

Tarification de l'accès et de l'utilisation des réseaux  
de transport et de distribution d'électricité: la  
perspective académique

Thomas-Olivier Léautier\*

February 24, 2015

---

\*Professeur de gestion, IAE Toulouse et directeur de recherche, Toulouse School of Economics. Ce rapport est préparé pour la Commission de Régulation de l'Énergie. Les conclusions présentées dans ce rapport engagent seulement l'auteur.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	Objectifs de la tarification des réseaux . . . . .	5
1.2	Incidations à court et à long terme . . . . .	5
1.3	Couverture des coûts des réseaux . . . . .	7
1.4	Structure de ce rapport . . . . .	8
<b>I</b>	<b>Efficacité à court terme</b>	<b>8</b>
<b>2</b>	<b>Les prix nodaux, signaux de consommation et de production optimales</b>	<b>9</b>
2.1	La tarification heure de pointe sur un marché unique . . . . .	9
2.1.1	Technologie de production unique . . . . .	10
2.1.2	Nombre fini de technologies de production . . . . .	11
2.1.3	Continuum de technologies de production . . . . .	12
2.2	La tarification heure de pointe étendue aux réseaux interconnectés : les prix nodaux . . . . .	12
2.2.1	Introduction aux prix nodaux . . . . .	12
2.2.2	L'exemple de deux marchés reliés par une interconnexion	15
2.3	Redispatching/counter-trading . . . . .	17
2.4	Gains pour la collectivité générés par les prix nodaux . . . . .	18
2.5	La mise en oeuvre des prix nodaux . . . . .	18
2.5.1	Assurance: les droits financiers de transport . . . . .	18
2.5.2	Redistribution . . . . .	19
2.5.3	Compatibilité avec l'architecture de marché "européenne"	20
2.5.4	Compatibilité avec le couplage des marchés européens	20
2.6	Prix nodaux, droits de transport, et pouvoir de marché . . . . .	21
2.6.1	Pouvoir de marché local . . . . .	21
2.6.2	Droits de transport en présence de pouvoir de marché	22
<b>II</b>	<b>Efficacité à long-terme</b>	<b>22</b>
<b>3</b>	<b>Les prix nodaux, signaux d'investissement optimaux</b>	<b>24</b>
3.1	Investissement en production . . . . .	24
3.1.1	Programme d'optimisation . . . . .	24

3.1.2	Application à deux marchés reliés par une intercon-	27
	nexion . . . . .	
3.2	Investissement optimal dans les réseaux . . . . .	28
3.2.1	Caractérisation de l’optimum de long terme . . . . .	29
3.2.2	Application à deux marchés reliés par une intercon-	30
	nexion . . . . .	
3.2.3	Difficultés pratiques . . . . .	30
3.2.4	Demande non réactive au prix . . . . .	31
3.3	Incitations à l’investissement optimal . . . . .	31
3.4	Parc de production et réseau optimal . . . . .	31
3.5	Internalisation des coûts de réseau par les producteurs . . . . .	32
<b>4</b>	<b>Tarif d’injection différencié incitatif</b>	<b>34</b>
4.1	Observations préliminaires . . . . .	34
4.2	Réseau optimal . . . . .	35
4.2.1	Congestion gérée par les prix nodaux . . . . .	35
4.2.2	Congestion gérée par le couplage de marché/redispaching	35
4.3	Aucune contrainte réseau autorisée . . . . .	37
4.3.1	Optimalité du tarif d’injection différencié incitatif . . . . .	38
4.3.2	Mise en oeuvre du tarif d’injection différencié dans le	38
	cas le plus simple . . . . .	
4.3.3	Mise en oeuvre d’un tarif de retrait différencié dans le	39
	cas le plus simple . . . . .	
4.3.4	Mise en oeuvre de la prime d’injection et de retrait	40
	différenciée dans un cas réaliste . . . . .	
4.4	“Deep connection charge” . . . . .	40
<b>III</b>	<b>Couverture de la totalité des coûts des réseaux</b>	<b>40</b>
<b>5</b>	<b>Allocation des coûts résiduels</b>	<b>41</b>
5.1	Partage des coûts résiduels entre producteurs et consommateurs	41
5.1.1	Justification théorique . . . . .	42
5.1.2	Aspects pratiques . . . . .	44
5.2	Assiette du tarif . . . . .	45
5.2.1	Cadre analytique . . . . .	45
5.2.2	Mise en oeuvre du tarif optimal . . . . .	46

5.3	Partage des coûts d'une interconnexion . . . . .	47
<b>6</b>	<b>Conclusion</b>	<b>47</b>
6.1	La tarification nodale de l'électricité est la bonne solution . .	47
6.1.1	Pourquoi adopter les prix nodaux . . . . .	47
6.1.2	Pourquoi ne pas adopter les prix nodaux? . . . . .	49
6.1.3	Implications de la mise en oeuvre des prix nodaux en France . . . . .	51
6.2	Un timbre d'injection différencié ne semble pas constituer une bonne solution intermédiaire . . . . .	51
6.2.1	Timbre d'injection différencié exprimé en €/MW/an .	52
6.2.2	Timbre d'injection différencié exprimé en €/MWh . .	54
6.3	Un tarif de raccordement "profond" est probablement légitime sur le réseau de distribution . . . . .	55
6.4	Partage des coûts résiduels . . . . .	55
6.4.1	La nécessité de couvrir les coûts résiduels . . . . .	56
6.4.2	In fine les consommateurs payent tous les coûts . . . .	56
6.4.3	Partage optimal des coûts entre consommateurs . . . .	57
6.5	Réponses aux questions posées par la consultation . . . . .	57
6.5.1	Réponses aux questions posées à l'article 5 de la con- sultation: . . . . .	57
6.5.2	Réponse aux objectifs de l'étude fixée par la consulta- tion: . . . . .	58
	<b>References</b>	<b>60</b>

# 1 Introduction

En préparation de l'élaboration des prochains Tarifs d'Utilisation des Réseaux Publics de l'Electricité (TURPE), la Commission de Régulation de l'Energie (CRE) souhaite engager une réflexion sur la structure de ces tarifs, en particulier sur l'opportunité "de faire payer une partie des coûts de transport et de distribution de l'électricité aux injections et de mettre en place sur le réseau de transport un signal-prix de localisation des capacités d'injections" (Commission de Régulation de l'Energie, 2014). Ce rapport présente la réponse académique à cette question, qui s'inspire de nombreux articles et rapports publiés par des académiques, mais aussi d'analyses originales conduites par l'auteur.

## 1.1 Objectifs de la tarification des réseaux

La tarification des réseaux doit remplir trois objectifs:

1. générer les bons signaux économiques d'utilisation des réseaux, en particulier la gestion de la congestion, i.e., conduire à la maximisation du surplus net à court terme,
2. générer les bons signaux économiques d'investissement pour les moyens de production (technologie et localisation), la localisation des centres de consommation, et l'expansion des réseaux, i.e., conduire à la maximisation du surplus net à long terme, et
3. couvrir l'ensemble des coûts des réseaux.

## 1.2 Incitations à court et à long terme

La réflexion académique sur la tarification du réseau de transport s'est tout d'abord concentrée sur le premier point. La théorie économique suggère en effet qu'il faut commencer par s'assurer que les prix spots reflètent effectivement le coût d'opportunité de produire un bien.

Dans le cas des réseaux électriques, le problème semble a priori difficile, pour deux raisons:

1. la capacité de transit sur les lignes de transport étant fixe, celles-ci sont congestionnées une partie du temps. Lorsque cela se produit, il

est impossible d'échanger plus de kilowattheures entre les extrémités de la ligne, et

2. les kilowattheures se déplacent sur les réseaux en suivant les lois de la physique, et non pas les contrats commerciaux. Par exemple, une partie de kilowattheures produits en France et vendus en Allemagne transite par la Belgique, une autre par la Suisse.

Heureusement, la réponse académique est extrêmement simple: la mise en oeuvre de prix nodaux de l'électricité (en anglais, Locational Marginal Prices, LMPs) permet d'atteindre l'efficacité de court terme, i.e., de générer des prix spots qui capturent le coût d'opportunité de l'électricité, incluant les congestions et les pertes sur les réseaux (e.g., Schweppe et al., 1988, et Hogan, 1992). Les prix spots ainsi déterminés fournissent les bons signaux économiques pour l'investissement dans les moyens de production et la localisation des centres de consommation. De plus, ils permettent de produire les bons signaux d'investissement dans les réseaux (Léautier, 2000, et Hogan et al., 2010). Atteindre le premier objectif permet ainsi d'attendre le deuxième.

Plusieurs analyses académiques ont documenté les gains importants pour la collectivité générés par la mise en oeuvre de prix nodaux. De plus, la mise en oeuvre de prix nodaux est désormais techniquement aisée, les marchés électriques des Etats Unis ayant adopté les prix nodaux depuis la fin des années 1990. Pour ces raisons, le consensus de la littérature académique est qu'il faut mettre en oeuvre des prix nodaux (e.g., Baldick et al. 2011, Perez-Arriaga et al., 2013). Une partie importante de ce rapport leur est donc consacrée. Une autre approche pour inciter à la localisation des moyens de production, les timbres d'injection différenciés et incitatifs, est aussi présentée, dont les propriétés sont comparées aux prix nodaux, qui constituent l'optimum de premier rang.

L'importance que ce rapport accorde aux prix nodaux peut surprendre certains lecteurs. La congestion sur réseau de transport français est très faible. Historiquement, EDF puis RTE ont conçu le réseau pour "évacuer" toute la production des centrales installées sur le territoire. Les flux d'énergie étant relativement bien anticipés, les réseaux étaient conçus pour éliminer ou très sévèrement limiter la congestion, i.e., la France était en première approximation une plaque de cuivre. Même si les prix nodaux constituent

l'optimum théorique, leur mise en oeuvre ne serait-elle pas trop coûteuse par rapport aux gains réalisés?

L'auteur de ce rapport pense que la congestion sur les réseaux français et européen va grandement augmenter dans les prochaines années. Le déploiement de production renouvelable sur le territoire va profondément modifier les flux d'énergie, et les rendre plus variables. Il faudrait donc modifier profondément la topographie du réseau pour accommoder ces nouveaux flux. Mais, l'acceptabilité des infrastructures en général et des lignes de transport en particulier étant bien plus faible que dans les années 1970s-1980s, il sera très coûteux voire impossible de mettre en oeuvre l'ensemble de ces modifications. Le réseau français sera donc plus souvent congestionné. Cette évolution se produira probablement plus rapidement et à plus grande échelle chez nos voisins (Neuhoff et al., 2013). La gestion efficace de la congestion deviendra alors un enjeu économique majeur, et les prix nodaux seront mis en oeuvre.

### **1.3 Couverture des coûts des réseaux**

Comme dans de nombreux problèmes en économie, la tarification au cout marginal (de long-terme) conduit à prendre les bonnes décisions d'utilisation des réseaux et d'investissement, mais ne garantit pas couverture de l'ensemble des coûts si les technologies exhibent des rendements d'échelle croissants, ce qui est le cas pour les réseaux de transport et de distribution d'électricité. Que des prix nodaux ou des tarifs d'injection différenciés et incitatifs soient mis en oeuvre pour atteindre le deuxième objectif, il faut aussi mettre en place un tarif complémentaire pour atteindre le troisième objectif, i.e., couvrir les coûts résiduels (Perez-Arriaga et al., 2013, Bell et al., 2011). Il faut donc deux instruments différents pour atteindre les deuxième et troisième objectifs.

Dans le secteur du transport d'électricité, le partage des coûts résiduels est souvent extrêmement simple: dans de nombreux pays, ces coûts sont simplement couverts par les consommateurs au prorata de la puissance maximale qu'ils soutirent ou de leur consommation (ou d'une combinaison des deux). Dans d'autres, les coûts résiduels sont partagés entre producteurs et consommateurs selon des règles mal justifiées économiquement, puis au sein de chaque groupe au prorata de la puissance installée/soutirée et/ou de la production/consommation. Ces pratiques semblent très simples comparées

à la très riche réflexion académique sur le partage des couts communs (voir par exemple la présentation de Boyer et al., 2006).

La couverture des coûts résiduels modifie les décisions des agents. Par exemple, si les consommateurs couvrent une portion des coûts au prorata de leur consommation, ils réduisent celle-ci. La règle optimale de couverture des coûts résiduelle minimise les distorsions par rapport aux incitations optimales développées plus haut. Ce problème a été analysé par Ramsey (1927) pour l'imposition optimale, puis par Boiteux (1956) précisément pour la couverture des coûts fixes d'une compagnie électrique (voir par exemple le Chapitre 8 de Boyer et al., 2006). La tarification optimale est telle que les utilisateurs du réseau (producteurs et consommateurs) payent un prix inversement proportionnel à leur élasticité.

Ce rapport examine comment ces arguments s'appliquent aux réseaux de transport et de distribution d'électricité.

#### **1.4 Structure de ce rapport**

Ce rapport est structuré en trois parties. La partie I (Section 2) présente les problématiques d'efficacité à court terme, i.e., de production, consommation, et utilisation optimale d'un réseau présentant des contraintes. La partie II examine les problématiques d'efficacité à long terme: investissement en production et extension optimale du réseau. Elle examine successivement les incitations créées par les prix nodaux (Section 3) puis par des tarifs d'injection différenciés incitatifs (Section 4). La partie III (Section 5) examine la problématique de couverture des couts fixes. La Section 6 présente les conclusions de ce rapport et répond aux questions posées par la consultation.

## Part I

# Efficacité à court terme

## 2 Les prix nodaux, signaux de consommation et de production optimales

### 2.1 La tarification heure de pointe sur un marché unique

Dès l'origine de l'industrie électrique à la fin du XIXe siècle, les ingénieurs-économistes ont examiné la problématique de la tarification optimale de l'électricité. La première analyse formelle est celle de Marcel Boiteux (1949), qui montre qu'à l'optimum le prix de l'électricité s'établit à son coût marginal.

Un modèle simplifié permet de présenter la tarification heure de pointe. L'incertitude sur l'état du monde, dénoté  $t \geq 0$ , en est le premier ingrédient. Ensuite, considérons un ensemble de consommateurs, dont la masse est normalisée à 1, qui ont la même demande  $D(p, t)$ , i.e., chaque consommateur est prêt à consommer  $D(p, t)$  dans l'état du monde  $t$  lorsque le prix est  $p$ . La demande décroît en fonction du prix, i.e.,  $D_p = \frac{dD}{dp} < 0$ , et varie en fonction de l'état du monde: par exemple, pour un même prix, la demande est plus élevée le jour que la nuit, la semaine que le weekend, l'hiver que l'été en France. Par simplicité, les états du monde sont ordonnés par demande croissante, i.e.,  $D_t = \frac{dD}{dt} \geq 0$ .

Le surplus marginal  $P(Q, t)$  correspond au prix que les consommateurs sont prêts à payer pour une quantité  $Q$  dans l'état du monde  $t$ . Il est défini par  $D(P(Q, t), t) = Q$ . Le surplus des consommateurs  $U(Q, t)$  est l'intégrale sous la courbe du surplus marginal, défini par

$$U(Q, t) = \int_0^Q P(q, t) dq.$$

L'analyse la plus simple suppose que les consommateurs ajustent leur demande au prix de gros, c'est à dire qu'ils sont sur un tarif "temps réel". Léautier (2014) montre que les résultats sont inchangés si une fraction des consommateurs fait face à un tarif constant, et que des délestages sont parfois nécessaires pour équilibrer la demande à l'offre<sup>1</sup>.

---

<sup>1</sup>Dans ce cas, la demande est plus "verticale". Si des délestages sont nécessaires dans certains états du monde pour équilibrer la demande à l'offre, le prix s'établit à la valeur

Le formalisme utilisé ici est donc très général. Les “états du monde” recouvrent à la fois différentes heures de l’année, mais différents scénarii d’évolutions de la demande sur plusieurs années. Les critères de fiabilité, par exemple la valeur de la défaillance, peuvent être incorporés à l’analyse au travers de la fonction de demande.

### 2.1.1 Technologie de production unique

L’analyse la plus simple considère qu’il existe une seule technologie de production. Le coût variable de production couvre essentiellement les coûts de combustible. Il est supposé constant et égal à  $c$  €/MWh. Le coût de capacité est aussi supposé constant et égal à  $r$  €/MWh (i.e.,  $r$  est l’annuité de capacité, exprimée en €/MW/an divisée par 8760 heures par an).

L’analyse la plus simple suppose que toute la capacité de production est disponible. Introduire une capacité de production aléatoire ne modifie pas les intuitions.

