

2009

concurrence et
politiques publiques dans les
marchés électriques

Larsen

Concurrence et politiques publiques dans les marchés électriques

Larsen - 2009

Avant-propos

9 les politiques de promotion des énergies renouvelables

10 Photovoltaïque : l'inadéquation du tarif d'achat
Dominique Finon CIRED & LARSEN

50 Large-scale wind power in European electricity markets : time for revisiting support schemes and market designs ?
Céline Hiroux University Paris Sud (ADIS-GRJM) & LARSEN
Marcelo Saguan University Paris Sud (ADIS-GRJM), Supelec & LARSEN

79 les particularités de la concurrence de détail

80 The dynamic of competition in presence of switching costs.
The case of British Gas (1997-2007)
Christophe Defeuilley LARSEN
Matthieu Mollard University Paris Sud (ADIS-GRJM) and LARSEN

105 Retailers' risk management and vertical arrangements in electricity markets
Raphaël H. Boroumand LARSEN, Georg Zachmann LARSEN

127 les politiques de la concurrence et l'organisation des marchés

128 HHI, an irrelevant market indicator without a relevant market
Delphine Perrot Voisard LARSEN, Georg Zachmann LARSEN
139 Legal uncertainty and competition policy in deregulated network industries. The case of long term contracts
Adrien de Hauteclercque University Paris Sud (ADIS-GRJM) & LARSEN

177 Signaux-prix et équilibre de long-terme.
Reconsidérer l'organisation des marchés électriques
Dominique Finon CIRED & LARSEN,
Christophe Defeuilley LARSEN
Fredéric Marty (CNRS, Université Nice Sophia-Antipolis), OFCE & LARSEN
210 Working papers 2006-2009 du Gis Larsen

Avant propos

Dominique Finon, directeur du Gis LARSEN

Les recherches développées dans le cadre coopératif du Gis LARSEN en 2009 se sont développées dans quatre directions principales :

- l'analyse des limites des réformes en termes d'efficience de long terme, notamment en relation avec l'enjeu de l'investissement en capacité de production,
- l'adaptation des politiques de la concurrence aux particularités des marchés de l'électricité et le long terme,
- l'analyse de la dynamique de la concurrence de détail en termes de comportement stratégiques pour les concurrents et de bénéfices pour les consommateurs,
- l'efficacité de la promotion des nouvelles technologies non carbonees et de leurs contraintes dans les systèmes électriques.

Les working papers du LARSEN (disponibles sur le site gis-larsen.org) rassemblent les productions spécifiques des chercheurs de l'Université Paris Sud (ADIS/GRJM), du CIRED, d'EDF R&D sur ces questions. Ils sont ensuite pour la plupart publiés dans des revues à comité de lecture (Environmental and Resource Economics, Energy Policy, Utilities Policy, Economies et Société, Competition and Regulation in Network Industries, etc.) ou de revue professionnelles comme la Revue de l'Energie.

Plusieurs séminaires et des workshops internationaux ont été organisés par le LARSEN sur ces thèmes, donnant lieu à publications de numéros spéciaux de revues ou bien d'ouvrages :

- Le workshop international sur 'Capacity mechanisms for long term supply security in liberalised electricity markets' organisé le 16 mars 2007, suivi de la publication d'un numéro spécial d'Utilities Policy, avec les auteurs de référence sur le sujet (Paul Joskow, Steve Stoft, Ignacio Perez Arriaga, etc.) au côté d'articles du LARSEN.
- Le workshop international 'Large scale wind Power, market power and market design' organisé le 21 et 22 Juin 2008, qui donnera lieu à publication d'un numéro spécial de la revue Energy Policy.

- Le colloque international 'Efficiency, Competition and Long Term Contracts in Electricity Markets' co-organisé avec la chaire de Politique énergétique européenne de l'Institut universitaire européen de Florence, les 15-16 Janvier 2009, qui sera suivi par la publication en 2010 d'un ouvrage chez l'éditeur Edward Elgar.
- Le workshop 'Du marché électrique décentralisé à l'oligopole vertical' co-organisé avec le GREDEG (CNRS et Université Sophia-Antipolis) et l'OFCE, le 6 novembre 2009 à Sophia Antipolis.

Comme nous l'avons fait l'an passé dans un recueil intitulé 'les imperfections des marchés électriques libéralisés' on a choisi ici de rassembler sept des plus récents working papers du LARSEN dans ce nouveau recueil sous le titre 'Concurrence et politiques publiques dans les marchés électriques' sous trois rubriques.

Les politiques de promotion des énergies renouvelables

Dominique Finon.

Photovoltaïque, l'inadéquation du tarif d'achat.

Ce working paper critique la politique de promotion à tout prix du photovoltaïque par l'usage de tarifs d'achat très généreux superposés à d'autres subventions publiques, en raison de l'éloignement de la technologie de sa maturité commerciale. L'argument est que les effets d'apprentissage n'ont pas suffisamment joué et les baisses futures de coûts et de prix réels sont insuffisantes. Il critique l'usage d'un tel instrument non seulement pour l'objectif de politique climatique (dont le coût est très élevé), mais aussi pour l'objectif de développement industriel d'une filière nationale (la diffusion actuelle et future se reposant sur des achats d'équipement étrangers).

Céline Hiroux, Marcelo Saguan.

Large-scale wind power in European electricity markets : time for revisiting support schemes and market designs ?

Ce papier examine les évolutions nécessaires des règles de marché pour orienter les choix des producteurs éoliens au fur et à mesure de l'accroissement de la part des productions intermittentes dans les systèmes électriques dérégulés. Il montre qu'il est nécessaire de

faire payer les producteurs pour les coûts qu'ils entraînent en les faisant participer aux marchés (énergie, services auxiliaires, congestion). Il souligne aussi la nécessité de trouver le bon équilibre entre responsabilisation et limitation des risques en définissant certaines règles spécifiques.

Les particularités de la concurrence de détail

Christophe Defeuilley, Matthieu Mollard.

The dynamic of competition in presence of switching costs. The case of British Gas (1997-2007).

Les modèles de concurrence avec coûts de changement montrent qu'en dynamique les commercialisateurs sont incités à adopter des stratégies en deux temps. Ils consentent des sacrifices pour acquérir des parts de marché en première période (phase d'investing) parce qu'ils anticipent de pouvoir rentabiliser leurs engagements en seconde période (phase d'harvesting). L'article montre que cette stratégie en deux temps n'est pas toujours couronnée de succès. Notamment lorsque certaines firmes ont des difficultés à fidéliser leurs clients en seconde période parce que ceux-ci bénéficient d'effets d'apprentissage leur permettant de réduire les coûts de changement qu'ils supportent.

Raphaël H. Boroumand, Georg Zachmann.

Retailers' risk management and vertical arrangements in electricity markets.

Dans le modèle de référence des marchés libéralisés, le business model type du commercialisateur est celui du fournisseur pur d'électricité ne détenant ni actifs de production, ni activités de réseaux. Mais les difficultés rencontrées par ce type de fournisseur reflètent l'in incapacité de ce modèle organisationnel à gérer efficacement les risques liés à l'achat et à la revente d'électricité dans un contexte de forte concurrence en prix. Certes, l'achat d'électricité par des contrats peut réduire l'exposition des fournisseurs à la volatilité des prix de gros mais elle constraint simultanément leur flexibilité alors qu'ils font face à une demande structurellement aléatoire. Un fournisseur ne peut donc pas définir ex ante son niveau d'exposition aux risques. Celui-ci ne sera révélé qu'ex post, une fois la demande réelle connue.

Par une approche portefeuille, cet article montre que, contrairement au modèle théorique de référence des réformes concurrentielles, les portefeuilles de contrats financiers sont des substituts imparfaits à la détention d'actifs physiques dans une perspective de couverture de risques.

Les politiques de la concurrence et l'organisation des marchés

Delphine Perrot-Voisard, Georg Zachmann.

HHI, an irrelevant indicator without a relevant market.

Ce working paper propose de revisiter l'utilisation des indicateurs de concentration servant à évaluer le risque d'exercice de pouvoir de marché sur les marchés électriques en partant de la réalité constamment changeante de l'intégration physique des marchés et de la dimension des marchés pertinents.

Adrien de Hauteclercque.

Legal uncertainty and competition policy in deregulated network industries. The case of long term contracts.

Le cas des contrats de long terme sur les marchés électriques de l'Union Européenne est emblématique des difficultés rencontrées par les autorités de concurrence avec la libéralisation des industries de réseaux. L'ambiguité des effets de ces contrats sur la structure concurrentielle, l'investissement, et le bien-être à long terme des consommateurs les positionne logiquement comme une priorité de la législation antitrust. Cependant, en raison de l'absence de référentiel et du processus en-cours de modernisation de la législation concurrentielle, l'incertitude légale perçue par les acteurs du marché est forte. Cet article se propose d'explorer les implications de la stratégie de la Commission pour faire face à ces enjeux.

Dominique Finon, Christophe Defeuilley, Fredéric Marty.

Signaux-prix et équilibre de long-terme. Reconsidérer l'organisation des marchés électriques.

Les industries électriques ont été libéralisées en référence au modèle d'équilibre de marché. La structure de marché dé-intégrée verticalement est supposée permettre l'optimalité des décisions

d'investissement. Ce modèle présente plusieurs défauts. D'une part le marché ne permet pas d'assurer un équilibre physique en pointe et en toute circonstance. D'autre part il ne permet pas d'orienter correctement les investissements pour aboutir à un parc optimal de production. Il y a « défaillances de marché ». Il faut dès lors accepter que les intervenants sur les marchés puissent utiliser une gestion de risques que le marché ne prend pas en charge. Ce qui nécessite un aménagement du cadre d'organisation du secteur électrique allant vers une reconnaissance de l'utilité de cette coordination hors-marché. La concurrence imparfaite pourrait être vue non plus comme le symptôme d'une réforme inachevée mais comme le palliatif à des défaillances de marché.



les politiques de promotion des énergies renouvelables

Photovoltaïque : l'inadéquation du tarif d'achat

Dominique Finon

CIRED & LARSEN

La politique actuelle de tarifs d'achat garantis appliqués au photovoltaïque manque de rationalité : non dégressivité des tarifs, durée trop longue des engagements de subventions par projets, cumul difficilement lisible des dispositifs d'aide (crédit d'impôts, subventions directes, prêts bonifiés, actions des régions.) et donnant des temps de retour trop avantageux sur beaucoup de projets. L'absence de différenciation entre filières peut gêner aussi l'émergence de nouvelles technologies du PV. Du coup on doit légitimement se poser la question de l'importance des dépenses futures annuelles associées aux aides engagées à mettre en regard avec les avantages potentiels de la politique qui vise à la fois à la réduction des émissions de CO₂ par la production électrique et le développement d'une filière nationale. On doit s'interroger en particulier sur la capacité de la France en position de second mover de rattraper les first movers (Allemagne, Japon, Etats-Unis). On ne critique pas en soi une politique fondée sur les tarifs d'achat, politique tout à fait adaptée à des techniques proches de la maturité commerciale en leur garantissant un signal stable et proche des prix de gros du marché électrique. On critique son inadéquation au niveau de maturité de la technique PV qui conduit à avoir des tarifs d'achat cinq fois plus élevés que pour l'éolien. Nous proposons de recibler le dispositif au regard de l'état de maturité de la technique (marqué par une stagnation des coûts et des prix) tout en prenant en compte les enjeux importants que pourra représenter le PV dans l'offre mondiale d'énergie. On devrait réduire rapidement les incitations coûteuses à l'installation de cellules PV et de centrales au sol, et accroître les budgets de démonstration industrielle pour favoriser la consolidation de nouvelles filières (silicium cristallin amélioré et couches minces). On devrait cibler plus particulièrement la phase de démonstration et le développement industriel où se joue une grande partie des apprentissages porteurs de l'essentiel des baisses de coût.

Introduction

La France veut se situer dans le courant des pays qui veulent faire une place au PV dans l'offre d'énergie en vue de la réduction des émissions de carbone et pour être présent industriellement dans le compétition mondiale autour de technologies vertes. Elle a adopté pour ce faire des tarifs d'achat comparable à ceux de l'éolien mais fixés à un niveau quatre à cinq fois plus élevé. Alors que beaucoup de voix s'élèvent pour critiquer les tarifs de l'éolien¹ et alors qu'il suscite beaucoup de scepticisme feutré dans les ministères et les entreprises, la critique ne suscite que des murmures, sans doute parce que la production électrique par des systèmes PV est encore confidentielle et que son image croise utilement progrès *high tech* et développement durable, comme le montrent les objectifs du récent Grenelle de l'environnement (MEDDAD, 2008). Il faudra sans doute attendre que sa production atteigne le TWh dans dix ans pour se rendre compte du coût excessivement élevé de cette politique alors que, dans une économie globalisée, ce marché très subventionné aura peu d'effets sur la constitution d'une filière nationale compétitive.

La diffusion des systèmes photovoltaïques dans le monde connaît une croissance très rapide (40% par an depuis 2006) qu'accompagne une compétition intense entre industries des pays de pointe et à présent de pays émergents. En Allemagne, Espagne, Californie, comme en France et en Italie maintenant, on observe une croissance exponentielle des contrats d'achat d'électricité PV par les entreprises électriques soumises à l'obligation de rachat² ainsi que le développement de projets de centrales au sol à base de PV (22 projets en France, début 2009), ce qui laisse anticiper une diffusion rapide et multiforme de cette technique dans un proche futur dans le monde et en France. Des annonces spectaculaires de baisse de coût des cellules par les industriels laissent croire à l'accès rapide de la technologie PV à ma-

¹ C'est à tort car il donne à l'investisseur la possibilité de revenus stables et prévisibles pour engager ses projets et son niveau de 85€/MWh n'est pas si éloigné du prix du marché de gros de l'électricité qui est croissant sous l'effet de la hausse des prix des combustibles et du carbone. Parmi les autres instruments, l'appel d'offres pour des contrats de long terme ne donnerait pas de coûts inférieurs pour atteindre la même cible de capacité ENR, tandis que l'obligation de certificats échangeables ne fait qu'ajouter des risques de marché au risque habituel d'investissement en unité ENR, ce qui entraîne un prix du capital beaucoup plus élevé et des réalisations inférieures à la cible d'installations visées (Finon, 2007).

² En France, en juin 2008, 36000 demandes sont déposées, alors que 5200 installations sont connectées.

turité commerciale. Depuis le niveau de 4500 €/kWc en 2000, le coût s'établirait, nous disent les industriels et les technologues de l'Agence Internationale de l'Energie (IEA, 2008) à 3000 €/kWc en 2010 et devrait baisser vers 1000 €/kWc, le niveau de parité du prix de l'électricité vendue en sortie du réseau de distribution, considéré par les promoteurs du PV comme le niveau de compétitivité de la filière pour les installations raccordés au réseau. Les annonces de la mise en place d'usines de fabrication de grande surface photovoltaïque (plus de 1000 MW) au Japon, en Allemagne, en Californie et depuis peu en Chine semblent conforter ces anticipations. On nous présente à souhait la décision prise à la mi-2009 d'une société américaine First Solar de construire une petite usine de 100 MW en France comme le début de rattrapage de retard français. De même le business model de compagnies promotrices de projets photovoltaïques en France est présenté comme l'exemple même de l'audace managériale alors que ces modèles d'affaires ne reposent que sur un empilement de subventions multiples, dont les tarifs d'achat.

Toutefois des questions se posent devant la réalité beaucoup moins brillante de la progression de l'innovation photovoltaïque et de son stade réel d'avancement technologique. Comme le montre un rapport très fouillé du Lawrence Berkeley Laboratory, les coûts et les prix des systèmes du silicium cristallin stagnent en termes de baisse de prix et de coût depuis 2000, tandis que les incertitudes de fonctionnement des filières concurrentes dans la durée persistent. Les marchés ne se développent qu'en raison des subventions très élevées venant des tarifs d'achat adoptés dans plusieurs pays.

On se propose ici de reprendre le débat sur les instruments de promotion des énergies renouvelables en production électrique à propos du photovoltaïque, et leur application en France. Entre subvention à la R&D, subvention à l'investissement et subvention à la production de type tarifs d'achat, ces instruments doivent être choisis en adéquation avec le niveau de maturité de la technique. Quand on choisit les tarifs d'achat, on fait le pari que des effets d'apprentissage consécutifs au développement cummulé des capacités installées feront baisser les coûts unitaires. Le pari peut devenir très coûteux si on se trompe sur la pente de cette courbe d'apprentissage et si on confond les effets de la R&D, les effets d'apprentissage et les effets de production en taille industrielle. Il est donc important de comprendre

les facteurs de baisse de coût et la place des effets d'apprentissage spécifiques aux technologies PV pour s'interroger sur la pertinence de l'instrument utilisé.

On procèdera en quatre étapes. Dans un premier temps on pose les principes du soutien public au développement des technologies ENR et on identifie les types d'instrument à utiliser selon les phases de leur évolution technologique et industriel. Dans un second temps, on caractérise les phases où se situent les différentes filières du PV. Dans un troisième temps on discute le choix du type et du niveau de subvention aux techniques photovoltaïques. Enfin dans un dernier temps, on discute de l'intérêt de l'objectif de construire une filière nationale pour un pays qui s'est laissé distancer sur les techniques de première génération.

Le soutien public au développement et au déploiement des innovations ENR

Le soutien à une nouvelle technologie ENR se justifie en théorie économique par l'existence de défaillances de marché. En dehors des défauts possibles d'internalisation des coûts environnementaux des technologies dominantes de production électrique, la première est propre aux technologies qui sont à un stade initial de développement et d'expérimentation et nécessitent des efforts de R&D et de démonstration importants. La subvention publique est justifiée, compte tenu du caractère de bien public des connaissances scientifiques et technologiques³. La seconde est propre aux barrières à l'entrée de toute innovation dans un espace technologique occupé. Amorcer le déploiement commercial de la technologie en la tirant par la demande permet d'enclencher les effets d'apprentissage accompagnés de l'introduction d'innovations incrémentales pour améliorer les procédés de fabrication et la technique elle-même qui peut arriver ainsi à concurrencer les techniques en place. La troisième est propre aux

³ Le raisonnement économique est le suivant : Sans subvention, les agents privés qui ne peuvent récupérer toute la valeur économique créée par les connaissances qu'ils ont générées vont sous-investir en R&D par rapport à l'optimum social.

effets industrialisants de la diffusion d'une innovation émanant de l'industrie nationale. Quand un pays est bien placé au départ dans la compétition technologique, il y a un véritable intérêt à accélérer la dynamique d'apprentissage en amplifiant les débouchés par une politique de *demand pull* par subvention à l'adoption.

Pour identifier à présent le type de soutien le mieux adapté à chaque stade de développement d'une technologie ENR, on se réfère généralement au modèle linéaire de l'innovation en plusieurs étapes depuis l'invention, la R&D appliquée, la démonstration, la diffusion pré-commerciale et à la diffusion commerciale. Dans cette représentation du développement technologique et de diffusion, le décollage de la technologie commence doucement pour s'accélérer sous l'effet des rendements croissants d'adoption et des baisses de coûts et de prix avant de ralentir à l'approche de la saturation du marché. Les instruments de soutien doivent être adaptés à chacun de ces stades.

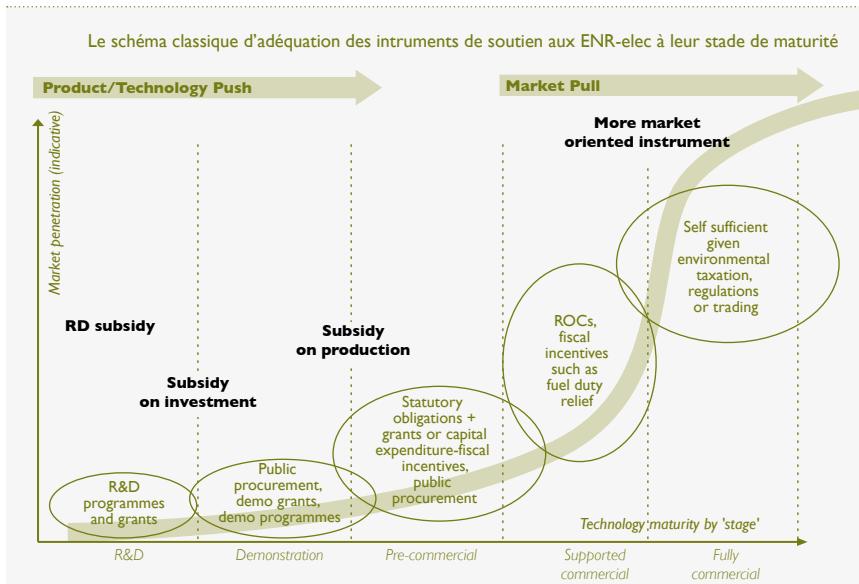
La phase de R&D porte sur le développement des savoirs scientifiques et technologiques. Ce développement est de nature et d'amplitude très différentes d'une technologie ENR à une autre, comme on l'imagine aisément entre l'éolien qui recourt aux techniques améliorées de l'électromécanique et de l'aéronautique, et le PV qui doit recourir aux connaissances les plus nouvelles et les plus en pointe en photoélectronique. A ce stade interviennent les laboratoires publics et les subventions publiques dans le cadre de partenariats public-privé.

La phase de démonstration se caractérise par la réalisation de quelques unités prototypes de taille croissante jusqu'à la réalisation de prototype de taille commerciale lorsqu'il s'agit de techniques de grande taille (ce qui est le cas par exemple des centrales solaires à cycle thermodynamique). Dans le cas du PV où les installations sont pour les plus nombreuses de taille réduite (3 à 100kW) et fréquemment d'usage individuel, la phase de démonstration est marquée par la construction d'un marché-niche via des aides importantes accordées aux utilisateurs pour permettre aux innovateurs de développer des procédés de fabrication pour des unités de taille industrielle autour des filières dominantes. C'était l'objectif des programmes dits 10 000 ou 100 000 toits photovoltaïques du Japon et de l'Allemagne dans les années quatre vingt-dix et du Royaume Uni actuellement. Cette

phase est financée par des subventions à l'investissement, surtout quand la technologie ENR est très capitaliste et se heurte aux contraintes d'accès au financement des adopteurs. C'est grâce à cette création de marché que de petites entreprises peuvent se constituer au cours de cette phase.

La phase de déploiement pré-commercial est l'étape où se développent les apprentissages *by doing* et *by using* et où les usines de fabrication peuvent passer à une échelle supérieure de production. Cette période est donc dominée par le développement des savoir-faire industriels et leur diffusion. Elle s'accompagne de l'ajustement des règles institutionnelles pour faciliter la diffusion de la technologie⁴. Les investissements des fabricants comme ceux des utilisateurs sont particulièrement risqués à ce stade. Des acteurs plus grands entrent au côté des petits innovateurs. Si l'industrie locale a démarré en retard du fait de supports plus timides que dans les autres pays, on peut imaginer des importations de technologie par achat de licences et une éventuelle entrée d'industriels locaux sur le marché par des innovations mineures sur la technologie principale. L'Etat à ce stade or-

Figure 1



ganise la création de marchés à plus grande échelle pour générer une demande d'ampleur et enclencher la dynamique d'effets cumulatifs. Deux voies sont possibles : soit une subvention à l'investissement des utilisateurs de technologie, soit une subvention à la production par un dispositif garantissant les revenus des nouveaux équipements pendant une grande partie de leur durée de vie. C'est l'esprit des tarifs d'achat garantis sur 15-20 ans⁵. Le choix entre *subvention à l'investissement* et *subvention à la production* dépend à ce stade des caractéristiques de coût de la technique innovante, de la taille financière des adopteurs et du stade de développement de la technique.

La subvention à l'investissement est nécessaire en début de déploiement pré-commercial lorsque la structure de coûts de la technique est dominée par les coûts d'investissement (*upfront costs*) comme le PV et l'éolien et quand les adopteurs sont de petits acteurs. Elle peut prendre différentes formes : subventions directe, bonification d'emprunts, crédit d'impôts, etc. Mais la subvention à l'investissement présente deux défauts en phase de maturité. L'aide à l'investissement n'encourage pas les utilisateurs à rechercher le meilleur équipement en performance, ce qui ne contribue pas à une sélection rapide des fabricants les plus fiables. Elle n'incite pas non plus à l'entretien des équipements et peut conduire à arrêter l'équipement à la première difficulté importante dès lors que leur apport en fonds propres est remboursé.

La subvention à la production devient un dispositif plus efficient au stade de la maturité pré-commerciale, car elle est basée sur la performance en production des unités installées. Elle incite donc à recher-

-
- 4 L'Etat doit ajuster les réglementations et les procédures d'autorisation pour limiter les obstacles administratifs, comme dans le cas de l'éolien ou l'adaptation de règlements d'urbanisme et du droit des propriétés dans le cas du photovoltaïque.
 - 5 On signalera aussi deux autres dispositifs de subvention à la production. D'abord les *contrats à prix garantis sélectionnés par appel d'offres* sur la base des prix offerts, qui constituent une variante du dispositif du tarif d'achat garanti. Cette variante utilisée pour les centrales PV en France présente une incitation à serrer les prix demandés, mais ne garantit pas forcément que ces prix seront inférieurs à des tarifs d'achat, sachant que ceux-ci sont établis pour garantir la rentabilité des projets sur des technologies encore dans l'enfance, et décroissent pour suivre les baisses de coût. Second dispositif, les *obligations de certificats verts échangeables* imposées aux fournisseurs (auquel plusieurs pays et la Commission européenne sont favorables) : il se situe dans le même esprit en apportant un revenu complémentaire aux investisseurs en ENR électriques qui sont supposés vendre normalement l'électricité d'un côté et les certificats sur un marché spécifique de l'autre. Mais ce revenu est très variable et incertain par rapport au tarif d'achat garanti. Certains pays incluent le PV dans les technologies éligibles en donnant à chaque kWh vendu une quantité quadruple ou quintuple de certificats

cher de bonnes performances car l'investisseur en unité ENR reçoit un paiement sur la durée de vie de l'unité en relation directe avec sa production. Elle contribue au recouvrement de l'investissement et à la rémunération du capital, mais, ne le faisant qu'ex post, elle ne présente pas les inconvénients de la première. Elle incite l'investisseur à choisir une technologie performante parmi celles proposées sur le marché. Elle incite aussi l'exploitant à procéder à l'entretien nécessaire au maintien des performances. La subvention sur chaque équipement est assurée sur une durée suffisamment longue pour permettre la rentabilisation normale. Elle évolue au fur et à mesure de la baisse des coûts par les effets d'apprentissage, pour s'effacer par la suite.

