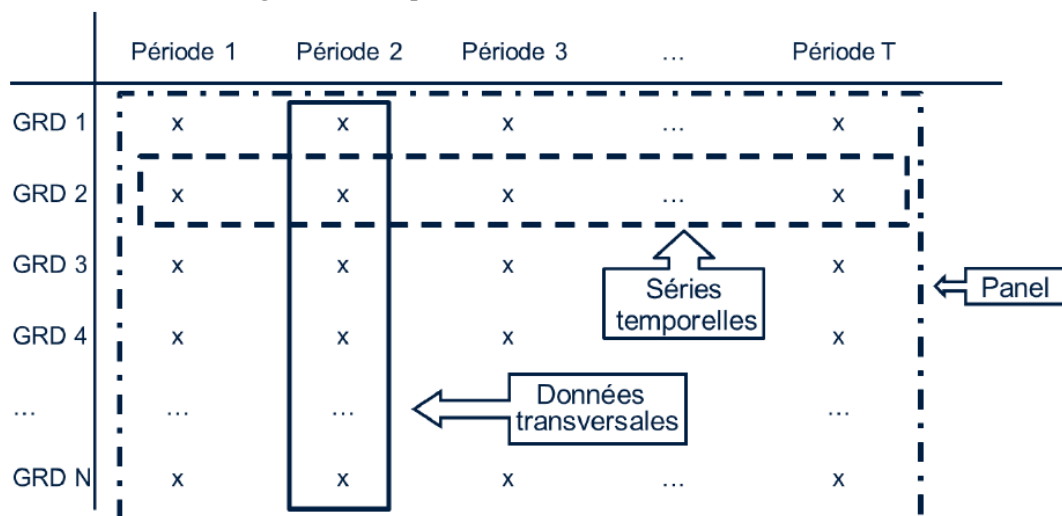
4.5.1. Remarques d'Oxera

Des analyses de benchmarking réglementaire ont été menées sur des bases de données de différentes tailles. Par exemple, le régulateur Bundesnetzagentur a accès aux données de 185 GRD gaziers allemands, tandis que l'Ofgem est limité à seulement huit GRD gaziers britanniques. Lorsque les régulateurs ont des échantillons de plus petite taille, il est courant d'envisager l'utilisation de « données de panel » pour augmenter la « taille » des données. Les données de panel peuvent également permettre des approches d'estimation plus sophistiquées qui peuvent éviter et/ou atténuer certaines hypothèses fortes formulées dans l'analyse transversale.

¹⁹ Lignes directrices relatives aux indicateurs de performance des gestionnaires de réseau de gaz et d'électricité actifs en région wallonne annulant et remplaçant les lignes directrices référencées cd-19i10-cwape-0025, paragraphe 3.3.1 page 12, avril 2020.



Figure 2. Exemples de structures de base de données



Source : Oxera

Par exemple, la méthode advanced COLS proposée par Schwartz and Co, si elle est appliquée sur des données transversales, ne peut pas faire la distinction entre les erreurs statistiques, l'inefficacité du GRD et l'hétérogénéité du GRD (c'est-à-dire les différences dans les caractéristiques structurelles et opérationnelles) qui ne sont pas capturés dans la spécification du modèle. Des estimateurs économétriques sophistiqués, tels que les modèles à effets aléatoires et les panels SFA, peuvent tenir compte de la structure de données du panel et de l'hétérogénéité non observée des GRD.

De plus, il est bien établi que les dépenses des entreprises peuvent être volatiles d'une année à l'autre, en particulier les dépenses de grands projets d'investissements. Ainsi, la performance d'une entreprise dans les modèles transversaux peut dépendre de sa position dans le cycle d'investissements (c'est-à-dire des niveaux d'activité), plutôt que de son niveau d'efficacité. L'utilisation de données de panel peut atténuer ce problème, car les dépenses de chaque entreprise sont évaluées sur un certain nombre d'années de sorte que les effets des cycles d'investissements disparaissent en moyenne.

ORES a demandé si Schwartz and Co utilisera des données de panel ou des données transversales pour estimer son modèle. En réponse à cette question, Schwartz and Co déclare :

« Les données de la même année devraient être utilisées afin d'assurer la cohérence temporelle et la robustesse de l'analyse. » Il est important de noter que la « cohérence temporelle » peut être assurée sur les données de panel, à condition que les mêmes années de données soient disponibles pour tous les GRD de l'échantillon. Par conséquent, la « cohérence temporelle » n'est pas un argument suffisant pour rejeter l'utilisation des données de panel. En effet, la maturité des GRD et le processus de régulation dans les juridictions sélectionnées pourraient fournir des facteurs contextuels importants dans l'identification de l'échantillon de données et du benchmark appropriés.