Le problème est de déterminer le prix  $p(t)$  dans chaque état du monde et la capacité  $k$  qui maximisent le surplus net, sous la contrainte que la production (égale à la demande) demeure inférieure à la capacité installée. Une formulation alternative est de déterminer la production  $q(t) \leq k$  et la capacité  $k$  qui maximisent le surplus net. Le coût virtuel de la contrainte de capacité dans l’état du monde  $t$  est dénoté  $\lambda(t)$ . Formellement, le problème d’optimisation s’écrit

$$\begin{aligned} \max_{q(t), k} \mathbb{E} [ (U(q(t), t) - cq(t)) ] - rk \\ \text{st : } q(t) \leq k \quad \forall t \quad (\lambda(t)) . \end{aligned} \tag{1}$$

Pour résoudre le problème, il faut distinguer deux périodes. Hors pointe, la production  $q(t)$  est inférieure à la capacité  $k$ . Le prix s’établit au coût marginal, i.e.,  $p(t) = c$ , donc la production est égale à la demande  $D(c, t)$ . En pointe, la demande est égale à la capacité  $k$ , le prix est déterminé par l’intersection de la demande et de l’offre, i.e.,  $p(t) = P(k, t)$ .

La capacité optimale est telle que la marge opérationnelle pour une unité marginale est égale au coût de capacité de cette unité:

$$\mathbb{E} [(p(t) - c)^+] = r,$$

de défaillance, déterminée administrativement.

où  $(x)^+ = \max(x, 0)$ . La marge opérationnelle est positive seulement durant les heures de pointe, durant lesquelles  $p(t) = P(k, t)$ . La capacité optimale est donc définie par

$$\Psi_0(k, c) = \mathbb{E}[(P(k, t) - c)^+] = r. \quad (2)$$

Une unité de capacité marginale produit un surplus marginal net  $\Psi_0(k, c)$ , et coûte  $r$ . A l'optimum, le surplus marginal est égal au coût marginal.

Les technologies de production présentant dans notre modèle des rendements d'échelle constants, les profits réalisés durant les heures de pointe permettent de financer entièrement le parc de production optimal.

L'analyse de Marcel Boiteux était développée à l'origine dans le cadre d'un monopole bienveillant. Elle s'étend à un marché en concurrence pure et parfaite. La condition d'optimalité (2) devient alors une condition d'entrée libre.

### 2.1.2 Nombre fini de technologies de production

Supposons maintenant qu'il existe un nombre fini de technologies de production, indexées par  $n = 1, \dots, N$ . Dénotons  $c_n$  le coût marginal de production et  $r_n$  le coût marginal de capacité de la technologie  $n$ . Sans perte de généralité, les technologies sont ordonnées par ordre de coût marginal de production croissant, i.e.,  $n \leq m \Leftrightarrow c_n \leq c_m$ . Par définition, les coûts marginaux de capacité sont donc décroissants:  $n \leq m \Leftrightarrow r_n \geq r_m$ . La technologie 1 est la technologie de base, centrale nucléaire ou charbon, la technologie  $N$  est la technologie de pointe, centrale gaz ou fioul. Dénotons  $k_n$  la capacité installée de technologie  $n$ , et  $K_n = \sum_{m=1}^n k_m$  la capacité cumulée installée jusqu'à la technologie  $n$ .

Formellement, il s'agit de résoudre le problème d'optimisation

$$\begin{aligned} \max_{q_n(t), k_n} \mathbb{E} \left[ U \left( \sum_{n=1}^N q_n(t), t \right) - \sum_{n=1}^N c_n q_n(t) \right] - \sum_{n=1}^N r_n k_n \\ \text{st : } q_n(t) \leq k_n \forall (t, n) \quad (\lambda_n(t)) \end{aligned} \quad (3)$$

L'optimum social est caractérisé par la courbe d'offre en marches d'escalier: lorsque que la technologie  $n$  est marginale, le prix est  $p(t) = c_n$ , ce qui correspond à la partie horizontale de la marche. Lorsque la technologie  $n$  produit à capacité, mais avant que la technologie  $(n + 1)$  ne produise, le prix est

déterminé par l'intersection de la demande et de l'offre, fixée à la capacité  $K_n$ , i.e.,  $p(t) = P(K_n, t)$  pour  $p(t) \in [c_n, c_{n+1}]$ , ce qui correspond à la partie verticale de la marche.

La capacité cumulée totale  $K_N$  est déterminée par

$$\Psi_0(K_N, c_N) = r_N, \quad (4)$$

et, pour  $1 \leq n \leq N$ , la capacité cumulée  $K_n$  est déterminée par

$$\beta_0(K_n, c_n, c_{n+1}) = \Psi_0(K_n, c_n) - \Psi_0(K_n, c_{n+1}) = r_n - r_{n+1}. \quad (5)$$

La capacité cumulée totale ne dépend que du coût de la technologie de pointe. La capacité cumulée  $K_n$  résulte d'un arbitrage entre les coûts des technologies  $n$  et  $(n+1)$ : à la marge, substituer une unité de technologie  $n$  à une unité de technologie  $(n+1)$  substitue le surplus net  $(\Psi_0(K_n, c_n) - r_n)$  au surplus net  $(\Psi_0(K_n, c_{n+1}) - r_{n+1})$ . A l'optimum, les deux valeurs sont égales.

### 2.1.3 Continuum de technologies de production

Paul Joskow et Jean Tirole (2007) étendent l'analyse à un continuum de technologies de production, ce qui permet de mieux représenter la réalité des courbes d'offres empiriques, pour lesquelles différentes centrales utilisant la même technologie ont des performances, donc des coûts, légèrement différents. Les intuitions sont inchangées.

## 2.2 La tarification heure de pointe étendue aux réseaux interconnectés : les prix nodaux

La deuxième contribution fondamentale est celle de Fred Schweppe et al. (1988), qui étend à un réseau interconnecté les résultats de Marcel Boiteux.

### 2.2.1 Introduction aux prix nodaux

**Notations** Considérons  $M$  marchés, reliés par  $I$  interconnexions. Comme précédemment, l'état du monde est dénoté  $t \geq 0$ .

Pour chaque marché  $m = 1, \dots, M$ ,  $Q_m^S(t)$ ,  $Q_m^D(t)$ , and  $Q_m(t) = Q_m^S(t) - Q_m^D(t)$  représentent la production, la demande et l'injection nette, et  $c_m(q)$  est le coût marginal de long-terme de produire la quantité  $q$ .

Cette introduction met l'accent sur l'efficacité de court terme, donc n'examine pas explicitement le problème d'investissement en production, qui sera exposé dans la Section 3.

Pour chaque ligne  $i = 1, \dots, I$ ,  $\varphi_i(t)$  est le flux l'interconnexion  $i$ , orienté par exemple selon la convention que  $\varphi_i(t)$  se déplace du marché au nombre le plus bas vers le marché au nombre le plus haut.  $\Phi_i^+$  et  $\Phi_i^-$  sont les capacités maximales de transfert dans chaque direction, avec la convention que  $\Phi_i^+$  correspond à  $\varphi_i(t) > 0$ . Le vecteur des flux orientés est dénoté  $\varphi(t) \in \mathbb{R}^I$ .

**Equilibre global** A chaque instant, la production sur l'ensemble du réseau est égale à la demande sur l'ensemble du réseau, plus les pertes thermiques  $\tilde{L}(\varphi(t))$ :

$$\sum_{m=1}^M Q_m^S(t) = \sum_{m=1}^M Q_m^D(t) + \tilde{L}(\varphi(t)). \quad (6)$$

Les  $M$  injections nettes étant liées par la relation précédente, seulement  $(M - 1)$  injections nettes sont indépendantes. Nous définissons alors un marché comme la référence (par exemple, le marché 1), et écrivons les équations des flux en fonction du vecteur  $\underline{\mathbf{Q}}(t) \in \mathbb{R}^{M-1}$  des  $(M - 1)$  injections nettes restantes.

**Approximation du courant continu** Cette approximation permet de linéariser les flux d'énergie:

$$\varphi(t) = \mathbf{H} \cdot \underline{\mathbf{Q}}(t)$$

où  $\mathbf{H} \in \mathbb{R}^I \times \mathbb{R}^{M-1}$  est la matrice d'admittance-transfert.

Sous l'approximation du courant continu, les pertes sont une fonction quadratique des flux, donc des injections nettes

$$\tilde{L}(\varphi(t)) = \underline{\mathbf{Q}}^T(t) \cdot \mathbf{B} \cdot \underline{\mathbf{Q}}(t) = L(\underline{\mathbf{Q}}(t))$$

où la matrice  $\mathbf{B} \in \mathbb{R}^{M-1} \times \mathbb{R}^{M-1}$  est symétrique. L'équilibre global s'écrit donc

$$\sum_{m=1}^M Q_m^S(t) = \sum_{m=1}^M Q_m^D(t) + L(\underline{\mathbf{Q}}(t)). \quad (7)$$

Dénotons  $\mu_e(t)$  le multiplicateur de Lagrange associé à cette contrainte.

$\mu_e(t)$  représente la valeur d'un MWh marginal dans l'état  $t$ .

**Contrainte de capacité sur les interconnexions** Chaque ligne ayant une capacité de transit maximale,

$$-\Phi_i^- \leq \varphi_i(t) \leq \Phi_i^+ \Leftrightarrow -\Phi_i^- \leq \sum_{m=2}^M H_{im} Q_m(t) \leq \Phi_i^+.$$

A l'optimum, la congestion est dans une direction seulement sur chaque ligne. Pour simplifier la notation, dénotons  $\Phi_i$  la contrainte active dans l'état  $t$ , et  $\eta_i(t)$  le multiplicateur de Lagrange associé.

**Problème d'optimisation** L'optimum social est donc caractérisé par

$$\begin{aligned} & \max_{Q_m^S(t), Q_m^D(t)} \mathbb{E} \left[ \sum_{m=1}^M \{U_m(Q_m^D(t)) - c_m(Q_m^S(t))\} \right] \\ \text{st : } & \begin{cases} \sum_{m=1}^M Q_m^S(t) = \sum_{m=1}^M Q_m^D(t) + L(\mathbf{Q}(t)) \quad \forall t & (\mu_e(t)) \\ \sum_{m=2}^M H_{im} Q_m(t) \leq \Phi_i \quad \forall (i, t) & (\eta_i(t)) \end{cases} \end{aligned} \quad (8)$$

**Prix nodaux** Les prix  $p_m(t)$  à l'optimum sont définis par les conditions d'équilibre local

$$p_m(t) = P_m(Q_m^D(t), t) = c_m(Q_m^S(t)) \quad (9)$$

pour tout  $m \geq 1$ , et les conditions d'équilibre global

$$\begin{cases} p_1(t) = \mu_e(t) \\ p_m(t) = \mu_e(t) \left(1 - \frac{\partial L}{\partial Q_m}\right) - \sum_{i=1}^I H_{im} \eta_i(t) \text{ pour } m > 1 \end{cases} \quad (10)$$

Les conditions d'équilibre local (9) imposent que, pour chaque marché  $m$ , le prix est égal au coût marginal de la production et au surplus marginal de la consommation.

Les conditions d'équilibre global (10) prennent en compte des externalités de réseau: les pertes thermiques et la congestion. Elles sont définies relativement au marché de référence, i.e., aucune externalité n'est prise en compte dans le prix du marché 1, qui est le prix de référence global  $\mu_e(t)$ .

Pour un marché  $m > 1$ , considérons tout d'abord la perspective du consommateur. Il paye prix de référence global  $\mu_e(t)$ , plus sa contribution

marginale aux pertes  $\frac{\partial L}{\partial Q_m^d} = -\frac{\partial L}{\partial Q_m}$ , où les pertes sont valorisées au prix de référence global  $\mu_e(t)$ , plus pour chaque ligne  $i$  sa contribution à la congestion sur cette ligne  $\frac{\partial \varphi_i}{\partial Q_m^d} = -\frac{\partial \varphi_i}{\partial Q_m} = -H_{im}$ , valorisée au coût (virtuel) de la congestion  $\eta_i(t)$ . De façon symétrique, un producteur vend son énergie au prix de référence  $\mu_e(t)$ , moins sa contribution marginale aux pertes  $\frac{\partial L}{\partial Q_m^s} = \frac{\partial L}{\partial Q_m}$ , valorisée à  $\mu_e(t)$ , moins sa contribution à la congestion sur chaque ligne  $i$   $\frac{\partial \varphi_i}{\partial Q_m^s} = \frac{\partial \varphi_i}{\partial Q_m} = H_{im}$ , valorisée à  $\eta_i(t)$ .

Chaque acteur fait ainsi face aux externalités marginales qu'il crée par ses décisions. En concurrence parfaite, la mise en oeuvre de prix nodaux décentralise l'optimum social. En d'autres termes, si les acteurs font face à d'autres prix que les prix nodaux, le surplus net est diminué.

**Surplus de l'opérateur de marché** Les prix étant différents d'un marché à l'autre, l'opérateur de marché réalise un profit. Chaque producteur vend et chaque consommateur achète à son prix local  $p_m(t)$ . L'opérateur de marché "achète" donc la production et "vend" la consommation à  $p_m(t)$ . Son surplus est donc

$$MS(t) = \sum_{m=1}^M (Q_m^D(t) - Q_m^S(t)) p_m(t) = - \sum_{m=1}^M Q_m(t) p_m(t).$$

Schweppe et al. (1988) montrent que

$$MS(t) = \mu_e(t) L(\mathbf{Q}(t)) + \sum_{i=1}^I \eta_i(t) K_i.$$

La mise en oeuvre des prix nodaux génère donc un surplus toujours positif, égal à la valeur des pertes (valorisés à  $\mu_e(t)$ ) plus la valeur de la congestion sur l'ensemble des lignes du réseau.

### 2.2.2 L'exemple de deux marchés reliés par une interconnexion

Considérons deux marchés,  $m = 1, 2$ . Pour simplifier, supposons que (i) la production dans le marché 1 est toujours moins chère que dans le marché 2:  $c_1(q) \leq c_2(q)$ , et (ii) les pertes thermiques sont négligeables. L'électricité se déplace donc du marché 1 vers le marché 2. La capacité de transfert de l'interconnexion dans le sens 1 vers 2 est dénotée  $\Phi^+$ .

Si  $\Phi^+$  est très large, l'interconnexion n'est jamais contrainte, ainsi que

représenté sur la Figure 1. Le prix  $p(t)$ , identique dans les deux marchés est donc:

$$c_1(Q_1^s(t)) = P(Q_1^d(t), t) = c_2(Q_2^s(t)) = P(Q_2^d(t), t).$$

La loi des noeuds (équilibre offre-demande) du marché 1 donne

$$Q_1^s(t) = Q_1^d(t) + \varphi(t) \Leftrightarrow \varphi(t) = Q_1^s(t) - Q_1^d(t) = Q_1(t).$$

La loi des noeuds appliquée au marché 2 donne

$$\varphi(t) = -Q_2^s(t) + Q_2^d(t) = -Q_2(t).$$

En prenant le marché 1 comme marché de référence,  $H_2 = -1$ . L'interconnexion n'est pas congestionnée tant que  $\varphi(t) \leq \Phi^+$ . Lorsque l'interconnexion est congestionnée (Figure 2), le plan de production est contraint par la capacité de transfert. Le prix  $p_1(t)$  est défini par

$$p_1(t) = c_1(Q_1^s(t)) = P(Q_1^d(t), t) \text{ et } Q_1^s(t) - Q_1^d(t) = \Phi^+,$$

et le prix  $p_2(t)$  par

$$p_2(t) = c_2(Q_2^s(t)) = P(Q_2^d(t), t) \text{ et } Q_2^d(t) - Q_2^s(t) = \Phi^+.$$

L'équation (10) donne

$$\begin{cases} \mu_e(t) = p_1(t) \\ p_2(t) = p_1(t) + \eta(t) \end{cases}.$$

La valeur de la congestion est donc

$$\eta(t) = p_2(t) - p_1(t) = c_2(Q_2^s(t)) - c_1(Q_1^s(t)).$$

Augmenter la capacité de la ligne d'un MW permet de substituer un MW de production utilisant la technologie 1 à un MW de production utilisant la technologie 2, donc d'économiser la différence  $(c_2(Q_2^s(t)) - c_1(Q_1^s(t)))$ .

### 2.3 Redispatching/counter-trading

Une autre méthode de gestion de la congestion est le “redispatching”, parfois appelé “counter-trading”. Cette approche produit le même plan de production que les prix nodaux, avec des transferts différents.

Pour simplifier, reprenons l'exemple simple de deux marchés reliés par une interconnexion. L'opérateur de marché reçoit toutes les offres (achats et ventes), et vérifie la compatibilité de ces offres avec la capacité de l'interconnexion. Si les flux résultants sont inférieurs à la capacité de l'interconnexion, toutes les offres sont acceptées (Figure 1).