On notera que dans le cas des subventions d'investissement, les coûts des dispositifs sont payés par le budget de l'Etat, tandis que, dans le cas des subventions à la production (tarifs d'achat), le coût est généralement remboursé par l'ensemble des consommateurs, via une taxe sur l'ensemble des kWh transportés. Vu cette différence de mode de remboursement, on notera que le dispositif de subventions à l'investissement (ou le crédit d'impôt à l'investissement) est exposé au risque de « stop and go » de telles politiques car elles dépendent de la conjoncture économique et électorale. C'est ce qu'on observe par exemple dans le cas des politiques de subvention directe à l'éolien aux Etats-Unis dans les années 80.

Les difficultés de séparer les étapes du développement technologique des filières PV

Avec le photovoltaïque, il existe une difficulté pour définir l'instrument et le niveau de la subvention car les frontières entre les différentes étapes ne sont pas claires alors que dans le cas de l'éolien, le progrès des techniques s'effectue plutôt par *learning by doing* et *by using*. Le progrès des connaissances scientifiques, des savoir-faire technologiques et les effets de taille des fabrications industrielles sont étroitement mêlés à chaque étape de progression de la technologie PV. Le besoin de R&D reste très important pour les techniques en concurrence, même celles en tête, alors qu'on les traite comme des

techniques proches du stade commercial. De plus contrairement aux autres technologies ENR où il y a généralement une technique qui s'affirme rapidement comme *dominant design*, la trajectoire technologique du PV n'est pas encore bien établie car plusieurs filières sont en compétition, mais à des stades de développement différents (Menantueau, 2000 ; Lincot, 2007).

Les diverses trajectoires technologiques

Les techniques PV se font concurrence en jouant de deux caractéristiques, le rendement de transformation et le coût de fabrication et d'installation, pour parvenir à une maturité commerciale qui leur permettra dans le futur de concurrencer les moyens classiques de production électrique. Les filières dont les matériaux permettent d'atteindre de très bon rendement (15 à 30%) – les techniques au silicium cristallin - reposent sur des procédés de fabrication a priori plus coûteux et ont des coûts hors modules plus onéreux que les autres filières dits en couches minces -- et vice versa. Le développement d'une telle variété est nécessaire pour qu'émerge un standard qui puisse rivaliser un jour avec les techniques électriques matures. Deux effets, l'effet de R&D et l'effet d'apprentissage jouent et vont jouer conjointement sur les baisses de coût.

Les systèmes au silicium cristallin occupent actuellement 90% du marché. C'est la première technique à avoir été étudiée et développée pour deux raisons : l'une contingente qui était de profiter de la base de connaissance de l'industrie électronique en matière de silicium pur et des disponibilités de matières premières (les rebuts de l'industrie électronique). La seconde était l'attrait des potentialités d'accès à de bons rendements. Le silicium cristallin est utilisé sous deux formes: le mono-cristallin et le polycristallin. Le premier est celui qui a d'abord été développé. Les systèmes ont de bon rendement (15% avec des perspectives de croissance jusqu'à 25%), mais ils présentent des coûts de fabrication élevés car les procédés doivent opérer en discontinu (procédés d'étirement de lingots et de découpage de plaquette, encapsulage, etc.). Le polycristallin présente des rendements moins bons, mais des possibilités de fabrication moins chères (notamment de procédés en continu). L'un des paramètres clés de la maîtrise des coûts est la réduction du taux de rebut à la fabrication qui se montait encore à plus de 50% au début des années quatre vingt-dix. Le début

de passage à des technologies de tirage de rubans à partir de bains fondus (3% en 2006) plutôt qu'en lingots ouvre des perspectives de baisse de coût en évitant l'étape de découpe des plaquettes et en passant à l'automatisation.

Les *filières en couches minces* se développent de leur côté dans deux directions : le silicium amorphe et les filières non silicium. Le *silicium amorphe* possède une possibilité d'absorption des photons plus efficace, ce qui a permis de passer aux couches minces (40-60 microns) sur des supports bon marché et d'utiliser des procédés de fabrication beaucoup moins chers (procédés de dépôt chimique notamment), ce qui compense en partie le moindre rendement des modules (6 à 7%). Etudiée depuis les années 90, cette filière représente seulement 4 % de la production en 2006-2007. Elle est en fait bloquée dans son développement par l'instabilité des rendements dans la durée.

Les filières couches minces non-silicium utilisent des matériaux adaptés à des utilisations en couche mince tout en présentant des promesses de rendement plus élevé. Deux filières se sont imposées progressivement, l'une basée sur le *tellure de cadmium* (appelée *filière CdTe*) et l'autre sur les *alliages cuivre indium/gallium sélénium* (appelée *filière CIS*). Elles permettent d'atteindre des rendements de 8-9% pour les premières et de 11-13% pour les secondes, ainsi que la réalisation des couches ultra-minces (moins d'un micron), associée à la possibilité de contrôle de l'usage des matériaux, de fabrication de modules de grandes surfaces par des procédés automatisés et sur des matériaux de support bon marché. La séquence de fabrication des modules en couches minces Silicium amorphe, CdTe et CIS est totalement différente de celle des modules classiques au silicium cristallin avec des étapes moins lourdes et prometteuses en termes de baisse de coûts.

L'éventail des technologies de production et d'étalement du CIS est ouvert : on teste l'électrolyse, ou bien la sérigraphie au lieu des procé-

Tableau 1

Potentialité de progression des performances des différentes filières PV

Technologies	Cellules au Si monocrystallin	Cellules au Si polycristallin	Couches minces Si-a, CIS, CdTe, Si-nano
Rendements actuels	13-15%	12-14%	6-10%
Rendements potentiels	25 %	20%	13-16%

Source: Données adaptées de tableaux II.1 et II.2 in IEA, 2008, p.374.

dés classiques de dépôt sous vide. Dans les deux cas, elles sont particulièrement bien adaptées aux grandes surfaces grâce à une tolérance du CIS aux défauts structuraux et chimiques des supports, contrairement au silicium. Ces deux filières présentent une bien meilleure stabilité de long terme que le silicium amorphe. Demeure une inconnue de taille sur cette stabilité à plus de 10-15 ans, qui peut se traduire dans le futur par un rejet commercial, notamment auprès des particuliers couverts par les règles de garantie décennale ou bien dans les contrats de garantie ordinaire avec les propriétaires de plus grande installation. Parties en retard, les filières non silicium qui ne couvraient en 2006 que 2% du marché mondial, mais connaissent depuis une accélération de leur diffusion en bénéficiant du pari de l'industrie américaine sur ces filières. Les couches minces peuvent présenter un avantage pour les centrales au sol et pour les installations grandes toitures, ce qui pourrait leur permettre de bénéficier du développement de ces marchés spécifiques, comme on le verra plus loin.

La troisième génération de technologies PV pourrait émerger à partir des années 2030 en empruntant deux voies technologiques. La première est celle des cellules PV à coût très bas mais à rendement faible (10%) de type nanocristallin, dont il existe déjà des pilotes au Japon ou celles de type organique (mais qui présentent d'important problème de stabilité). La seconde voie est celle des cellules à très haute efficacité (30 à 50%) dont la conception utilisera la physique du solide (électrons chauds, trous quantiques) et des nanostructures (IEA, 2008 ; Lincot, 2007). Mais on ne peut pas anticiper de date de maturation technique pour aucune d'entre elles.

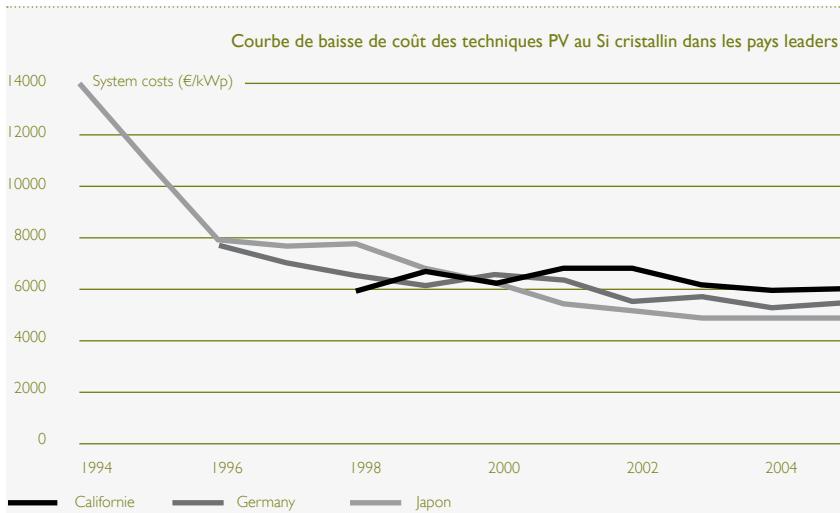
Les difficultés de progression des différentes filières

Les prospectives de l'AIE (2008) et les « roadmaps » des organismes officiels des différents pays considèrent que toutes les technologies vont progresser et qu'aucune ne va définitivement dominer les autres dans le futur, avec par exemple une répartition possible du marché par moitié entre le silicium cristallin et les couches minces en 2020. Elle considère que le coût du système basé sur les cellules SiC qui s'établissait à 4100 €/kWc (6250 \$) en 2006 devrait baisser vers 3400 €/kWc (4400\$) en 2010, puis vers 1500-1700 €/kWc (1900-2200 \$/kWc) en 2020 pour parvenir grosso modo au fameux 1000 €/kWc (1300 \$/kWc) de la fameuse « parité réseau » vers 2040-2050. Cette pro-

gression sur la période 2010-2050 s'effectuerait par la reproduction de l'effet d'apprentissage de 18%-20% observé entre 1985 et 2000 (Neij, 1997 ; Rabl et Van der Zwaan, 2007). Rappelons que l'effet d'apprentissage se mesure par la baisse des coûts à chaque doublement de capacité. Les couches minces connaissent des niveaux de coût voisins, notamment en raison des progressions plus rapides en procédé de fabrication que le SiC (LBL, 2009). Malgré une stabilité des performances peu assurée dans la durée, elles se développent notamment dans les grandes installations au sol ou en toitures.

Toutefois ce scénario se discute. On observera qu'il ignore à la fois les difficultés de progression rencontrées par les techniques au Si cristallin et les contraintes de développement des couches minces. Le ralentissement de la bausse des coûts est visible sur les données rassemblées par Lopez-Polo, Haas et Suna (2007) dans les trois pays leaders (Japon, Allemagne, Etats Unis) (voir figure 1). Le constat de ce ralentissement recoupe les résultats de l'étude de Papineau (2006) à partir de données sur l'Allemagne, la Suisse et les Etats-Unis entre 1992 et 2000. Ce ralentissement intervient alors que cette technologie n'a pas encore atteint le stade de maturité, ce qui pose une véritable interrogation car, comme l'histoire des techniques le montre, au stade du déploiement pré-commercial, les baisses de coût devraient s'accélérer (Ayres et Martinas, 1992).

Figure 2



Croissance de taille des lignes de production selon les filières (en MWc/an)

	1995 -2005 13-15%	2006-2010 12-14%	Après 2010 6-10%
Silicium cristallin	20-30 MWc	400 MWc	1000 MWc*
Couches minces	15 MWc après 2000	15-100 MWc (Japon, All.)	400-600 MWc/an**

* Annonces de construction au Japon, en Chine, aux Etats-Unis et en Allemagne.

** Annonce de Nanosolar avec CIS en Californie et en Allemagne.

Source: Lincot (2007) et divers.

Tableau 2

Du côté des couches minces, en dehors des obstacles technologiques de stabilité des performances à surmonter, elles pourraient être bloquées par des contraintes de ressources. D'une part les couches minces de CIS et de CdT peuvent se heurter à des contraintes de ressources de tellurium ou d'indium dont la production mondiale pourrait ne pas décoller. L'indium par exemple qui est aussi abondant dans la nature que l'argent et le cadmium n'est produit à l'heure actuelle que comme le sous-produit de l'extraction d'un autre minéral. Sa production est cent fois moins importante que celle de cadmium (Academie des technologies, 2008). D'autre part la toxicité du cadmium pourrait aussi être un sérieux paramètre de sélection des technologies, même si les composés du cadmium utilisés n'ont pas de dangerosité avérée. Les réglementations sur le cadmium ont déjà dissuadé des entreprises innovantes comme BPSolar de s'engager dans la filière CdT .

On pourrait aussi imaginer qu'un verrouillage technologique puisse s'opérer sur le SiC avec des performances médiocres (selon le schéma classique d'Arthur et Cowan), sachant qu'il s'est développé plus rapidement que les couches minces et qu'il bénéficie du même type d'aide que les couches minces. Il bénéficie donc avant les couches minces de la possibilité de passer à de grande taille qui constitue dans le cas du SiC une puissante incitation à améliorer considérablement les procédés⁶. En d'autres termes les effets combinés d'apprentissage

6 Il y a des progrès à attendre au niveau des coûts de fabrication des cellules SiC sur les procédés de fabrication : montage de lignes de production automatisées sur le procédé discontinu en lingots sur des productions de 400 MWc après le saut de taille depuis des unités de 20 à 30 MWc (95% des procédés sur le Si-cristallin sont en discontinu en 2008), passage au procédé en continu de tirage de rubans qui conduit à une meilleure utilisation de la matière première. (IEA, 2008, Lincot 2007). Des progrès sont aussi à attendre au niveau de la conception (multi-jonction) et avec le passage probable du SiC aux couches minces qui est considéré comme possible pour 2020 (IEA, 2008).

et de l'augmentation de taille des usines que permet la croissance des marchés tirés par des subventions élevées tendent à favoriser le silicium cristallin (voir tableau 2). Mais trois éléments infirment dès à présent ce scénario du verrouillage sur le SiC à rendements médiocres. D'abord les couches minces connaissent depuis 2006 un engagement certain des industriels, et bénéficient de pouvoir être tirés par des marchés privilégiés, les centrales au sol et les grandes toitures, (ce qui est aussi une solution classique de déblocage d'un verrouillage technologique). Ensuite la plupart des Etats cherchent à encourager en RD&D toutes les technologies (AIE, 2007).

Enfin si verrouillage il y avait, le SiC présente des promesses d'amélioration au niveau des procédés, mais aussi plus immédiatement au niveau des coûts. Une partie de la stagnation des coûts entre 2003 et 2006 s'explique par la hausse des coûts des matériaux du fait de la pénurie de silicium pur résultant de l'insuffisance du gisement des rebuts de l'industrie électronique, mais cette barrière vient d'être surmontée par le démarrage d'usines spécialisées dans la fabrication de silicium cristallin dédié au PV dans le monde (Académie des technologies, 2008). Mais nouvel avatar d'un développement technologique en situation de concurrence imparfaite, joue à présent sur les prix le pouvoir de marché des vendeurs de cellules PV qui ont profité des tensions sur les capacités de production du fait de la croissance trop rapide d'un marché mondial tiré par les subventions généreuses accordées aux acheteurs pour ne pas baisser leur prix.

Les spécificités du processus de développement technologique du PV

En résumant, dans le domaine du PV, le processus de développement présente trois spécificités qui complexifient le choix d'un instrument de subventions. D'une part la R&D joue en continu un rôle moteur en ne portant pas seulement sur la cellule elle-même et son rendement, mais sur le progrès des techniques de fabrication. On vient de voir que toutes les techniques PV, même celles de première génération au Si cristallin, ont besoin de progresser encore par des efforts de R&D focalisés en majorité sur les techniques de fabrication et d'assemblage. On a besoin de subventions et de partenariat public privé autour de plateformes industrielles à ce niveau. Est ce pour autant le seul moyen ? L'existence d'une subvention élevée à la production de type tarifs d'achat qui autorise les constructeurs à pratiquer des prix

élevés et à dégager des marges satisfaisantes, notamment parce que leurs débouchés se développent aussi, est un moyen de permettre un réinvestissement en RD&D industrielle sur les procédés. Mais est ce en soi une justification de choisir cet instrument ? On revient par la suite sur cette question.

D'autre part dans le *learning by doing* particulier au PV, on doit souligner l'importance du saut dans les échelles de fabrication car c'est ce saut de taille qui à la fois incite et permet l'amélioration des procédés de fabrication. Il va jouer aussi dans la concurrence entre filières en donnant un avantage à la filière qui a bénéficié des premiers efforts dans un nouveau champ technologique, effet classique de *lock-in* que l'on retrouve dans différentes technologies, comme cela a été le cas des filières nucléaires (Cowan, 1990).

Enfin l'effet d'apprentissage joue à tous les niveaux des procédés de fabrication et d'installation des systèmes pour chaque filière. Il concerne les coûts hors modules (convertisseur, câblage, etc). Il va concerner aussi le développement de composants standardisés de la construction pour l'intégration du PV dans le bâti. Plus en aval encore, l'apprentissage concerne les savoir-faire dans la commercialisation et l'installation des systèmes photovoltaïques. Le caractère individuel ou décentralisé de la production par photovoltaïque ajoute cette dimension dans les apprentissages par rapport au processus d'innovation des éoliennes, car il nécessite la mise en place de véritables « réseaux d'innovation » à l'échelle de l'industrie régionale et de l'artisanat en installateurs PV dans chaque pays.

Etant donné la structure des coûts d'un système PV -- le coût du module proprement dit couvre 60% du coût et celui hors module qui comprend l'onduleur, le câblage, la structure de montage, l'installation et la commercialisation des systèmes couvre 40% du coût -- il est pertinent de différencier les effets d'apprentissage sur les différentes parties du coût. Les travaux économétriques (Duke, 2002 notamment) montrent que l'apprentissage sur les coûts des modules est un phénomène agissant plutôt à l'échelle de l'industrie mondiale tandis que l'effet d'expérience sur les coûts hors module relève plutôt de l'échelle locale. Sur le cas du Japon, Duke montre même que ce dernier est plus important que l'effet d'expérience sur les modules.

Cette observation invite à ne pas perdre de vue l'intérêt de développer des compétences sur le « hors module » et sur l'installation

(montage de projets, formation, labelisation des installateurs, etc.) à un moment pertinent de progression de la technique et du développement de la filière PV. Tiré par un tarif d'achat généreux et un empiement d'autres aides, la filière se développera d'abord par la mise sur pied de l'aval (métier de développeur de projet, métier d'installateur, etc.) mais ce peut être en décallage complet avec l'état du développement des productions de cellules et de systèmes dans un pays et entraîner des importations d'autres pays plus avancés. C'est cette situation qui se dessine en France sous l'effet de cette politique de tarifs d'achat qui recherche d'abord un effet de vitrine.

Le choix du tarif d'achat : un choix dispendieux pour le PV

Le type d'instrument choisi doit être défini par rapport aux caractères de la technologie (taille, production à l'échelle individuelle ou industrielle, part du coût initial dans la structure de coût) et du stade de progression que reflètent le niveau de coût et les incertitudes sur les performances.

Les différents types de subvention

L'évolution du type d'aides et de leur niveau suit en principe la progression de la technique et les baisses de coûts y afférant. Dans les justifications habituelles des promoteurs du PV, il ne semble y avoir aucun doute à ce que la progression de la technique va s'opérer en permettant des baisses de coût et de prix rapides. Détaillons ces étapes ainsi que celle de la compétitivité supposée de la filière à parité réseau.

Après la phase initiale de la R&D qui aboutit à la mise au point des premières cellules PV au SiC et au test du silicium amorphe, les politiques de promotion du PV se sont appuyées sur les subventions à l'investissement dans le cadre des programmes 10/100 000 toits photovoltaïques jusque vers 2000. Ces programmes avaient pour but de créer un petit marché-niche pour les producteurs nationaux et

susciter la création d'unités de fabrication de taille croissante pour tester et développer les procédés. Les politiques se focalisent depuis sur la création de marché et s'appuient sur deux types d'instruments : les subventions à la production, voie empruntée par presque tous les pays à l'heure actuelle sous forme de tarifs d'achat et de programme de *net metering* (rachat aux prix de détail de toute l'électricité produite par comptage inversé (voir encadré 1)), et les subventions à l'investissement, voie empruntée par la seule Californie actuellement. Comme les coûts des systèmes PV sont très loin d'approcher des seuils de compétitivité, le niveau de subventions à la production doit être extrêmement élevé. Les tarifs d'achat adoptés sont ainsi cinq à neuf fois supérieurs au prix de gros de l'électricité selon l'installation de PV⁷. En principe les tarifs devraient être décroissants pour arriver au moins au niveau du tarif de détail en basse tension pour les installations individuelles et en moyenne tension pour les installations en grande toiture et les centrales au sol. On passerait alors au *net metering* lorsque la décroissance des coûts aura permis d'arriver à la « parité réseau » ou bien pour les installations de producteurs purs à l'achat obligatoire par le réseau de toute la production par les fournisseurs historiques aux prix de vente en sortie du réseau MT ou BT.

Avant d'entrer dans la critique du tarif d'achat et de sa conception dans le cas du PV dans le monde et en France, on doit faire deux remarques sur ce processus à étapes. D'abord tous les gouvernements, en France autant qu'ailleurs, utilisent plusieurs instruments à la fois, en ajoutant aux subventions à la production que sont les tarifs d'achat, de multiples aides à l'investissement : subvention directe, bonification d'emprunts, crédit d'impôt, etc., sans s'interroger sur l'efficacité de cet empilement d'aides et leur cohérence (voir tableau 3). La nécessité d'ajouter aux tarifs d'achat déjà très généreux d'autres types d'aide montrent bien que le tarif d'achat n'est pas l'instrument approprié par rapport à ses principes d'usage précisés précédemment.

Seconde remarque qui porte sur l'aboutissement du processus lors de l'accès supposé de la filière à la parité-réseau, la parité réseau

⁷ Il y a aussi la voie des revenus qu'assure le dispositif d'obligation de certificats verts (CV) échangeables. Dans le système des CV adoptés en GB, en Belgique et au Japon un investisseur en PV tire des revenus de la vente de certificats verts à des acheteurs obligés que sont les fournisseurs d'électricité, avec un comptage des certificats pour les renouvelables.

n'est une référence utile que pour le producteur qui consomme, mais pas pour le marché de l'électricité sur lequel serait déversé cette électricité dans les industries libéralisés. Le net-metering qui émane de cette référence de la parité réseau est fondamentalement une subvention. Dès lors que le PV est raccordé au réseau qui sert à la fourniture en dehors des périodes d'ensoleillement, il n'y a très peu économie en réseau de transport et en production du fait de la non-stockabilité de l'électricité. Une exception comme le montrent Borenstein (2008) et Kahn (2008), dans les pays à fort ensoleillement où il y a une corrélation entre les usages de climatisation et la production PV il y a économie à partir d'une certaine échelle de production photovoltaïque bien répartie sur le territoire d'un opérateur de système. En général il faut rajouter des marges de réserve et des équipements de conduite de réseau supplémentaires lorsque la part de PV dans la production atteint des niveaux très importants et dont il faudrait en fait faire payer le coût aux producteurs PV. Il n'y en a pas du tout d'économie non plus en distribution, en dehors des pertes en distribution évitées. Justifié par la parité-réseau lorsqu'elle serait atteinte, le net metering reviendrait à considérer à tort que la production d'un kWh PV qui est revendue au réseau, même si elle est autoconsommée, permettrait d'économiser tous les coûts en réseau de distribution, de transport et en équipement de production centralisé en ignorant les coûts cachés. C'est une subvention car elle revient à faire bénéficier le producteur individuel d'un tarif de rachat aligné sur le prix d'un kWh BT ou MT (environ 125 €/MWh en France, 250 € au Japon), alors qu'il faudrait que le tarif de rachat à maturité soit aligné sur le prix de gros ou s'en rapproche comme dans le cas de l'éolien, puisqu'encore une fois, il n'y a aucune ou très peu d'économie de réseau et on a autant besoin de capacité de production que s'il n'y avait pas de PV dans le système.

On examinera l'adéquation de l'instrument au niveau de progression technologique du PV en discutant du choix des tarifs d'achat effectué par plusieurs pays dont la France, et ce sous trois angles :

- La première question porte sur la conception des tarifs d'achat eux-mêmes - niveau, durée des contrats, différenciation régionale selon l'ensoleillement, décroissance - en admettant que ce soit cet instrument qui doit être choisi.

- La seconde question porte sur le bien-fondé de la subvention à la production et le retour à des subventions à la R&D que prônent certains critiques comme solution plus efficace pour l'économie.

Tableau 3

Les politiques d'aide en phase de déploiement pré-commercial en 2008

France	
Type d'aide	Subvention à la production
Niveau de l'aide	32,8 c€/kWh* 60,2 c€/kWh intégré au bâti
Durée de l'aide	Constant et sur 20 ans
Autres dispositif d'aide	Appel d'offres pour contrats de 20 ans (centrales au sol)
Autres appuis	1. Crédit d'impôt : 50% du coût du matériel (i.e. 50% de 65% du coût d'inv.) 2.TVA allégée à 5,5% (ou nulle pour <3kW) 3.Bonification de 2% des intérêts d'emprunts en région
Allemagne	
Type d'aide	Subvention à la production
Niveau de l'aide	39,6 à 43 c€/kWh de 30 kW à 1MW 54,21 c€/kWh intégré au bâti
Durée de l'aide	Constant et sur 20 ans
Autres dispositif d'aide	-
Autres appuis	1. Bonification du taux d'intérêt nominal de 3,9 à 4,4%** 2. Subventions à l'investissement par les länder 3. Subventions de l'UE et du gouvernement aux usines de fab. de PV dans länder de l'est
Espagne	
Type d'aide	Subvention à la production (avec limite sur installations/an)*'
Niveau de l'aide	34 c€/kWh si < 20kW 32 c€/kWh si >20 kW
Durée de l'aide	Constant et sur 20 ans
Autres dispositif d'aide	-
Autres appuis	1. Bonification de taux d'intérêt 2. Subvention à l'investissement des régions
Japon	
Type d'aide	Subvention à la production Ou système de certificats verts
Niveau de l'aide	43 c€/kWh
Durée de l'aide	-
Autres dispositif d'aide	-
Autres appuis	Maintien de subvention à l'investissement dans non résidentiel
Californie	
Type d'aide	Subvention à l'investissement "Performance based incentives" & Net metering
Niveau de l'aide	2500\$/kW si <100kWc 39c\$/kWh*** si >100kWc sur 5 ans
Durée de l'aide	>100 kWc 5 ans
Autres dispositif d'aide	Net metering
Autres appuis	Crédit d'impôt fédéral

*' Depuis 2008 limitation sur capacité annuelle installée en Espagne

* 40 c€/kWh dans les DOM.

** Programme de 2005 Solarstromerzeugen

*** Subvention dégressive en fonction de la taille de l'installation

- La troisième question porte sur l'efficacité d'une option qui combinerait des subventions à l'investissement et des aides à la R&D industrielle et qui apparaît mieux adaptée au stade de maturité technologique du PV.