Recommandation d'Oxera :

- La taille de l'échantillon proposée est relativement petite, et une erreur de données dans n'importe quel GRD peut avoir un impact significatif sur le modèle global. Cette question est particulièrement pertinente, car COLS est une méthode déterministe d'estimation de l'efficacité. Par conséquent, Schwartz and Co devrait s'assurer que les résultats ne sont pas sensibles aux petits changements des données et du modèle.
- Schwartz and Co devrait utiliser des données de panel pour estimer son modèle d'analyse comparative. Lors de l'utilisation de données de panel, Schwartz and Co devrait explorer des approches d'estimation plus sophistiquées telles que les modèles à effets aléatoires et les panels SFA qui peuvent éviter certaines hypothèses de modélisation fortes qui sont requises dans son approche préférée.
- Si Schwartz and Co ne peut recueillir que des données transversales, il devrait s'assurer que son modèle (et les scores d'efficacité qui en résultent) ne soient pas influencés par l'hétérogénéité (observée et non observée) des entreprises. De même, les scores d'efficacité des entreprises ne doivent pas être influencés par leurs niveaux d'activité ou l'étape du cycle réglementaire. Ces aspects pourraient être pris en compte dans l'élaboration d'un échantillon de données approprié, dans le processus de développement du modèle (par exemple par l'incorporation de drivers de coûts), dans le processus de validation (par exemple en termes de points de référence - ou benchmark - appropriés) et dans la prise en compte d'une « marge d'erreurs » pour atténuer le biais potentiel.

4.5.2. Remarques de RESA, AIEG, AIESH et REW

Nous avons également constaté que certains régulateurs appliquaient leur modèle sur plusieurs années différentes ('panel data') pour vérifier la corrélation temporelle entre les paramètres d'inefficience.

De même la robustesse du modèle est testée en modifiant l'échantillon. Nous souhaiterions que le rapport du consultant intègre ces bonnes pratiques.

4.5.3. Réponse aux remarques d'Oxera et de RESA, AIEG, AIESH et REW

Utilisation de données de panels

L'utilisation de données de panels n'est pas une pratique courante comme l'affirme Oxera chez les régulateurs disposant d'échantillons de taille modeste. Chez les régulateurs que nous avons analysés en détail dans le cadre de cette mission (Allemagne, Autriche, France et Finlande), aucun ne s'est appuyé sur des données de panels, à l'exception du régulateur finlandais qui utilise la moyenne des données 2011 à 2014 pour la quatrième période de régulation débutant en 2016 et qui utilise la moyenne des données 2015 à 2018 pour la cinquième période de régulation débutant en 2020²⁰.

²⁰ Regulation methods in the fourth regulatory period of 1 January 2016 – 31 December 2019 and the fifth regulatory period of 1 January 2020 – 31 December 2023, Energiavirasto, November 2015



Par ailleurs, comme présenté au paragraphe 2.4.2.2, si la qualité statistique de la modélisation se révélait non satisfaisante en raison de la taille de l'échantillon de GRD considéré initialement (jusqu'à une vingtaine), nous intégrerions à l'échantillon les GRD supplémentaires requis pour obtenir un modèle de bonne qualité statistique.

Vérification de la robustesse du modèle économétrique obtenu

RESA, AIEG, AIESH et REW souhaitent que nous intégrions à la méthode proposée le test de la robustesse du modèle « en modifiant l'échantillon ». Par ailleurs Oxera indique que « Schwartz and Co devrait s'assurer que les résultats ne sont pas sensibles aux petits changements des données et du modèle ».

Nous comprenons que les GRD font référence aux tests de ce que l'on appelle l'effet de levier et des points d'influence qui en résultent. Il s'agit d'un point tout à fait pertinent, que nous avons abordé partiellement dans le rapport intermédiaire dans le paragraphe 3.2.4 sur les points aberrants, et qui mérite des compléments qui seront apportés dans le rapport final.

La méthode Advanced COLS que nous proposons, comprend en effet une analyse des effets de levier qui vise à mesurer si chaque GRD est un point d'influence du modèle, c'est-à-dire si sa présence dans l'échantillon modifie sensiblement le modèle économétrique obtenu.

Cette analyse est réalisée à l'aide de deux méthodes complémentaires, qui vont d'ailleurs au-delà de ce que demandent RESA, AIEG, AIESH et REW :

- Le calcul de la diagonale de la matrice HAT^{21} permet de mesurer l'effet de levier de chacun des GRD, et de déterminer si un GRD est un point d'influence à travers un test du coefficient calculé.
- Le calcul de la statistique $DFFTS^{21}$, qui permet de détecter les points d'influence en calculant l'impact sur le modèle du retrait de chacun des GRD sur les N GRD de l'échantillon, à travers un test.

La méthode advanced COLS intègre également une détection des observations anormales, à travers le calcul du résidu standardisé²¹.

4.5.4. Précisions apportées dans le rapport final

Le rapport final sera complété avec les réponses apportées par Schwartz and Co.

²¹ Pour une description détaillée voir la section 3 du livre « Économétrie », Dunod, Collection Module économique, 10ème édition 2018 de Régis Bourbonnais, Docteur en économie mathématique et économétries, Maître de conférences à l'Université Paris Dauphine et co-développeur de la méthode Advanced COLS