Si les flux résultants excèdent la capacité de la ligne, l'opérateur de marché doit modifier les transactions pour ajuster les flux. Il achète de l'électricité aux producteurs du marché 2 (et aux consommateurs, si c'est possible), afin d'augmenter la production dans le marché 2 (et réduire la demande), et réduire les importations. En parallèle, il réduit la production dans le marché 1 (et augmente la demande si c'est possible), afin de réduire les exportations du même volume que la réduction des importations. L'opérateur effectue ces ajustements jusqu'à ce que les flux soient égaux à la capacité de la ligne. Le plan de production est alors identique à celui généré par la mise en oeuvre des prix zonaux (Figure 3).

En revanche, les transferts sont différents. Si les prix zonaux sont mis en oeuvre, l'opérateur achète et les producteurs vendent au prix  $p_1$  dans le marché 1, alors que les consommateurs achètent à  $p_2 > p_1$  dans le marché 2. L'opérateur reçoit donc un profit positif, la rente de congestion ou le “merchandizing surplus”. Si les marchés sont couplés, l'opérateur doit compenser les producteurs (et les consommateurs) pour leurs ajustements. Cette compensation est socialisée au travers d'une charge payée par tous les consommateurs.

Le “counter trading” est l'approche retenue pour gérer la congestion en France. Rte effectue dans le marché d'ajustement le redispatching indiqué ci-dessus, planifié dès le marché ( $J - 1$ ). Le “counter trading” est aussi l'approche retenue pour gérer la congestion dans la plupart des pays européens.

## 2.4 Gains pour la collectivité générés par les prix nodaux

Plusieurs analyses académiques ont documenté les gains importants pour la collectivité générés par la mise en oeuvre de prix nodaux. Green (2007) estime que la mise oeuvre de prix zonaux accroîtrait le surplus de 1.3% en Angleterre comparé au système actuel (redispatching, décrit plus loin) si le marché est parfaitement concurrentiel, et de 3.1% si les producteurs exercent leur pouvoir de marché. Plus récemment, Neuhoff et al. (2013) examinent l'impact de la mise en oeuvre de prix nodaux sur la plaque européenne, en considérant différents scénarii de pénétration des renouvelables. L'étude estime les gains réalisés comparés au status quo (prix zonaux par pays et redispatching au sein de chaque pays, décrit plus loin). Elle produit trois résultats principaux. Le passage au prix nodaux permet de:

1. accroître les échanges aux frontières (en MW) jusqu'à 34%,
2. réduire les coûts d'opération de 0.8 à 2.0 milliards d'Euros annuellement (1.1 à 3.3%), dépendant de la pénétration des renouvelables, et
3. réduire le prix moyen dans 60 à 75% des pays considérés.

## 2.5 La mise en oeuvre des prix nodaux

Ainsi que mentionné précédemment, les prix nodaux ne sont pas l'approche utilisée pour gérer la congestion en France (et dans la plupart des pays européens). Leur mise en oeuvre requiert donc un changement important dans l'opération du marché de gros. Les prix nodaux ayant été mis en oeuvre aux Etats Unis depuis bientôt 20 ans, cette transition ne présenterait pas d'importantes difficultés techniques et informatiques.

En revanche, la mise en oeuvre des prix nodaux requerrait un certain nombre de modifications à l'organisation des marchés, décrites dans cette Section.

### 2.5.1 Assurance: les droits financiers de transport

Le prix payé par les agents (producteurs et consommateurs) est incertain, pour deux raisons: premièrement, le prix de référence  $\mu_e(t)$  est incertain, comme l'est aujourd'hui le prix de gros sur le marché français. Les acteurs peuvent s'assurer contre ce risque en achetant (ou en vendant) des produits à terme, physiques ou financiers.

Deuxièmement, la différence entre le prix de leur marché local et le prix de référence est elle aussi incertaine, ce qui crée un “basis risk”, i.e., un risque différentiel. Dans une contribution séminale, William Hogan (1992) suggère que les acteurs peuvent se protéger contre cette incertitude en vendant (ou achetant) des produits financiers à terme, qui payent la différence entre les prix nodaux pour chaque unité d’assurance achetée. Ces “droit financiers de transport” (Financial Transmission Rights, FTR) sont mis aux enchères par le l’opérateur du marché, et parfaitement transférables.

Ils assurent par exemple un producteur du marché 1 qui vend à un client dans le marché 2 que son prix de vente sera bien celui du contrat. La transaction se décompose en effet en (i) production à coût  $c$  dans le marché 1, (ii) vente dans le marché 1 au prix  $p_1(t)$ , (iii) achat dans le marché 2 au prix  $p_2(t)$ , et (iv) vente au client final au prix  $p$ . Le profit unitaire du producteur/fournisseur est donc

$$\pi = p_1(t) - c + p - p_2(t) = p - c + p_1(t) - p_2(t).$$

Le producteur/fournisseur est donc exposé à la différence entre les prix nodaux. S’il dispose d’un FTR, qui paye  $(p_2(t) - p_1(t))$ , son profit unitaire devient

$$\pi = p - c + p_1(t) - p_2(t) + p_2(t) - p_1(t) = p - c.$$

Le profit est donc assuré contre les fluctuations des prix nodaux.

L’opérateur du marché paye aux détenteurs la valeur de leurs FTRs, et, ainsi que montré plus haut, reçoit un surplus résultant de l’achat et de la vente d’énergie à des prix différents. Hogan (1992) et Bushnell and Stoft (1996) montrent que, si la somme des FTRs mis aux enchère est faisable (i.e., si elle est cohérente avec la topographie du réseau), le second est toujours plus important que le premier: l’opérateur de marché peut toujours couvrir les paiements associés aux FTRs.

### 2.5.2 Redistribution

Les prix nodaux maximisent le surplus total. Passer d’un modèle de gestion de la congestion par du redispatching aux prix nodaux crée de réelles différences: les consommateurs dans les zones contraintes à l’import et les

producteurs dans les zones contraintes à l'export verront respectivement leurs factures augmenter et leurs prix de vente baisser. C'est précisément l'effet recherché, qui génère les bonnes incitations. Pour faciliter la transition, donc l'adoption, des transferts directs peuvent être mis en place, qui s'atténuent progressivement.

### 2.5.3 Compatibilité avec l'architecture de marché "européenne"

Les prix nodaux sont mis en oeuvre dans les marchés des Etats Unis, qui sont centralisés. L'opérateur de marché est aussi l'opérateur du réseau. Il reçoit toute les offres (production et consommation), et calcule en  $(J - 1)$  le plan de production optimal.

L'architecture des marchés européens est différente: les opérateurs de marché (et les responsables d'équilibre) produisent des plans de production/consommation équilibrés, qu'ils soumettent au GRT. Ce dernier ne contrôle donc pas entièrement le plan de production.

Cette architecture ne constitue pas un obstacle insurmontable à la mise en oeuvre de prix nodaux. Aujourd'hui, lorsqu'il reçoit les offres, le GRT vérifie leur compatibilité avec les contraintes réseau, et, le cas échéant, modifie le plan de production pour gérer les contraintes (ainsi que présenté dans la Section 2.3). Si les prix nodaux étaient mis en oeuvre, le GRT vérifierait la compatibilité des offres avec les contraintes réseau, et, le cas échéant, modifierait en  $(J - 1)$  le plan de production pour gérer les contraintes, et déclarerait des prix nodaux différents. Ainsi qu'expliqué précédemment, les utilisateurs qui le désirent pourraient acheter des droits financiers de transport pour se protéger de la congestion.

### 2.5.4 Compatibilité avec le couplage des marchés européens

La congestion sur les réseaux de transport européens est gérée par la méthode des "prix zonaux". Pour simplifier, chaque pays est une zone. Entre pays, la congestion est gérée selon l'algorithme présenté dans la Section 2.2. Des prix "zonaux" sont donc calculés. Les prix dans deux marchés adjacents sont identiques tant que l'interconnexion entre eux n'est pas contrainte, ils diffèrent dès que l'interconnexion est contrainte<sup>2</sup>. Au sein de chaque zone, la

---

<sup>2</sup>Dans certains cas, les lois de la physique font que prix entre deux marchés différent même si l'interconnexion n'est pas contrainte.

congestion est gérée selon la méthode du “counter-trading”, présentée dans la Section 2.3. Cette approche, dénommée couplage de marchés, s’étend progressivement à l’ensemble des pays de l’Union Européenne (Figure 4).

Il est possible pour un pays de modifier son approche de gestion interne de la congestion. Dans ce cas, il lui faut constituer un prix synthétique qui correspond à sa zone dans l’algorithme de couplage externe.

## **2.6 Prix nodaux, droits de transport, et pouvoir de marché**

La littérature académique a examiné les interactions entre le pouvoir de marché des producteurs, les contraintes sur les réseaux, et les approches mises en oeuvre pour les gérer.

### **2.6.1 Pouvoir de marché local**

La présence de congestion sur le réseau de transport génère un pouvoir de marché local. Un producteur peut être trop petit pour influencer le marché national (ou régional), mais lorsque le réseau est contraint et que le marché est divisé en sous-marchés, il peut se trouver en position d’exercer un pouvoir de marché dans son sous-marché local. Les producteurs peuvent exercer leur pouvoir de marché local quelle que soit la méthode de gestion de la congestion.

Si la congestion est gérée par les prix nodaux, l’analyse économique classique s’applique: un producteur situé dans un sous-marché contraint à l’import peut réduire sa production (ou augmenter son prix de vente) pour accroître son profit. Le problème (pour le producteur) est la prise en compte des effets réseaux: en modifiant sa production, il modifie les flux sur le réseau, donc la congestion. Cet effet a été étudié dans de nombreux articles académiques. Par exemple Borenstein et al. (2000) examinent l’équilibre dans un réseau radial. Cardell et al. (1997) montrent que dans certains cas, les producteurs augmentent leur production pour augmenter la congestion et accroître leur profit.

Si la congestion est gérée par le redispatching/counter trading, un producteur situé dans un sous-marché contraint à l’import peut lui aussi annoncer un prix plus élevé que son coût marginal, s’il anticipe d’être inclus dans l’ordre de mérite sans contrainte. Qu’il soit appelé à produire, ou que le GRT lui rachète sa production, il bénéficiera ainsi d’un prix plus élevé.

Si la congestion est gérée par le redispatching/counter trading, une autre manipulation est possible. Un producteur dans le marché contraint à l'export a intérêt à annoncer un coût très bas. Si le GRT doit lui racheter sa production, il recevra la différence entre prix sans contrainte et le coût qu'il a annoncé. Ces stratégies, connues sous le nom de jeu de réduction ("Dec game", Green, 2007, Holmberg et Lazarczyk, 2012) ont été mises en oeuvre par certains producteurs durant la crise Californienne.

Quelle méthode de gestion de la congestion (prix nodaux ou redispatching/counter trading) est la plus robuste à l'exercice du pouvoir de marché? L'analyse académique suggère que les prix nodaux sont préférables. L'intuition pour ce résultat est que le redispatching/counter-trading crée une opportunité de manipulation supplémentaire, le "Dec game". Ce résultat est illustré sur quelques exemples simple par Harvey et Hogan (2000), et démontré dans une analyse théorique récente (Holmberg et Lazarczyk, 2012), sous des hypothèses très générales. L'analyse empirique de la Grande Bretagne (Green, 2007) déjà citée confirme ce résultat théorique. Elle montre que, si le pouvoir de marché des producteurs est pris en compte, passer du redispatching au prix zonaux augmente le surplus net de 3.1% des revenus des producteurs: 1.3% correspond au gain en concurrence parfaite, et 1.8% correspond à la réduction du pouvoir de marché des producteurs.

### **2.6.2 Droits de transport en présence de pouvoir de marché**

Une autre branche de la littérature académique a examiné l'interaction entre le pouvoir de marché des producteurs et les droits financiers de transport. Joskow et Tirole (2000) démontrent qu'un producteur en monopole (ou un consommateur en monopsonne) peut augmenter son pouvoir de marché en utilisant des droits de transmission. Ils montrent également que les droits financiers sont généralement préférables aux physiques de transmission : un producteur détenant des droits physiques pourrait augmenter ses profits en retenant le marché une partie de la capacité acquise. Ce ne serait pas possible avec des droits financiers. Léautier (2001) montre comment les incitations à augmenter la capacité du réseau (discutées ci-après) sont modifiées par le pouvoir de marché des producteurs.

## Part II

# Efficacité à long-terme

Les problématiques d'investissement dans les marchés de l'électricité sont extrêmement complexes, en particulier la coordination entre investissement dans les moyens de production (et centres de consommation) d'une part, et dans les réseaux d'autre part.

La première difficulté est la temporalité: l'investissement en production précède-t-il ou suit-il l'investissement dans les réseaux? La réalité est multiforme. Les ajouts majeurs aux réseaux demandent une longue période de gestation, de l'ordre de la dizaine d'année, alors que la mise en service de centrales au gaz demande trois ou quatre ans à peine. D'un autre côté, le développement d'une centrale nucléaire prend plus de dix ans, alors que de nouveaux moyens de production peuvent être raccordés en quelques mois.

La littérature académique de la fin des années 1990s (e.g., Léautier, 2000) ne mettait pas l'accent sur la problématique de l'investissement en production. Elle considérait implicitement que l'investissement dans les réseaux était plus rapide que celui en production. Cette perspective ne semble plus valide aujourd'hui. Depuis le début des années 2000, de très importants volumes d'investissement (et le dés-investissement) ont profondément transformé les parcs de productions dans plus marchés européens et nord américains, alors que le réseau évoluait assez lentement. De nombreux académiques considèrent donc désormais que le plan d'expansion du réseau est fixe, et que le parc de production s'adapte (Baldick et al., 2011).

Cette note considère un processus d'investissement dynamique: le parc de production s'adapte au réseau existant, puis le réseau s'adapte au parc de production, puis le parc de production s'adapte aux évolutions du réseau, etc. Si la concurrence et la régulation sont parfaites, l'équilibre de long-terme obtenu est précisément l'optimum social (défini plus bas).

Deuxième difficulté, les investissements en production et en transport sont parfois complémentaires (la valeur de la production augmente avec la capacité de transport), parfois substituts (la valeur de la production diminue avec la capacité de transport).

La difficulté finale est la couverture des coûts du réseau. En première

approximation, la production d'électricité exhibe des rendements d'échelle constants: un système de prix couvrant le coût marginal couvre aussi le coût moyen. Ca n'est pas le cas pour les réseaux, qui exhibent des rendements d'échelle croissants. La couverture des coûts marginaux ne garantit pas la couverture des coûts moyens (e.g., Bell et al., 2011). Une tarification résiduelle, qui couvre les coûts fixes est à déterminer. Cet aspect est abordé dans la partie III.

### 3 Les prix nodaux, signaux d'investissement optimaux

#### 3.1 Investissement en production

La théorie de la tarification heure de pointe s'applique naturellement aux marchés locaux. Pour simplifier, considérons qu'une seule technologie est disponible dans chaque marché  $m$ , caractérisée par un coût marginal de production  $c_m$  et le coût marginal de capacité  $r_m$

##### 3.1.1 Programme d'optimisation

L'optimum social est donc caractérisé par

$$\begin{aligned} & \max_{Q_m^S(t), Q_m^D(t), k_m} \mathbb{E} \left[ \sum_{m=1}^M \{U_m(Q_m^D(t)) - c_m Q_m^S(t)\} \right] - \sum_{m=1}^M r_m k_m \\ \text{st : } & \begin{cases} Q_m^S(t) \leq k_m \quad \forall (m, t) & (\lambda_m(t)) \\ \sum_{m=1}^M Q_m^S(t) = \sum_{m=1}^M Q_m^D(t) + L(\mathbf{Q}(t)) \quad \forall t & (\mu_e(t)) \\ -\Phi_i^- \leq \sum_{m=2}^M H_{im} Q_m(t) \leq \Phi_i^+ \quad \forall (i, t) & (\eta_i(t)) \end{cases} \end{aligned} \quad (11)$$

Ce programme d'optimisation est la combinaison du programme d'optimisation (8) conduisant aux prix nodaux, et du programme de tarification heure de pointe pour multiples technologies de production (3). Les coûts et les contraintes de capacité de production présentes dans le programme (3) ont été ajoutés au programme (8). La structure de la solution est inchangée. La seule différence est que la forme du coût marginal de long-terme  $c_m(Q_m^S(t))$  est précisée.

**Prix nodaux** Les prix nodaux sont définis par les conditions d'équilibre local (9). Les coûts marginaux étant constants, les conditions (9) s'écrivent

$$p_m(t) = \begin{cases} c_m & \text{if } Q_m^S(t) < k_m \\ c_m + \lambda_m(t) & \text{if } Q_m^S(t) = k_m \end{cases},$$

où  $\lambda_m(t)$  est le multiplicateur associé à la contrainte  $Q_m^S(t) \leq k_m$ .