Très cher tarif d'achat

La conception générale d'un tarif d'achat d'unités de production ENR présente les caractères suivants :

- Le niveau : les différents tarifs sont en principe calculés de sorte que les adopteurs potentiels soient incités à investir car ils sont censés offrir des temps de retour normaux et des taux de rentabilité corrects.
- La durée d'engagement: le tarif d'achat est accordé pour une durée longue pour toute nouvelle installation en énergies renouvelables. Il présente donc cette qualité de signal stable, propice à l'investissement en technologie capitaliste. Pour chaque contrat il décroît généralement au bout de cinq ou dix ans le temps de recouvrement du coût d'investissement du projet.
- La baisse des tarifs : Les tarifs d'achat décroissent entre les nouveaux contrats pour ne pas laisser aux acheteurs de l'année n une rente par rapport à ceux de l'année de départ du dispositif du fait de la diminution des coûts par l'effet des apprentissages.
- La différenciation entre type de projets : Ils sont en principe différenciés selon deux critères : les tailles de projet pour prendre en compte les effets de taille des projets et les effets d'expérience particulier, et l'importance de la ressource selon la localisation (ici la durée d'ensoleillement).

Dans le cas du PV on différencie ainsi le tarif à la fois par les tailles et les cibles (système sur châssis, intégration au bâti, centrales au sol, etc.) pour tenir compte de ces effets de taille importants pour les grandes installations sur toit ou au sol, mais aussi des besoins d'apprentissage sur les matériaux pour les installations supposées intégrées au bâti. Pour des centrales au sol, le dispositif de l'appel d'offres adopté en France qui conduira à des contrats d'achat sur 20 ans aboutira nécessairement à des prix voisins de ce qui est considéré comme nécessaire pour aboutir à des taux de rentabilité corrects aux yeux des prêteurs. En l'état actuel, un prix offert de 31 c€/kWh serait suffisant pour y parvenir avec les prix de catalogue offerts par les vendeurs de systèmes.

Un défaut de ce dispositif est son exposition au risque de « capture du régulateur ». L'influence des groupes d'intérêts et des industriels qui veulent se développer dans ce domaine technologique se remarque en France sur tous les caractères du dispositif : a) le niveau du tarif et la longueur de la période d'attribution qui est de 20 ans au lieu de 10 à 15 ans pour les autres ENR, comme l'éolien⁸, b) la non-dégressivité sur la durée de chaque contrat en France, comme en Allemagne, en Espagne et en Italie, alors que pour l'éolien il y a le plus souvent une baisse de 100% au bout de 10 ans à la fin de la période de recouvrement; c) l'absence de différenciation selon l'ensoleillement des régions d) l'absence de dégressivité en France sur les nouveaux contrats d'une année sur l'autre depuis le décret de 2005. Enfin les tarifs d'achat sont partout combinés à des aides importantes à l'investissement (crédit d'impôts, allègement de TVA, subventions directes à l'investissement et prêts bonifiés de 2% par les régions, etc.) de façon plus importante que pour les autres ENR électriques.

Le niveau de la subvention et sa durée posent trois problèmes : un problème redistributif en créant une rente indue pour les vendeurs de système PV et pour les acheteurs, rente financée par tous les consommateurs d'électricité ; un problème d'irréversibilité du coût de la politique avec des engagements fixes pris pour chaque projet ; et un problème d'efficience de long terme en favorisant trop les filières au Si cristallin. Considérons ces différents problèmes sur le cas des tarifs d'achat français.

Le problème redistributif

Les tarifs d'achat de 300 à 500 €/MWh, qui sont de même niveau que dans les différents pays européens, participent à l'emballage du marché du PV en Europe qui a pour effet l'offre de prix élevés par les différents vendeurs. Profitant de leur pouvoir de marché, les vendeurs ne répercutent pas depuis 2008 la baisse des coûts dans leur offre de prix, qui a pu reprendre depuis l'ouverture des usines de production de silicium cristallin spécialisé pour la filière PV.

⁸ On peut citer l'exemple des promoteurs de centrales au sol dans les zones du sud de la France qui ont engagé leurs investissements sur la base du tarifs d'achat de 31 c€/kWh, et qui n'en réclament pas moins un relèvement du tarif à 44 c€/kWh. Selon eux, le tarif serait peu incitatif en permettant des temps de retour trop longs. L'effet de taille d'un équipement de quelques MW ne jouerait donc pas si on écoutait ces industriels. Voir Le Figaro, 22 octobre 2008.

Les tarifs d'achat de 31 c/kWh avec les aides d'investissement conduisent à un temps de retour sur investissement d'environ 13,5 ans pour des équipements PV à 4500 €/kWc et une durée d'ensoleillement de 1500 h par an (zone ensoleillée) dans le cas français (voir tableau 4). Avec les revenus stables assurés sur 20 ans, le taux de rentabilité interne du capital engagé sur le projet est de 8,3% sans crédit d'impôts et de 14,5% avec crédit d'impôts. La suppression des aides à l'investissement pour ne maintenir que le tarif d'achat conduirait à un temps de retour de 19 ans, en conformité avec la durée de l'engagement du tarif de 20 ans. Avec le tarif à 56 c/kWh (intégration dans le bâti), pour un même investissement de 4500€/kW, le temps de retour de 7 ans et le taux de rentabilité interne de 14% sans crédit d'impôt sont déjà très attractifs dans une zone ensoleillée en moyenne nationale (1200h). L'adjonction du crédit d'impôt place le temps de retour à 5 ans et le taux de rentabilité interne à 20%. On nuancera cependant ces résultats en soulignant le fait que les installations avec cellules intégrées dans le bâti coûtent souvent un peu plus cher. Pour une unité au sol de 3500 €/kW dans une zone ensoleillée et sans aucune aide que le tarifs d'achat, on tombe aussi sur des bons taux de rendement de 12%, ce qui signifie que toutes les aides supplémentaires (subventions des régions, taux bonifiés, etc.) ne peut qu'améliorer très rapidement le bilan financier des capitaux investis. (Voir tableau 5)

Tableau 4

Bilan financier d'une unité PV individuelle en France (avec mesure de suppression de la TVA ou TVA à 5,5%)		
Projet	4500 €/kW	4500€/kW
	1500h ensoleillement	1200 h. ensoleillement
		Intégration dans le bâti
	Tarifs d'achat sur 20 ans : 31 c/kWh	Tarifs d'achat sur 20ans : 56 c/kWh
Ensemble d'aides	Crédits d'impôts 50%	Crédit d'impôts 50%
Valeur actuelle nette	sans crédit d'impôt : 180 €/kW avec crédit d'impôt : 1320	sans crédit d'impôt : 730 €/kW avec crédit d'impôt : 1840
Temps de retour sur investissement	sans crédit d'impôt : 19 ans avec crédit d'impôt : 13,5 ans	sans crédit d'impôt : 7 ans avec crédit d'impôt : 5,2 ans
Taux de rentabilité interne*	sans crédit d'impôt : 8,3% avec crédit d'impôt : 14,5%	sans crédit d'impôt : 14% avec crédit d'impôt : 20%

* Le taux de rentabilité interne est le taux d'actualisation qui annule la valeur actuelle nette.

Hypothèse de financement : emprunts à 80%, taux d'intérêt nominal à 7%, taux de rendement sur fonds propres recherché de 12%, taux moyen de rendement du capital de 8%.

Hypothèse de dépenses : remplacement de l'onduleur au bout de 8 ans pour une dépense de 1000€ pour une installation de 3 kW.

Hypothèse sur le crédit d'impôt de 50% : il porte sur les dépenses de matériel qui correspondent à 65% du coût total.

Bilan financier d'une centrale au sol sans subvention

Projet	3500 €/kW
	1500h ensoleillement
Aides	Tarifs d'achat sur 20 ans : 31 c/kWh
Valeur actuelle nette	1100 €/kW
Temps de retour sur investissement	11 ans environ
Taux de rentabilité interne	12,20%

Hypothèse . Financement par 80% d'emprunts (taux de la dette de 7%)

Tableau 5

Les niveaux actuels des tarifs d'achat pour une nouvelle installation assure donc des flux de revenus nets très importants après la période de recouvrement de l'investissement. Le propriétaire individuel de PV ou d'une centrale au sol va bénéficier au bout d'une dizaine d'années (13,5 ans dans le premier cas avec crédit d'impôts, moins dans le second cas) d'un revenu qui, ramené en valeur actuelle, conduit à des taux de rentabilité élevés.

Cette perspective de rentabilité est sans doute une incitation efficace à s'équiper de PV ou à installer et exploiter des équipements en grande toiture ou une centrale au sol quand on a les capitaux nécessaires ou un accès facile à l'emprunt. C'est aussi une incitation à l'entrée pour les opérateurs qui installeraient, financerait et possèderaient les installations des particuliers dans le cadre de modèles d'affaires innovants, comme celui que développe EDF Energies Nouvelles. Mais l'absence de décroissance annuelle du tarif pour les nouveaux contrats renforcera la rente des adopteurs de cette technologie dans le futur si les coûts baissent dans les prochaines années.

En s'inspirant en partie des propositions du récent rapport parlementaire présenté le député Serge Poignant (2009) qui se permet des remarques sur le niveau des tarifs d'achat, ces défauts pourraient être rectifiés de trois façons : la décroissance du tarif pour chaque projet, la dégressivité annuelle du tarif d'achat pour les nouveaux contrats, et la différenciation entre zones.

La dégressivité du tarif d'achat pour une installation doit être rapide pour le ramener vers un niveau précis. La pérennité du tarif de départ pour chaque projet devrait n'être assurée que le temps de recouvrir l'investissement, pour tomber au bout de 7 ans.

La baisse du tarif d'achat entre nouvelles installations doit être programmée à des rythmes élevés : en Allemagne, cette décroissance a été

accélérée de 5% à 8% en 2008 et 2009, puis 10% en 2010, puis ensuite 9% pour s'aligner ensuite sur le niveau de la parité réseau pour les individuels, ou celui des tarifs d'achat éolien pour les centrales au sol⁹.

Une dissociation régionale des tarifs. Le tarif est désormais considéré comme trop généreux dans les zones méridionales, et les développeurs ont visiblement bien identifiés les rentes à retirer dans celles-ci.

Une dernière possibilité serait l'adoption du système californien de la subvention à l'investissement de type « performance-based » avec un paiement étalé seulement sur une période courte de cinq ans et en fonction des kWh produits, ce qui revient à une aide de 2000 €/kWc. Ceci évite clairement de perpétuer les rentes après le recouvrement des coûts d'investissement du côté des adopteurs.

Le contrôle du coût de la politique par tarifs d'achat.

Le contrôle du coût de la politique de promotion d'une ENR basée sur le tarif d'achat est une difficulté classique des politiques d'environnement basées sur un « instrument-prix » (ici les tarifs d'achat) plutôt qu'un « instrument-quantité » (ici les obligations des fournisseurs d'acheter un montant d'électricité verte) en situation d'incertitude sur les fonctions de coût et de dommages évités (Menneteau & Finon, 2005). L'instrument-prix ne garantit pas d'atteindre l'objectif visé, soit par défaut, soit par excès, selon la croissance de la courbe de coûts marginaux de production ENR, tandis que l'instrument-quantité peut se révéler très coûteux si les coûts sont beaucoup plus élevés que prévu.

Dans le cas du PV, en définissant un niveau très généreux du tarif, le problème posé n'est pas celui de ne pas atteindre la cible de capacité visée par l'Etat, mais celui d'attirer de trop nombreux adopteurs par les taux de rendement élevés. Dans ce cas cela entraîne un gonflement du coût du mécanisme par le cumul progressif de subventions à verser annuellement sur 20 ans. De ce point de vue, le succès spectaculaire des incitations en Allemagne invite à se poser la question du niveau du coût total de l'engagement en France. Ce succès se manifeste en Allemagne par le doublement annuel de la capacité en place chaque année de 2000 à 2005, puis une croissance de 50% par an depuis 2006 pour aboutir à une capacité de 4500 MW début 2009. Or une croissance de 1000 MW comme en 2007 ou 2008 entraîne

Comparaison de l'évolution des aides pour les nouvelles installations

	Durée de l'aide par installation	Taux de décroissance annuelle entre nouveaux contrats
France	Constant sur 20 ans	Avant 2000, dégressivité de 5% par an Depuis 2006, dégressivité abandonnée.
Allemagne	Constant sur 20 ans	Dégressivité de 8%, puis 10% en 2010, puis 9% par an en 2011 Dégressivité de 6,5% pour les systèmes intégrés au bâti. Vers la parité-réseau (25 €/MWh) en 2015
Japon	Durée du dispositif des certificats	
Californie	5 ans (Subvention à l'investissement)	Dégressivité de 10% à partir d'une certaine capacité cumulée.Limite de dégressivité pour les >100kW : 30\$/MWh

Note : L'Espagne a aussi un système de tarifs d'achat qui est non dégressif sur les nouveaux contrats

Tableau 6

un engagement de 600 millions d'€ à payer aux nouveaux adopteurs de PV pendant 20 ans. En termes de coût cumulé, si l'Allemagne avait arrêté en 2008 le dispositif du tarif d'achat pour les nouvelles installations PV, le coût net par an de l'engagement (après soustraction du prix de gros de l'électricité) serait de 2,5 milliards avant de décroître très progressivement pour ne s'annuler qu'en 2030 et le coût cumulé net serait de 30,8 milliards d'euros (voir Frondel et al., 2008). Si le dispositif est arrêté en 2010, le coût cumulé net serait de 64,2 milliards au rythme actuel de croissance, avec environ 7200 MW installés en PV en 2010.

On a fait le même type d'estimation pour le programme français en posant le principe selon lequel le coût de la subvention est à estimer par rapport aux prix de gros, et non par rapport au tarif intégré en sortie de réseau puisqu'il n'y a pas d'économie à attendre en réseau (à part les pertes) (voir l'encadré 1)¹⁰. On montre que, même moins ambitieuse la politique PV française pourrait coûter 2,25 milliards de € par an en 2020 aux consommateurs d'électricité si les objectifs du Grenelle de l'environnement (5200 MW en 2020) sont atteints alors que l'aide à l'éolien terrestre qui produirait sept fois plus en 2020 ne leur coûterait que 680 millions d'€. Ceci s'explique par le fait que l'aide par kWh produit par PV est 20 fois plus élevée. L'engagement

⁹ La dégressivité proposée par le ComOp ENR de 5% est insuffisante de ce point de vue.

¹⁰ C'est d'ailleurs avec cette méthode que procèdent ou ont procédé le ComOp ENR et la Commission de régulation de l'Électricité

porterait sur 25 milliards d'€ jusqu'en 2040 pour une production dérisoire de 6,5 TWh par an. Le coût de cette option de politique climatique n'en est pas moins impressionnant : le coût de la tonne de CO₂ évitée en production cycle combiné gaz serait de 1500 €/t environ et de 780 €/t en production charbon¹¹

Pour contrôler le coût de la politique du tarif d'achat, trois solutions sont possibles. Une première solution est de limiter les capacités installées bénéficiant du tarif dès qu'un niveau de capacité est atteint annuellement. Débordé par le développement de centrales au sol, le gouvernement espagnol a décidé ainsi en 2008 de limiter la capacité nouvelle bénéficiant du tarif d'achat au niveau de 44c/kWh à une capacité nouvelle installée de 380 MWc/an (avec 100MWc sur systèmes au sol et le reste sur systèmes sur bâtiment). En principe la loi sur l'énergie de 2005 donne une telle possibilité au gouvernement en complément du processus de planification plurianuelle d'investissement (PPI), mais depuis 2005 à aucun moment n'a été annoncé le début d'une réflexion sur la révision des tarifs pour l'après 2010 ;

Une seconde solution est de programmer une baisse très rapide du tarif comme celle du tarif allemand. Une dernière solution voisine est de prévoir dans le décret de mise en œuvre des révisions programmées dès que la capacité installée atteint un niveau précis. C'est une bonne façon de sortir du tarif d'achat si les baisses de coût ne sont pas au rendez-vous parce qu'on aurait trop misé sur le *learning by doing* et la diffusion des savoir-faire en fabrication : autant en prendre acte dans ce cas et changer d'instruments pour mieux miser sur l'effort de démonstration industrielle et de R&D au stade des procédés de fabrication.

Mais s'il est difficile de situer le niveau de coût auquel les filières au Si cristallin pourrait baisser, il est aussi difficile d'apprécier les chances de progression des challengers « couches minces ». De plus les dynamiques d'entrée des industriels sont générées aussi par les barrières à l'entrée inhérentes au processus de production des cellules et des modules couches minces (Académie des technologies, 2008, p.55). La chaîne d'activités est beaucoup plus intégrée que dans le SiC, car elle est centrée sur le procédé d'élaboration des couches

¹¹ Le surcoût serait de 2,24 milliards pour 6,2TWh de production évitée à base de cycle combiné à gaz qui émet 230g/kWh. (ou à base de centrale charbon moderne qui émet 480 g/kWh)

Encadré 1

Le coût futur de la politique d'appui au PV

La promotion des systèmes PV devrait contribuer à l'effort de la France pour respecter ses futurs engagements à développer les ENR pour atteindre en 2020 une part de 20% de son bilan primaire, selon les engagements de la directive Renouvelables du 3ème paquet Energie et Climat. Dans le domaine de la production électrique par ENR, les perspectives de développement du PV resteront modestes, même dans les scénarios les plus ambitieux : l'objectif de 5200 MW en 2020 proposé au Grenelle de l'Environnement ne contribuera qu'à 1,3% des besoins électriques français s'il est atteint et 5% de la production électrique à base d'ENR.

Tableau a. Objectifs de développement des capacités PV en France à horizon 2020

	2012	2015	2020
Objectifs pré-Grenelle (PPI 2006 et Arrêté de 2006 sur le PV)	160	500	950
Objectifs proposés au Grenelle	1100*	2900	5200

* Dont 40% dans les DOM

Si la subvention à la production reste à un niveau très élevé sans dégressivité rapide sur les nouveaux contrats, le coût pour les consommateurs sera important et durable du fait du niveau très élevé de la subvention par kWh. Si l'on suppose que les tarifs de rachat pour le PV restent à ce niveau, le coût annuel net pour les consommateurs s'élèvera à 1,12 milliards en 2015 et à 2,24 milliards en 2020.

On sort donc des niveaux confidentiels actuels où la production de 30 GWh par les 25 MW installés en 2007 ne coûte que 12 millions d'€ d'aide. On notera qu'en 2020 l'aide à l'éolien terrestre qui produira sept fois plus (45 TWh environ) selon l'objectif du Grenelle de l'environnement ne leur coûtera que 680 millions d'€ parce que l'aide ne sera que de 15 €/MWh (85 €/MWh le tarifs d'achat moins 70 €/MWh le prix moyen du marché de gros) au lieu de 310 €/MWh pour le PV (380€/MWh en moyenne moins 70€/MWh pour le prix de gros). Le total cumulé sur 2010-2020 sera de 12,3 milliards. Le coût cumulé des engagements du dispositif si on l'arrêtait en 2020 avec 5400 MW installés serait grossièrement de l'ordre de 25 milliards d'€., ses effets se faisant sentir jusqu'en 2040.

Tableau b. Coût de la politique de développement selon les objectifs du Grenelle

	2012	2015	2020	Cumul 2010-2020
Capacité installée (MW)	1000	2700	5400	
Production (TWh)	1,2	3,24	6,48	
Surcoût de la production PV	0,39	1,12	2,24	12,3

Hypothèses de calcul du tableau b. Calcul de la production avec hypothèse de 1200h d'ensoleillement moyen à 1kW/m². Tarifs d'achat : 305 €/MWh pour les 2/3 de l'électricité PV et 550 €/MWh pour le 1/3 de l'électricité PV par intégration dans le bâti. On ne dissocie pas les tarifs selon la taille et le type des installations (moins cher pour les grandes installations, plus cher pour les centrales au sol) et dans les DOM (plus cher). On fait l'hypothèse de constance du tarif d'achat pour les nouveaux contrats sur toute la période (terme de l'arrêté de 2006). Le surcoût par kWh s'évalue par rapport à un prix moyen annuel élevé du marché d'électricité de gros de 70 €/MWh (qui correspond au coût variable de la centrale à gaz qui serait marginale toute l'année sur les marchés horaires avec un prix du gaz de 6\$/Mbtu et un prix du carbone de 30€/t de CO₂).

A ce coût supporté par les consommateurs via le prélèvement de la CSPE, il faudrait ajouter le coût des crédits d'impôt pour le budget de l'Etat et la compensation des prêts à taux bonifiés. Pour les crédits d'impôt (qui porte sur 50% du coût de l'équipement, hors coût d'installation), la subvention s'élèverait à 0,7 milliards € par an entre 2015 et 2020 si un tel crédit d'impôt est maintenu, ce qui augmente le coût annuel pour la collectivité de 50 % en moyenne sur les cinq années*.

* On calcule ce coût avec une hypothèse de calcul simple : 600 MWc par an d'installation entre 2015 et 2020, 60% en installations individuelles et 4500 €/kWc pour toutes ces installations.

de CdTe ou de CIS combiné à l'étalement. Les procédés au SiC sont parcellisés entre la production de silicium pur, la fabrication des lingots, puis celle des cellules et l'assemblage des modules ; un industriel peut donc se spécialiser sur une partie de la chaîne d'activités. L'entrée dans le domaine des couches minces, même avec des petites tailles (20 MWc), nécessite un investissement initial plus élevé, et le risque associé à la maîtrise technologique de l'outil de production et à la fiabilisation de techniques nouvelles et complexes est important.

Faut-il retourner aux seules subventions à la RD ?

Considérons le point de vue des partisans du retour aux seules subventions à la R&D (voir par exemple Frondel et al., 2008). Les différentes filières ne sont pas assez avancées et doivent d'abord recevoir des subventions de R&D, notamment sur les technologies moins bien établies de couches minces (les nouvelles formes de silicium, les matériaux CIS et CdTe, les nouveaux supports, les PV organiques, etc.). De ce même point de vue la politique du Japon est exemplaire puisque ce pays a maintenu un effort constant en financement public de R&D depuis les années 80 et en le diversifiant alors que les budgets de R&D sur le PV des autres pays leaders (Allemagne, Etats-Unis) connaissaient des fluctuations importantes comme d'ailleurs sur l'ensemble des techniques ENR. Un pays qui ne ferait pas partie du premier peloton des pays innovants pourrait adopter une stratégie de long terme en finançant les technologies appelées à contester la position du *dominant design* avec des techniques avancées, ou à s'imposer sur de nouveaux marchés particuliers, comme certains pays (Japon, Finlande) ont su le faire avec leurs industries électroniques pour rivaliser avec l'industrie américaine.

Ce point de vue fonde la position de ceux qui prônent l'abandon des tarifs d'achat et un retour à un subventionnement de la R&D tels que Frondel et al. (2008) de l'institut RWI à propos du programme allemand. Ces critiques soulignent le fait que les politiques de tarifs d'achat élevés se focalisent à tort sur les effets d'apprentissage en fonction des capacités cumulées en pariant sur la pente de la baisse de coût du kWh PV. La base de leur argument est de considérer que le progrès des rendements est un facteur plus important que l'effet d'apprentissage. D'après eux, même si une politique de pénétration

subventionnée du marché déclenche des effets d'apprentissage dans les coûts de fabrication et d'implantation des cellules PV, ce ne serait pas ces effets d'apprentissage qui amélioreraient substantiellement les performances économiques des modules¹².

Le mérite d'un tel point de vue est d'être clair et d' inviter à la réflexion. Mais le prémissé de départ selon lequel il n'y a que l'amélioration du rendement qui conduit à la baisse significative de coût du kWh en photovoltaïque est très discutable. Les progressions sur les coûts de fabrication, notamment par la limitation des re-buts, l'amélioration de l'assemblage des cellules en modules en SiC, l'amélioration des procédés d'étalement dans le cas des couches minces, celle des coûts hors module dans toutes les filières sont des facteurs aussi importants des baisses passées et le seront pour les baisses futures. Ils résultent d'un côté des effets de R&D industrielle et de démonstration et de l'autre de *learning by doing* intrinsèquement lié à la croissance de taille des usines qui stimule le progrès des techniques de fabrication, comme déjà dit. Il ne faut donc pas les négliger.

Pour ce faire il faut chercher à combiner les efforts de R&D publique et privée et de démonstration industrielle et les effets d'apprentissage associés à des fabrications de taille élevée qui suivent le développement d'un marché minimal. Ces différents constats sur l'enchevêtrement des savoirs et des savoir-faire liés à l'innovation PV invitent à maintenir des soutiens à la R&D industrielle et à maintenir un dispositif de *demand pull* par la subvention à l'adoption, à condition que ce dispositif soit bien ajusté à ce que l'on recherche. C'est-à-dire développer un marché local minimal sur lequel les industriels nationaux puissent en partie s'appuyer pour développer des efforts de R&D industrielle et de démonstration, notamment par le passage des usines de fabrication à des grandes tailles.

¹² Frondel et al. (2008) mettent en avant plusieurs arguments. D'abord ils soulignent le lien entre les progressions des performances de rendement dès que on consacre beaucoup de fonds à la R&D : Le rendement des cellules est monté de 5% autour de 1975 à 15% en 2001 parce que selon eux il y a eu un effort significatif de R&D. Le rapport de l'IEA (2004) sur les technologies énergétiques souligne que le progrès des rendements du PV entre 1983 et 1990 avait été précédé d'un très important effort de R&D au plan mondial de 1,5 milliard de \$. De plus le recensement des ruptures innovantes sur les cellules qui ont été porteuses de ces améliorations depuis 1980 montre que 10 sur 16 ont été permises par la R&D (Nemet, 2006). Ceci suggère, selon eux, que c'est par l'investissement en R&D que les performances techniques et économiques de la technologie PV vont d'abord croître, plutôt que par une politique de pénétration extrêmement subventionnée où les améliorations techniques ne sont que des sous-produits.

Un dernier facteur ignoré par les tenants d'un retour à la seule R&D est l'apprentissage au niveau des métiers de commercialisateur et d'installateur, les coûts de vente et d'installation constituant une part significative des coûts de l'acheteur. Le développement de telles compétences fait partie du processus de baisse de coûts accompagnant le déploiement d'une technologie décentralisée comme le PV. Le développement rapide d'un marché tiré par des tarifs d'achat très généreux a pour effet de développer ces savoir-faire. Mais la question demeure de savoir quand il faut rechercher ces effets. Faute de maturation technologique suffisante des filières, chercher à enclencher le développement de tels métiers par des tarifs élevés – ce que l'on entend en France comme argument de défense actuellement – n'est pas une bonne solution. Les baisses de coût sur cette partie de la chaîne d'activités ne seront pas suffisantes si les autres coûts ne baissent pas suffisamment pour justifier d'une politique d'appui relevant du stade précommercial et qui s'appuierait sur un tarif d'achat plus modéré de 10 c€/kWh par exemple au lieu de 30 à 50 c/kWh.

L'intérêt d'un retour aux subventions à l'investissement

On a vu que les subventions à l'investissement sont l'instrument adéquat en phase de démonstration et au début du déploiement pré-commercial pour les techniques ENR à fort coût d'investissement comme le PV et l'éolien et quand les adopteurs sont de petits acteurs sans surface financière importante. Dans le cas du PV, elles prennent trois formes principales de qualité voisine: a) la subvention directe jusqu'à deux tiers de l'investissement, b) le crédit d'impôt sur l'investissement en équipement, c) la bonification des taux d'emprunts. Ce dernier instrument a été utilisé dans les programmes 100 000 toits japonais et allemands, dans lesquels il était combiné avec une subvention directe à l'investissement¹³. On notera son effet de réduction de coûts : un prêt bonifié à 0% remboursable sur 15 ans pour un équipement de 22 500€ conduit à une économie de 15 100 € par rapport à un prêt ordinaire à 7%

¹³ Le programme japonais (1994-2005) combinait des prêts bonifiés et des subventions à destination des propriétaires de maisons individuelles, des promoteurs et des organisations publiques. La subvention (pour des équipements jusqu'à 4 kWc de 200 000 yens/kWc en 2005 soit 1500 €/kWc) avait d'abord été définie en % du coût de l'installation. Le programme allemand lancé en 1999 pour des installations de 1 à 3 kWc a reposé aussi sur des prêts bonifiés à 0% au départ, (puis 1.8%) et une subvention permettant un temps de retour de 10 ans. (Lopez-Polo et al., 2007).

et remboursable en 15 ans ; un prêt à 2% après bonification permettra une économie de 11 700 € par rapport au même prêt à 7%. C'est à l'évidence l'instrument le mieux adapté au stade de développement du PV pour créer un petit marché favorable à la démonstration des différentes techniques PV et de leurs procédés de fabrication. Il vaut mieux en rester à la subvention à l'investissement, qui a au moins l'avantage de ne pas installer un régime de rente de long terme pour les vendeurs et les heureux adopteurs de technologie.

Pour pallier les défauts de la subvention à l'investissement qui ont été précisés avant (notamment le manque d'incitations à choisir la technique et le fabricant les plus performants et à exploiter de façon efficace l'équipement), l'intérêt du dispositif du programme californien actuel de la CPUC pour les équipements de plus de 100kW doit être rappelé. C'est un instrument intermédiaire de type « performance-based » qui consiste à étaler la subvention publique sur cinq ans en la définissant de façon proportionnelle à la production annuelle effective pour inciter à choisir les techniques performantes. Cela revient à un tarif d'achat raccourci à une durée de 5 ans, qui permet le recouvrement sur cinq ans de la partie du coût d'investissement que finance la subvention. Insistons sur un des intérêts d'une telle approche : l'engagement de subvention par projet PV se fait sur une durée bien moins longue que les 20 années de l'engagement du tarif d'achat.

Tableau 7

Variantes de subventions à l'investissement

Paramètres de l'instrument	Avantages	Inconvénients
Subvention directe		
Pourcentage du coût d'investissement ou bien forfait		Peu incitatif à rechercher la meilleure technique Peu incitatif à la performance d'exploitation
Crédit d'impôts		
Soit sur le coût du matériel. Soit sur la partie fonds propres du financement de l'investissement	Important pour technologies à large « upfront cost »	Peu incitatif à choix efficace. Exposition au risque de fraude
Prêts bonifiés		
Taux à 0% en début de déploiement. Puis taux bonifié de 3-4%		Incitation à maintenir l'équipement par la durée de remboursement du prêt
Subventions basées sur performance		
Etalement de l'appui sur 5 ans en fonction de la production (exemple californien)	Favorables aux techniques à large coût initial. Incitatif à performance d'exploitation	

L'inefficience des politiques de rattrapage basées sur le développement à tout prix d'un marché national

On a rappelé en introduction que la politique française de promotion du PV a un double objectif : contribuer aux objectifs de la politique climatique, et développer une filière nationale. Il est temps de se demander si le coût de cette politique par rapport à ce objectif n'est pas incroyablement disproportionné. Si le premier objectif était le seul, il serait particulièrement coûteux par rapport à d'autres moyens, comme on l'a vu dans l'encadré I. Le second objectif pourrait mériter un tel engagement de moyens si le jeu en valait la chandelle. L'argument selon lequel, pour construire une filière nationale, il faut au moins développer un marché national est fréquemment évoqué. Mais y a-t'il vraiment des résultats industriels probants à attendre dans le contexte de l'intégration internationale des marchés ?

Comme dans toutes les politiques de pays leaders (Japon, Allemagne, Etats-Unis) et d'autres *second movers*, le soutien au déploiement du PV en France vise principalement des objectifs industriels et d'innovation, à savoir être présent dans la compétition technologique et consolider l'industrie nationale. Mais on doit s'interroger sur l'efficacité d'une telle stratégie quand elle est mise en œuvre dans un pays *second mover*. L'industrie française est pour l'heure à la traîne : pour l'essentiel elle ne peut compter que sur l'industriel Photowatt qui s'est fait une place dans les systèmes au silicium polycristallin avec 1,8% du marché mondial et une capacité de fabrication de 60 MWc. D'autres industriels intéressés par le PV préfèrent développer leurs projets d'unités en Allemagne, tel Saint Gobain qui, allié avec Shell, installe une unité de fabrication 20Wc de la filière CIS en raison des subventions généreuses accordées à la localisation d'industries dans les länder de l'est (projet Avantis). Aucun projet d'usine de grande taille (400MW et plus) n'est annoncé même si on fait grand cas en 2009 de l'installation d'une petite usine de 100MW en couches minces de First solar en France. Le marché français qui est en train de se développer ne fait que s'adresser partiellement aux producteurs français.

Capacité de production de cellules PV dans le monde (en MWc)

	2006	2007
Japon	920	950
Chine	380	1200
Allemagne	500	900
USA	190	300
Autres pays en Europe	200	200
Autres pays dans le monde	Taiwan : 200 Australie : 60	Taiwan : 380

Sources : Solarbuzz et Karberger (Swedish Energy Board) 2008

Tableau 8

L'industrie française est-elle en mesure de rattraper son retard d'installation dans le domaine des filières au Si cristallin qui vont rester fort probablement les techniques dominantes dans les quinze prochaines années ? La vivacité de la concurrence entre firmes et entre pays leaders, confirmée par l'arrivée de l'industrie chinoise dans le peloton de tête en 2007 et l'entrée récente de « capital-risqueurs » dans le business du PV aux Etats-Unis, ne doit pas faire oublier que les débouchés dans les pays industrialisés (qui sont le marché principal de l'industrie mondiale du PV) sont très subventionnés. Des ruptures dans les programmes de subventionnement dans les nouveaux grands marchés, comme cela a été le cas en Espagne en 2008 peuvent entraîner un tel ralentissement de la demande mondiale que l'industrie mondiale peut se retrouver en surcapacité comme en 2009. Dans un tel contexte, les pays du second peloton (Australie, Espagne, France, Italie, Royaume Uni, etc.) doivent se poser trois questions.

Un programme volontariste de développement d'un marché fortement subventionné pour appuyer l'industrie nationale peut-il permettre de rattraper le retard ? Deux voies seraient possibles : l'une qui consisterait à privilégier l'industrie nationale pour les marchés français, l'autre qui serait de tabler sur un créneau de débouchés spécifiques à une des technologies PV. La première voie est irréalisable. Le marché subventionné ne peut pas permettre de privilégier les firmes françaises car on encourrait la condamnation des instances européennes de la concurrence ou de l'OMC. Vu au plan mondial on observe que les marchés fortement subventionnés tendent à favoriser

les industries des pays *first movers* et maintenant des pays émergents qui n'ont même pas besoin de marchés internes subventionnés pour se développer. Une politique alternative qui serait a priori beaucoup plus efficace pour ce seul objectif industriel pourrait être une aide directe et conséquente à une installation d'une usine de grande taille (400 MWc, voire 1000 MWc) par une firme française (ou à majorité française). Le projet pourrait être généreusement articulé avec la politique de coopération de la France avec les pays africains, dans le domaine de l'électrification rurale décentralisée où les besoins restent importants. Mais cet objectif n'est pas non plus compatible avec les principes de la politique de la concurrence et à l'article 82 du Traité sur les aides d'Etat qui n'envisage pas ce cas de figure.

Ne faut-il pas adopter un système d'appui différencié entre filières de maturité différente pour positionner l'industrie française?

Le choix d'instruments non différenciés favorise principalement le silicium cristallin et comme ce choix est général à tous les pays, il peut contribuer à un *lock-in* autour de cette filière. L'objectif principal devrait être de positionner aussi l'industrie nationale sur les filières couches minces dont elle est absente. La différenciation de l'appui entre filières SiC et couches minces serait une forte incitation à l'entrée et à la diversification des entreprises françaises présente sur le silicium vers cette voie des couches minces, ce qu'un petit nombre d'entreprises (Sharp, Q-Cells, Mitsubishi, etc.) ont déjà choisi de faire d'elles-mêmes. Ces entrées sur le marché mondial comme ces exemples de différenciation montrent que les couches minces sont perçues comme présentant les mêmes chances commerciales que le SiC par les industriels innovants au stade actuel de développement technologique des deux filières, même avec des systèmes d'appui comparables.

La démarche française s'apparente en partie à cette préoccupation de différenciation vis-à-vis des autres industries parties en tête sur le SiC mais en cherchant à mettre l'accent sur les cellules intégrées dans le bâti. L'idée, explicitée dans un récent rapport parlementaire (Pognant, 2009) est de faire de ce marché un objectif prioritaire pour le développement des techniques spécifiques propres à ce créneau (tuiles PV à systèmes SiC, colorisation, intégration entre modules et structures, etc.).

First Solar (USA)	50%
Uni-Solar (USA)	11%
Kaneka (Japon)	10%
Mitsubishi (Japon)	7%
Sharp (Japon)	5%
Wurth (Allemagne)	2%
Autres*	15%

* Dont Nanosolar, Solar Cells, etc.

Source : EUPD Research, 2008, cités dans Usine Nouvelle, n°3117, 2 octobre 2008, p.38

Tableau 9

Le tarif est donc beaucoup plus important que le tarif du PV normal, avec un différentiel de tarif supérieur au surcoût du PV intégré au bâti pour marquer cette priorité. Ceci dit, le créneau n'est pas aussi simple à exploiter techniquement (problème d'augmentation de la température sans ventilation, stabilité des rendements, exposition) et juridiquement (définition d'installation réellement intégré au bâti) au point qu'un tarif intermédiaire a dû être défini pour les installations partiellement intégrées dans le bâti. L'innovation particulière associée à ce créneau donnera-t-elle un avantage décisif aux firmes françaises pour se positionner sur le marché mondial ? Peu d'éléments le laisse penser.

Les firmes nationales n'ont-elles pas d'autres opportunités pour entrer et se développer sur le marché mondial du PV dans un contexte globalisé ? L'expérience des politiques d'innovation montre l'intérêt de constituer une base de compétences sur une technologie dans un nouveau domaine de pointe, même si on part en retard et même si on développe cette base de compétences scientifiques et technologiques sur des filières qui éventuellement ne survivront pas à la sélection technologique (Guellec, 1999). Ceci dit, une stratégie reposant sur un programme de développement d'un marché national du PV est elle utile ? N'y a-t'il pas des stratégies plus simples pour les firmes françaises ? Elles peuvent en fait préférer s'allier à, ou racheter des entreprises étrangères et transférer ensuite les savoir-faire industriels en France. On voit les exemples de Saint Gobain allié à Shell qui installe son usine de fabrication de couches minces en Allemagne, On voit également EDF Energies Nouvelles acheter des parts dans

Nanosolar pour profiter de ses futures fabrications de couches minces en Europe.

L'action des industriels français acteurs de l'aval (installation et exploitation de systèmes individuels ou de grandes unités de production par PV) peut également avoir un fort pouvoir industrialisant. Ainsi, sans incitations publiques, EDF Energies Nouvelles s'est engagé contractuellement pendant l'été 2009 avec le producteur américain First Solar pour acheter la production d'une usine de fabrication de couches minces de 100MW/an qui sera installée rapidement en France. La mise en place d'une filière ne dépend donc pas uniquement de la seule action publique, mais aussi des stratégies des entreprises qui prennent leur risque. L'entrée dans cette filière en cours de constitution n'est d'ailleurs pas sans risque comme dans toute nouvelle industrie puisqu'on a vu le consortium de fabrication de polycristallin mis sur pied par différents acteurs de l'amont et de l'aval pour installer une usine en Haute Provence faire faillite en août 2009 avant l'achèvement du projet.

Les politiques amont ne sont elles pas plus efficaces au stade actuel de développement de la technologie ? Une autre ligne d'action devrait s'articuler autour d'un effort de financement public de RD & D industrielle sur les techniques innovantes en SiC et sur les couches minces les plus prometteuses. A côté de l'acquis de Photowatt en SiC, la puissance publique donne actuellement une impulsion au développement de savoir-faire industriels sur le SiC via la nouvelle plateforme technologique de l'INES, l'appui à l'alliance SolarNano-Crystal regroupant les industriels français (Photowatt, ESSIL, etc.) par un programme mobilisateur de l'OSEO-ANVAR, et probablement à un soutien au développement d'une unité de production de 100 MWc à la société PV-Alliance. L'industrie française devrait aussi se positionner sur la trajectoire des couches minces par un appui majeur. Pour l'heure on assiste principalement au développement d'une plateforme EDF-CNRS-ENSCP centrée sur le procédé électrolytique d'étalement dans la filière CIS. Mais l'intérêt d'équilibrer l'effort de R&D industrielle des acteurs français sur les deux filières justifierait la constitution d'un consortium comparable à PV-Alliance et appuyé par un programme mobilisateur.

En guise de conclusion : quelques recommandations

La France veut se situer dans le mouvement général de diffusion de la technologie PV pour installer une filière industrielle nationale en développant à marche forcée un marché intérieur du PV sur la base de tarifs d'achat très élevés et d'un empilement d'aides diverses à l'investissement. En même temps le gouvernement veut se servir de son choix de soutenir le PV comme technologie verte high-tech et étandard de sa politique climatique dans le domaine de l'énergie.

Mais en fait, devant la faiblesse de l'apport énergétique d'une politique PV volontariste et son coût (l'objectif du Grenelle de l'environnement de 5,2 TWh en 2020 pour un coût annuel de 2,2 milliards d'€, soit 700 à 1400 €/tCO₂ évitée), on ne peut que souligner le coût immense de l'effet de propagande recherché et ce d'autant plus que les effets d'entraînement industriel attendus sont limités. Les promoteurs de cette filière qui ont raison de souligner les progrès récents des deux filières technologiques (SiC, couches minces) ont cependant tort à la fois de surestimer la progression des coûts et de faire croire que l'accès à parité-réseau, la future terre promise pour les promoteurs de la filière, sera équivalent à l'accès à la compétitivité du PV en termes d'optimum collectif. Une capacité de production électrique en PV n'entraîne aucune économie de réseau et d'investissement en production électrique centralisée du fait de l'intermittence des productions et la non-stockabilité de l'électricité produite. Le coût du PV pour la collectivité doit se comparer au prix de gros de l'électricité et non pas au prix en sortie de réseau.

Le choix et la conception du tarif d'achat comme instrument de soutien pour atteindre les deux objectifs industriels et énergétiques posent trois problèmes. En premier lieu le tarif d'achat n'est pas le bon instrument par rapport au stade de progression de la technologie PV qui n'est pas encore au niveau pré-commercial, ce qui oblige à avoir des tarifs quatre à cinq fois plus élevés que ceux de l'éolien. Il serait préférable de renforcer l'effort de R&D industrielle sur toutes les filières et en particulier les filières couches minces bien moins avancées, et de limiter le développement du marché à des programmes de type 10 000 toits PV par des subventions à l'investissement pour

permettre un effort de démonstration sur les procédés industriels. La subvention à l'investissement engage bien moins sur la durée que le tarif d'achat qui est une subvention à la production sur 20 ans.

En second lieu, même en imaginant qu'il serait justifié de retenir cet instrument parce que l'on considèrerait que la technologie se situerait en phase pré-commerciale, le dispositif français actuel est mal conçu sur quatre plans :

- la durée de l'engagement du tarif pour chaque nouvelle installation et l'absence de décroissance du tarif d'achat sur les 20 ans de l'engagement,
- l'absence de différenciation entre régions d'ensoleillement, ce qui favorise indument les installations dans le midi,
- l'inutile ajout des aides à l'investissement avec crédits d'impôts, prêts bonifiés et allègement de TVA,
- l'absence de décroissance du tarif d'une année sur l'autre pour les nouvelles installations, alors que l'on pourrait déjà limiter les dépenses futures de ce dispositif par des baisses plus rapides que l'évolution des coûts,

En troisième lieu, l'objectif industriel de construire une filière photovoltaïque en France en tirant le développement d'un marché interne par ces tarifs ne sera jamais atteint dans le contexte de concurrence globalisée, sachant que les industries de quelques pays se sont placées en tête grâce à des politiques d'appui plus précoces. Le dispositif va engager de façon exponentiellement croissante l'argent des consommateurs français d'électricité sur une période de 20 ans pour chaque nouveau contrat sans faire décoller l'industrie française sur le Si cristallin ou les couches minces. Le dispositif servira plutôt à subventionner les industries des autres pays (Japon, Allemagne, Etats-Unis et maintenant Chine) s'il n'est pas recyclé.

Le tarif d'achat élevé, combiné à plusieurs autres subventions ne sert en réalité qu'à impulser le développement de l'aval de la filière. On peut certes attacher une valeur élevée au développement d'un réseau de monteurs de projets et d'installateurs que permettra le dispositif actuel. Mais rien ne démontre que ce soit efficient socialement dans la mesure où la technologie reste chère et où l'on importe et importera la grande majorité des systèmes PV.

On ne critique donc pas en soi une politique fondée sur les tarifs d'achat, politique tout à fait adaptée à des techniques proches de la

maturité commerciale comme l'éolien car, sans coût excessif, il leur garantit un signal stable et proche des prix de gros du marché électrique. On critique son inadéquation au niveau de maturité de la technique PV qui conduit à avoir des tarifs d'achat cinq fois plus élevés que pour l'éolien. Alors que, sur le plan technologique, le futur est encore ouvert entre les filières améliorées au silicium cristallin et les filières couches minces, on propose de recibler le dispositif au regard des enjeux énormes que représente le PV. On doit réduire rapidement les tarifs d'achat, incitations coûteuses à l'installation de cellules PV et de centrales au sol qui engage pour 20 ans le versement de subventions pour chaque projet en préférant les subventions à l'investissement (combinées à du net metering) qui engagent sur très peu de temps. On doit accroître les budgets de R&D et de démonstration industrielle pour favoriser la consolidation de nouvelles filières (couches minces, mais aussi Si cristallin amélioré). On doit cibler plus particulièrement la phase de démonstration et le développement industriel où se joue une grande partie des apprentissages porteurs de l'essentiel des baisses de coût.

Références

- Academie Des Technologies (2008) *Les perspectives de l'énergie solaire*, rapport de Yves Maigne, Jean-Pierre Causse, Maurice Claverie, Bernard Equer, Groupe de travail Energie solaire, Commission Énergie et changement climatique, Paris, Académie des technologies.
- Ayres, R.U. and Martinas, K., (1992), "Learning from experience and the life cycle: some analytic implications", *Technovation* Vol.12 n°7
- Borenstein S. (2008), "The market value and cost of solar photovoltaic electricity production", UCEI , CSEM Working Paper 176.
- Cowan R., Kline D. (1996), "The implications of potential lock-in in markets for renewable energy", National Renewable Energy Laboratory, NREL/TP-460-221
- Duke R. Williams R.H., et Payne A. (2005), "Accelerating residential PV expansion: demand

- analysis for competitive electricity markets", *Energy Policy*, Vol.33, n°15
- Finon, D. et Perez,Y. (2006). "The social efficiency of instruments for the promotion of renewable in the electricity industry: A transaction cost perspective". *Ecological Economics*, n°62
- Finon D. et Menanteau P. (2004). "La promotion des énergies renouvelables dans les nouveaux marchés électriques concurrentiels", *Annales des Mines, Gérer et Comprendre*, Septembre
- Foxon T.J. and al. (2005), "UK innovation systems for new and renewable energy technologies: drivers, barriers and systems failures", *Energy Policy*, Vol 33
- Frondel M., Ritter N., Schmidt C.M. (2008), "Germany's solar cell promotion: dark clouds on the horizon", *Energy Policy*, Vol. 36
- Guellec D., (1999), *L'économie de l'innovation*, Paris Edition la Découverte, Collection Repères.
- Haas R., Held A., Finon D., Meyer N., Lorenzoni A., Wiser R., Nishio (2007), Promoting electricity from renewable energy sources – lessons learned from the EU, US and Japan, in Shioshansi P. (2007), *Competitive Electricity Markets: design, Implementation, Performance*, London: Elsevier
- IEA (2008), *Energy Technology Perspectives – In support of the G8 Plan of Action, Scenarios & Strategies to 2050*, OECD
- Kahn, E., (2008), "Avalable transmission cost is a substantial benefit of solar PV", *The Electricity Journal*, Vol.21, n°5
- Kimura, O. and Suzuki,T. (2006),Thirty years of solar energy development in Japan: coevolution process of technology, policies and the market, paper presented in Resource Policies: Effectiveness, Efficiency, and Equity, Berlin
- Lincot D. (2007), "La conversion photovoltaïque de l'énergie solaire", *Découverte*, n°344-345, janvier-février
- Lopez-Polo A., Haas R., Suna D. (2007), *Promotional drivers for PV*, Report to the Program "PV Upscale"
- Meddad (2008), *Grenelle Environnement : réussir la transition énergétique / 50 mesures pour un développement des énergies renouvelables à haute qualité environnementale*, Novembre (Dossier de Presse)
- Menanteau, P., Finon, D., Lamy, M. L. (2003) "Prices versus quantities : environmental policies for promoting the development of renewable energy". *Energy Policy*, vol.31, n°8
- Menanteau P. (2000), "La constitution des technologies photovoltaïques : apprentissage et dépendance du sentier" in Bourgeois B., Finon D., Martin J.M., (2000), *Energie et Changement Technologique : une approche évolutionniste*, Paris : Economica
- Nemet G.F. (2006), "Beyond the learning curve: factors influencing cost reductions in photovoltaics", *Energy Policy*, Vol.34, n°17
- Neij L. (1997), "Use of experience curves to analyse the prospects for diffusion and adoption of renewable technology", *Energy Policy*, Vol.23, n°13
- Pappineau M.,(2006), "An economic perspective on experience curves to analyze the prospects for diffusion and adoption of renewables energies technologies", *Energy Policy*, Vol. 34, n°4
- Poignant, S., 2009, Assemblée nationale, *Rapport d'information sur l'énergie photovoltaïque*, Rapport AN n° 1846
- Rabl A. et van der Zwaan H (2007), "The learning potential of photovoltaics: implications for energy policy", *Energy Policy*, vol.32, n°13

Large-scale wind power in European electricity markets

time for revisiting support schemes and market designs ?

Céline Hiroux

University Paris Sud (ADIS-GRJM) & LARSEN

Marcelo Saguan

University Paris Sud (ADIS-GRJM), Supelec & LARSEN

This paper questions whether current renewable support schemes and electricity market designs are well-suited to host a significant amount of wind energy. Our analysis aims at finding the right equilibrium between market signals received by wind generators and their intrinsic risks. More market signals are needed to give the right incentives for reducing wind integration costs but should not undermine the effectiveness of support schemes. Although several alternatives combining support schemes and market signals could improve the current situation in terms of market signals and risks, feed-in premium support scheme seems actually to be the more balanced option. Furthermore, an adequate sharing of wind generation technical responsibility between the System Operator and wind power producers can help to control wind integration costs even in the absence of accurate market signals.

Introduction

Wind power technologies in Europe have benefited from renewable promotion policies for more than ten years. Renewable policies have been implemented through different support schemes such as feed-in tariff, feed-in premium or green certificates. Successful support schemes have been acknowledged by their ability to boost the initial deployment of desirable renewable technologies (Butler and Neuhoff, 2008; Del Rio et al, 2007; Ragwitz et al, 2007; Mitchell et al, 2006). Nevertheless these support schemes kept wind power technologies aside from day-to-day operation of electricity markets and henceforth act as an isolating device to complete and transparent market signals. This has been judged as acceptable stage to the initial development. Indeed when wind power capacity in the system is relatively low, the wind integration costs due to the characteristics of this technology (variability and low predictability) are also low and can be absorbed by the rest of market participants.

Following the EU Renewables Directive and the 3x20 target, wind power capacity in Europe has entered in a new large-scale development phase. Among European countries Spain, Germany and Denmark stand out by their impressive growth over the last decade. In these countries wind power capacity represents between 18% to 24% of the total installed capacity. Resch et al (2008) consider that Germany would have about 34 GW, France about 24 GW and Spain around 28 GW in 2020. With the increasing share of wind energy in the energy mix, wind integration costs inevitably rise and their impacts on both the power system and the electricity markets are becoming of utmost importance. In this second stage of large-scale wind power development, support schemes and electricity market designs need to be assessed and adjusted in order to give the right incentives to all generators (renewable and conventional) while maintaining fair benefits for any other renewable technology but also in reducing the social costs paid by end-users. The following question then arises: how to identify what constitutes the right equilibrium between right market signals in one hand and the associated increased risks generated by these signals in the other hand ? More market signals are needed to give right incentives for reducing wind integration costs but should not undermine the effectiveness of support scheme (encouraging investment and limiting capital costs).

Recent studies have investigated the participation of renewable technologies on electricity markets and the impact of market designs on power systems with large scale of renewables. Klessmann, et al (2008) analyzes pros and cons of exposing renewables to electricity market signals and make a comparison between support schemes through the analysis of three countries: Germany, Spain and the UK. They conclude that exposing renewables to market signals, particularly non-intermittent technologies (e.g. biomass), would be beneficial for the total social cost on the condition that this does not dramatically increase the renewable producers' risk and therefore the support payment. However, they do not find considerable benefits to expose wind power technology to market signals (or risks) because of the lack of short-run responsiveness of this technology. A limit of this work stands in the lack of consideration of electricity market design as a possible variable to better adjust market signals and to promote efficient large scale wind power integration. Barth et al (2008) and Green (2008) consider how electricity markets should be designed to improve the incentives given to market participants in order to maximize their efficiency by reducing wind integration costs. These last two papers, however, do not consider the interaction between support schemes and electricity markets.