**Capacité de production optimale** Dans le marché  $m$ , la capacité de production optimale satisfait

$$\mathbb{E}[(p_m(t) - c_m)^+] = r_m.$$

En général,  $p_m(t)$  dépend du vecteur des capacités  $\mathbf{K} \in \mathbb{R}^M$ . Le vecteur  $\mathbf{K}^* \in \mathbb{R}^M$  des capacités optimales est donc défini de manière unique par les conditions:

$$\mathbb{E}[(p_m(\mathbf{K}^*, t) - c_m)^+] = r_m \quad \forall m. \quad (12)$$

Ces relations correspondent à l'optimum social, mais aussi à la condition d'entrée libre. La mise en oeuvre de prix zonaux conduit donc au parc optimal dans chaque marché (compte tenu des technologies disponibles dans chaque marché). Ce résultat est une application d'un résultat plus général en économie: l'optimalité de court terme conduit à l'optimalité de long-terme, i.e., si les prix reflètent correctement les conditions de court terme, ils produisent les bons signaux d'investissement.

**Conditions nécessaires à l'investissement optimal** Il n'est pas inutile de rappeler les conditions nécessaires à ce résultat. Premièrement, la concurrence entre investisseurs dans les moyens de production doit être parfaite. Cette hypothèse est rarement vérifiée en pratique. Toutefois, dans de nombreux marchés, suffisamment d'investisseurs sont présents pour que l'industrie soit considérée quasi-concurrentielle. De plus, les autorités de la concurrence sont créées précisément pour s'assurer que les industries sont concurrentielles.

Deuxièmement, la construction de nouveaux moyens de production doit être libre. Là encore, les autorités de la concurrence et les régulateurs sectoriels sont responsables de garantir qu'aucune barrière n'est érigée à l'entrée de nouveaux moyens de production.

Troisièmement, les investisseurs doivent avoir une bonne connaissance du réseau et du parc de production existant, et une bonne anticipation des développements futurs planifiés dans les réseaux et dans le parc de production. C'est le cas aujourd'hui en Europe: les transporteurs publient leur plan décennal d'investissement, coordonnés par l'Entsoe<sup>3</sup>. Les producteurs annoncent à leurs investisseurs leurs plans de développement. Des firmes de conseil spécialisées consolident ces annonces.

**Décisions d'investissement en pratique** En pratique, chaque investisseur se forge sa propre opinion sur la probabilité que ces investissements aient effectivement lieu, i.e., sa propre anticipation du réseau et du parc de production futurs. Il incorpore ces anticipations dans des modèles de dispatching et d'investissement multi-régions (i.e., des versions plus sophistiquées du programme (11)), qui lui permettent de forger son anticipation des prix futurs dans plusieurs géographies. Il utilise ensuite ces anticipations de prix pour décider (ou non) d'investir, selon un critère ressemblant à l'équation (12).

Le processus décrit ci-dessus est le processus "standard" de détermination des prix futurs, qui s'applique dans n'importe quelle industrie. La mise en oeuvre des prix nodaux permet aux investisseurs d'augmenter la granularité géographique de leurs décisions, par exemple, de distinguer une centrale en Auvergne d'une centrale dans les Alpes Maritimes.

Finalement, il n'existe aucune garantie que l'investissement optimal *ex ante* le soit *ex post*: les anticipations conduisant à la décision d'investissement peuvent s'avérer fausses. Par exemple, durant les années 2005-2010, les producteurs européens ont investi massivement dans des centrales cycles combinés gaz, anticipant une forte croissance de la demande en Europe, et des prix des matières premières (gaz, charbon,  $CO_2$ ) favorables. La crise économique de 2008 et le développement massif de renouvelables ont invalidé ces anticipations. *Ex post*, ces n'ont pratiquement aucune valeur économique, et doivent donc être fermées.

---

<sup>3</sup>European network of transmission system operators for electricity. Le plan de développement décennal est disponible à <https://www.entsoe.eu/about-entso-e/system-development/ten-year-network-development-plan/Pages/default.aspx>

### 3.1.2 Application à deux marchés reliés par une interconnexion

Léautéier (2013) propose une solution analytique pour deux marchés séparés par une interconnexion. Supposons que la technologie disponible dans le marché 1 (caractérisée par  $(c_1, r_1)$ ) est moins chère que la technologie disponible dans le marché 2, i.e.,  $c_1 < c_2$ , et  $r_1 > r_2$ . Supposons aussi que les consommateurs dans les deux marchés sont identiques, et dénotons  $\theta_m$  la masse de consommateurs dans le marché  $m$ . Finalement, dénotons  $K_m^U$  la capacité optimale cumulée jusqu'à la technologie  $m$ , lorsque le réseau n'est jamais contraint. Les résultats de la Section 2.1 montrent que  $K_2^U$  est défini de manière unique par l'équation (4) et  $K_1^U$  par l'équation (5). Supposons finalement que  $\theta_1 K_2^U \leq \theta_2 K_1^U$ . Cette hypothèse technique simplifie la présentation, sans modifier l'intuition.

La capacité sur l'interconnexion dans le sens marché 1 vers marché 2 (resp. 2 vers 1) est dénotée  $\Phi^+$  (resp.  $\Phi^-$ ).

Léautéier (2013) montre les résultats suivants:

1. Si la capacité de l'interconnexion est suffisamment large, spécifiquement si  $\Phi^+ \geq \frac{\theta_2}{\theta_1 + \theta_2} K_1^U$ , l'interconnexion n'est jamais contrainte.
2. Si la capacité de l'interconnexion est intermédiaire, spécifiquement si  $\Phi^+ < \frac{\theta_2}{\theta_1 + \theta_2} K_1^U$  et  $\Phi^+ + \Phi^- \geq \frac{\theta_1}{\theta_1 + \theta_2} K_2^U$ , l'interconnexion est contrainte dans le sens marché 1 vers marché 2 pour certains états du monde. La congestion disparaît lorsque la demande est très élevée. La capacité cumulée optimale est celle sans congestion:  $K_2^* = K_2^U$ . La capacité optimale dans le marché 1 est

$$K_1^* = \frac{\theta_1}{\theta_1 + \theta_2} K_1^U + \Phi^+$$

3. Si la capacité de l'interconnexion est très faible, spécifiquement si  $(\Phi^+ + \Phi^-) < \frac{\theta_1}{\theta_1 + \theta_2} K_2^U$ , l'interconnexion est successivement en contrainte dans les deux directions. Comme dans le cas précédent, l'interconnexion est contrainte dans le sens marché 1 vers marché 2, puis cette contrainte disparaît lorsque la demande augmente. Cependant, lorsque la demande est très élevée, l'interconnexion est contrainte dans le sens 2 vers 1.

Ce résultat illustre bien l'impact de la contrainte. Pour la zone intermédiaire, la capacité installée à l'équilibre concurrentiel dans le marché 1 croit avec la capacité de l'interconnexion. La contrainte pénalise la production installée dans le marché 1, ( $K_1^* < K_1^U$ ) et avantage donc la production installée dans le marché 2, puisque la capacité cumulée est constante.

Cet exemple illustre aussi la nature changeante des flux et contraintes sur les interconnexions. Les analystes pensent parfois que, si le marché 1 bénéficie d'une technologie moins chère, l'interconnexion sera contrainte en permanence dans le sens marché 1 vers marché 2. Cette intuition est fautive, qui ignore les coûts fixes. Le prix dans le marché 1 doit augmenter pour couvrir le coût marginal de capacité  $r_1$ . Il dépasse donc le coût marginal de production  $c_2$ . La technologie 2 devient alors marginale pour les deux marchés, et la congestion disparaît.

Si la somme des capacités d'interconnexion est très faible, l'interconnexion peut aussi être contrainte dans le sens marché 2 vers marché 1. Cet exemple n'est pas simplement théorique: l'interconnexion France-Angleterre est contrainte dans les deux sens à différents moments (Figure 5).

### 3.2 Investissement optimal dans les réseaux

La première difficulté dans la définition du réseau optimal est la détermination des coûts du réseau. L'augmentation de la capacité sur une ligne donnée modifie la capacité de transfert de cette ligne, mais aussi de nombreuses lignes. Donc, de nombreuses solutions permettent d'augmenter la capacité sur une interface, certaines requérant des actions sur d'autres lignes. La fonction de coût résultante est le coût minimum sur toutes les solutions possibles (Boyer et al., 2006 donnent un exemple très clair pour un réseau de gaz naturel). Elle ne présente pas toujours les propriétés mathématiques désirées (Hogan et al., 2010). Par exemple, dans certains cas, augmenter la capacité d'une ligne *réduit* la capacité d'une autre. La fonction de coût résultante n'est donc pas nécessairement convexe, et l'optimum n'est pas forcément unique. Dans cette introduction, cette difficulté n'est pas prise en compte. On suppose que la fonction  $\Gamma(\Phi)$  est bien définie et satisfait aux hypothèses standard.

### 3.2.1 Caractérisation de l'optimum de long terme

L'optimum social est donc caractérisé par

$$\begin{aligned}
 & \max_{Q_m^S(t), Q_m^D(t), k_m, \Phi_i} \mathbb{E} \left[ \sum_{m=1}^M \{U_m(Q_m^D(t)) - c_m Q_m^S(t)\} \right] - \sum_{m=1}^M r_m k_m - \Gamma(\Phi) \\
 & \text{st : } \begin{cases} Q_m^S(t) \leq k_m \quad \forall m & (\lambda_m(t)) \\ \sum_{m=1}^M Q_m^S(t) = \sum_{m=1}^M Q_m^D(t) + L(\mathbf{Q}(t)) & (\mu_e(t)) \\ \sum_{i=1}^m H_{im} Q_m(t) \leq \Phi_i \quad \forall i & (\eta_i(t)) \end{cases} .
 \end{aligned} \tag{13}$$

Le programme d'optimisation est le programme (11), auquel a été ajoutée la détermination des capacités  $\Phi_i$ . En poursuivant l'analyse des prix nodaux, Léautier (2000) montre que le réseau optimal satisfait

$$\mathbb{E}[\eta_i(t)] = \frac{\partial \Gamma(\Phi^*)}{\partial \Phi_i} \quad \forall i. \tag{14}$$

La valeur d'une ligne est l'espérance du coût virtuel de la contrainte sur cette ligne, i.e., le coût espéré de la congestion. Le résultat est standard en optimisation. On le retrouve par exemple dans la détermination la capacité de production optimale dans le problème de la tarification heure de pointe présenté dans la Section 2.1.

La relation (14) indique qu'à l'optimum, le coût espéré de la congestion sur une ligne est égal au coût marginal de la capacité sur la ligne. Donc, à l'optimum, toute ligne doit être contrainte dans certains états du monde. Comme le coût marginal de la capacité d'interconnexion est positif, une ligne qui ne serait contrainte dans aucun état du monde aurait une capacité trop élevée. Cela ne signifie pas que chaque ligne doit être contrainte quelques heures chaque année. Cela signifie qu'une ligne qui ne serait contrainte pour aucun des scénarii envisagés lors de l'analyse du développement du réseau serait surdimensionnée.

Pour comprendre les relations (14), considérons le processus d'investissement séquentiel et dynamique, ainsi que décrit plus haut. Pour une configuration de réseau  $\Phi$  donné, les producteurs investissent les capacités  $K_m(\Phi)$ . Les prix nodaux sont donc une fonction du vecteur  $\mathbf{K}^*(\Phi) \in \mathbb{R}^M$  de ces capacités. Les coûts virtuels  $\eta_i(t)$  étant une combinaison linéaires de ces prix nodaux, ils sont aussi une fonction de  $\mathbf{K}^*(\Phi)$ . Le vecteur  $\Phi^*$  décrivant le

réseau optimal est donc caractérisé par les équations (14)

$$\mathbb{E} [\eta_i ((\Phi^*), t)] = \frac{\partial \Gamma (\Phi^*)}{\partial \Phi_i} \quad \forall i.$$

### 3.2.2 Application à deux marchés reliés par une interconnexion

Supposons que le coût de l'interconnexion est

$$\Gamma (\Phi) = F + \gamma \Phi,$$

où  $F$  représente les coûts fixes de déploiement de l'interconnexion, et  $\gamma$  le coût variable de la capacité, supposé constant. En négligeant les pertes, l'équation (14) devient:

$$\mathbb{E} [(p_2 (t) - p_1 (t))^+] = \gamma$$

Supposons que les coûts sont tels que la capacité de l'interconnexion est dans la zone intermédiaire à l'optimum. Léautier (2013) montre, que pour une capacité  $K_1$  fixée, l'équation (14) devient

$$\beta_0 \left( \frac{\theta_1 + \theta_2}{\theta_2} \Phi^+, c_1, c_2 \right) - \beta_0 \left( \frac{\theta_1 + \theta_2}{\theta_1} (K_1 - \Phi^+), c_1, c_2 \right) = \gamma. \quad (15)$$

Si la capacité de production dans le marché 1 est optimale,  $K_1 = K_1^* (\Phi^+) = \frac{\theta_1}{\theta_1 + \theta_2} K_1^U + \Phi^+$ , l'interconnexion optimale  $\Phi^{+*}$  est l'unique solution de

$$\beta_0 \left( \frac{\theta_1 + \theta_2}{\theta_2} \Phi^{+*}, c_1, c_2 \right) = \beta_0 (K_1^U, c_1, c_2) + \gamma. \quad (16)$$

### 3.2.3 Difficultés pratiques

En pratique, la mise en oeuvre de cette règle simple est difficile. Premièrement, les accroissements de capacité ne sont pas continus, mais discrets (“lumpy”). Il est donc possible qu'une ligne soit trop importante, car il n'est pas possible de dimensionner précisément la ligne. Par exemple, supposons que l'optimum soit une ligne de 2 650 MW, alors que le GRT a le choix entre une ligne de 2 000 MW et une ligne de 4 000 MW. Il installera probablement la ligne de 4 000 MW, qui ne sera pas assez congestionnée.

Deuxièmement, ainsi que mentionné précédemment, la fonction  $\Gamma (\Phi)$  ne possède pas nécessairement de bonnes propriétés de régularité.

### 3.2.4 Demande non réactive au prix

Historiquement, les planificateurs de systèmes électriques considéraient que les demandes  $Q_m^D(t)$  étaient données, et non pas dépendantes du prix. Le programme de maximisation du surplus se simplifie donc en un programme de minimisation des coûts:

$$\begin{aligned} & \min_{Q_m^S(t), k_m, \Phi_i} \hat{\mathbb{E}} \left[ \sum_{m=1}^M c_m Q_m^S(t) \right] + \sum_{m=1}^M r_m k_m + \Gamma(\Phi) \\ \text{st : } & \begin{cases} Q_m^S(t) \leq k_m \quad \forall m & (\lambda_m(t)) \\ \sum_{m=1}^M Q_m^S(t) = \sum_{m=1}^M Q_m^D(t) + L(\mathbf{Q}(t)) & (\mu_e(t)) \\ \sum_{i=1}^m H_{im} Q_m(t) \leq \Phi_i \quad \forall i & (\eta_i(t)) \end{cases} \end{aligned} \quad (17)$$

En pratique la demande ne peut pas être servie dans tous les états du monde. Les pouvoirs publics établissent donc un critère de défaillance, i.e., limitent les états du monde sur lesquels la demande doit être servie. Dans le programme ci-dessus, les contraintes et l'espérance sont donc limitées à ces états du monde.

### 3.3 Incitations à l'investissement optimal

Léautier (2000) montre que si l'opérateur de réseau est responsable de la somme du (i) coût des pertes (valorisées à  $\mu_e(t)$  dans l'état du monde  $t$ ) et (ii) du coût du redispatching, il investit optimalement, i.e., le vecteur qui maximise son profit compte tenu de son exposition au coût des pertes et de la congestion est exactement  $\Phi^*$ . L'intuition pour ce résultat est que le coût du redispatching est une bonne mesure du coût des contraintes sur le réseau.

Hogan et al. (2010) proposent une autre approche pour inciter l'opérateur de réseau à entreprendre l'investissement optimal, qui s'appuie sur les FTRs créés par un expansion du réseau, donc sur les prix nodaux.

### 3.4 Parc de production et réseau optimal

En combinant les analyses précédentes, l'optimum social est caractérisé par les conditions (12) and (14). C'est le résultat obtenu par la planification centralisée, mais aussi par la concurrence libre en production, si les prix nodaux sont utilisés, et le réseau régulé de façon optimale. En effet, pour

un vecteur  $\Phi$  donné, les producteurs en concurrence parfaite dans le marché  $m$  investissent collectivement  $K_m^*(\Phi)$ . L'opérateur de réseau optimalement régulé adapte alors le réseau, ce qui engendre une modification du parc de production, jusqu'à ce que le réseau soit caractérisé par  $\Phi^*$  défini par les équations (14).

Ce résultat suppose évidemment que l'environnement (la demande, les coûts de production et d'investissement) sont constants dans le temps, ce qui n'est pas vérifié en pratique. Toutefois, il constitue un résultat théorique important.

### 3.5 Internalisation des coûts de réseau par les producteurs

Une propriété souhaitable de l'optimum est que les producteurs prennent en compte les coûts de réseau que leur installation engendrent. La tarification nodale/zonale vérifie cette propriété.

**Deux marchés reliés par une interconnexion** Considérons deux marchés reliés par une interconnexion, de capacité  $\Phi$ . Un producteur choisit d'investir (à la marge) en production dans le marché 1 ou le marché 2. Comme précédemment, produire (une unité) dans le marché 1 est moins cher:  $c_1 + r_1 < c_2 + r_2$ . D'un autre côté, ses revenus espérés sont  $\mathbb{E}[p_2(t)]$  s'il entre dans le marché 2,  $\mathbb{E}[p_1(t)]$  s'il entre dans le marché 1. Supposons que l'interconnexion est moyennement contrainte, donc  $p_2(t) \geq p_1(t)$  et  $\mathbb{E}[p_2(t)] > \mathbb{E}[p_1(t)]$ .

En achetant un FTR, le producteur peut obtenir les revenus  $\mathbb{E}[p_2(t)]$  et dépenser  $(c_1 + r_1)$  pour produire dans le marché 1. A l'équilibre, le prix unitaire du FTR est  $\mathbb{E}[p_2(t) - p_1(t)]$ . Tel que démontré précédemment, si l'interconnexion est optimalement dimensionnée, le prix unitaire du FTR est égal au coût marginal de capacité sur l'interconnexion  $\gamma$ .

Donc, afin d'obtenir les revenus  $\mathbb{E}[p_2(t)]$  le producteur choisit entre (i) dépenser  $(c_2 + r_2)$  pour produire dans le marché 2, et (ii) dépenser  $(c_1 + r_1 + \gamma)$  pour produire dans le marché 1 et transporter financièrement dans le marché 2. Si le réseau est optimal, la mise en oeuvre de prix nodaux conduit les producteurs à internaliser le coût marginal du réseau dans leur décision d'investissement.

**Configuration quelconque** Ce résultat se généralise à un réseau quelconque. Supposons que le coût du réseau est

$$\Gamma(\Phi) = F + \sum_{i=1}^I \gamma_i \Phi_i.$$

Un producteur choisit d'investir dans le marché  $n$  (coût de production  $c_n + r_n$ ) ou dans le marché  $m$  (coût de production  $c_m + r_m$ ). Supposons que le marché  $m$  est toujours plus attractif que le marché  $n$ :  $p_m(t) \geq p_n(t) \forall t$ . Si l'investisseur veut recevoir en moyenne le prix du marché  $m$ , il peut investir (i) dans le marché  $m$ , ou (ii) dans le marché  $n$  et acheter un droit financier de transport entre les marchés  $m$  et  $n$ . A l'équilibre, le prix unitaire de ce FTR est

$$\begin{aligned} \mathbb{E}[p_m(t) - p_n(t)] &= \mathbb{E} \left[ \mu_e(t) \left( \frac{\partial L}{\partial Q_n} - \frac{\partial L}{\partial Q_m} \right) \right] + \sum_{i=1}^I (H_{in} - H_{im}) \mathbb{E}[\eta_i(t)] \\ &= \mathbb{E} \left[ \mu_e(t) \left( \frac{\partial L}{\partial Q_n} - \frac{\partial L}{\partial Q_m} \right) \right] + \sum_{i=1}^I (H_{in} - H_{im}) \gamma_i \end{aligned}$$

en utilisant la définition (10) des prix nodaux, et les conditions d'optimalité du réseau (14). Une injection unitaire marginale dans le marché  $n$  et un retrait unitaire dans le marché  $m$  génère un flux marginal  $z_i = H_{in} - H_{im}$  sur la ligne  $i$ . Si la ligne  $i$  n'est pas contrainte, ce flux marginal ne crée pas de coût. Si la ligne est contrainte, ce flux marginal a un coût d'opportunité  $\mathbb{E}[\eta_i(t)] = \gamma_i$  si le réseau est optimal. Le coût marginal pour le réseau de cette injection unitaire donc est  $\sum_{i=1}^I (H_{in} - H_{im}) \gamma_i$ .

Pour recevoir  $\mathbb{E}[p_m(t)]$ , le producteur peut soit investir ( $c_m + r_m$ ) dans le marché  $m$ , soit investir ( $c_n + r_n$ ) dans le marché  $n$  et acheter un FTR qui couvre le coût incrémental sur le réseau de transport plus le différentiel du coût des pertes.

Si le réseau est optimal, la mise en oeuvre de prix nodaux conduit bien les producteurs à internaliser le coût marginal du réseau dans leur décision d'investissement.

## 4 Tarif d'injection différencié incitatif

La consultation pose spécifiquement la question de l'opportunité d'un tarif d'injection différencié, i.e. de l'opportunité de faire payer un tarif d'injection aux producteurs, et de différencier ce tarif en fonction de leur localisation sur le réseau. Cette approche est mise en oeuvre en Suède, où le tarif d'injection croît avec la latitude, et en Grande Bretagne, sur le réseau de transport et sur les réseaux de distribution (Bell et al. 2011).

Cette approche est justifiée par le désir de faire contribuer les producteurs aux coûts sur le réseau que leurs décisions de localisation engendrent. Considérons les deux marchés examinés précédemment. L'argument des tenants du tarif d'injection différencié est le suivant: un producteur s'installant dans le marché 1 (exportateur) augmente la congestion sur le réseau: suite à son installation, l'opérateur de réseau va devoir renforcer celui-ci. Il est donc normal que ce producteur paye une partie des coûts de renforcement. Cela permet de comparer l'installation d'un nouveau moyen de production dans le marché 1, qui coûte moins cher en énergie mais engendre des coûts de réseau, à l'installation d'une centrale dans le marché 2, qui coûte plus cher en énergie mais n'engendre pas de coûts de réseau.

### 4.1 Observations préliminaires

Rappelons tout d'abord le résultat précédent: si le réseau est optimal, la mise en oeuvre de prix nodaux conduit les producteurs à internaliser le coût marginal du réseau dans leur décision de localisation.

Deuxièmement, le tarif d'injection différencié incitatif couvre le coût marginal de renforcement du réseau engendré par la connexion d'un producteur. Il n'a pas vocation à couvrir l'ensemble des coûts du réseau. Cette observation est standard en économie: la tarification au coût marginal ne permet pas de recouvrir l'ensemble des coûts en présence de rendements d'échelle croissants, ce qui est le cas pour les réseaux. Cette différence peut sembler secondaire, mais elle est en fait essentielle: il faut bien distinguer la couverture des coûts incrémentaux de la couverture des coûts totaux. Cette Section traite donc seulement des effets incitatifs d'un tarif d'injection différencié. La partie III examine la couverture des coûts résiduels, i.e., des coûts non couverts par un tarif d'injection différencié incitatif (ou par les rentes de congestion si les prix nodaux sont utilisés).

Finalement, le tarif d'injection différencié incitatif doit être déterminé en fonction de la capacité (€/MW/an ou €/kW/an) et non pas de l'énergie (€/MWh ou €/kWh) pour minimiser les distorsions de production et de consommation. En effet, l'objectif de ce tarif est de modifier les décisions d'investissement. Il est donc légitime de mettre en oeuvre un tarif en fonction de la capacité. Un producteur verra ainsi son coût d'investissement augmenter (décroître) en fonction de sa position sur le réseau.

Un tarif fonction de l'énergie modifierait d'abord les décisions d'utilisation du réseau (production et consommation). Il n'y a pas de raison de penser que ces distorsions améliorent l'efficacité du système. A priori, il est donc souhaitable de les limiter. Donc, un tarif d'injection différencié incitatif doit être déterminé en fonction de la capacité, et non pas de l'énergie.

## 4.2 Réseau optimal

Supposons tout d'abord que le GRT investit optimalement dans le réseau. Comme la capacité est déterminé optimalement, une congestion résiduelle demeure sur chaque ligne. La première question est donc: comment est gérée cette congestion? Deux possibilités: prix nodaux ou couplage de marché/redispatching.

### 4.2.1 Congestion gérée par les prix nodaux

Dans le cas, les acteurs reçoivent/payent les prix nodaux. L'analyse précédente montre que les prix nodaux conduisent à l'investissement optimal. Ajouter un tarif d'injection (positif ou négatif) au coût du capital conduit donc mécaniquement à un investissement sous-optimal.

### 4.2.2 Congestion gérée par le couplage de marché/redispatching

Considérons à nouveau l'exemple simple de deux marchés reliés par une interconnexion. La condition d'entrée libre en concurrence parfaite implique que la capacité installée dans le marché  $i$  est telle que

$$\mathbb{E} [(\hat{p}_m(t) - c_m)^+] = r_m + \delta_m$$

où  $\hat{p}_m(t)$  est le prix reçu par les producteurs dans le marché  $m$ , et  $\delta_m$  le tarif d'injection (positif ou négatif). Dans le cas général, la première difficulté

est que différents producteurs reçoivent des prix différents. Les producteurs non-redispatchés reçoivent le prix  $p^U(t)$  qui prévaudrait en l'absence de congestion, les producteurs redispatchés (à la hausse et à la baisse) reçoivent leur coût marginal. Les incitations sont donc différentes pour les producteurs redispatchés et pour les autres. Si les coûts marginaux sont constants, ce problème ne se pose pas, et  $\hat{p}_1(t) = \hat{p}_2(t) = p^U(t)$ .

**Tarif d'injection optimal** Connaissant la structure du prix  $p^U(t)$ , dans le marché 2, la condition d'entrée libre est donc

$$\mathbb{E} \left[ (p^U(t) - c_2)^+ \right] = \mathbb{E} \left[ (P(K_2, t) - c_2)^+ \right] = \Psi_0(K_2, c_2) = r_2 + \delta_2.$$

En comparant avec l'équation (4), la capacité cumulée est optimale si (et seulement si)  $\delta_2 = 0$ .

Dans le marché 1, la condition d'entrée libre est

$$\mathbb{E} \left[ (p^U(t) - c_1)^+ \right] = \beta_0(K_1, c_1, c_2) = r_1 + \delta_1 - r_2.$$

L'analyse précédente a montré que la capacité optimale est  $K_1^*(\Phi^+) = \frac{\theta_1}{\theta_1 + \theta_2} K_1^U + \Phi^+$ , donc la capacité dans le marché 1 est optimale si et seulement si

$$\begin{aligned} \delta_1 &= \beta_0 \left( \frac{\theta_1}{\theta_1 + \theta_2} K_1^U + \Phi^+, c_1, c_2 \right) - (r_1 - r_2). \\ &= \beta_0 \left( \frac{\theta_1}{\theta_1 + \theta_2} K_1^U + \Phi^+, c_1, c_2 \right) - \beta_0(K_1^U, c_1, c_2). \end{aligned}$$

Supposons que l'interconnexion est dimensionnée optimalement, i.e.,  $\Phi^+ = \Phi^{+*}$ , et que le coût du réseau est

$$\Gamma(\Phi) = F + \gamma\Phi.$$

Le coût marginal d'une unité est donc  $\gamma$ .

En insérant l'équation (16), le tarif d'injection optimal est caractérisée par

$$\delta_1^* = \gamma + \beta_0 \left( \frac{\theta_1}{\theta_1 + \theta_2} K_1^U + \Phi^{+*}, c_1, c_2 \right) - \beta_0 \left( \frac{\theta_1 + \theta_2}{\theta_2} \Phi^{+*}, c_1, c_2 \right).$$

On vérifie que

$$\Phi^{+*} < \frac{\theta_2}{\theta_1 + \theta_2} K_1^U \Leftrightarrow \frac{\theta_1}{\theta_1 + \theta_2} K_1^U + \Phi^{+*} > \frac{\theta_1 + \theta_2}{\theta_2} \Phi^{+*} \Rightarrow \delta_1^* < \gamma.$$

Le tarif d'injection optimal est donc inférieur au coût marginal de l'interconnexion.

**Coût incrémental sur l'interconnexion** En général, le surcoût marginal sur l'interconnexion engendré par une entrée marginale de production dans le marché 1 n'est pas égal à  $\delta_1^*$ . Différenciation totale de l'équation (15) par rapport à  $K_1$  conduit à

$$\begin{aligned} \frac{d\Phi^{+*}(K_1)}{dK_1} &= \frac{\frac{\theta_1 + \theta_2}{\theta_1} \beta'_0 \left( \frac{\theta_1 + \theta_2}{\theta_1} (K_1 - \Phi^+), c_1, c_2 \right)}{\frac{\theta_1 + \theta_2}{\theta_2} \beta'_0 \left( \frac{\theta_1 + \theta_2}{\theta_2} \Phi^+, c_1, c_2 \right) + \frac{\theta_1 + \theta_2}{\theta_1} \beta'_0 \left( \frac{\theta_1 + \theta_2}{\theta_1} (K_1 - \Phi^+), c_1, c_2 \right)} \\ &= \frac{1}{1 + \frac{\theta_1}{\theta_2} \frac{\beta'_0 \left( \frac{\theta_1 + \theta_2}{\theta_2} \Phi^+, c_1, c_2 \right)}{\beta'_0 \left( \frac{\theta_1 + \theta_2}{\theta_1} (K_1 - \Phi^+), c_1, c_2 \right)}} < 1. \end{aligned}$$

Augmenter la capacité de production dans le marché 1 augmente le coût de  $\gamma \frac{d\Phi^+}{dK_1}$ . Le tarif d'injection correspondant au surcoût (marginal) sur l'interconnexion est donc  $\gamma \frac{d\Phi^+}{dK_1}$  par unité. En revanche il n'y a aucune raison que  $\delta_1^*$  soit égal à  $\gamma \frac{d\Phi^+}{dK_1}$ .

Ce résultat est extrêmement important. Contrairement à l'intuition, facturer aux producteurs dans le marché 1 le surcoût sur l'interconnexion  $\gamma \frac{d\Phi^+}{dK_1}$  ne conduit pas au parc de production optimal. Il n'existe donc aucune justification théorique robuste au tarif d'injection différencié.

Ce résultat peut sembler surprenant. L'intuition est que le prix  $p^U(t)$  est par construction indépendant du niveau de congestion optimal. Il n'existe donc aucune raison d'obtenir l'investissement optimal à partir de ce prix.

### 4.3 Aucune contrainte réseau autorisée

Supposons que l'opérateur de réseau ne soit pas autorisé à laisser perdurer des contraintes importantes et durables sur le réseau. Cela semble être le cas pour les réseaux de distribution en France. Observons que cette politique n'est pas optimale: le réseau est sur-dimensionné.

### 4.3.1 Optimalité du tarif d'injection différencié incitatif

Considérons à nouveau deux marchés reliés par une interconnexion de capacité  $\phi^+$ . Ainsi que discuté précédemment, Léautier (2013) montre que la capacité de l'interconnexion doit être supérieure à  $\bar{\phi} = \frac{\theta_2}{\theta_1 + \theta_2} K_1$  pour que la ligne ne soit jamais contrainte. Le coût d'investissement réel pour la collectivité d'un MW de production situé dans le marché 1 est donc  $\left(r_1 + \frac{\theta_2}{\theta_1 + \theta_2} \gamma\right)$ , le coût du capital de la production plus le coût de renforcement de l'interconnexion. La capacité optimale dans le marché 1 est donc définie par

$$\beta_0(K_1, c_1, c_2) = \Psi_0(K_1, c_1) - \Psi_0(K_1, c_2) = r_1 + \frac{\theta_2}{\theta_1 + \theta_2} \gamma - r_2.$$

Si les producteurs dans le marché 1 font face au tarif d'injection  $\frac{\theta_2}{\theta_1 + \theta_2} \gamma$  par MW-an, la capacité installée à l'équilibre concurrentiel est donc optimale. Sur cet exemple très simple, le tarif d'injection différencié conduit à l'investissement optimal.

### 4.3.2 Mise en oeuvre du tarif d'injection différencié dans le cas le plus simple

Même sur cet exemple simple, la mise en oeuvre du tarif d'injection différencié efficace est délicate.

Considérons tout d'abord un producteur dans le marché 2. L'intuition est qu'en s'installant dans le marché 2, le producteur réduit la demande pour l'interconnexion, donc les coûts de réseau, donc devrait recevoir une prime. L'analyse précédente montre que cette intuition est erronée. Un producteur s'installant dans le marché 2 n'a aucun impact sur la contrainte, donc il ne doit recevoir aucune prime.

Quels producteurs situés dans le marché 1 doivent payer le tarif d'injection différencié? La logique économique suggère que, l'objet du tarif d'injection différencié étant de modifier les décisions de localisation, seuls les nouveaux entrants doivent le payer. Les producteurs déjà installés ayant par définition déjà pris leur décision d'investissement, imposer un tarif d'injection différencié ne va pas les amener à relocaliser leur production.