The contributions of this paper are three fold. First this paper studies electricity market designs in the light of the potential and necessary improvements required for the large scale integration of wind power. Second it adds a new layer to the actual research on the interaction of support schemes and electricity markets through a deeper study of market signal impacts on wind generators' behavior. Finally this paper uses this new framework to determine what could be a good balance between market signals and risks and to discuss how the current support schemes and market designs can be adapted.

The paper is organized as follows. The first section briefly presents the support schemes used for wind power in Europe. Then the integration costs related to large-scale development of wind power are identified and the impacts of the electricity market design on these costs are discussed. Incentives and market signals provided by short-term forward markets (day ahead, intraday), balancing markets, congestion (and losses) pricing and connection and network tariff

are analyzed. The following section focuses on the design of different support schemes to assess their adequacy with market participation and market rules. We first analyze to what extent wind power producers should participate in electricity markets. Then we analyze market participation for three support schemes: fixed feed-in tariff, feed-in premium and green certificates. The last section concludes with some policy recommendations concerning the trade-off between support mechanisms and market responsibility to ensure well functioning electricity markets and system.

Large-scale development of wind energy

Wind power support schemes in Europe

Renewable support schemes drive the actual and future deployment of wind energy technology. After ten years of renewable support policy experience, three main support mechanisms can be distinguished: the fixed feed-in tariffs, the feed-in premium and the green certificates¹. Table n°1 shows support schemes used in some countries in Europe. Feed-in tariff is the most used support mechanism in Europe (European Commission, 2006). Feed-in tariff scheme guarantees a fixed price for the total wind energy amount fed into the grid. This price is usually higher than the electricity market price and the difference represents a premium for the positive environmental externalities generated by windmills.

Table 1

Support schemes in selected countries	
Countries	Support schemes
Denmark	Feed-in premium added to the market price
Spain	Either a feed-in tariff indexed on the regulated price for 20 years or a feed-in premium + market price for 20 years
Germany	Fixed feed-in tariff for 5 years then 15 years with decreasing tariff
France	Fixed feed-in tariff for 10 years then 5 years with decreasing tariff
Netherlands	Feed-in premium to add to the market price or reference price (SDE) since 2008
UK	Renewable obligation certificate (ROC) price to be added to the market price

A variant of feed-in tariffs is the feed-in premium scheme. Under this scheme, wind power producers receive the electricity market price and a fixed regulated premium for producing renewable energy. This feed-in premium scheme may include a cap-and-floor limit that guarantees minimum and maximum tariffs independent of the electricity market price thus reducing the overall risk. For instance, fixed feed-in tariffs and feed-in premiums are possible options in Spain but more than 97% of wind power producers have chosen the last option. Feed-in premium has been preferred because the total income from market price and premium is usually higher than the fixed feed-in tariff.

Green certificate scheme is based on the level of renewable generation obligations generally imposed on suppliers. To fulfill these obligations, suppliers can either produce (internally or externally) "green electricity" or buy the equivalent in green certificates. Green certificates are produced each time an accredited renewable energy source generates. For instance, in UK, each MWh produced by wind power plants generates one ROC (Renewable Obligation Certificate). Wind generators then sell their production on electricity markets and green certificates on a specific certificates market. If suppliers do not fulfill their renewable obligations, they must pay a penalty: the buy-out price. The level of this buy-out price is of importance for the effectiveness of the support scheme since the incentives to achieve renewable targets can be distorted² (Mitchell et al, 2006).

Implementation of support schemes facilitates a high penetration of wind power in several European countries. A higher level of wind power capacity brings many benefits in terms of reduction of GHG emissions, increasing diversification and security of supply, developing new sustainable technologies for the future, developing

-
- 1 Other types of support schemes as tendering procedures or investment subsidies are not considered here since they are not often used in Europe.
 - 2 If the level of the buy-out price is too low, suppliers could find cheaper not to respect renewable obligation rather than buy ROCs on markets. In theory, the buy-out price should set the cap of ROCs market. However, in practice, the ROC price is higher than buy-out price because the amount of money paid by suppliers not respecting obligations is recycled to suppliers respecting obligations. ROC prices are therefore quite difficult to predict; they can create distortions and are sensible to opportunistic behavior.
 - 3 The cost of support schemes corresponds to the difference between the renewable and conventional technology generation costs. In the long run, this difference should decrease since the learning curve evolution of renewable technologies becomes more competitive. In terms of generation costs, wind power generation cost is around 56€/MWh whereas the CCGT cost is around 45€/MWh (UKERC, 2007)

new industries, etc (Lamy, 2004). These benefits are notably higher than the costs of support schemes³ and other negative externalities (use of the land, landscape, etc.). As wind power technology has specific generation characteristics (variability, low predictability, wind resources far from consumption sites, etc.), other costs are added into the power system; costs that have to be taken into account in order to adjust adequately support schemes and electricity market designs : the integration costs.

Integration costs of wind power

Integration costs represent additional system-induced costs due to the integration of large-scale wind energy. These integration costs can be separated into i) balancing costs, ii) reliability costs iii) congestions (and losses) costs and iv) network connection and reinforcement costs (Gross et al 2006). Additional balancing costs come from wind power intermittency and will affect both the unit commitment of the conventional power plants and the increasing need for balancing the system. Balancing costs increase because the system needs more reserves and balancing services, all of which will be used more frequently. Using additional quick start capacity and conventional power plants running part-load are the main reasons of cost increase⁴.

Additional reliability costs are associated to the weak contribution of wind power to peak situations and to the corresponding variability of wind power generation during these periods. When intermittent wind generation replaces conventional generation, an additional installed generation capacity is needed to get the same level of reliability (e.g. a given value Loss of Load Probability). Additional congestion (and losses) costs are due to higher and different use of the network mostly when wind power generators are located in remote areas that are usually far from the load. Moving cheap electricity over large distance can generate potentially more losses and more frequent occurrences of bottlenecks, which increase losses and congestion costs. The latter increase because both the lead time of construction of wind plants is much lower than the time required to increase the network capacity.

⁴ Several studies have demonstrated that increasing share of wind energy in system load results in higher balancing costs (Holttinen et al, 2007; Gross et al, 2006; DENA 2005). For wind energy penetration from 5% to 20% of gross energy demand, system operating costs increase due to wind variability and uncertainty amount for about 1-4€/MWh depending on the observed system (Holttinen et al, 2007).

The integration of wind power into the power system implies also connection costs and network (transmission and distribution) reinforcement costs. Connection costs are due to the additional installations (underground cable, etc.) required to connect the wind power plants to the existing transmission and distribution network. In addition, the connection of new wind farms can need upgrades and reinforcements on the bulk network. Reinforcements on the network to accommodate wind energy flows reduce additional congestions and losses costs; minimizing the net sum of both integration costs is necessary to improve the efficiency of the system.

Electricity market architecture as the key for the distribution of integration costs

As the amount of intermittent generation increases on the system, taking additional integration costs into consideration become more and more relevant as the incentives to reduce them have to be incorporated to the analysis of efficient support schemes. The analysis of electricity market architectures helps to clearly identify how integration costs can be controlled and how they are distributed using different market design options to give signals to market participants.

The electricity market architecture. Electricity is a complex good, largely constrained by physical and technical laws for its production and transmission on the grid (Stoft, 2002). The introduction of competition needs the design of specific market architecture(s) or market design(s) (Wilson, 2002). Given system operation constraints, a "standard" market, where supply matches directly demand, cannot manage the complexity of a power system in real time. A centralized authority, called the "System Operator" (SO), is responsible for real-time management. Therefore, "standard" markets, where generators, intermediaries (traders, brokers) and large consumers trade electricity each other are normally "forward" markets⁵ i.e. they take place before the moment of delivery.

In the short run, forward energy can be traded in the day-ahead market (functioning 24 hours before delivery)⁶, in intraday markets (functioning within the day) or even in real-time in the balancing market.

⁵ Forward markets are financial markets that trade electricity ahead of its delivery.

⁶ Forward markets have maturities that go from 3 years to a few hours before delivery. We focus on Day-ahead market because day-ahead prices are a major benchmark for all forward trades.

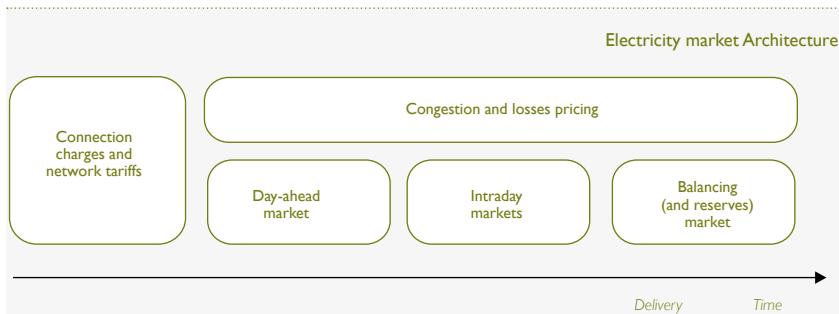
Other services may be defined in the market architecture in order to better represent the operation of the power system⁷. An efficient use of transmission networks requires the implementation of market mechanisms to reduce congestions and losses. These mechanisms and related costs send economic signals valuing the use of scarce network capacity resources and other externalities (Ehrenmann and Smeers, 2005). Transmission and distribution activities present features of natural monopolies mainly due to network costs mostly composed by fixed costs. Hence, scarcity (marginal) pricing framework does not enable the full network costs recovery.

Therefore regulated connection charges and network tariffs are applied to network users in order to ensure the viability of the network services. Figure 1 sums up typical electricity markets that, in a general meaning, have been organized as follows: short-term forward energy markets (day-ahead and intraday), congestion (& losses) pricing, balancing market and connection and network tariffs.

Each of these markets or mechanisms could be defined by different rules and designs. Not all designs have the same economic properties concerning incentives and efficiency (Wilson 2002). The market design will determine the “quality” or “accuracy” of market signals. Furthermore, depending on the market architecture and SO arrangements, the responsibilities of system operation can be shared between the market signals or SO authority.

Day-ahead and intraday markets. The most liquid market in electricity is the day-ahead market. As the traded electricity products correspond to 1 hour or $\frac{1}{2}$ hour periods, market signals given through this market allow to value electricity differently according to the delivery times

Figure 1



during the day⁸. Intraday markets allow trading electricity from day-ahead market closure to few hours before delivery.

The degree of centralization of day-ahead and intraday markets is the key element to distinguish different market designs (Wilson, 2002). These markets can be organized from a centralized auction (e.g. Spain) to a completely decentralized bilateral market (e.g. UK). More centralized markets concentrate trades and increase market liquidity. They can also optimize generation power scheduling by incorporating uncertainties and intertemporal and network constraints⁹. Therefore, more centralized designs can help improve coordination and dispatch efficiency, and reduce wind power integration costs (Green, 2008). Most of European markets are based on bilateral transactions and on a power exchange accounting for less than 10% of the power consumption. They are not well equipped to minimize wind integration costs since decentralized coordination does not allow for an optimized generation scheduling.

The temporal position of the "Gate Closure" is another important design parameter; mainly in more decentralized markets. The Gate Closure determines the closure of the forward (intraday) markets and the opening of the balancing market. A Gate Closure that closes intraday markets near real-time help to decrease individual imbalances and wind integration balancing costs. When market participants are able to trade electricity near real-time, it implies that better information concerning the actual wind power generation is available to all the market participants (Barth et al, 2008). Müsgens and Neuhoff (2006)

7 Reserves and long term capacity mechanisms are not considered explicitly in this paper.

- Some (fast) generation capacities have to be prepared as "reserves" before real-time in order to prevent blackouts. The reserve mechanisms represent tools that the SO runs in short term to manage the security of the network. In electricity markets, reserves are market driven through the balancing market. As reserves and balancing are the two faces of the same coin, we consider both under the term balancing market.

- Long term capacity mechanisms (e.g. capacity markets, capacity payments) are sometimes added to the market architecture in order to ensure an adequate level of generation capacity (Joskow, 2006). As these mechanisms are not broadly used in Europe, they are not considered in this paper.

8 For instance electricity produced during off-peak hours (during the night) is cheaper than electricity produced during peak hours. Other products including aggregated hours are also commercialized in forward markets (e.g. base or peak blocks).

9 Centralized day-ahead market in US use optimization tools to clear the market respecting very detailed power plants constraints (e.g. start up costs, ramping constraints, etc.) implying a better use of generation resources.

show that a gate closure near real time would reduce the balancing costs, since fewer thermal power stations would be started up, only to be ready to replace previously unexpected wind power outputs.

Balancing market. The balancing market is managed by the SO. On the one hand, the SO balances the system using balancing offers/bids and reserves. On the other hand, the SO computes imbalances (measuring the actual injections/withdrawals of energy in real-time and comparing it with forward contract positions) and settles them using imbalance prices. Balancing market signals indicate the real-time (spot) value of electricity for each settlement period at the time of delivery. Forward signals are only based on expectations while real-time signals integrate both new information coming after forward markets and the value of flexibility (Barth et al, 2008). Balancing signals should be cost-reflective in order to induce efficient behavior. Balancing signals indicating the marginal cost of balancing the system and enable individual market participants to compare their own cost to reduce imbalance against the cost of buying balancing services from the system. This helps market participants to take efficient decisions. If this is not the case, two different inefficient situations can appear. Either inappropriate high imbalances charges will lead market participants to use individual balancing options (e.g. own back up thermal plant) increasing the total balancing cost, or excessively cheap balancing charges will give incentives to market participants to over-use balancing services and this will undermine efficiency. One of the most important design parameter of balancing markets is the definition of imbalance prices. This will determine how the total balancing costs are distributed and how incentives are given to market participants. There are basically two types of design : a dual-price design and a single price design¹⁰. The dual price imbalance design is reputed to be less cost-reflective than the single price

¹⁰ The dual-price balancing mechanism design uses two different prices for negative and positive imbalances. These prices are often computed using average prices of accepted bids/offers. The single-price real-time market design uses only one price for all types of imbalances and this price correspond to the price of the marginal accepted offer/bid.

¹¹ It is important to note that defining optimal balancing rules is not straightforward given costs allocation problems (e.g.no proper imbalance measure, non-convexities, inter-period costs and fixed costs allocation). However good practice guidelines exist in literature and have to be applied in order to give proper incentives (see for instance Littlechild, 2007;Vandezande et al, 2009).

design¹¹ (Newbery, 2005). As a matter of fact dual imbalances prices are usually computed using average prices and artificial penalties added to imbalance costs. As the system balancing cost does not depend on individual imbalances but on the total net imbalance, positive or negative individual imbalances need to have the same price. On the other hand, single price design is reputed to give more volatile signals since imbalance price is computed using the proposed price of marginal offer or bid and this can change for each settlement period. Dual-price settlement system is used in several European countries and some of them are renowned not to be cost-reflective (ILEX, 2002; Littlechild, 2007). However many countries are improving market design in order to get balancing signal right (Vandezande et al, 2009).

Short term and long term locational signals. The goal of congestion (and losses) short-term pricing and connection and network tariffs is to give locational signals indicating the costs/benefits for the whole system of the production/consumption of energy in each node of the network. Short-term congestion (and losses) pricing is often integrated with the energy markets¹². There are different possible designs of pricing depending on the aggregation of signals: i) nodal pricing, ii) zonal pricing and iii) redispatching. Nodal pricing establishes one energy price for each node in the network while redispatching gives the same price of energy no matter where it is injected or withdrawn. In the latter case, the system operator takes action to solve congestions (and minimize losses) and the costs of this action are socialized among network users. Most of European countries use a redispatching design that gives no short-term locational signal¹³. The choice of the degree of differentiation depends on two issues. On the one hand, more differentiated pricing gives more locational signals indicating where and when to produce or consume. On the other hand, less differentiated pricing reduces the transaction costs and the induced risks on the revenue of market participants (Ehrenmann and Smeers, 2005).

¹² This is called “implicit auctions” and consists in computing energy electricity prices taking into account transmission constraints and transmission losses.

¹³ Italy and Scandinavian countries are two exceptions using zonal pricing.

Potential market signals		Potential integration costs reductions	Market design options	Accuracy of market signals*		Risk induced by market signals*
Day-ahead and intraday markets	Temporal differentiation of electricity	Balancing and reliability costs	Degree of centralization	Decentralized	0	0
			Centralized	+	-	
Balancing market	Value of electricity at delivery/	Balancing and reliability costs	Imbalance price	Far real-time	0	0
	Value of flexibility			Close real-time	+	-
Congestions (and losses) pricing	Locational/temporal differentiation	Congestion and reinforcement costs	Zonal aggregation	Redispatching	0	0
			Zonal	+	+	
			Nodal	++	++	
Connection and network tariffs	Locational/temporal differentiation and cost recovery	Congestion and reinforcement costs	Connection and network tariff	Shallow	0	0
			Deep	+	++	
			Zonal tariff	++	+	

Table 2

Furthermore, this trade-off concerns the degree of delegation of power to the transmission system operator (TSO) and the extent of system operation authority. The lesser short term locational market signal are given to the market participants, the more the TSO has to intervene ("out of the market") to solve congestions problems.

Network connection charges and network tariffs are complementary payment mechanisms used to cover the total cost of transmission and distribution infrastructure and, potentially to give long-term locational signals to generators. Several designs are possible for network connection charges going from "shallow cost" design (where the new connected installation only pays the cost of the own connection to the system) to "deep cost" design (where the new connected installation pays all the extra network cost due to its connection)¹⁴. Network connection charges designs can be combined with different designs for transmission (distribution) network tariffs which can include long-term locational signals for generators. Using

¹⁴ For a more detailed assessment of connection cost policy: Barth, et al, 2008; Rious et al, 2008; Swider et al, 2006).

different combinations of network tariffs and connection charges designs, the locational signals, the burden allocated to generators and the associate risk can be balanced.

At one extreme, “shallow cost” and weak (or no) network tariffs do not provide any locational signals but minimize the risks and the extra costs for generators. At the other extreme, “deep cost” design gives locational signal to generators but, as proper and transparent deep cost signals are very difficult to estimate, this might make generators to over react choosing not efficient locations or increasing the risk.¹⁵ Zonal network tariff design gives good locational signals and involves less risk than “deep cost” solution (Rious et al, 2008).

Charging congestion (and losses) and reinforcement costs to the responsible parties incite market participants, both wind power and conventional, to reduce as such the integration cost of wind power. Wind power characteristics imply that the patterns and the frequency of congestions changes constantly and hence the role of short-term signals becomes very important for the efficiency of the system. The absence of locational signals in the presence of high amounts of wind power can considerably increase congestion (and losses) costs.

In fact, without locational signal, conventional generators continue to schedule production and to plan their location without taking into account transmission network impacts and the System Operator has to make a greater effort to control congestion (and losses) implying a higher integration cost. Introducing nodal/zonal pricing could help to reduce these integration costs (Weigt et al, 2008; Leuthold et al, 2008). However, as more accurate market signals as nodal/zonal pricing are more volatile, this can increase the risks born by market actors. Table 2 summarizes the main options of market design and the accuracy of market signals and related risks.

Different designs send different signals to market participants. Thus, distribution and control of wind power integration costs depend

¹⁵ Deep connection cost design involves more risk than other design options. At the beginning of an investment project, the producer has high uncertainties concerning connection charges. These charges could be high and depend generally on unclear TSO/DSO rules. Investments can be deterred if the financial cost provoked by this risk is too high.

on different market designs. Efficient designs need accurate market signals to give the right incentives. In Europe there is considerable room to improve market design and accuracy of market signals. However, accurate signals imply more volatile revenues and therefore an increase of risks. If the risks are too high, this creates a negative investment effect mostly on technologies with high capital cost.

The adequate trade-off has to be done to find a good balance between increased market responsibility and risks. This is true for all types of markets participants (conventional generators, renewable and demand) but particularly for wind power technology given the most important differences with conventional technology which is their ability to manage risks. In the case of wind energy, investors are quite sensitive to risks which could prevent such investment.

Market signals under different support schemes

The choice and implementation details of support scheme have implications on the way that wind power producers participate into markets and are exposed to market signals. We first analyzed why wind power producers should be exposed to more market signals and then how each type of support scheme efficiently transfers (or not) these signals.

Do wind power producers have to be exposed to market signals ?

Enforcing wind power producers to participate in electricity markets or exposing them to market signals entails several positive effects which are balanced by few negative effects. It has been argued that, as wind power technology has no means to react to market signals, it is not useful to expose them to it (Makarov et al, 2005, Klessmann et al, 2008). This is partly true in the short-run because wind power production has a high incentive to produce whenever wind is blowing and without regarding to the electricity price given its null (or very weak) marginal cost.

However, there is still a set of long term positive effects that can be found if one analyzes market signals and their effects more deeply. With the objective of a significant increase of wind energy in Europe by 2020, longer term effects cannot be ignored. Positive effects of exposing wind power producers to markets signals can be summarized as follows:

Optimal selection of wind sites related to temporal generation pattern. Forward prices and balancing signals differentiate time-delivery periods i.e. time periods when energy is highly valued (peak periods) have higher prices. As wind sites have different wind generation patterns, adequate selections of wind sites should take into account the different temporal value of energy expressed in forward and balancing market signals.

Optimal selection of wind sites related to congestion costs and losses. Wind power producers subject to locational signals should choose the wind sites that are more advantageous not only in terms of wind resource but also of capability of the network (Forsund et al, 2007, Barth et al, 2008, Di Castelnouvo et al, 2008). Installing wind power plants in highly windy areas may not bring high benefits if there is not enough transmission capacity to transport all the produced energy or if the extra losses induced by wind power production reduce considerably the useful energy. For instance, short-term congestion (and losses) pricing should give an indication of the zones where new power production can be accepted and therefore incite investors/producers to make an arbitrage between more congestions (and losses) costs and lower wind resource sites.¹⁶ This is particularly important for wind power given the leadtime gap between the build of wind mills and reinforcing transmission networks and congestions can remain for long periods. Short term locational signals can be replaced or completed by reinforcement network investments. Locational network tariffs or cost-reflective connection costs may act in the same direction than short-term locational signals; although network tariffs loose some of the (temporal) accuracy of short-term signals, they reduce considerably the congestion (and losses) cost risk.

¹⁶ In some cases wind power projects can also help to reduce congestion, losses or to postpone network investments but as they are not remunerated for this side-benefit, they prefer to choose another wind site.

Improvement of maintenance planning. Market participation of wind energy on forward and balancing markets implies higher responsiveness to price levels when implementing maintenance planning¹⁷. If wind power producers do not receive the right signals for planning maintenance, they might operate the maintenance when there is a lot of wind or when the wind energy is more valuable for the system.

Improvement of technology combinations and portfolio effects. Time differentiated electricity prices give signals for the optimal combination of geographically distributed wind power installations and the optimal combination of wind power production and other production technologies (intermittent such as solar or geothermal or storable such as hydro power plants).

These signals also allow to assess the short-term flexibility and storage options of the different technologies. Technology combination and innovation can result from an improvement of coordination between different technologies (wind/storage, wind/hydro, etc.) or from a new “firm organization” structure (size of the firm, portfolio type, etc.). Indeed, exposing all market participants to equal market signals for all technologies allows market actors to create efficient portfolios combining different kinds of technologies.

To control (reduce) production for extreme case of imbalance and network constraints. With high amounts of wind power in a system, it could be possible to have negative prices (Weigt, 2006); i.e. power producers are paid to reduce their production. This can happen during extreme congestion periods (under nodal pricing) or when the system has too much energy and there is not enough flexibility to reduce production of conventional units (for instance at night, coal plants produce at minimal capacity because they do not want to stop and start). In these cases, if wind power producers are exposed to adequate market signals (balancing or nodal pricing), they will reduce their production in their own interest and will contribute to the operation of the system. Note that even with the absence of market signals, other centralized and mandatory command-control alternatives can be implemented to improve

¹⁷ Suppose that a wind power producer has to select one moment of the day to stop its wind turbine and undertake the maintenance tasks and suppose also that forecasted wind production is constant over the day. Under a feed-in tariff scheme, the wind power producer has no preference in selecting maintenance hours during the day.

wind power response (e.g. mandatory obligation of connecting wind farms to local dispatch controls).

Improving controllability by innovation. Participation of wind power energy in markets can provide good incentives for more controllability and innovation. By controllability & innovation we mean all the actions that can be implemented to make wind power technology more similar to a conventional technology (e.g. new control system, IT installation, more centralized dispatch, etc) (Verhaegen et al, 2006). Innovations may appear also in windmill design by favorising more constant and controllable generation against only a maximal output objective.

Improving individual forecasting & system balancing efficiency. One of the reasons given to encourage the market participation of wind energy and to support equal balancing rules for all market participants is that this encourages the wind power producers to provide accurate predictions for system operation (Mitchell et al, 2006). If wind power producers have to pay for their imbalances, they will invest in forecast tools in order to reduce their balancing costs and therefore to maximize their profit. Wind farm owners can provide more accurate forecasts of their own production since they know the machines' availability and could run downscaling programs with detailed information from the field the terrain in order to increase the predictions' accuracy. A detailed prediction of each farm is particularly important in some special cases, for instance, when considering grid constraints violation. Although wind power centralized forecasting is needed¹⁸, improvements on individual forecasting can be translated into system balancing efficiency. Two conditions are needed for that: i) wind power producers give good forecasting information to the System operator and ii) the System Operator uses all the scheduling information to reduce the cost of system balancing.

Transparency of the support schemes. If wind power producers participate in electricity markets as other conventional technologies, they have to support (a part of) the integration costs from their intermittent production (balancing, congestions, etc.). These extra costs have to be included in some way in the support scheme in order to avoid creating a barrier for wind power. Including integration

¹⁸ This is because balancing cost depends on total forecasting error (demand, conventional generation and wind power) and the accuracy of an overall forecast is much higher than those of an individual wind farm due to "the large numbers" effect. The large numbers effect is very important in wind forecasting. The error reduction induces the wind farms to concentrate themselves on an only bid or schedule, or on a few ones.

costs in the support scheme clearly identify the real costs of the subsidy for each of the proposed new technologies and avoid cross-technology subsidies and consequent distortions. Excluding wind power producers of system-induced charges may imply extra costs for the transmission and distribution system operators and finally for the network users since these costs are socialized. This could lead to an under-evaluation of the necessary subsidy for the development of wind power energy and a problem of acceptability.

Negative effects of exposing wind power producers to market signals are :

Increase of risks of wind power producers. The revenue of wind power producers that need participating in forward and balancing markets is more risky than the revenue ensured by the feed-in tariff. As for other market participants, wind power producers will face market risks since they face volume and price risks for their output. Wind technology costs are mainly composed by fixed costs, and particularly by capital costs. Such market risks incur a huge investment risk that could deter investment in wind farms.