Cependant, la logique juridique suggère qu'il est délicat d'imposer un tarif d'injection différent à deux producteurs situés au même endroit.

### 4.3.3 Mise en oeuvre d'un tarif de retrait différencié dans le cas le plus simple

Considérons maintenant un nouveau consommateur s'installant dans le marché 1. Comme il augmente la demande locale dans le marché 1, il réduit la production exportée, donc la demande sur l'interconnexion, donc il devrait recevoir une prime. L'intuition suggère que cette prime revue est égale à la prime payée par un producteur dans le marché 1. L'analyse montre que cette intuition est erronée.

La demande maximale étant  $K_2$ , la demande maximale d'un consommateur est  $\frac{K_2}{\theta_1+\theta_2}$ . Une augmentation de  $d\theta_1$  de la masse de consommateurs dans le marché 1 augmente la demande dans ce marché de  $\frac{K_2}{\theta_1+\theta_2}d\theta_1$  MW. L'augmentation de capacité d'interconnexion résultant d'un MW de demande supplémentaire dans le marché 1 est donc

$$\frac{\theta_1 + \theta_2}{K_2} \frac{\partial \bar{\phi}}{\partial \theta_1} = -\frac{\theta_2}{\theta_1 + \theta_2} \frac{K_1}{K_2}.$$

Un consommateur s'installant dans le marché 1 doit recevoir une prime égale à  $\frac{\theta_2}{\theta_1+\theta_2} \frac{K_1}{K_2} \gamma$ , inférieure à la prime payée par un producteur.

Observons que le résultat est différent si le consommateur quitte le marché 2 pour s'installer dans le marché 1. Dans ce cas,  $d\theta_2 = -d\theta_1 < 0$ , et la demande d'interconnexion est réduite. L'analyse précédente montre que le consommateur doit recevoir une prime égale à  $\frac{\theta_1}{\theta_1+\theta_2} \frac{K_1}{K_2} \gamma$ .

Considérons ensuite un nouveau consommateur s'installant sur le marché 2. Comme il augmente la fraction  $\frac{\theta_2}{\theta_1+\theta_2}$ , la demande d'interconnexion augmente, et il doit payer une prime, égale à  $\frac{\theta_1}{\theta_1+\theta_2} \frac{K_1}{K_2} \gamma$ . Si un consommateur quitte le marché 1 pour s'installer dans le marché 2, la prime est égale à  $\frac{K_1}{K_2} \gamma$ . Dans les deux cas, les consommateurs dans le marché 2 payent une prime, alors que les producteurs n'en payent aucune.

Cet exemple, le plus simple possible, illustre à quel point la prime peut être contre-intuitive, donc difficile à justifier. Les nouveaux consommateurs reçoivent une prime dans le marché exportateur, différente de celle payée par les producteurs. En revanche, si les consommateurs quittent le marché 2 pour le marché 1, ils reçoivent une prime qui n'est pas toujours plus élevée que celle des nouveaux consommateurs. De même, les consommateurs dans le marché importateur payent une prime, qui varie s'ils sont de nouveaux

consommateurs, ou s'ils se déplacent du marché exportateur vers le marché importateur. Il semble difficile à l'opérateur de réseau de moduler la prime en fonction de la provenance des consommateurs.

#### **4.3.4 Mise en oeuvre de la prime d'injection et de retrait différenciée dans un cas réaliste**

La réalité des marchés électriques est plus complexe. Les réseaux sont maillés, donc la capacité requise sur une ligne pour éliminer les contraintes dépend des injections sur l'ensemble des noeuds du réseau, et des caractéristiques électriques de chaque ligne (en particulier leur admittance). De plus, les consommateurs ont des profils de charge différents, donc des impacts différents sur la capacité requise. La prime reçue ou payée doit donc varier en fonction du profil de charge. En conséquence, la littérature académique (e.g., Olmos et Perez-Arriaga, 2009, Bell, 2011, et Perez-Arriaga et al. 2013) indique qu'il est très difficile de calculer précisément l'impact d'un MW de production sur le coût de long-terme du réseau de transport.

#### **4.4 “Deep connection charge”**

L'analyse précédente montre qu'il n'existe pas en général de justification théorique à la mise en oeuvre de tarif d'injection et de retrait différenciés.

Toutefois, l'analyse précédente suggère que si l'opérateur de réseau n'est pas autorisé à laisser perdurer des contraintes après la connexion de nouveaux moyens de production, il est optimal que le producteur paye les coûts (marginaux) de renforcement du réseau. Cela correspond à “la deep connection charge”. Ainsi que mentionné précédemment, les difficultés pratiques de calcul de ces charges sont importantes.

## Part III

# Couverture de la totalité des coûts des réseaux

## 5 Allocation des coûts résiduels

Les réseaux présentent des rendements d'échelle croissants (i.e., le coût moyen diminue). Les profits réalisés durant les heures de congestion (ou des tarifs d'injection différenciés incitatifs) ne permettent pas de financer le réseau optimal. Il faut donc ajouter une tarification de l'accès, afin de compléter la couverture des coûts du réseau.

Déterminer le tarif de couverture des coûts résiduels, c'est répondre à trois questions différentes:

1. Comment partager les coûts résiduels entre consommateurs et producteurs? Dans certains pays, les consommateurs couvrent la (quasi) totalité des coûts des réseaux, dans d'autres ces coûts sont partagés entre producteurs et consommateurs,
2. quelle assiette de tarif adopter? Tarif assis sur la puissance ( $\text{€}/\text{MW}/\text{an}$  ou  $\text{€}/\text{kW}/\text{an}$ ), où la puissance correspond à la capacité installée pour les producteurs et à la demande de pointe (par exemple la moyenne des 3 ou 10 heures de demandes les plus élevées) pour les consommateurs? Ou tarif assis sur l'énergie, exprimé en  $\text{€}/\text{MWh}$ ? Ou une combinaison des deux? et
3. comment partager les coûts des nouvelles interconnexions, en particulier entre états? Cette question se pose pour les nombreuses interconnexions nécessaires à l'approfondissement du marché unique

### 5.1 Partage des coûts résiduels entre producteurs et consommateurs

En France aujourd'hui, la quasi totalité des coûts des réseaux est couverte par les consommateurs. Il est tentant de trouver cette situation "injuste". Certains coûts sont clairement attribuables à certains producteurs. Ces derniers ne devraient-ils pas porter une partie des coûts?

L'analyse économique montre que cette question n'est en général pas pertinente. Transférer la couverture d'une partie des coûts de réseaux (ou des coûts résiduels) aux producteurs aurait pour effet d'augmenter les coûts des producteurs, donc les prix qu'ils factureront aux consommateurs. Quelle que soit la répartition initiale, les consommateurs couvrent l'ensemble des coûts de réseau à l'équilibre. Pour cette raison, l'impact de la répartition initiale des coûts de réseau est en général faible.

### 5.1.1 Justification théorique

**Tarification en €/MWh** Supposons que les consommateurs payent un coût constant  $\gamma$  €/MWh pour couvrir les charges (résiduelles) du réseau. Considérons le cas le plus simple: une seule technologie, tous les consommateurs réagissent au prix spot, pas de contrainte sur le réseau de transport, et un marché "energy only". Le surplus marginal net est  $(P(Q(t), t) - \gamma)$ . Hors pointe, la demande est donc déterminée par

$$P(Q(t), t) = c + \gamma,$$

tandis que la capacité à l'équilibre est déterminée par l'équation (2) modifiée:

$$\mathbb{E} [(P(K, t) - (c + \gamma))^+] = r.$$

Supposons maintenant que les producteurs payent  $\gamma$  €/MWh pour couvrir les charges du réseau. Le coût marginal est  $(c + \gamma)$ . La demande hors pointe et la capacité à l'équilibre, donc les charges de réseau couvertes, sont inchangées.

La répartition initiale des coûts de réseau entre producteurs et consommateurs n'a aucun impact. L'intuition est que les producteurs et consommateurs ont en fait la même demande.

**Tarification en €/MW** Supposons que les producteurs payent le coût constant  $\gamma$  €/MW/an pour couvrir les charges (résiduelles) du réseau, la capacité à l'équilibre est déterminée par

$$\mathbb{E} [(P(K, t) - c)^+] = r + \gamma.$$

Supposons maintenant que les consommateurs payent  $\gamma$  €/an par MW

de leur consommation maximale, déterminée comme leur consommation moyenne pendant un ensemble  $V_0$  des heures de l'année, et définie analytiquement par  $\frac{\mathbb{E}[Q(t)\mathbb{I}_{\{t \in V_0\}}]}{v_0}$  où  $\mathbb{I}_{\{\cdot\}}$  est la fonction indicatrice<sup>4</sup> et  $v_0 = \mathbb{E}[\mathbb{I}_{\{t \in V_0\}}]$  est la mesure de l'ensemble  $V_0$ . Par exemple, si  $V_0$  correspond aux 10 heures durant lesquelles la demande est la plus élevée, la consommation maximale est la somme des consommations durant ces 10 heures, divisée par 10.

Par construction  $V_0$  est inclus dans l'ensemble des heures de pointe<sup>5</sup> dénoté  $V$ , et défini par

$$t \in V \Leftrightarrow P(K, t) \geq c.$$

Le comportement des consommateurs n'est pas modifié hors de  $V_0$ . Durant les heures  $V_0$ , le surplus net est

$$S(Q(t), t) - (c + \lambda(t))Q(t) - \gamma \frac{\mathbb{E}[Q(t)\mathbb{I}_{\{t \in V_0\}}]}{v_0}.$$

La consommation optimale est donc caractérisée par

$$P(Q(t), t) = (c + \lambda(t)) + \frac{\gamma}{v_0}\mathbb{I}_{\{t \in V_0\}},$$

donc

$$\lambda(t) = P(K, t) - c - \frac{\gamma}{v_0}\mathbb{I}_{\{t \in V_0\}},$$

et la capacité à l'équilibre est caractérisée par

$$r = \mathbb{E} \left[ \left( P(K, t) - c - \frac{\gamma}{v_0}\mathbb{I}_{\{t \in V_0\}} \right) \mathbb{I}_{\{t \in V\}} \right].$$

L'ensemble  $V_0$  étant inclus dans l'ensemble des heures de pointe, cette relation est équivalente à

$$\begin{aligned} r &= \mathbb{E}[(P(K, t) - c)\mathbb{I}_{\{t \in V \setminus V_0\}}] + \mathbb{E} \left[ \left( P(K, t) - c - \frac{\gamma}{v_0} \right) \mathbb{I}_{\{t \in V_0\}} \right] \\ &= \mathbb{E}[(P(K, t) - c)\mathbb{I}_{\{t \in V \setminus V_0\}}] + \mathbb{E}[(P(K, t) - c)\mathbb{I}_{\{t \in V \cap V_0\}}] - \mathbb{E} \left[ \frac{\gamma}{v_0} \mathbb{I}_{\{t \in V_0\}} \right] \end{aligned}$$

<sup>4</sup>Dans ce cas,  $t$  indexe les heures et non plus les états du monde

<sup>5</sup>Formellement, l'ensemble des heures de pointe est une fonction de la capacité  $K$  et du coût marginal  $c$ . Cet aspect n'est pas important pour le raisonnement, et n'est donc pas pris en compte dans cette discussion.

$\Leftrightarrow$

$$r + \gamma = \mathbb{E} [(P(K, t) - c) \mathbb{I}_{\{t \in V\}}] = \mathbb{E} [(P(K, t) - c)^+].$$

A nouveau, la répartition initiale des coûts de réseau entre producteurs et consommateurs n'a aucun impact.

**Robustesse du résultat** Ce résultat est robuste si plusieurs technologies de production et plusieurs classes de consommateurs sont introduites, si des contraintes sont présentes sur le réseau, et si la concurrence entre producteurs est imparfaite.

Ce résultat est robuste si un marché de capacité parfait est introduit. Si le marché de capacité est imparfait, par exemple si l'espérance de défaillance utilisée pour dimensionner le parc est inférieure à l'optimum économique, des distorsions additionnelles sont possibles. Des distorsions additionnelles sont possibles si les paiements sont portés par les fournisseurs et non pas les clients, et si la concurrence entre producteurs est imparfaite.

Ce résultat est-il robuste dans une économie ouverte? Supposons qu'un pays (par exemple la France) transfère une partie des coûts de réseau à ses producteurs, alors que les autres pays ne le font pas. Deux situations extrêmes sont possibles: premièrement, les producteurs français restent compétitifs par rapport à leurs concurrents étrangers, malgré le coût supplémentaire. Une partie des coûts de réseau est donc couverte par les voisins. Deuxièmement, le surcoût rend les producteurs français non compétitifs par rapport à leurs concurrents étrangers. L'ensemble des coûts de réseau reste à la charge des consommateurs français. De plus, la profitabilité des producteurs français est réduite, ce qui peut conduire à une augmentation du prix payé par les consommateurs français.

Cette analyse suggère qu'il n'existe pas d'argument théorique fort pour déterminer le partage initial des coûts du réseau entre producteurs et consommateurs. Le choix semble être plutôt dicté par des aspects pratiques, décrits ci-dessous.

### 5.1.2 Aspects pratiques

Transférer une partie des coûts résiduels des réseaux aux producteurs augmente les coûts de ceux-ci. Cependant, il semble que la plupart des autres pays européens font payer une partie importante des coûts de réseau aux con-

sommateurs. Si la France décidait unilatéralement de transférer une partie de ces charges aux producteurs, elle pénaliserait ceux-ci par rapport à leurs concurrents étrangers ainsi que décrit précédemment. Une telle décision doit donc être harmonisée au niveau européen.

## 5.2 Assiette du tarif

La question est complexe, qui va au delà de la problématique de cette consultation. Le tarif d'utilisation (ou d'injection différencié) génère a priori les comportements et incitations optimales. Introduire un tarif supplémentaire pour couvrir les coûts fixes génère donc des distorsions par rapport à cet optimum. L'objectif est donc de minimiser ces distorsions.

Ce problème a été analysé par Ramsey (1927) pour l'imposition optimale, puis par Boiteux (1956) précisément pour la couverture des coûts fixes d'une compagnie électrique (voir par exemple le Chapitre 8 de Boyer et al., 2006).

### 5.2.1 Cadre analytique

A l'optimum, les utilisateurs du réseau (producteurs ou consommateurs) payent un prix  $p$  déterminé par un règle dite de Ramsey-Boiteux, du type

$$\frac{p - c}{p} = \frac{\lambda}{\lambda + 1} \frac{1}{\eta(p)}$$

où

1.  $c$  est le coût marginal et le ratio  $\frac{p-c}{p}$  est l'indice de Lerner
2.  $\eta(p)$  est l'élasticité prix du comportement des utilisateurs, i.e., la réduction en pour-cent de la demande correspondant à une augmentation du prix de 1%, et
3. le multiplicateur  $\lambda$  est le coût virtuel de la contrainte de couverture des couts.

Si le problème exhibait des rendements d'échelle non croissants, la tarification au coût marginal assurerait la couverture des coûts, le coût virtuel serait  $\lambda = 0$ : la tarification au coût marginal serait bien optimale. Les réseaux de transport et de distribution d'électricité exhibent des rendements d'échelle croissants, donc la tarification au coût marginal ne couvre pas l'ensemble

des coûts, et  $\lambda > 0$ . Le prix doit excéder le coût marginal pour couvrir les coûts.

### 5.2.2 Mise en oeuvre du tarif optimal

L'expression précédente est volontairement générale. Le prix  $p$  peut être exprimé en €/MWh ou en €/MW/an. Dans le premier cas,  $c$  correspond au coût marginal de transporter un MWh supplémentaire, et  $\eta(p)$  à l'élasticité prix de la consommation (ou de la production). Dans le second cas,  $c$  correspond au coût marginal de connecter un MW de consommation (ou de production) supplémentaire, et  $\eta(p)$  à l'élasticité prix de l'installation de centre de consommation (ou de production).

Théoriquement, il est possible de déterminer le tarif binôme optimum, i.e., le prix  $p_1$  exprimé en €/MWh et  $p_2$  exprimé en €/MW/an qui minimise les distorsions (i.e., maximise le surplus).

Pour ce faire, il faut premièrement déterminer comment les coûts résiduels varient en fonction de (i) l'énergie qui transite sur les réseaux et (ii) la puissance installée sur les réseaux. Il s'agit ici d'un problème économétrique assez sophistiqué, qui requiert une bonne coopération avec les opérateurs. Observons simplement qu'il n'est pas évident a priori que le partage soit le même pour les réseaux de transport et de distribution, et pour les réseaux de distribution urbains et ruraux.

Deuxièmement, il faut déterminer comment les utilisateurs répondent aux changements de prix, i.e., quelles sont les élasticités de la demande (ou de la production) et de l'installation de centre de consommation (ou de production).

Il est probable que les coûts et les élasticités varient selon les classes et la localisation des clients. Le tarif optimal est donc par construction différencié.

Si le problème est bien posé théoriquement, il semble très difficile à résoudre en pratique. Les études académiques qui ont sérieusement examiné ce problème insistent sur la difficulté de mise en oeuvre pratique de méthodes sophistiquées de partage des coûts (Perez-Arriaga et al., 2013, et Olmos et Perez-Arriaga, 2009).

### **5.3 Partage des coûts d'une interconnexion**

Le problème ici est légèrement différent. Les pouvoirs publics (ou les GRTs) de chaque état doivent décider de financer (ou non) une nouvelle interconnexion. Cette question n'étant pas explicitement incluse dans le périmètre de la consultation, elle n'est pas traitée en détail. Cette discussion préliminaire a pour but de mettre en évidence les particularités d'une interconnexion.

L'analyse précédente suggère que l'optimum de premier rang est (i) d'utiliser les prix nodaux de l'électricité pour gérer la congestion en temps réel et envoyer les bon signaux d'investissement, et (ii) de faire porter aux consommateurs les coûts résiduels du réseau. Une interconnexion entre deux pays n'est pas différente d'une interconnexion entre deux marchés au sein du même pays. L'optimum de premier rang serait d'utiliser des prix nodaux sur une plaque recouvrant plusieurs pays, et de faire porter l'ensemble les coûts du réseau de transport également sur l'ensemble des consommateurs.

Cet optimum est probablement difficile à atteindre. Les tarifs de transport ne sont pas encore harmonisés sur la plaque européenne. Une approche pragmatique est de partager les coûts en fonction des bénéfices espérés pour chaque collectivité, selon les modalités décrites par Boyer et al. (2006).

## **6 Conclusion**

Cette section résume les principales conclusions de l'analyse académique, puis répond précisément aux questions posées par la consultation.

### **6.1 La tarification nodale de l'électricité est la bonne solution**

#### **6.1.1 Pourquoi adopter les prix nodaux**

L'analyse académique confirme que tarification nodale de l'électricité demeure la meilleure approche, qui permet de gérer à la fois la congestion en temps réel, mais aussi d'envoyer les signaux d'investissement optimaux pour les moyens de production et la localisation des centres de consommation. La tarification nodale permet donc de maximiser simultanément le surplus à court- et à long-terme.

Si le réseau de transport est optimal, les producteurs incluent les coûts de renforcement du réseau dans leur décision de localisation. En revanche,

un tarif d'injection différencié construit à partir des coûts de renforcement du réseau ne conduit pas à l'investissement optimal.

Cette supériorité économique se traduit par des gains importants pour la collectivité. Green (2007) estime que la mise oeuvre de prix zonaux accroîtrait le surplus de 1.3% en Angleterre comparé au système actuel (redispatching/counter-trading) si le marché est parfaitement concurrentiel, et de 3.1% si les producteurs exercent leur pouvoir de marché. Plus récemment, Neuhoff et al. (2013) examinent l'impact de la mise en oeuvre de prix nodaux sur la plaque européenne, en considérant différents scénarii de pénétration des renouvelables. L'étude estime les gains réalisés comparés au status quo (prix zonaux par pays et redispatching/counter-trading au sein de chaque pays). Elle produit trois résultats principaux. Le passage au prix nodaux permet de:

1. accroître les échanges aux frontières (en MW) jusqu'à 34%,
2. réduire les coûts d'opération de 0.8 à 2.0 milliards d'Euros annuellement (1.1 à 3.3%), dépendant de la pénétration des renouvelables, et
3. réduire le prix moyen dans 60 à 75% des pays considérés.

De plus, l'analyse académique et la pratique montrent que la tarification nodale de l'électricité réduit l'impact du pouvoir de marché des producteurs. Ce résultat peut sembler a priori surprenant. L'intuition est que, si la congestion est gérée par le redispatching/counter-trading, une nouvelle opportunité de manipulation des prix est possible. Un producteur dans un marché contraint à l'export a intérêt à annoncer un coût très bas. Si le GRT doit lui racheter sa production, il recevra la différence entre le prix sans contrainte et le coût qu'il a annoncé. Donc, plus bas est le coût annoncé, plus élevé est le paiement reçu par le producteur. Ces stratégies, connues sous le nom de jeu de réduction ("Dec game", Green, 2007, Holmberg et Lazarczyk, 2012) ont été mises en oeuvre par certains producteurs durant la crise Californienne.

Même en l'absence de congestion, la mise en oeuvre de prix nodaux permet de faire porter aux utilisateurs (producteurs et consommateurs) l'impact des pertes marginales qu'ils occasionnent. En France aujourd'hui, les pertes sont payées par les consommateurs finaux, et aucun acteur n'a d'incitation financière à réduire leur volume. La mise en oeuvre de prix nodaux con-

duirait les fournisseurs à minimiser le coût des pertes associées à leur approvisionnement, donc à s’approvisionner au plus près (au sens électrique) de la demande.

Finalement, deux aspects pratiques contribuent à l’attrait de la tarification nodale de l’électricité:

1. c’est une technologie bien maîtrisée, qui est mise en oeuvre depuis une quinzaine d’années aux Etats Unis. Des logiciels robustes de calcul des prix zonaux existent, et la majorité des problèmes de développement de l’approche ont été étudiés et réglés, et
2. la méthodologie de calcul des prix nodaux est transparente pour les participants aux marchés, et difficilement discutable. Une ligne est contrainte ou ne l’est pas, un producteur produit ou ne produit pas. Les prix nodaux sont donc robustes institutionnellement.

### **6.1.2 Pourquoi ne pas adopter les prix nodaux?**

La question pour les pouvoirs publics est donc: pour quelles raisons priver la collectivité du gain associé à la tarification nodale de l’énergie? Deux arguments sont souvent invoqués: (i) la faible congestion en France, et (ii) les effets redistributifs, en particulier l’incompatibilité supposée des prix nodaux avec la péréquation tarifaire.

Le premier argument est que le coût de la congestion est très faible en France (Léautier et Thelem (2009) l’estiment à 0.36% du coût de l’électricité produite, 10 fois moins que dans certains marchés nord américains). Mettre en oeuvre les prix nodaux nécessiterait une transformation profonde et coûteuse des règles de marché et des systèmes informatiques. Il ne semble pas évident que les bénéfices dépassent les coûts.

Cet argument est valide dans le court-terme, mais probablement pas dans le moyen- long-terme. En effet, il est probable que la congestion sur les réseaux de transport français et européens augmente grandement dans les prochaines années. Historiquement, EDF puis Rte ont conçu le réseau pour “évacuer” toute la production des centrales installées sur le territoire. Les flux d’énergie étant relativement bien anticipés, les réseaux étaient conçus pour éliminer ou très sévèrement limiter la congestion, i.e., la France était en première approximation une plaque de cuivre. Les deux

principales poches de congestion sont les régions Bretagne et Provence Alpes Côte d'Azur (PACA).

Le déploiement de production renouvelable sur le territoire va profondément modifier les flux d'énergie, et les rendre plus variables. Certaines interfaces aujourd'hui non congestionnées le deviendront. Il faudrait donc modifier profondément la topographie du réseau pour accommoder ces nouveaux flux. Mais, l'acceptabilité des infrastructures en général et des lignes de transport en particulier étant bien plus faible que dans les années 1970s-1980s, il sera très coûteux voire impossible de mettre en oeuvre l'ensemble de ces modifications. Par exemple, l'interconnexion France-Espagne récemment inaugurée a dû être enterrée, et a donc coûté 7 à 8 fois plus cher qu'une ligne aérienne<sup>6</sup>. Le réseau français sera donc plus souvent congestionné. Ainsi que montré par l'étude de Neuhoff et al. (2013) citée plus haut, cette évolution se produira probablement plus rapidement et à plus grande échelle chez nos voisins. La gestion efficace de la congestion deviendra alors un enjeu économique majeur, et les prix nodaux seront mis en oeuvre.

Le deuxième argument n'est pas très convaincant. Premièrement, la péréquation tarifaire est économiquement inefficace. Le dispositif est très récent dans l'industrie électrique française: il a été imposée à EDF au début des années 1960s, alors que celle-ci défendait une tarification géographiquement différenciée et alignée sur les conditions d'équilibre offre-demande locales<sup>7</sup>.

Deuxièmement, si les pouvoirs publics souhaitent maintenir la péréquation, les prix zonaux peuvent être mis en oeuvre sur les marchés de gros, sans être nécessairement répercutés sur les prix au détail. Par exemple, la péréquation peut perdurer dans le Tarif Réglementé de Vente pour les particuliers (tarif bleu). Avec la disparition des tarifs jaunes et verts, les fournisseurs vont développer des offres de marchés différenciées selon le profil des consommateurs. Les pouvoirs publics peuvent autoriser ou non les fournisseurs à différencier géographiquement leurs offres.

---

<sup>6</sup>[http://www.lesechos.fr/journal20150218/lec2\\_industrie\\_et\\_services/0204144388300-electricite-france-et-espagne-doublent-leurs-capacites-dimport-export-1094334.php](http://www.lesechos.fr/journal20150218/lec2_industrie_et_services/0204144388300-electricite-france-et-espagne-doublent-leurs-capacites-dimport-export-1094334.php)

<sup>7</sup>On peut consulter à ce sujet l'audition du Président Boiteux devant la commission d'enquête sur les tarifs de l'électricité de l'Assemblée Nationale le 5 novembre 2014 <http://videos.assemblee-nationale.fr/video.6047.tarifs-de-l-electricite--m-robert-durdilly-pdt-de-l-ufe--m-marcel-boiteux-pdt-d-honneur-d-edf-5-nov>

### **6.1.3 Implications de la mise en oeuvre des prix nodaux en France**

Etant donnée la faible congestion sur le réseau français aujourd’hui, la mise en oeuvre de prix nodaux aurait un impact très limité dans les 3 à 5 prochaines années (i.e., jusqu’à ce que la transformation du parc de production européen modifie les flux d’énergie). Quelques heures par an, les prix de gros seraient plus élevés dans les régions Bretagne et PACA, tandis que les prix nodaux sur le reste du territoire resteraient très proches les uns des autres (en excluant les pertes). Les producteurs et opérateurs d’effacement investiraient en priorité vers la Bretagne et la PACA, ce qui est exactement l’effet désiré.

Si les fournisseurs étaient autorisés à différencier géographiquement leurs offres, les consommateurs de Bretagne et PACA paieraient probablement un peu plus, ce qui réduiraient vertueusement leur consommation. Sinon, certains fournisseurs achèteraient des droits financiers de transport pour se protéger de la congestion.

Dans 5 à 10 ans, les flux d’énergie en Europe seront modifiés. Etant données les coûts et les difficultés de construction, peu de lignes seront construites, et la congestion sera plus importante. Les prix nodaux permettraient d’accompagner facilement cette évolution. Ils seraient progressivement plus différenciés, et enverraient les bons signaux d’investissement.

A l’horizon 2030, les flux d’énergie en Europe seront profondément modifiés. La congestion sera trop importante pour être gérée autrement que par les prix nodaux. L’ensemble des pays européens se rangera à cette évidence, et les prix nodaux seront adoptés sur l’ensemble du continent.

## **6.2 Un timbre d’injection différencié ne semble pas constituer une bonne solution intermédiaire**

La consultation s’interroge sur l’opportunité de “mettre en place sur le réseau de transport un signal-prix de localisation des capacités d’injections”, i.e., de mettre en oeuvre un timbre d’injection différencié. Deux objectifs sont possibles pour un tel timbre d’injection différencié: modifier les décisions d’investissement des producteurs ou modifier leurs décisions opérationnelles. Dans le premier cas, le timbre est exprimé en €/MW/an, dans le second en €/MWh.

### 6.2.1 Timbre d'injection différencié exprimé en €/MW/an

L'intuition justifiant un timbre d'injection différencié exprimé en €/MW/an est la suivante: si les producteurs payent un timbre d'injection différencié égal au coût marginal pour le réseau entraîné par leur localisation, ils incluent dans décisions d'investissement non seulement les coûts d'installation d'une centrale, mais aussi les coûts de réseau induits. Des timbres d'injection différenciés égaux au coût marginal de renforcement de réseau guideraient efficacement les investissements.

Cette approche est a priori séduisante, qui semble envoyer les bons signaux d'investissement, sans modifier l'architecture du marché. Elle semble donc un compromis raisonnable entre la situation actuelle (aucun signal de localisation n'est envoyé) et l'optimum économique des prix nodaux. Malheureusement, l'analyse montre que cette intuition est erronée.

Premièrement, le calcul de timbres d'injection différenciés est extrêmement complexe. L'efficacité économique suggère que le timbre en un point donné soit égal au surcoût marginal pour le réseau d'un MW de capacité installée en ce point. Comment ce surcoût est-il calculé en pratique?

La première question est: le surcoût pour quel réseau? Le plus simple est de considérer le réseau actuel. Mais celui-ci n'est pas forcément optimal. Le coût des renforcements requis pour accommoder un MW de capacité supplémentaire est-il supérieur à ce qu'il serait si le réseau était optimal? Si c'est le cas, pourquoi un producteur serait-il pénalisé, et devrait payer ce surcoût?

Considérons maintenant le réseau optimal. Par définition, il n'est pas observable. Quelles hypothèses retenir pour le définir? A quel horizon de temps le réseau optimal est-il défini? Aujourd'hui? Dans 5 ans? dans 10 ans? Quel parc de production et quels profils de demande sont pris en compte? L'analyse technico-économique présentée plus haut suggère que le réseau optimal est tel que la valeur marginale de la congestion sur chaque ligne est égal au coût marginal d'expansion de la ligne. Il faut donc calculer la congestion, ce qui requiert des hypothèses sur les coûts de production des différentes technologies, et sur les profils de demande aux différents noeuds. Quelles hypothèses sont utilisées? Il faut aussi calculer les coûts d'expansion du réseau. Quelles structures de coûts sont utilisées?

Pour simplifier, pourquoi ne pas considérer que le réseau est non con-

traint, i.e., que le GRT augmente la capacité de transit sur le réseau pour “évacuer” toute la production d’un MW marginal de capacité? Dans ce cas, le producteur devrait payer pour plus de capacité de transit qu’il est optimal, i.e., il devrait payer pour une capacité de transit non justifiée.

La discussion précédente est loin d’être exhaustive. Elle a pour objet d’illustrer les difficultés pratiques du calcul de timbres d’injections différenciés. De nombreux choix méthodologiques et hypothèses doivent être arbitrés. Etant donné l’impact d’une telle réforme sur les profits de certains producteurs, il est très probable que le processus de consultation sera long, contradictoire, et conflictuel. Les opérateurs, qui possèdent une bien meilleure information que le régulateur sur de nombreux points, essaieront d’influencer les règles à leur avantage. Le résultat de ce processus de consultation sera probablement une mécanique administrative complexe, sans garantie d’efficacité économique.

Deuxièmement, en supposant qu’il soit possible de calculer des timbres d’injection différenciés, il faut déterminer qui le paye. A cette question, les logiques économiques et juridiques donnent des réponses différentes.

La logique économique suggère que, l’objet du timbre d’injection différencié en €/MW/an étant de modifier les décisions de localisation, seuls les nouveaux entrants doivent le payer. Les producteurs déjà installés ayant par définition déjà pris leur décision d’investissement, imposer un tarif d’injection différencié ne les amène pas à relocaliser leur production.

Cependant, la logique juridique semble limiter la possibilité d’offrir des conditions différentes à deux producteurs situés au même endroit. Cette dernière l’emportant souvent, il est probable que tous les producteurs soient astreints à payer un timbre d’injection différencié. Les producteurs déjà installés pourront alors légitimement s’interroger sur l’efficacité économique d’une telle disposition, ce qui rendra encore plus difficile la mise en oeuvre du dispositif.

Troisièmement, des timbres d’injection différenciés auront un impact très limité sur les décisions d’investissement des nouveaux entrants.

La quasi-totalité des nouveaux moyens de production d’ici à 2020 seront des renouvelables, qui recevront une subvention, couvrant (en moyenne) leurs coûts<sup>8</sup>. Un timbre d’injection différencié serait inclus dans le coût

---

<sup>8</sup>La mécanique de la subvention peut évoluer dans le temps: tarif de rachat, prime,

des renouvelables, donc dans la subvention requise. Si les projets sont en concurrence, ce dispositif semble vertueux: un projet bien localisé requiert une subvention inférieure au même projet moins bien localisé.