Transaction costs increase. Participation in markets implies more transaction costs than those incurred in the feed-in system. First, the producers have to understand the complex electricity market architecture and have to be able to understand and react to different signals sent by markets. The incurred transaction costs are of importance for small players. This is particularly the case for wind farms since they are usually small size plants (up to 50 MW for onshore wind farms)¹⁹. Nevertheless learning-by-using should lower these transaction costs since complex operations can become routines.

Considerable potential gains exist from exposing wind power producers to market signals. However it can increase risks and transactions costs for this particular segment of generators and to the system and society

¹⁹ Note that the transaction cost increase depends mainly on the industrial organization of wind power producers. If wind power plants belong to incumbents or big electricity companies, the transaction costs increase should be low since these companies hold all skills to participate on markets and they will just add wind energy to their generation portfolio. Conversely, if wind power plants belong to independent power producers, the transaction costs should be higher since producers should learn how to participate on markets and how to react to market signals.

in the end. These two opposite factors have to be balance in order to integrate large amount of wind power in a socially efficient way. There exist several intermediate solutions for exposing wind power producer to market signals, depending on the support scheme implemented and if there is specific market rules applied for wind power producers.

The adequacy between support schemes and electricity markets

In order to figure out the interactions of wind power producers with electricity markets and market design we build up our analysis following Klessmann et al (2008). They study how wind power producers are exposed to market signals (or risks) under three different support schemes: i) fixed feed-in tariff, ii) feed-in premium and iii) green certificate. For each support schemes, we analyze how wind power producer are exposed to “market” signals respectively in forward markets (day-ahead and intraday), balancing markets, congestion (and losses) pricing and connection and network tariffs.

Feed-in tariffs scheme

Forward markets signals. With feed-in tariff wind power producers do not realize electricity transactions as other conventional producers, i.e. wind power producers do not participate in day-ahead/intraday markets and are not exposed to short-term forward market signals. Wind power producers are set aside of markets; both the price and the sold volume are guaranteed for their output.

Balancing market signals. Concerning the balancing market, we distinguish two types of feed-in tariff implementations: i) feed-in tariff without balancing responsibility and ii) feed-in tariff with balancing responsibility. In the first case, the wind power producer does not bear the balancing responsibility since the electricity production and injection on the grid is made without any special obligation automatically (e.g. France, Germany). Integration costs due to intermittency are generally born by the system operator then spread over network users. This reduces completely the risk of balancing costs of wind power producers while it does not give any balancing signals to them. In the second case, balancing responsibility for wind power producers is included as a feed-in tariff rule. Wind power producer has to provide a load-profile before the time of delivery and the imbalances are computed following this load-profile (e.g. Spain). A feed-in tariff with

balancing responsibility scheme can be combined with specific rules for imbalance charges applied to wind power in order to arbitrate between the balancing signals and balancing cost risk.²⁰

Short term and long term locational signals. Feed-in tariff scheme isolates wind power producers of eventual short term locational market signals since they do not participate on electricity market. However, feed-in tariff scheme may be combined with different designs of connection and network tariff. "Shallow cost" design of connection policy minimizes the risk and the cost burden of wind power producers but does not give any long-term locational signal while a more "deep cost" approach gives long-term locational signal at the expense of higher risks and cost burden for wind power producers. The impact of different connection and network tariff designs depends on how these costs have been taken into account on the definition of the level of feed-in tariff or at the regulatory level. In some cases renewable energy producers do not pay for connection charges since it has been decided at the legislative level.

Feed-in tariff scheme is well-known to reduce uncertainty in revenue of wind power participants (Dinica, 2006; Mitchell et al, 2006). However, as feed-in tariffs give the same price for the electricity produced in all hours, this scheme does not give any signals of temporal valuation of energy²¹. Suppose that two locations are candidate to install a wind farm. Both locations have the same average wind speed levels but a different temporal production pattern. One site would produce more energy during night and the other site would produce more energy during the day. Under a feed-in tariff scheme, the investor has no preference for any site because he will have the same remuneration in both sites. From a system point of view, the optimal choice is the wind site producing more energy during peak hours. Therefore wind sites contributing to reduce the reliability system cost could be not chosen at first. Moreover, this scheme does not give signals

²⁰ For instance, wind power producers under feed-in tariff scheme in Spain have particular balancing rules. They have a fixed (regulated) imbalance price (7.8 €/MWh) and this price applies to the deviation between scheduled and actual delivered volume beyond fixed tolerances. For wind and solar energy, the tolerance margin is 20% (Rivier, 2008).

²¹ Recently, a few implementations of feed-in tariff have included different payment for different time of the day (Klein et al, 2008). This introduces some differentiation in the value of energy at different hours. If this differentiation is related to the real effects of injections into the system at different hours, this can give some signals to wind power producers.

for an optimal maintenance organisation. Concerning balancing, some implementations of feed-in tariff can give some signals which will imply some improvements in individual generation forecasting, controllability and transparency of the support scheme costs.

As wind power producers do not participate on energy (locational) markets, they do not have to support any short term locational signal²². Hence wind sites originating congestions (and losses) may be selected indistinctly by wind power producers. This implies higher additional congestion (and losses) costs. However, a feed-in tariff policy scheme combined with connection and network tariff designs whose focus on providing adequate locational signal²³ could help to lead optimal location of large-scale wind power generation.

Feed-in tariff scheme alone gives weak incentives to control (reduce) generation under extreme congestion or imbalance situations. The opportunity cost for wind power producers to reduce production is the value of the feed-in tariff, which is based on the average total cost of wind not related to the expected conditions of the system. However specific rules concerning the right of disconnection/shortage can be implemented to deal with congestions or other stability problems in the short-term (e.g. Spain, Portugal²⁴, Germany²⁵). This could happen in case of emergency and this decision could be taken exclusively by the SO in order to ensure the well-functioning of the system. This fact can illustrate the trade-offs between the authority of the SO and the market responsibility of participants.

Feed-in premium support scheme

Forward markets signals. Under feed-in premium support scheme, wind power producers have to participate in forward markets as

²² Note that although some feed-in tariff implementations give different prices for different regions, it does not correspond to locational signals indicating where is better for the system to install wind power farms but to reduce windfall profits of zones with very good wind resources.

²³ "Over-sized" locational signals can avoid the development of wind projects. This is the case for instance with the "deep cost" connection rule where the wind power developer has to pay the cost of all reinforcements made in the network after the new installation.

²⁴ In the case of technical problems, the system operator is allowed to interrupt wind farms production during valley hours (50h/year) (Peças Lopes, 2008).

²⁵ In Germany, the congestion management implies a curtailment rule: the network operator is allowed to curtail the output of renewable energy if the network is already congested. This rule has huge implications for the producer's income. In 2005, this curtailment rules would decrease the wind power producer revenue by 5% (Klessmann et al, 2008).

other market participants. The energy produced at each hour is valued differently and proportionally to the hourly electricity market price. Furthermore, wind power producers receive a regulated premium for each kWh sold (and produced) that represents the value of the positive externality of the renewable use.

Balancing market signals. Participation in the balancing market is also required since wind producers follow the same rules than the other market participants and therefore, they receive balancing market signals. They have to pay balancing charges if their production in real time differs from their contractual position in the forward markets. It is possible however to apply specific balancing rules for wind power (Rivier, 2008). These rules may consist in setting up tolerance values of balancing volume under which there is no cost of imbalance or in applying a fixed (regulated) imbalance price to reduce the uncertainty related to market based imbalance prices²⁶.

Short term and long term locational signals. Wind power producers may have the same obligations and signals than other participants or particular rules. Feed-in premium scheme may also be combined with different designs of connection and network tariffs.

The total revenue under this scheme is relatively more volatile than the fixed feed-in tariff. Nevertheless, the income risk can be limited in using a cap-and-floor mechanism. This mechanism implies that the market price plus the premium has to be set between a lower and an upper limit²⁷. This cap-and-floor mechanism brings a revenue warranty and limits seriously the price risk. Floor limits are particularly appropriate for investors who can continue to invest since they are able to determine a minimum return on their investments²⁸.

Under feed-in premium, wind power producers are exposed to the market signals so that they can adopt more efficient behavior. On forward markets, they have the possibility to value wind power depending on the particular time they are producing. Thus an investor will choose wind sites with more potential production during peak

²⁶ For instance, in Spain, the wind power producers under the premium scheme do not have to pay any charge for secondary reserve.

²⁷ For instance, in Spain, the market price plus the premium must be contained between at least 71.27€/MWh (lower limit) and 84.94€/MWh (upper limit).

²⁸ Note that this income warranty is only a price guarantee given that there is not a volume guarantee since wind power producers have to find a counterparty on markets.

hours in order to increase his revenue. This contributes to reducing the extra reliability system cost induced by wind power²⁹. Wind power producers should plan efficiently the maintenance of wind farms since their revenues depend on the moment when they are disconnected. Concerning balancing signals wind generators have incentives to improve forecasts of wind energy and to improve controllability of wind farms as they will have to support balancing costs. Furthermore, as the opportunity cost of not producing is mostly based on the "environmental premium" and on the expected condition of the system, wind power producers are more likely to reduce production when the system operator needs it. Finally, wind power producers can be exposed to the same short term locational signals than other market participants and this can contribute in the efficient selection of wind sites. It will be consistent with existing long term locational signal whatever the design.

Green certificates support scheme

Forward markets signals. Under green certificates, wind power producers have to participate in forward markets as other market participants. The energy produced at each hour is valued differently and proportionally to the hourly market price. Furthermore, wind power producers receive a market-based premium for each kWh sold (and produced) that represents the value of the positive externality of the renewable use. This market-based premium is the green certificate price which is de facto more volatile than the administered one. This introduces in the one hand a more accurate signal of the scarcity of green production but in the other hand news risks and transaction costs.

Balancing market signals. Participation in the balancing market is also required since wind producers follow the same rules than the other market participants and therefore, they receive balancing market signals. They have to pay balancing charges if their production in real time differs from their contractual position in the forward markets. It is possible however to apply specific balancing rules for wind power as in the premium feed-in price (Sioshansi et al, 2008, Makarov et

²⁹ This is strongly related to the so called "capacity credits" of wind generation. Selecting wind power sites with high production in peak hours corresponds to maximize the capacity credits of the wind generation and to minimize the over-cost of adequacy (see for instance Gross et al, 2006).

al, 2005). These rules may consist in setting up tolerance values of balancing volume under which there is no cost of imbalance or in applying a fixed (regulated) imbalance price to reduce the uncertainty related to market based imbalance prices³⁰.

Short term and long term locational signals. Wind power producers may have the same obligations and signals than other participants or particular rules. Green certificates scheme may also be combined with different designs of connection and network tariff.

The total revenue of wind power producers under green certificates support scheme is considerably more volatile with respect to feed in tariff and premium. Indeed the revenue depends on the green certificates prices which can vary considerably during the lifetime of a wind power plant (about 20 years). Furthermore, the existence of a buy-out price (the penalty that has to be paid by suppliers not fulfilling their green quantity obligations) complicates the estimation of the certificates price and increases again the risk. On the top of these support scheme risks, wind power producers have to bear the risk of participating in the electricity market. Even if wind power producers are exposed to electricity market signals which improve the incentives, the relative very high risk can increase considerably the cost of capital of wind power investments. Therefore only projects backed by long term contracts or under vertical arrangements will be undertaken which may limit considerably the wind power development (Finon and Perez, 2007).

Table 3 summarizes the main points of the interaction between support schemes and electricity markets. There are many possibilities of support schemes and market designs and combinations to expose wind power producers to market signals with different sharing of integration costs. Table 3 presents a subset showing the accuracy of market signals, risks induced by market signals and main revenue risk for a combination of support scheme and a benchmark of market design. This benchmark is characterized by a more centralized market, a gate closure near real time, one single imbalance price, zonal pricing

³⁰ For instance, the case of California is one example of specific balancing rules for wind power producers. Wind power producers included in the PIRP (Participating Intermittent Resources Program) have to schedule their energy in the forward market without incurring hourly or daily imbalance charges when the delivered energy differs from the scheduled amount. They are instead subject to imbalances charges accounted for monthly imbalances (Makarov, et al 2005). In Belgium, where a green certificates scheme is applied, balancing responsibility for wind is also limited as there is a tolerance margin of 30%.

		Feed-in Tariff		Feed-in Premium		Green Certificates	
	Market design options	Accuracy of market signals	Risk induced by of market signals	Accuracy of market signals	Risk induced by of market signals	Accuracy of market signals	Risk induced by of market signals
Day-ahead and intraday markets	Centralized	0	0	+	-	+	-
	Gate Closure Close real-time	0	0	+	-	+	-
Balancing market	single imbalance price	0/+*	0/+*	+**	+**	+**	+**
Congestions (and losses) pricing	Zonal pricing	0	0	+	+	+	+
Connection and network tariffs	Zonal network tariff	++	+	++	+	++	+
Main Revenue Risks		Low		Medium		High	

* feed-in tariff with balancing responsibility
 ** possible specific rules for wind power.

Table 3

and zonal network tariffs which correspond to a realistic efficient market design for Europe.

Table 3 shows also that from the policy maker's perspective, there is a trade-off between exposing the markets participants to a more accurate signals approach and a "low risks/transaction costs" approach. When renewables face high market risks and transactions costs, a higher level of financial support is required to stimulate renewable development than in a low risk and transaction cost environment. But the exposure to market signals may also give an incentive to make efficient use and development existing infrastructures and recent innovations, thus limiting the indirect costs to society.

Furthermore, support schemes have to take into account these extra costs resulting from combinations that expose wind power producers to market signals. The support schemes have to include some "normal/

efficient” subsidy for integration costs (e.g. efficient wind balancing costs). Wind power producers will participate in the market and face with a part of their integration costs but without stopping development. Theoretically, green certificate scheme includes naturally this extra subsidy because the price of certificates adjusts itself in order to give enough revenues to wind power producers on condition that capacity investment is in line with the target capacity (this is not true if the penalty or the buy-out price is fixed too low). Conversely, feed-in tariff and feed-in premium schemes need to consider in their tariff definition a normal/efficient target of integration cost. In this case, participants with improved behavior earn extra profit and participants having worse results than normal/efficient have a loss.

Policy recommendations and conclusions

Policy recommendations to promote an efficient integration of large amounts of wind power into the system can be summarized in three points:

- Readjust support schemes in order to increase the participation of wind power producers in markets and their exposition to market signals;
 - Improve market designs and market signals to avoid distortions;
 - Counter-balance recommendation 1) and 2) with the potential increase of transaction costs and generated risks for market investors.
- Support schemes should be designed to give more market signals to wind power producers. Market signals to wind power producers can be beneficial to improve the selection of wind sites (considering temporal patterns, congestions and losses), to improve maintenance planning, to improve the combination with other technologies, to incorporate portfolio effects and to add transparency concerning the total cost of promotion policy, etc. The feed-in premium seems to be the best trade-off solution because this option allows enjoying the benefits of exposing wind power producers to market signals without creating considerable new risks and transaction costs.

While European Directives set ambitious targets concerning the penetration level of renewable energies, promotion policies have to be

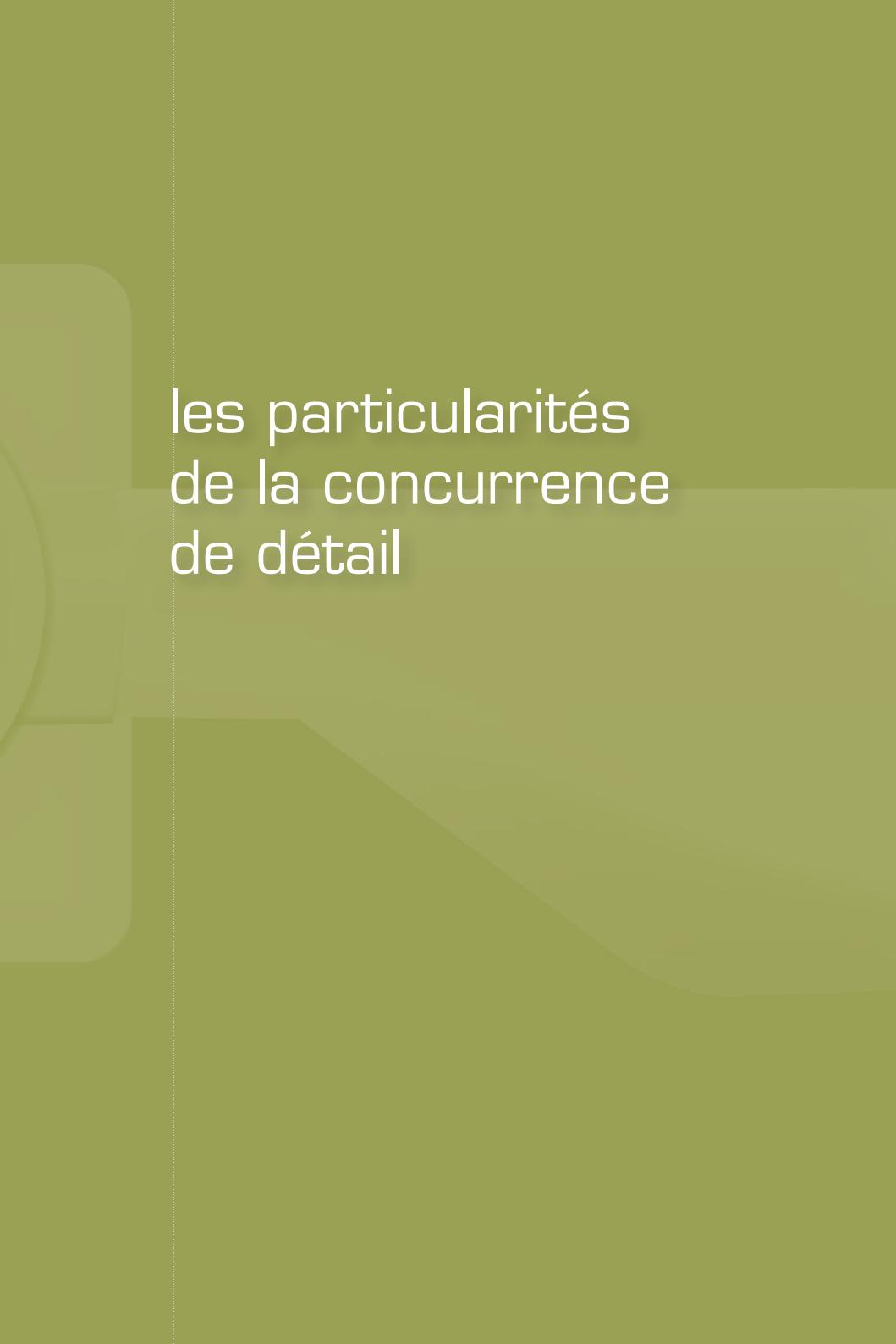
adapted in order to make it possible for renewable energy producers to be more sensitive on market signals. With high penetration levels, wind power cannot be set anymore aside of market signals and market participation. Nevertheless, market participation does not have to create entry barriers for such investments. A trade-off between support schemes and market participation has to be done at the regulatory level. Nowadays, two main issues have to be addressed. The first one concerns the “adaptability” way to smooth the passage between a low risk approach (the common feed-in tariff without balancing responsibility) to a more risky support scheme where producers can participate to markets and react on market signals. Several intermediate steps can be possible. One way to ensure investment while encouraging producers to act on markets is to let the choice to wind power producers to benefit from either a fixed feed-in tariff with balancing responsibility or a feed-in premium scheme while indirectly encouraging the latter.

Softer “specific rules” for wind power producers (e.g. setting up of tolerance margins for balancing settlement) can be used to limit the risk of being exposed to market signals. The second issue is to find the accurate equilibrium between what has to be managed by the SO and what can be managed through wind power producers’ market responsibilities. Transferring specific wind generation responsibility to the SO (by means of dispatch centers) can contribute to more wind energy deployment in the absence of accurate market signals. What happened in Spain represents one of the best compromises between market signals, low risk and adequate sharing of wind generation responsibility between market wind power producers and the SO.

References

- Barth, R., Weber, Ch., Swider, D., 2008. "Distribution of costs induced by the integration of RES-E power" *Energy Policy*, Volume 36, Issue 9
- Butler, L., Neuhoff, K., 2008. "Comparison of feed-in tariff, quota and auction mechanisms to support wind power development", *Renewable Energy*, Volume 33, Issue 8n
- Del Rio, P., Gual, M., 2007. "An integrated assessment of the feed-in tariff system in Spain", *Energy Policy*, Volume 35, Issue 2
- Deutsche Energie-Agentur (DENA), 2005. *Planning of the Wind Integration of Wind Energy in Germany Onshore and Offshore up to the Year 2020* Berlin, Final Report
- Dinica, V., 2006. "Support schemes for the diffusion of renewable energy technologies – an investor perspective", *Energy Policy*, Vol. 34, Issue 4
- Di Castelnuovo, M., Leach, M. and Pearson, P., 2008. "An analysis of spatial pricing and renewable generation in the British electricity system", *International Journal of Global Energy Issues*, Vol. 29, Issues 1/2
- Ehrenmann, A., Smeers, Y., 2005. "Inefficiencies in European Congestion Management Proposals", *Utilities policy*, Volume 13, Issue 2
- European Commission, 2006. *Communication de la Commission, du 7 décembre 2005, relative à l'aide en faveur de l'électricité produite à partir de sources d'énergie renouvelables* [COM(2005) 627 final - Journal officiel C 49 du 28.2.2006]
- Finon, D., Perez, Y., 2007, "Transactional Efficiency and Public Promotion of Environmental Technologies: The Case of Renewable Energies in the Electric Industry", *Ecological Economics*, vol. 62
- Forsund F., Singh B., Jenson T., Larsen C., 2007. "Wind power, network congestion and hydro resource utilization in the Norwegian power market". Working paper No. 33/07 - Institute for research in economics and business administration, Bergen
- Green R., 2008. "Electricity Wholesale Markets: Designs Now and in a Low-carbon Future", *The Energy Journal*, Special Issue.
- Gross R., Heptonstall P., Anderson D., Green T., Leach M. and Skea J., 2006. *The Costs and Impacts of Intermittency, An Assessment of the Evidence on the Costs and Impacts of Intermittent Generation on the British Electricity Network*, available online at <http://www.ukerc.ac.uk/content/view/258/852>
- Holttinen H., Lemström B., Meibom P., Binder H., Orths A., van Hulle F., Ensslin C., Tiedemann A., Hofmann L., Winter W., Tuohy A., O'Malley M., Smith P., Pierik J., Olav Tande J., Estanqueiro A., Ricardo J., Gomez E., Söder L., Strbac G., Shakoor A., Smith J.Ch., Parsons B., Milligan M., and Wan Y., (2007), "Design and operation of power systems with large amounts of wind power: State-of-the-art report", VTT working paper, available on <http://www.vtt.fi/inf/pdf/workingpapers/2007/W82.pdf>
- ILEX, 2002, *Cost reflectivity of imbalance charges*, A note from ILEX to Slough Heat and Power, November
- Joskow, P., 2006, "Competitive Electricity Markets and Investment in New Generating Capacity", MIT WP-2006-009
- Klein, A., Pfluger, B., Held, A., Ragwitz, M., Resch, G. and Faber, Th., 2008. "Evaluation of different feed-in tariff design options – Best practice", paper for the International Feed-In Cooperation, October
- Klessmann, C., Nabe, Ch., Burges, K., 2008. "Pros and cons of exposing renewable to electricity market risks –A comparison of the market integration approaches in Germany, Spain and the UK", *Energy Policy*, Volume 36, Issue 10
- Lamy, M-L., 2004. *Efficacité des politiques environnementales d'incitation à l'adoption de nouvelles techniques: le cas des énergies renouvelables*, Ph.D Dissertation, University Pierre Mendès-France, Grenoble - France
- Leuthold, F., Weigt, H., von Hirschhausen, Ch., (2008). "Efficient pricing

- for European electricity networks – The theory of nodal pricing applied to feeding-in wind in Germany", *Utilities Policy*, vol. 16, n°4.
- Littlechild, S., 2007. *Electricity cash-out arrangements*, Memo - 9 March 2007, available at <http://www.cornwall-energy.com>
- Makarov, Y., Blatchford, J., Alarian, H., DeMarse, K., O'Hara, M., Scholz, M., Jericich, S., Vidov, J., Leuze, E., Abernathy, R., Hawkins, D., 2005. *Incorporation of wind power resources into the California energy market - CAISO*
- Mitchell, C. D. Bauknecht, P.-M. Connor, 2006. "Effectiveness through risk reduction : a comparison of the renewable obligation in England & Wales and the feed-in system in Germany", *Energy Policy*, Vol. 34, Issue 3
- Mügens, and Neuhoff, K., 2006. "Modelling Dynamic Constraints in Electricity Markets and the Costs of Uncertain Wind Output", Cambridge Working Paper in Economics 0610, Faculty of Economics, University of Cambridge, February
- Newbery D., 2005. *Refining Market Design*, Discussion paper of the European Union research project SESSA, available on www.sessa.eu.com
- Peças Lopes, J., 2008. *Wind Power Technology*, Presentation Joint Cambridge-MIT Conference Electricity Markets, Paris, EdF - 4 July
- Ragwitz, M., Held, A., Resch, G., Faber, T., Haas, R., Huber, C., Morthorst, P.E., Jensen, S.G., Coenraads, R., Voogt, M.,
- Reece, G., Konstantinaviciute, I., Heyder, B., 2007. *Assessment and optimization of renewable energy support schemes in the European electricity market*, OPTRES Final Report, Karlsruhe
- Resch, G ; Held, A ; Faber, Th ; Panzer, Ch ; Toro, F ; Haas, R. (2008) "Potentials and prospects for renewable energies at global scale", *Energy Policy*, Vol. 36, Issue 11
- Rious, V., Glachan, t JM., Perez, Y., Dessante, P., 2008. "The diversity of design of TSOs". *Energy Policy*, vol. 36, issue 9
- Rivier, J., 2008. "Spanish experience in market participation of wind parks". *Submitted to Energy Policy*
- Sioshansi, R., Hurlbut, D., 2009. "Market Protocols in ERCOT and their effect on wind generation". *Energy Policy*, in press
- Stoft, S., 2002. *Power system economics*, IEEE Wiley
- Swider, DJ., Voß, A. (Eds.), Barth, R., Brand, H., Weber, C., Vogel, P., Meibom, P., Strbac, G., Black, M., Figueiredo, V., Auer, H., Obersteiner, C., Weissensteiner, L., Prugger, W., Faber, T., Resch, G., 2006. *Disaggregated system operation costs and grid extension cost caused by intermittent RES-E grid integration*, Deliverable 5b of the EU-project GreenNetEU27 ("Guiding a least cost grid integration of RES-Electricity in an extended Europe"), Stuttgart
- UKERC, 2007. *A review of electricity unit cost estimates*, UKERC, Working Paper
- Vandezande, L., Meeus, L., Belmans, R., Saguan, M., Glachant, J.-M., 2009. "Well-functionning balancing markets as a prerequisite for wind power integration". *Energy Policy*, in press
- Verhaegen, K.; Meeus, J.L. & Belmans, R, 2006. "Development of balancing in the internal electricity market in Europe" in Proceedings 2006 European Wind Energy Conference, Athens, Greece, March 2006
- Weigt, H., 2006. *A Time-Variant Welfare Economic Analysis of a Nodal Pricing Mechanism in Germany*. Dresden University of Technology - Electricity Markets Working Paper No. EM-11, Chair of Energy Economics and Public Sector Management - Faculty of Business and Economics
- Weigt, H., Jeske, T., Leuthold, F., von Hirschhausen, Ch., 2009. "Take the Long Way Down" – Integration of Large-Scale Offshore Wind in the North Sea through HVDC-Lines". *Energy Policy*, in press
- Wilson, R., 2002. "Architecture of electricity markets", *Econometrica*, vol. 70, Issue 4



les particularités de la concurrence de détail

The dynamic of competition in presence of switching costs

the case of British Gas (1997-2007)

Christophe Defeuilley

LARSEN

Matthieu Mollard

University Paris Sud (ADIS-GRJM) and LARSEN

Theoretical models with switching costs have shown that, in a dynamic game, firms have incentives to adopt two-period strategies. Initially firms will 'invest' in markets at an early stage in their development to be able to then 'harvest' at a later stage when consumers are locked-in to the supplier they have previously patronized. This article illustrates, based on a case study: (the supply of electricity in Great Britain), that this two-period strategy is not always crowned with success. Specifically, certain firms (in this case British Gas) may have difficulties sustaining customer loyalty in the second phase because customers benefit from a 'learning effect', which in turn lowers their switching costs. Taking these learning effects into account could enrich the range of situations analyzed in the existing economic literature with respect to the understanding of the impact of switching costs on competition.

Introduction

Imperfect competition can take different forms : be it through the exercise of monopoly, oligopoly or monopsony market power, the production of non-homogeneous goods, or price discrimination (Tirole J., 1998). Most of the models that deal with the issue of competition in the presence of switching costs are usually found amongst those that deal with the more general framework of markets with differentiated goods. This differentiation can be created through different ways: localisation, availability, quality, or lack of information concerning the existence of other goods and their characteristics. Differentiation allows for the possibility of price discrimination if arbitrage opportunities are limited. This is the case when it is difficult to sell-on the good to a third party once it has been purchased, either because the characteristics of the good in question do not allow it or because the good in question is bespoke to the preferences of one and only one precise category of consumers.

There are situations in which the act of purchasing a good can give rise to certain specific costs that fall solely on the purchaser. These costs, qualified as switching costs, can lead to product differentiation by implicitly modifying the characteristics of the good that has been exchanged. From the perspective of the consumer that purchases the good, the fact that switching costs may not be fully known until the purchase has taken place, leads to the differentiation of goods that would normally be a priori substitutes. This differentiation can lead the seller of the good in question to execute strategies which –if successful- will in turn differentiate the seller from his/hers competitors. These strategies are deployed in ways that are different according to the case : increasing switching costs to exploit consumer inertia and thus extract the maximum amount of consumer surplus, or alternatively reducing switching costs in order to favour switching and attract additional demand (Diamond P., 1971; Salop S., 1977; Klemperer P., 1987a).

Dynamically, the existence of switching costs can create incentives for firms to adopt two-period strategies. After making efforts in terms of price and advertising expenses to enter a new market and attract new clients (willingness to sacrifice, at least temporarily, their profitability), firms will modify their behaviour in a manner favourable

to them (ending discounts, raising prices) counting on the existence of switching costs – which are borne by consumers - to act as a barrier to exit (Klemperer P, 1995). Thus, firms consent to make 'sacrifices' in the first period only because they anticipate being able to recoup their 'investment' in the second period.

This theory of dynamic competition in the presence of switching costs can be applied to several generic situations. Industrial sectors as varied as telecoms, the Internet, insurance or banking tend to support and corroborate its validity (OFT, 2003; Farrell J., Klemperer P. 2007). The present article seeks to make a contribution to this body of empirical work¹.

We examine a market with switching costs : in this case the British residential electricity supply market. We analyze the strategies of the main firms that compete with each other in this market, and in particular we take a look at the strategy executed by British Gas the market leader during the 1997-2007 period. Two issues come to the forefront from this analysis. British Gas, having first taken the lead amongst its competitors, executed a two-period strategy, acquiring market share by means of door-to-door selling, strong advertisement and targeted prices. Nevertheless, contrary to its expectations, the company did not succeed in developing strong customer loyalty, which leads to the company failing to meet its profitability targets and the subsequent weakening of its business model.

Switching costs constitute a strategic variable for firms within this market, yet it is clear that the second phase of this strategy (harvesting) can be difficult to execute, at least this seems to have been the case for British Gas. It is harder than initially thought to develop customer loyalty, because customers reinforce their mobility when they learn how to switch suppliers. At the same time firms have limited means to erect barriers to exit ('exit' in this context refers to customers). This consumer learning effect which allows them to exercise their freedom of choice more easily -because their switching costs diminish over time- is almost absent from the theoretical models dedicated

¹ This article falls within the literary methods used in economic analysis (see D. McCloskey, 1998). We try to produce a narrative that plays the role of a "plausibility test" of these theoretical models of dynamic competition with switching costs. In this way we confront the conclusions of these models with a real life example, that serves to identify in an exploratory manner the anomalies or the surprises that may exist - in relation to the existing literature - and thus to advance new hypotheses that could, if rejected, be used to enrich future theoretical work.

to analyzing dynamic competition in the presence of switching costs. Thus taking into account this effect could enrich the range of situations analyzed by the literature and their impact on the functioning of markets where there is imperfect competition.

This article is organized as follows. The first section contains a brief review of the literature dealing with the principal models that analyze dynamic competition when there are switching costs. The second part analyzes the case of the British electricity supply market over the period 1997-2007, insisting on the two-stages of competition. Following we present the main lessons drawn from this example. The third section concludes.

A brief literature review

Switching costs can be defined as the costs (real or perceived) that follow as a result of changing supplier; costs that would not have been borne by the consumer if he/she had remained loyal to his/her previous supplier. This means that a consumer that has bought a good from a firm may incur additional costs if he/she buys the same good from a different firm, even if the good in question is sold for exactly the same price. In other words, switching costs may arise in any commercial relationship when the parties involved have gains to make from renewing their relationship instead of starting a new one with a different supplier (OFT, 2003, Klemperer P, 1987a).

This gives a characteristic of specificity to the commercial relationship once it has been established. Thus, from a consumer's point of view switching costs introduce a differentiation of goods that are a priori identical as far as their technical, usage and price characteristics are concerned. This differentiation can have numerous consequences, including the price level paid by consumers, the structuring of the market and the strategic behaviour of firms.

Switching costs may be classified as follows (Klemperer P, 1987a):

- Information costs: these are related to research and information gathering activities conducted by consumers prior to the purchase of the good in question in order to identify the available offers and to compare them with each other so as to find the one that best corresponds to his/her needs.

- Transaction costs: they arise most notably from the opening and closing of a contractual relationship with a supplier (closing an account, communication of new account details, etc).
- Learning costs: they relate to the consumer learning how to use or manage a new product or service. Even in the case of products that are functionally identical the consumer may have to discover and learn how to use the functionalities of the new product (i.e. software associated to the acquisition of a new computer).
- Uncertainty: these costs relate to the uncertainty surrounding the consumption of the good, i.e. when the good in question sold is an experience. In this case, the consumer is confronted with a degree of uncertainty with respect to the real quality of the service that is being sold by the new supplier.

Numerous articles have tried to examine the impact of the presence of switching costs on the functioning of markets². In the next section we present an overview of certain key conclusions that have come out of these works in question.

Models without price discrimination. In these types of models, the underlying hypotheses are that all consumers face an identical price in each period; two firms (A and B) are in competition and they bear their own production costs. The existence of switching costs introduces a differentiation element amongst consumers, notably, those consumers already tied to a supplier (C1) and those that are not yet tied to a supplier (C2) (Klemperer P, 1995). C1 type consumers must bear an additional cost to change suppliers. Firm A, wishing to attract customers of Firm B, must reduce its price below Firm B's so as to compensate for the switching costs borne by consumers. As price discrimination has not yet been introduced in this model, Firm A must charge the same (lower) price to its existing customers. Firm A will thus have to proceed to make an arbitrage decision between the gains to be made from acquiring new customers and the foregone profits that will result from charging a lower price to its entire customer base. The choice will be a function of Firm A's historical market share: the larger it is the smaller the net gain to be derived from a strategy that targets the acquisition of new customers.

² This part follows largely from the literature review conducted by the OFT (2003) and Farrell J. et Klemperer P, (2007).

Let us now consider not just one period, but instead two. During the first period, the two firms compete with each other to attract consumers, in the second period; consumers will bear the additional costs if they wish to change suppliers (in this model, there are no new consumers). This type of setting can illustrate a market that is evolving from an expansion phase (with intense competition seeking to attract the maximum number of clients) to a mature phase where consumers are tied to a supplier. Farrel J. and Shapiro C. (1988) show that, in the second period, the firm that benefits from the largest market share will be able to charge a price higher than that of the rival firm without having to fear a loss of market share - to the extent that the firm in question benefits from a large base of captive customers not inclined to change suppliers-. Thus the firm in question has more to gain from increasing its prices and extracting additional profits from its own client base than in acquiring market share. This incentivises both firms to compete aggressively in the first period so as to gain the largest possible market share, even if this means charging prices below marginal cost³. These types of two period strategies are known under the terms "bargain then rip-off" or "investing then harvesting" (Klemperer P., 1987a, b ; Klemperer P., 1995 ; Padilla A., 1992 ; Farrell J. and Klemperer P., 2007).

The existence of new consumers in the second period, introduces the need for competing firms to make an arbitrage decision between, charging higher prices to their own clients or investing to gain new clients. This arbitrage will also be influenced by each firms' market share. The firm with the smaller market share will have a bigger incentive to offer lower prices. This can be interpreted in the following way: the existence of new consumers can partially protect captive consumers by modifying the price strategies of at least some of the firms competing in the market.

The overall impact of these new clients on the strategies followed by firms during the first period can nevertheless be ambiguous (OFT, 2003). Particularly when one considers that, when there are new consumers in the second period, firms will have less of an incentive to compete aggressively against each other in the first period if this

³ Overall even when consider the 'investing' period, most models are in agreement in concluding that prices in a market with switching costs will be higher than those that would prevail in the absence of switching costs.

implies that there will be an important disequilibria in the second period, with one firm having a large market share and another one a smaller market share. In this instance the latter will have an incentive to reduce its prices to attract new consumers, which would force the larger firm to in turn reduce its prices (both prices being influenced by switching costs, which we have assumed to be fixed). One could thus expect that in a market where there is a significant amount of new consumers in the second period, firms will not compete aggressively in the first period. In this way, firms signal to each other that they are ready to be aggressive with respect to their pricing policies in the second period. They are thus prone to modify their strategy from – in the words of D. Fudenberg and J. Tirole (1984) - one of "fat cat" to one of "lean and hungry".

Models with price discrimination. Introducing the possibility of firms discriminating amongst their respective client bases leads to the following results. Chen (1997) presents us with a model in which firms can discriminate between their old clients (captive) and their new clients by charging different prices to each group. New clients pay a lower price than the older (captive) clients for the same good, which allows the firm to attract these new clients –as they are compensated for their switching costs-. The model is based on two periods and assumes that two firms produce an identical good. In the first period, the firms compete with each other on price, and in the second period they may offer discounts. Consumers face additional costs if they switch suppliers. Both firms are assumed to have perfect knowledge of the purchasing practices of consumers in the first period.

In this model, firms adopt a strategy of 'investing' in the first period and then 'harvesting' in the second period. Moreover, the price strategies developed by each firm are independent of their respective market share. The firm with the higher number of clients will not necessarily offer a higher price to its clients than the firm with the smaller market share. This is due to the fact that the possibility of discriminating in the presence of perfect information increases the price elasticity of demand. During the second period, if the larger firm were to sharply increase the price it charges its captive clients, the smaller firm could seek to attract them by proceeding to offer targeted discounts.

These discounts will be all the more interesting the higher the price differential between the two firms becomes.

In fact, the dominant firm must moderate its price rises, even for its captive customers. The profits derived by firms in competition are lower, and the market shares of the two competing firms are not stable (loss of market share of the dominant firm in favour of its competitor). It must be noted that this model makes a strong hypotheses, which is that firms have perfect knowledge of customers' practices, this allows them to execute strategies conducive to efficient discrimination (for other models based on price discrimination see also, Nilssen P., 1992; Shaffer G. & Zhang Z., 2000 and Taylor C., 2003).

Models with endogenous switching costs. In the previous models described, switching costs were considered to be exogenous. In other words, switching costs were deemed to be independent of strategic actions taken by firms. We now turn our attention to consider switching costs as strategic variables: thus making them endogenous. Firms can have incentives to make switching costs rise in order to increase the difficulties borne by consumers when changing suppliers (thus erecting barriers to exit). In order to do that, firms are prone to seek to manipulate compatibility costs (by making their products less compatible with each other and their usage more specific) and transaction costs (by creating loyalty programmes). The possibility to make switching costs endogenous has been taken into account by certain models (P. Klemperer, 1995), in which firms decide in period 0 whether to increase or not switching costs, after considering the impact on their expected profits in periods 1 and 2. There has not been much research done on the issue of what measures consumers may engage in to mitigate or bypass the actions taken by firms in order to artificially raise switching costs. M. Riordan's and D. Sappington's (1989) article, suggests that a possible strategy that could be undertaken by consumers in order not to be too captive would be to diversify their purchases by seeking to source themselves from a second supplier selling the same goods.

The workings of markets with switching costs. Modeling the impact of switching costs on the way that markets work can lead to numerous

interesting theoretical conclusions. We consider the following to be the most significant:

The introduction of switching costs in the analysis of strategies executed by firms shows that the “investing/harvesting” strategy can be a dominant strategy for firms. It leads firms to compete (via prices) in the first period -even if it means pricing below marginal costs- in the expectation that they will reap the benefits in the second period by increasing prices charged on their captive client base.

Firms’ decisions will depend on their capacity to offer price discounts to certain –well identified- category of consumers (price discrimination with perfect information) and/or to artificially increase customers’ switching costs. The nature of the market in question (mature with little or no new consumers entering the market or fast growing) will also affect the strategic behaviour of firms.

These decisions will have an impact on market structure: the market share of the incumbent firms (their size and stability), churn rates, ease or difficulty of entering the market. We would like to highlight that the existence of switching costs does not necessarily imply higher barriers to entry into the market. The incumbent firm (benefiting from a captive client base), will not necessarily react in an aggressive way to a perceived threat of entry, specially if the latter is concentrated on a particular segment of the market i.e. on new customers in a growing market (Farrell J. and Klemperer P, 2007; Gabszewicz J. et al., 1992).

The British electricity supply market

The European electricity sector has been progressively liberalised to the point of allowing for competition in the residential supply segment of the value chain, which became mandatory in all EU member states on July 1st 2007. Ahead of other member states Great Britain began opening its residential electricity supply sector in September 1998⁴. In this section, we will examine how the competitive dynamics evolved in the British market during the period 1997-2007, thereafter we will

⁴ It was completely opened in May 1999.

contrast the conclusions arrived to from this analysis with the results derived from the theoretical models described in the previous section.

Description of the market

The residential electricity supply market is characterised by four elements that can have an impact on the dynamics of competition: a) the opening of the market and the opportunity to switch suppliers, brings to the forth the issue of switching costs that firms will now have to consider in their strategies aimed at either acquiring new clients or creating client loyalty; b) the product: 'electricity supply' is a homogeneous product in terms of its quality, it is thus difficult for suppliers to propose a differentiated good beyond that of price differentiation; c) the market is mature (weak growth in demand) and almost all customers are already tied-into a pre-existing supplier⁵, which rises new client acquisition costs for new entrants; d) it's a subscription market with the peculiarity that a consumer cannot seek to be supplied by two firms at the same time.

These characteristics create a set of constraints for firms in competition, which will have to consider these issues in their actions (choice to enter the market, pricing policy, level of discounts granted to certain categories of consumers). To these constraints, there are, in addition, specificities unique to the British market, which are a result of the industrial structure existing prior to the opening of the retail electricity supply market.

Before the reforms initiated by The Electricity Act in 1989, the British electricity sector was organised in the following manner: one state owned company was in charge of the generation and transmission of electricity (the Central Electricity Generating Board or CEGB), distribution and supply was guaranteed by 14 regional monopolies, also state owned, better known by their acronym, the RECs: Regional Electricity Companies (Helm D., 2004). These regional distribution-supply monopolies benefited initially from a franchise granting them exclusivity to supply residential customers in their respective historical

⁵ Between 2003 and 2007, almost one million new clients were added to the total client base in the British market thus bringing the total account number to 26,7 million at the end of 2007. This modest growth in account numbers (less than 4% over a five year period) was driven by clients that were already being serviced by a supplier but that opened a second account (due to the establishment of a secondary residence, separation, divorces, etc). New customers entering the market (arriving on British territory) were relatively few.

zones. The first phase of the reform was to privatise them between 1990 and 1991, and then to gradually make their geographical supply monopolies disappear (Surrey J., 1996)⁶.

A generalised investing phase (1997-2003)

As mentioned before the pre-existing industrial configuration of the retail market was characterised by the existence of the historical regional distribution/supply monopolies. Thus initially the competitive dynamics in the retail electricity supply segment were characterised first by an extension of the geographical footprint of several suppliers and thereafter by the creation of a unified market, epitomised by the establishment of national brands. For all the incumbent firms (and a fortiori for all those seeking to enter the incumbent firms' markets), the acquisition of clients outside of their historical (former monopoly) zone became a necessity.

Two elements characterise this first phase of competition. On a first instance, the former distribution monopolies enter the zones of their homologues and begin to take market share from them. By December 2003, the portfolio of the former monopolies is composed as follows: 75% of their clients are still former clients from their historical operation zones and 25% are clients located outside of their historical influence zone (Electricity Association, 2004; OFGEM, 2004)⁷. This loss of market share impacts all of the existing firms in all of their zones (nevertheless there are significant regional variations), indicating that competition is thus developing over the whole geographical market (see Table I).

The second element, is the appearance of a new entrant in the electricity supply market, British Gas, the former public gas supply monopoly, gathers over the course of a couple of years a client base of several million electricity supply clients, to the point of having a 24% market share by the end of 2003⁸.

⁶ The CEBG was also split into production and transmission activities. The production activities were distributed amongst several companies that were themselves privatised in 1991 and 1996.

⁷ Note that between the privatisation dates and 2003, the 14 former distribution-supply monopolies proceeded to vertically integrate upstream (M&A with generation companies –excluding nuclear- and by developing new production units) thereafter they merged amongst themselves creating 5 integrated groups (Npower, SSE, EDF Energy, Powergen, Scottish Power). Foreign large European energy groups control four of these five groups.

Supplier	Residential clients total		Within historical area customers	Out-of-area customers		
	%	Number		%	Number	%
Npower	15%	3 751 800	76,52%	2 870 750	23,48%	881 050
SSE	14%	3 501 680	81,93%	2 869 060	18,07%	632 620
EDF Energy	14%	3 501 680	88,96%	3 115 260	11,04%	386 420
Powergen	21%	5 252 520	77,90%	4 091 760	22,10%	1 160 760
Scottish Power	11%	2 751 320	66,90%	1 840 620	33,10%	910 700
Total	75%	18 759 000	79%	14 787 450	21%	3 976 550
British Gas	24%	6 002 880	0,00%	0	100,00%	6 002 880
Others	1%	250 120	0,00%	0	100,00%	250 120
Total	100%	25 012 000	56%	14 787 450	44%	10 224 550

Sources: Electricity Association, 2004; OFGEM, 2004. Figures as of end of 2003.

Table 1

British Gas becomes the main challenger of the former regional monopolies over all their zones by offering electricity to all its former historical gas customers. On the other hand, the other new entrants do not succeed in establishing a durable presence in the market, and are either bought or cease operations after a couple of years (Littlechild S., 2005). At the opening of the market there were 14 new entrants in the electricity supply market, today there are only four. Together they account for less than 0.3% of total electricity retail sales (OFGEM, 2008).

Five years following the opening of the market to competition around 44% of residential customers had changed suppliers, a situation not encountered in any other European country (Defeuilley C., 2009). British Gas accounted for 60% of this total, the rest being distributed amongst the five large groups that were created out of the M&A activity that occurred around the former distribution-supply monopolies. The market share of other players was insignificant.

B British Gas is the supply brand used by Centrica. Centrica was created in 1997 following the privatisation and dismantlement of the former public monopoly British Gas. Centrica regrouped the retail supply activities in the UK (thereby retaining the brand) and some upstream gas assets that belonged to the former British Gas monopoly. The retail gas supply market was partially opened to competition in 1996 and then completely in 1998 (before the complete opening of the electricity supply market).

Firms clearly implement an expansion strategy, characterised by the will to increase their market share and enlarge their client base (Mollard M., 2007). For this strategy to be successful, they are obliged to bear part of the switching costs that act as barriers to entry and slow down the switching rates of consumers. This kind of strategy is characterised by an 'investing' phase during which the competing firms will temporarily sacrifice part of their profits in order to gain new clients. The differentiation of the product 'supply of electricity' being limited (the good is a homogeneous one, its quality is defined by strict technical standards), firms will choose to differentiate their products through their pricing policies (and to a lesser extent by the types of contracts offered) and by their advertisement and marketing efforts.

Pricing policies. The former monopolies begin to adapt their strategies and they begin to discriminate amongst their client base by charging higher prices to their 'historical clients' (those clients located within their historical former monopoly distribution-supply zone) and lower prices to those clients situated outside their historical territory. Firms proceed to implement pricing policies -not at the national level- but on a region-by-region basis (the 14 former distribution-supply monopoly regions), they thus geographically segment the market (place of residence of the client)⁹. This strategy is implemented by all the former monopoly suppliers, and results in a significant price differential (above 10% on average) between the prices charged by the historical operators within their historical zones and the best alternative offer for the same zone (OFGEM, 2007).

Within their historical zones, the former monopolies fix their prices in reference to those charged by British Gas. British Gas being considered the most serious of the challengers by the former monopolies, that is to say, the one that could make them lose the largest number of clients. Two reasons lead to this conclusion. First, British Gas is the only firm with a national brand recognised by almost all electricity consumers. Over 80% of British households are

⁹ At least during the years immediately following the opening of the market. Later on as historical switching increases in the system, accounting for flows derived from moving, the historical relation between consumers and their suppliers becomes more difficult to manage.

connected to the natural gas distribution network (OFGEM, 2008). Second, British Gas suffered an unfavourable market opening of its historical market (gas supply) and an asymmetric regulatory regime that forced it to lose market share in its retail gas supply business from 1996 onwards, a fact that favoured electricity suppliers¹⁰. British Gas must therefore seek to reconstitute quickly its client base or alternatively risk being marginalised (Thomas S., 2003).

British Gas decides to respond to the entry of the electricity suppliers into the gas supply market by entering into the electricity supply market. British Gas leverages strongly the two assets that are conferred to it by its nationally recognised brand and its historical relation with consumers (its residential gas supply market share was 76% in 1999 versus 100% in 1996). The firm begins to propose a combined offer: electricity and gas (dual fuel), which in essence simply means offering to sell electricity to its own gas residential customers. The prices proposed by British Gas become frequently slightly lower than those of the former monopolies. The firm appears thus –in the eyes of consumers- as the least expensive ‘known alternative supplier’. Outside of their historical zones, the former monopolies begin to offer prices generally below those of the historical supplier and equally lower than those proposed by British Gas, who can, as we have seen, count on its notoriety¹¹.

One could thus consider British Gas as being the pivotal firm in the market, in reference to which the other firms determine or adjust their price offers (OFGEM, 2008, House of Commons, 2008). This does not necessarily mean that they do so also in terms of timing. British Gas is not generally the first firm to publicly announce its price changes. The former monopolies have a certain degree of flexibility to

¹⁰ In the gas market, the regulator forces British Gas to lose market share in the Industrial and Commercial segment of the market, which is rapidly penetrated by Oil & Gas firms. The firms must thus concentrate on supplying the residential market, which was being attacked by the electricity supply firms. Later on, the regulator also prohibits Centrica from lowering its prices in the residential gas supply sector until it judges that there is a sufficient degree of competition. This will eventually be the case in 2002.

¹¹ Note that the British regulator, OFGEM, decided to regulate until 2002 the offers made by the former monopolies to their historical clients. Constraints on their minimum prices are set (minimum price restraints), which allow competitors to offer more competitive prices and gain market share –within the context of a market where wholesale prices were falling– (Mollard M., 2007). These price controls are abolished in 2002.

pre-empt British Gas without too much risk (if they consider that it is in their interest to do for image reasons or for profitability reasons). In effect, they can anticipate what could be British Gas' future price movements by virtue of public information at their disposal¹².

Advertising and direct marketing. The second differentiating factor used by firms is advertisement and door-to-door sales. The former monopolies and the new entrants engage themselves in aggressive client acquisition campaigns. This results in increased advertisement expenditures¹³ and into massive direct marketing campaigns (specifically into door-to-door sales campaigns) ¹⁴. In June 2003, over 90% of the market is targeted through this latter method, with annual advertisement expenditures estimated to be in the region of £250 million (OFGEM, 2008). Advertising and direct marketing allow the supply firms to free themselves from having to compete on price to attract new customers, some even resort to the use of questionable practices (House of Commons, 2008).

Initial investments and profitability. The necessary investments required in order to acquire a client base (in terms of publicity, of direct marketing and of competitive price offers) are executed to the detriment of profitability. British Gas is an example of this practice. The firm decides to embark on a massive acquisition strategy and invests several hundred million pounds in advertisement, direct marketing, and the deployment and implementation of its IT client relationship management and billing systems. This results in significant initial investments that have a negative impact on gross margins during the first four years (1997-2000).

¹² British Gas gives every year provides abundant information in terms of its electricity and gas sales activities: number of clients, turnover, consumption, sourcing costs, distribution costs, operational costs gross margin, profit targets. From the information provided, it is possible to estimate with a reasonable margin of error what is the financial equation that British gas needs to solve and how it may change its retail price levels in function of changes to wholesale prices.