En pratique, l'efficacité économique de ce mécanisme augmente avec la granularité du calcul des timbres d'injection différenciés : le mécanisme permet de discriminer entre des projets situés dans des zones différentes, mais pas entre projets situés dans la même zone. Mais la complexité du calcul des timbres augmente avec leur granularité. Il est donc probable qu'un nombre limité de timbre serait effectivement calculé, ce qui limiterait l'efficacité du dispositif. De plus, il n'est pas certain que les projets soient effectivement en concurrence, mais que tous les projets soient financés. Dans ce cas, les timbres d'injection différenciés augmenteraient simplement les subventions payées.

Finalement, ainsi que mentionné précédemment, à l'horizon 2030, il sera nécessaire de mettre en oeuvre des prix nodaux en Europe pour gérer la congestion sur le réseau de transport, et donc mettre un terme aux timbres d'injection différenciés. Terminer un dispositif réglementaire et tarifaire en France est toujours difficile. Ce constat incite à la prudence avant de mettre en oeuvre un dispositif dont la durée de vie est limitée.

L'analyse présentée ci-dessus suggère que la mise en oeuvre de timbres d'injection différenciés en €/MW sera complexe, que leur efficacité économique sera limitée, et leur durée de vie limitée. Il semble donc opportun de ne pas s'engager dans cette voie.

### **6.2.2 Timbre d'injection différencié exprimé en €/MWh**

Une autre possibilité est de modifier les décisions d'exploitation des producteurs en mettant en oeuvre des timbres d'injection différenciés exprimés en €/MWh. L'objectif serait que les producteurs internalisent le coût pour le réseau de transport d'un MWh marginal injecté sur le réseau. L'impact marginal d'une injection variant très significativement d'une heure à l'autre, la première question est alors: quelle est le pas de temps considéré?

La solution la plus simple est de déterminer des timbres d'injection différenciés qui durent une période longue, par exemple une année ou une

---

appel d'offres, etc. Le résultat économique est le même: la rémunération des renouvelable couvre en moyenne leurs coûts.

période tarifaire, en prenant la moyenne des impacts marginaux sur l'ensemble de la période. Comme toutes les tarifications au coût moyen, cette approche est simple, mais peu efficace économiquement: le timbre n'est pratiquement jamais égal à l'impact marginal réel. L'autre extrême est de déterminer des timbres d'injection différenciés qui varient à chaque heure. Ce sont précisément les prix nodaux discutés plus haut. Entre ces deux extrêmes, toutes les combinaisons sont possibles: timbres été/hiver, jour/nuite, semaine/weekend.

Une fois le pas de temps déterminé, les timbres d'injection différenciés exprimés en €/MWh font face aux mêmes difficultés de mise en oeuvre que ceux exprimés en €/MW/an.

Pour ces raisons, il ne semble donc pas très judicieux de s'engager dans cette voie.

### **6.3 Un tarif de raccordement “profond” est probablement légitime sur le réseau de distribution**

Il semble légitime de mettre en oeuvre un tarif de raccordement “profond” pour les producteurs lorsque l'opérateur de réseau doit éliminer la congestion, ce qui semble être le cas pour les réseaux de distribution.

Lorsque les réseaux sont dimensionnés pour ne jamais être congestionnés (ce qui est le cas pour les réseaux de distribution), les prix nodaux ne peuvent pas être utilisés pour envoyer des signaux d'utilisation ou de localisation. En revanche, il est légitime de demander aux producteurs de contribuer aux coûts marginaux de renforcement de réseau engendrés par leur connexion. La plupart des difficultés évoquées précédemment demeurent, mais sont atténuées: le réseau étant par construction non congestionné, la définition du réseau “optimal”, et le calcul des coûts de renforcement peuvent être rendus plus transparents, et surtout, plus granulaires. Chaque Gestionnaire de Réseau de Distribution peut annoncer différents tarifs de raccordement sur son réseau, pour guider les décisions des entrants.

### **6.4 Partage des coûts résiduels**

La consultation s'interroge sur l'opportunité “de faire payer une partie des coûts de transport et de distribution de l'électricité aux injections”. L'analyse académique suggère que ce n'est pas justifié.

#### **6.4.1 La nécessité de couvrir les coûts résiduels**

Comme dans de nombreux problèmes en économie, la tarification au coût marginal (de long-terme) conduit aux bonnes décisions d'utilisation des réseaux et d'investissement, mais ne garantit pas couverture de l'ensemble des coûts si les technologies exhibent des rendements d'échelle croissants, ce qui est le cas pour les réseaux de transport et de distribution d'électricité. Que des prix nodaux (présentés dans la Section 6.1) ou des tarifs d'injection différenciés et incitatifs (Section 6.2) soient mis en oeuvre, il faut aussi mettre en place un tarif complémentaire pour couvrir les coûts résiduels (Perez-Arriaga et al., 2013, Bell et al., 2011).

Dans le secteur du transport d'électricité, le partage des coûts résiduels est souvent extrêmement simple: dans de nombreux pays, ces coûts sont simplement couverts par les consommateurs au prorata de la puissance maximale qu'ils soutirent ou de leur consommation (ou d'une combinaison des deux). Dans d'autres, les coûts résiduels sont partagés entre producteurs et consommateurs selon des règles mal justifiées économiquement, puis au sein de chaque groupe au prorata de la puissance installée/soutirée et/ou de la production/consommation. Ces pratiques semblent très simples comparées à la très riche réflexion académique sur le partage des coûts communs (voir par exemple la présentation de Boyer et al., 2006).

#### **6.4.2 In fine les consommateurs payent tous les coûts**

Quelle que soit le partage initial des coûts du réseau, les clients les payent in fine: si les producteurs paient une partie des coûts des réseaux, leurs charges augmentent, et ils incluent cette augmentation dans leur prix de vente. Même si un timbre d'injection est mis en oeuvre, d'un point de vue économique, les consommateurs et non pas producteurs couvrent les coûts du réseau.

De plus, dans la plupart des cas, faire payer les coûts résiduels aux producteurs qui les transmettent aux consommateurs ou directement au consommateurs n'a pas d'impact sur les décisions des acteurs. Pour cette raison, il ne semble pas justifié de transférer une partie des coûts résiduels des réseaux aux producteurs.

### 6.4.3 Partage optimal des coûts entre consommateurs

S'il n'est pas économiquement justifié de transférer une partie des coûts résiduels aux producteurs, il est possible de modifier les règles d'allocation entre classes de consommateurs.

La couverture des coûts résiduels modifie les décisions des agents. Par exemple, si les consommateurs couvrent une portion des coûts au prorata de leur consommation, ils réduisent celle-ci. La règle optimale de couverture des coûts résiduelle minimise les distorsions par rapport aux incitations optimales présentées plus haut. Ce problème a été analysé par Ramsey (1927) pour l'imposition optimale, puis par Boiteux (1956) précisément pour la couverture des coûts fixes d'une compagnie électrique (voir par exemple le Chapitre 8 de Boyer et al., 2006). La tarification optimale est telle que les utilisateurs du réseau (producteurs et consommateurs) payent un prix inversement proportionnel à leur élasticité. Si cet optimum est bien défini théoriquement, la plupart des analystes reconnaissent les difficultés de calculer cet optimum (Perez-Arriaga et al., 2013, et Olmos et Perez-Arriaga, 2009). De plus, l'importance du gain potentiel n'est pas documentée. Toutefois, cette avenue d'amélioration mérite d'être explorée.

## 6.5 Réponses aux questions posées par la consultation

### 6.5.1 Réponses aux questions posées à l'article 5 de la consultation:

1. *A quelles conditions peut-on considérer que les injections génèrent des coûts de réseau?* A court terme, les injections génèrent deux types de coûts de réseaux: elles contribuent à la congestion et aux pertes. Les producteurs qui vendent aux prix zonaux/nodaux internalisent pleinement ces deux externalités. De même, les soutirages génèrent aussi ces deux externalités, et les consommateurs qui achètent aux prix nodaux les internalisent pleinement. A long-terme, l'installation de nouveaux moyens de production conduit l'opérateur à augmenter la capacité de certains lignes du réseau.
2. *Un tarif d'injection permet-il de contribuer à un meilleur fonctionnement du marché de l'électricité, ou au contraire, est-il susceptible de porter atteinte au bon fonctionnement du marché?* Les prix nodaux

correspondant à l'optimum, leur mise en oeuvre optimise l'utilisation du réseau, donc contribue à un meilleur fonctionnement du marché. Un tarif d'injection exprimé en €/MW n'a aucun impact sur le fonctionnement du marché hors pointe, il a un impact sur le marché en pointe et sur les investissements. Toutefois, même s'il était possible de calculer précisément le coût de réseau marginal engendré par un MW de capacité de production à un noeud donné, il n'existe aucune garantie que ce timbre d'injection conduise à l'investissement optimal. Un tarif d'injection exprimé en €/MWh a un impact sur le fonctionnement du marché. Par construction, il serait moins efficace que les prix nodaux, qui correspondent à l'optimum de premier rang. Dépendant de la méthode de calcul, il peut améliorer ou détériorer la situation par rapport au status quo.

3. *A quelles conditions est-il pertinent de mettre en place un signal de localisation pour les injections sur le réseau de transport?* La littérature académique n'a identifié aucune contre-indication à la mise en oeuvre de prix nodaux sur le réseau de transport, qui constituent le signal de localisation optimal. A la question "à quelle condition est-il pertinent de mettre en place un timbre d'injection différencié pour guider les investissements des producteurs?", la réponse est: si le réseau est structurellement sur-dimensionné et n'est jamais congestionné. Dans ce cas, un timbre d'injection égal au coût marginal de réseau engendré par un MW de capacité de production à un noeud donné est optimal.

### **6.5.2 Réponse aux objectifs de l'étude fixée par la consultation:**

1. *l'efficacité économique en termes de surplus global:* par construction, la mise en oeuvre de prix nodaux maximise le surplus global. Toute autre approche dégrade le surplus.
2. *les effets redistributifs entre acteurs du marché de l'électricité* ne sont pas à prendre en compte dans les tarifs, qui guident les décisions de production, consommation et investissement. Si les pouvoirs publics le souhaitent, il peuvent traiter les effets redistributifs par des transferts hors marchés (taxes, subventions directes).
3. *les effets sur les échanges transfrontaliers et le fonctionnement du*

*marché européen.* Les prix nodaux sont compatibles avec le couplage des marchés Européens. Un timbre d'injection (différencié ou non) modifierait la compétitivité des producteurs français vis à vis de leurs concurrents étrangers. La mise en oeuvre d'un timbre d'injection doit donc être harmonisée entre pays européens.

4. *les effets incitatifs de la structure du tarif sur les investissements de production entre les différentes technologies:* pour une topologie fixée du réseau de transport, les prix nodaux conduisent au parc optimal, tant pour les technologies que pour la localisation des moyens de production. Tout autre structure tarifaire incite à un parc sous-optimal.
5. *les effets incitatifs de la structure du tarif sur l'utilisation efficace des actifs de production entre les différentes technologies:* les prix nodaux conduisent à la production et à la consommation optimale. L'autre approche de gestion de la congestion, le re-dispatching, conduit à une la même utilisation du parc de parc de production, mais à des transferts (donc des profits) différents.
6. *les limites théoriques des modèles étudiés et leur sensibilité à des changements d'hypothèses:* il n'existe pas vraiment de limite théorique. Les prix nodaux s'appuient sur les lois de la physique qui régissent les réseaux électriques. Les limites "standard", par exemple le pouvoir de marché, affectent toutes les méthodes. Les analyses académiques existantes suggèrent que les prix nodaux sont moins sensibles que le redispatching à l'exercice du pouvoir de marché.
7. *l'articulation entre les coûts de réseaux payés lors du raccordement et les coûts de réseaux payés au travers du tarif d'injection:* a priori, il n'existe aucune différence entre payer un coût lors du raccordement et le payer progressivement le long de la vie d'un ouvrage au travers d'un tarif d'injection. Dans le premier cas, le producteur finance le coût, dans le second cas, l'opérateur de réseau le finance.
8. *les difficultés pratiques de détermination des tarifs (disponibilité des données, complexité des calculs, etc.):* les difficultés de mise en oeuvre des prix nodaux sont importantes, mais elles ont déjà été résolues par les GRTs lors du développement des algorithmes de dispatch. Les

difficultés de mise en oeuvre de tarifs d'injection différenciés sont encore plus importantes, et ne semblent pas encore avoir reçu de solution satisfaisante (Bell et al., 2011).

## References

- [1] R. Baldick, J. Bushnell, B. Hobbs, and F. Wolak. Final Report Project TransmiT: Optimal Charging Arrangements for Energy Transmission Great Britain Office of Gas & Electricity Markets. Technical report, 2011.
- [2] K. Bell, R. Green, I. Kockar, G. Ault, and J. McDonald. Project transmit: Academic review of transmission charging arrangements. Technical report, Report produced on behalf of the Gas and Electricity Markets Authority, 2011.
- [3] M. Boiteux. La tarification des demandes en pointe: application de la théorie de la vente au coût marginal. *Revue Generale de l'Electricite*, pages 321–40, August 1949.
- [4] M. Boiteux. Sur la gestion des monopoles astreints à l'équilibre budgétaire. *Econometrica*, 24:22–40, 1956.
- [5] S. Borenstein, J. Bushnell, and S. Stoft. The competitive effect of transmission capacity in a deregulated electricity industry. *The RAND Journal of Economics*, 31(2):294–325, Summer 2000.
- [6] Marcel Boyer, Michel Moreaux, and Michel Truchon. Partage des coûts et tarification des infrastructures. Technical report, Cirano, 2006.
- [7] J. Bushnell and S. Stoft. Electric grid investment under a contract network regime. *Journal of Regulatory Economics*, 10:61–79, 1996.
- [8] Judith B. Cardell, Carrie Cullen Hitt, and William W. Hogan. Market power and strategic interaction in electricity networks. *Resource and Energy Economics*, 19(1-2):109–137, March 1997.
- [9] R. Green. Nodal pricing of electricity: how much does it cost to get it wrong? *Energy Journal*, 31:125–149, 2007.

- [10] S. Harvey and W. Hogan. Nodal and zonal congestion management and the exercise of market power, 2000.
- [11] W. Hogan. Contract networks for electric power transmission. *Journal of Regulatory Economics*, 4:211–242, 1992.
- [12] W. Hogan, J. Rosellon, and I. Vogelsang. Towards a combined merchant-regulatory mechanism for electricity transmission expansion. *Journal of Regulatory Economics*, 38(2):113–143, 2010.
- [13] P. Holmberg and E. Lazarczyk. Congestion Management in Electricity Networks : Nodal , Zonal and Discriminatory Pricing. 2012.
- [14] P. Joskow and J. Tirole. Transmission rights and market power on electric power networks. *Rand Journal of Economics*, 31(3):450–487, Autumn 2000.
- [15] P. Joskow and J. Tirole. Reliability and competitive electricity markets. *RAND Journal of Economics*, 38(1):60–84, Spring 2007.
- [16] T.-O. Leautier. Regulation of an electric power transmission company. *The energy journal*, 21(4):61–92, 2000.
- [17] T.-O. Leautier. Marcel boiteux meets fred schweppe: Nodal peak load pricing. Mimeo, Toulouse School of Economics, January 2013.
- [18] T.-O. Leautier. Is mandating smart meters smart? *The Energy Journal*, 35(4):135 – 158, October 2014.
- [19] Thomas Olivier Léautier. Transmission Constraints and Imperfect Markets for Power. *Journal of Regulatory Economics*, 19(1):27–54, 2001.
- [20] Thomas Olivier Léautier and Véronique Thelen. Optimal expansion of the power transmission grid: Why not? *Journal of Regulatory Economics*, 36(2):127–153, 2009.
- [21] K. Neuhoff, J. Barquin, J. W. Bialek, R. Boyd, C. J. Dent, F. Echavarren, T. Grau, C. von Hirschhausen, B. F. Hobbs, F. Kunz, C. Nabe, G. Papaefthymiou, C.h Weber, and H. Weigt. Renewable electric energy integration: Quantifying the value of design of markets for international transmission capacity. *Energy Economics*, 40:760–772, 2013.

- [22] Luis Olmos and Ignacio J. Pérez-Arriaga. A comprehensive approach for computation and implementation of efficient electricity transmission network charges. *Energy Policy*, 37(12):5285–5295, 2009.
- [23] I. Perez-Arriaga, L. Olmos, and M. Rivier. Transmission Pricing. In Juan Rosellón and Tarjei Kristiansen, editors, *Financial Transmission Rights*, volume 7 of *Lecture Notes in Energy*, chapter 2, pages 49–77. Springer London, London, 2013.
- [24] F. P. Ramsey. A contribution to the theory of taxation. *Economic Journal*, 37:47–61, 1927.
- [25] F. Schweppe, M. Caramanis, R. Tabors, and R. Bohn. *Spot Pricing of Electricity*. Kluwer, Norwell, MA, 1988.