¹³ More than £150 million are thus spent on advertisement campaigns between 2001 and 2004 (Electricity Association, 2003; 2004 ; 2005).

¹⁴ According to different estimates the costs of acquiring an electricity supply client are somewhere between £30-60 (Waddams-Price C., 2004 ; Littlechild S., 2005), to which one must add the operational costs, which are estimated to be £20- 30 per client.

¹⁵ The other firms are vertically integrated and do not provide as much detailed accounting information regarding their retail sales as British Gas does in the United Kingdom.

British Gas accumulates almost £400 million in operational losses during this period. Although British Gas manages to rebalance the situation in the subsequent 3-year period (2001-2003), the result is modest (around £150 million of cumulative operational profit), which represents a gross margin below 4%. Overall, during the period comprised between 1997 and 2003, British Gas does not succeed in kick-starting a positive operational profit on its retail electricity supply activities (see table 2).

This initial period, characterised by a strong investment phase, leading to recurring operational losses, had been anticipated by the firm, which had given itself up until 2000 to begin making progress in terms of producing positive results (Centrica 2001). British Gas barely manages to stick to this timetable; all in all it takes an additional year (compared to its original plans) to produce its first operating profit in this business segment (residential electricity supply).

What about the other firms operating in the market ? Due to a lack of specific business segment data, it is difficult to determine with the same precision the investments that they have realized and the results that they subsequently had during this period¹⁵. We can nevertheless guess that the expenses incurred by British Gas in terms of advertisement and direct marketing should have incited them do likewise in order to limit or compensate their resulting client loses. The pricing policies of British Gas, likewise, had an impact on the pricing policies of its main competitors. The latter begin to use British Gas prices' as *de facto price ceilings* not to be surpassed in the zones where they operate (excluding their historical operating zones).

Table 2

British Gas retail electricity sales figures in GB (1997-2003)						
	1997	1998	1999	2000	2001	2002
Client numbers (000)	0	850	2100	4000	5400	5795
Market share (%)	0%	3%	8%	15%	20%	22%
Turnover	0	5	240	792	1242	1380
Distribution and wholesale costs	0	5	156	686	972	1 037
Operational costs	25	85	84	106	270	289
Operational profit	-25	-85	-156	-107	40	54
Gross margin	n.s.	n.s.	-65%	-13,5%	3,2%	3,9%
						3,5%

In £million. For 2001, 2002 and 2003, data regarding operating costs, operational profit and gross margin are estimates (Morgan Stanley's). Sources: Centrica annual reports, Morgan Stanley, 2007.

Beyond that level, the offers of British Gas' competitors are fixed at levels that are influenced by geographical targets regarding market extension that its competitors have chosen, by their profitability goals and by their sourcing costs. British Gas' competitors rely to a large extent on self-supply given their portfolio of physical assets. In any case, the '*investing*' strategy undertaken by British Gas ends up having repercussions on the decisions taken by the former electricity distribution-supply monopolies. However, the latter were able to benefit from factors that softened the impact of British Gas' strategy (captive clients in their historical operating zones and partial or total physical coverage of their downstream retail requirements). These factors helped to limit the financial impact of British Gas' *investing* strategy phase on the historical operators.

The firms competing in the British market try thus to attract clients that have become less mobile due to existence of switching costs by clipping their margins. The Main new entrant, British Gas, plays a central role, in this 'investment' strategy, which we have chosen to analyze during the period 1997-2003 for the following reasons: 2003 marks the high watermark in terms of clients targeted by door-to-door salesmen and the highpoint for British Gas in terms of new client gains. In 2004 a new approach begins, characterised by important changes in terms of price offers, which could be considered as attempts to execute a 'harvesting' strategy.

British Gas : an aborted harvesting attempt

Beginning in 2004 and up to 2006, British Gas will try to secure a return from its initial investments. The group modifies its pricing policy for its residential customers. In 2004, British Gas announces a price increase above and beyond that of all its other competitors. The group passes on, in a more aggressive manner, the rise in wholesale prices to its final customers. The rise in wholesale prices having forced a rise in British Gas' sourcing costs. In 2005, British Gas is one of the last companies to announce a rise in prices, but when it does so, it is one of the highest. Finally in 2006, once again it is one of the suppliers that increases the most its retail prices. Overall taking into account the date of the announcement and the level of price rises, British Gas is the supplier that increased the most its prices between 2004 and 2006 (House of Commons, 2008).

2004	Price rise	Date
EDF Energy	3.8%	13th September
British Gas	9.4%	20th September
Npower	7.6%	1st October
Scottish Power	8%	4th October
Powergen et SSE	0%	
2005	Price rise	Date
EDF Energy	5.4%	17th January
SSE	6.7%	1st March
EDF Energy	12%	5th August
Powergen	11.9%	31st August
British Gas	14.2%	19th September
Scottish Power	12%	17th October
Npower	0%	
2006	Price rise	Date
Npower	12%	1st January
SSE	8.9%	1st January
Scottish Power	8%	1st March
British Gas	22%	1st March
Powergen	18.4%	10th March
EDF Energy	4.7%	13th March
Npower	13.4%	31st March
SSE	9.4%	1st May
Scottish Power	10%	1st July
EDF Energy	8%	31st July
Powergen	9.7%	21st August
British Gas	9.4%	1st September
Npower	9.9%	1st October

Source : Energywatch in House of Commons, 2008.

Table 3

A failed gamble. British Gas takes the gamble that it will be able to increase its prices faster than its competitors in order to re-establish its operational margins, without at the same time, loosing an important number of clients. British Gas being well established within the electricity supply segment, with a strong notoriety and selling both gas and electricity; the firm considers that it has the necessary trump cards to retain its clients and to render them less susceptible to switch in the face of the price gaps that begin to emerge between British Gas and its main competitors. Yet the result is precisely the opposite. British Gas begins to loose market share rapidly starting in 2004 and running up to 2006. Overall the net balance of customer gains and losses during this 3-year period is negative,

resulting in a net loss of 430,000 customers, that is to say 7% of its end of 2003 client base. To compensate for these losses British Gas decides to intensify its advertising campaigns and its direct marketing efforts. This raises British Gas operational costs (acquisition costs and the costs of managing new customer accounts). This in turn ends up weighing heavily on margins, which do not manage to recover in spite of the rise in prices (see table 4). Finally, the trend is reversed and in 2007 the group decides to offer price reductions of greater magnitude than those of its main competitors in a context of falling wholesale prices¹⁶.

In the end British Gas does not manage to meet its operating margin targets: the company's residential electricity supply financial results remain weak and below its long-term objectives (operating margin within a 5-10% range). Since 1997, the leader of the British market only manages to make a cumulative operating profit of slightly over £50 million, eventually it takes the firm 7 years (2001-2007) to cover its initial losses from entering the electricity supply market (1997-2000).

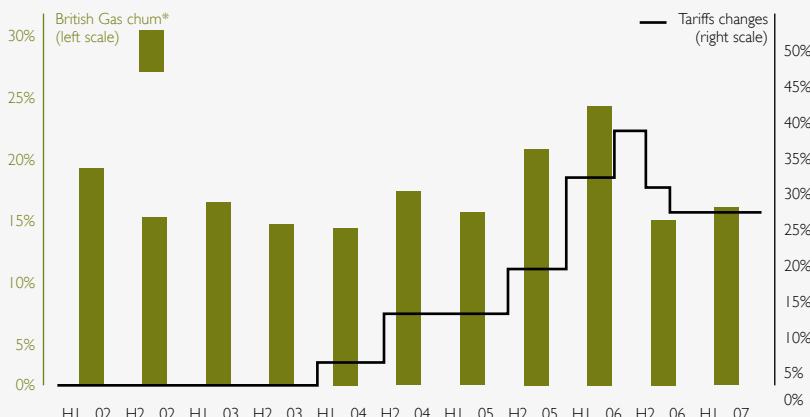
How to explain this phenomenon? The customers attracted by British Gas were by definition mobile (switchers): having switched suppliers at least once when quitting their historical electricity supplier. These clients are generally well informed regarding prices, offers and the gaps that may arise between the prices of different competing suppliers. They react accordingly exercising their freedom of choice. British Gas' clients switching rates remain high, and tends to increase during periods of price rises, when the price differential –perceived or real– between the prices offered by the firm and those of its competitors

Table 4

British Gas retail electricity sales figures in GB (2004-2007)

	2004	2005	2006	2007
Client numbers (000)	5 950	5 920	5 759	6 019
Market share (%)	23%	23%	21%	22%
Turnover	1731	1836	2 280	2 161
Sourcing and distribution costs (energy)	1 385	1 474	1 825	1 700
Operating costs	346	340	378	350
Operating Profits	68	22	77	111
Gross margin	3,9%	1,2%	3,4%	5,1%

Figures in £ millions. The data regarding operating costs, operational profit and gross margin are estimates (Morgan Stanley's). Sources: Centrica, Annual reports, Morgan Stanley, 2007.



* Cumulative number of losses and gains within British Gas' client base related to the group's total.

Sources : Centrica, Annual reports, Morgan Stanley, 2007.

Figure 1

increases (see figure 1). The behaviour of British Gas' clients did not change as desired by the firm: they did not become less mobile as time went by or more loyal.

This forces British Gas to have to align closely its tariff to changes in wholesale electricity prices and at the same time to continuously spend money on advertising and publicity campaigns in order to attract new customers in order to compensate for its customer losses. This makes the firm's costs structure increase and reduces its opportunities to bring down its charges in order to improve its operating margins. The company is nonetheless able to use certain tools at its disposal in order to try to improve this situation: by trying to reduce operational costs tied to client relationship management (outsourcing, reducing payroll, better management of IT systems

- 16** In 2007, the group succeeds in increasing its gross margins by choosing strategically the timing of its announcements related to price changes. The group benefits from a fall in wholesale prices –beginning of 2007- which it will not pass-on to its clients until the March. The group manages to gather profits and then significantly lowers its prices (more so than its competitors) which helps it to gain clients. This kind of situation is difficult to replicate following a period of price rises. Customers seem more inclined to change supplier following a price rise rather than following a price fall (they anticipate a reaction from their existing supplier and prefer to wait rather than switch). Aside from this during the 1st semester of 2008, characterised by a new episode of price rises, British Gas' operational margins on its retail sales return to more modest levels (around 3%).

and billing systems, etc.), the timing of announcements related to tariff changes, the introduction of capped price offers over a several year time horizon, etc. Measures that in essence seek to reduce churn rates amongst clients by offering them protection against a possible continuation in the rise of wholesale energy prices. (OFGEM, 2008).

On the other hand, British Gas's competitors are in a better position. They have at their disposal a historical client base, these 'historical' clients are generally less prone to switch and equally less sensitive to price changes. Furthermore, the competing firms, own more physical electricity production assets (generation capacity) that in turn render them less dependant on wholesale markets. They thus have larger room for manoeuvre when choosing their pricing policy, equally they more flexibility when trying to sustain or re-establish their operating margins.

Learning effects and endogenous switching costs. The difficulties encountered by British Gas in moving from an 'investing' phase into a 'harvesting' phase can be mostly explained by the behaviour of its client base. Switching costs are neither constant nor static. Once suppliers enter into the competitive game, their actions have an effect both on the level and the nature of certain components of switching costs. On the one hand, advertisement expenditures and direct marketing campaigns considerably and permanently reduce consumers' information and research costs. Consumers who have changed suppliers, at least once, maybe benefiting from learning effects. Their past experiences contribute to improving their knowledge and their information as well as their capacity to make an effective choice that corresponds to their preferences. This learning dynamic allows consumers to gain expertise and autonomy *vis-à-vis* suppliers. This ends up contributing to reduce the level of switching costs that consumers face and in turn favours their mobility¹⁷.

¹⁷ Nevertheless, not all the elements that make up switching costs are concerned. Transactional costs cannot be modified permanently either by operators or by customers switching experiences given that these are costs associated with entering or re-entering into the contractual relationship: they are thus recurrent. (Nilssen T., 2000). Only intervention by the public regulator can modify their level or their allocation amongst the different players.

New entrants are the most affected by this issue. They effectively attract those consumers that are the least risk averse and that have the lowest switching costs, and who consequently constitute the group that is the most sensitive to price changes. Clients that leave their historical supplier, reveal –through their behaviour- their propensity to be mobile and their risk taking profile. *A priori* they are the types of clients that will be the most difficult to retain and the most likely to set in motion learning dynamics.

This places new entrants in a more difficult situation compared to historical suppliers who dispose of a pre-existing client base. Additionally, suppliers have few tools with which to reinforce customer loyalty. Suppliers have not been able to effectively bundle the sale of high value added associated services to the sale of electricity. Doing so would have allowed them to (i) differentiate their offers in qualitative terms (ii) reduce the price elasticity of their customers. Whilst it maybe the case that retail competition may lead to the introduction of certain improvements (innovative contracts in terms of prices and tenure, dual fuel offers, energy efficiency advice, etc), it does not become automatically the source of radical innovation capable of enriching the content of the offers made by electricity suppliers. Suppliers have (so far) not been able to create strong customer loyalty via the proposal of new services or functionalities. What are the consequences of taking into account consumers 'learning effects' on the analysis of the functioning of the competitive dynamics in a market with switching costs?

- Switching costs are a strategic variable, not only for firms (who can raise them artificially in order to reduce mobility) but also for customers, who can reduce their level via dynamic learning effects. Consumers, or at least some consumers, are capable of making switching costs become an endogenous variable.
- This strategic variable (switching costs) will weigh more heavily on new entrants than on the incumbents. Incumbents benefit from having an existing client base whose behavioural characteristics are different. New entrants face two handicaps: they have additional costs to incur in order to acquire clients and these clients are harder to retain. If these constraints are known and observable they can have a negative effect not only on new

entrants expectations (regarding the potential opportunities in entering a market) but also it can have an effect on the expectations of incumbent firms (regarding the opportunities to compete aggressively with each other if it will be actually difficult to retain customers in the second period).

- The clients of existing (incumbent) firms are less mobile (less prone to switch) than those of new entrants, consumers have heterogeneous behaviour characteristics. This heterogeneity implies that new entrants have more difficulty in successfully price discriminating between those customers who make part of their existing client base and those customers that they wish to attract.

Imagine the case of Firm A, who is unable to price discriminate it must make an arbitrage decision between the gains to be made from acquiring new clients and the gains to be foregone due to the reductions in price that it will have to extend to its existing client base (due to it being unable to price discriminate). This arbitrage will be influenced by Firm A's market share: the bigger it is the lower the net gain from executing an aggressive acquisition strategy. This effect maybe reinforced by two factors: a) can Firm's A competitor price discriminate?, b) is the market 'mature' i.e. the number of clients not attached to an existing supplier is limited.

In such a situation Firm A, has no incentives to try to establish an important client base in the first period (and if Firm A is a new entrant it has no incentives to enter the market), this is because Firm A will be aware that in the second period it will not be able to price discriminate (and therefore it will not be able to recoup its investment efforts from the first period), whilst at the same time it is aware that its competitors are able to do so. In this way, one can arrive to a conclusion which is in essence opposed to that derived from the works of J. Farrell and C. Shapiro (1988).

Taking into consideration the learning effects of a segment of consumers creates a degree of heterogeneity in the choices faced by competing firms. Certain firms are able to price discriminate; others are not able to do so. This introduces a bias in favour of the incumbent firms and it tends to reduce the incentives that firms have to compete with each other to acquire a set of customers that in the end risks being (by its very nature) difficult to retain.

Conclusion

The analysis of the residential electricity supply market in the United Kingdom has allowed us to identify the important role played by consumers learning effects on the strategies executed by the firms competing in the market. Taking into account the heterogeneous nature of consumers' behaviour can enrich the conclusions derived from theoretical models that analyse competitive dynamics in the presence of switching costs. In this particular case study, it seems that the aborted attempt by British Gas to move into a 'harvesting' phase could have an important impact on the market equilibrium outcome. It may discourage new entry and additionally it may reduce suppliers incentives to compete with each other via prices.

References

- Centrica (2001), *Annual report 2000*, London
- Chen, Y. (1997), "Paying consumers to switch", *Journal of Economics and Management Strategy*, vol. 6
- Diamond P. (1971), "A model of price adjustment", *Journal of Economic Theory*, vol. 3
- Defeuilley C. (2009), "Retail competition in electricity markets", *Energy Policy*, vol. 36
- Electricity Association, (2003), *Electricity Industry Review 7*
- Electricity Association, (2004), *Electricity Industry Review 8*
- Electricity Association, (2005), *Electricity Industry Review 9*
- Farrell J., Shapiro C. (1988), "Dynamic competition with switching costs," *Rand Journal of Economics*, vol. 19
- Farrell J., Klemperer P. (2007), "Coordination and Lock-In: Competition with Switching Costs and Network Effects", in M. Armstrong and R. Porter (eds.), *Handbook of Industrial Organization*, North-Holland
- Fudenberg D., Tirole, J. (1984), "The Fat-cat effect, the puppy-dog ploy, and the lean and hungry look," *American Economic Review*, vol. 74
- Gabszewicz J., Papall L., Thisse J.F. (1992), "Sequential Entry with Brand Loyalty Caused by Consumer Learning-by-Using", *The Journal of Industrial Economics*, vol. 40
- Helm D. (2004), *Energy, the State and the market. British energy policy since 1979*, Oxford University Press, Oxford
- House of Commons (2008), *Energy prices, fuel poverty and OFGEM*, Business and Enterprise Committee, vol. 1 et 2
- Klemperer P. (1987a), "Markets with consumer switching costs", *The Quarterly Journal of Economics*, vol. 102
- Klemperer P. (1987b), "The competitiveness of markets with switching costs", *Rand Journal of Economics*, vol. 18

- Klemperer P. (1995), "Competition when consumers have switching costs : an overview with application to industrial organization, macro economics and international trade", *Review of Economic Studies*, vol. 62
- Littlechild S. (2005), *Smaller suppliers in the UK domestic electricity market : experience, concerns and policy recommendations*, Ofgem report, London
- McCloskey D. (1998), *The rhetoric of economics*, Madison, The Wisconsin Universiy Press
- Molland M., (2007), *Théorie des coûts de changement et analyse de la concurrence de détail dans l'électricité. Le cas britannique*. Thèse de doctorat, Université Paris-Sud 11
- Morgan Stanley (2007), *Centrica Business Changing Significantly - Justifies a More Positive Stance*, Morgan Stanley Research Europe, London, june 26.
- Nilssen T. (1992), "Two kinds of consumer switching costs", *Rand Journal of Economics*, vol. 23
- Nilssen T. (2000), "Consumer lock-in with asymmetric information", *International Journal of Industrial Organization*, Vol. 18
- OFGEM (2004), *Domestic Competitive Market Review : A Review Document*, OFGEM, London
- OFGEM, (2007), *Domestic Retail Market Report*, OFGEM, London
- OFGEM (2008), *Energy Supply Probe - Initial Findings Report*, OFGEM, London
- OFT (2003), *Switching costs*, Economic Discussion Paper 5, NERA report, OFT, London
- Padilla A. (1992), "Mixed pricing in oligopoly with consumer switching costs," *International Journal of Industrial Organization*, vol. 10
- Riordan M. and Sappington D. (1989), "Second sourcing," *Rand Journal of Economics*, vol. 20
- Salop S. (1977), "The noisy monopolist : imperfect information, price dispersion and price discrimination", *Review of Economic Studies*, vol. 44
- Shaffer G., Zhang Z. (2000), "Pay to switch or pay to stay: preference-based price discrimination in markets with switching costs," *Journal of Economics and Management Strategy*, vol. 9
- Surrey J. (ed.) (1996), *The British electricity experiment. Privatization : the record, the issues, the lessons*, Earthscan Publications, London
- Taylor, C. (2003), "Supplier surfing: Competition and consumer behaviour in subscription markets", *The RAND Journal of Economics*, vol. 34
- Tirole J. (1988), *The theory of industrial organization*, MIT Press, Cambridge, Massachusetts
- Thomas S. (2003), "Gas as a commodity. The UK gas market : from nationalism to the embrace of free market", in Arentsen M, Kunneke R. (ed.), *National reforms in European gas*, Elsevier, Oxford
- Waddams-Price C., (2004), "Spoilt for Choice? The Costs and Benefits of Opening UK Residential Energy Markets", University of Berkeley – CSEM WP n°123

Retailers' risk management and vertical arrangements in electricity markets

Raphaël H. Boroumand
LARSEN

Georg Zachmann
LARSEN

The failure of the asset light retailer's economic model is indicative of the incapacity of this organizational structure to manage efficiently the combination of sourcing and market risks in the current market environment. Because of the structural dimensions of electricity's market risks, a retailer's level of risk exposure is unknown ex ante and will only be revealed ex post when consumption is known. In contrast to the "textbook model" of electricity reforms, the paper demonstrates that in the current market context pure portfolios of contracts are incomplete risk management instruments compared to physical hedging. The latter is critical to overcome the asset light retailer's curse.

Introduction

The 'ideal textbook' model of competitive decentralized electricity markets required the vertical separation of generation, retail as well as network services (transmission and distribution). Introducing competition at the retail level was thought to imply the emergence and development of 'asset light retailers' who neither own generating nor distribution assets. By offering innovative retail contracts with attractive prices to electricity consumers those retailers were expected to stimulate a fierce competition in the retailing segment of the value chain (Hunt, 2002, Hunt and Shuttleworth, 1997). In stark contrast to this theoretical vision, asset light retail entry has never eventuated as expected. Asset light retailers bankrupted, left the market, were taken over, or evolved towards an upstream integration in all the retail markets opened to competition (UK, New Zealand, Australia, France,...). Even in the UK, presented as a benchmark for electricity deregulation (Thomas, 2006), twenty new entrants left the retail market since 2000 (Oxera, 2008). At their climax between 1999 and 2001, the total market shares of the new entrants into retail was less than 2% in the UK (Ofgem, 2007) despite high levels of net switching.¹

By studying the risk management constraints of a retailer, this paper explains why the asset light organizational model is not sustainable in decentralized electricity markets. We argue that in contrast to physical assets, purely contractual portfolios are not efficient risk management devices² for hedging uncertain delivery obligations of retailers. That is, the paper aims at demonstrating the critical necessity to manage electricity market risks through a *combination* of contractual and physical assets. Comparing the determinants, costs, and benefits of different institutional arrangements has been a strong focus in the last three decades of the New Institutional Economics framework and more specifically of the literature on Transaction Costs Economics (Coase, 1960, Williamson, 1985, Joskow, 1985, Shelanski and Klein,

¹ 40% of net switching for the residential segment in the UK in 2007 (Ofgem, 2007)

² Risk Management embodies the process and the tools used for evaluating, measuring and managing the market risks within a retailer's portfolio of contracts and plants. The value of energy trades change over time as market conditions and underlying price variables change. In electricity markets, effective risk management depends not only upon proper portfolio analysis tools but also on a solid forecasting of forward prices.

1995, Coeurderoy and al, 1997, Whinston, 2003). However, the originality of our paper is to compare vertical arrangements through the analytical lens of risk management (rather than transaction costs economics) taking the new perspective of an electricity retailer specific's intermediation function.

The paper is organized as follows. In the first section we put forward the market risks faced by a retailer. The following section demonstrates the limits of pure contractual hedging in liberalized electricity markets. Then we analyse the comparative benefits of physical hedging for managing risks. The following section is devoted to comparing quantitatively the risk profiles of different portfolios' structures. The last section concludes.

The market risks faced by retailers

By sourcing electricity for resale to final consumers, retailers are market intermediaries (Spulber, 1999). As market intermediaries, retailers have the contractual responsibility to balance on a real time basis their upstream and downstream portfolios of electricity. This real-time matching function exposes them to quantity and price risks.

An electricity retailer is specifically exposed to a quantity risk over a short term horizon (from a few days, a few hours, to real time risk exposure) due to unanticipated load variations, (e.g. related to the imperfect predictability of weather conditions).³ As demand for electricity is of stochastic nature, very inelastic, and characterized by strong short term variability and supply is rigid on the short term, spot prices are volatile (Stoft, 2002, Geman, 2005). Since electricity is not economically storable, all imbalances will have to be settled on the spot market at unforeseeable prices.⁴ This non storability exacerbates the consequences arising from the classical matching uncertainty problem between any market intermediary's sourcing (downstream) and selling (upstream) portfolios (Bailey, 1998, Gehrig, 1993, Hackett,

³ Weather uncertainty can be theoretically mitigated through *weather derivatives*. However, due to their speculative and illiquid features, difficulty of pricing, and lack of liquidity, *weather derivatives* are very seldom used by electricity retailers (Geman, 2005).

⁴ For further details on the links between the storage's level and price volatility of a commodity, please refer to Geman, 2005 and Wirsing, 1949

1992, Spulber; 1999). Furthermore, the positive correlation between price and demand in electricity wholesale markets (Stoft, 2002, Chao et al, 2005), is worsening the financial costs of any under-contracted positions settlement on the spot market. This load/price positive correlation is seen as an important incentive for contractual hedging (Mackay and Moeller; 2007). Finally, any change on the number of a retailer's customers (loss/gain of market shares in newly liberalized retail segments) will generate vertical imbalances forcing the retailer to sell or buy any over or under-contracted position at uncertain prices. These structural imbalances put retailers' margins under threat. Consequently, quantity risks systematically translate into price risks. The price risk is generated by the discrepancies between the selling price of electricity on the retail market (generally a fixed price) and the price of the complementary spot transactions to offset the structural disequilibria between a retailer sourcing and selling portfolios. To minimize quantity risk and price risk, retailers will aim at contractually hedging the main proportion of their aggregated load requirements through the purchase of hourly electricity blocks with a minimum physical capacity of 1 MW. Each individual demand being stochastic, retailers will define their contractual level of procurement based upon the imperfect segmentation method known as *load profiling*.

The expected role of long-term contracts and their structural limits for managing market risks

In the ideal theoretical paradigm of liberalized electricity markets, financial contracts⁵ (*forward* and *futures*) were predicted to be efficient instruments for managing quantity and price risks and assumed to be perfect substitutes to physical assets (Chao and Huntington, 1998, Hunt, 2002). These expectations have not been reached. In all electricity markets, *forward* and *futures* cover a limited fraction of the total demand (Anderson and al, 2007; Chao and al, 2008). Meanwhile, upstream vertical integration has been maintained or has arisen in

⁵ For the purpose of this paper, it is not necessary to distinguish between financial and physical contracts since potential physical delivery is not changing anything to our conclusion.

most of the markets (Cornwall, 2008, Mansur and Saravia, 2007, Kuhn and Machado, 2004). In this section, we examine the structural limits of contractual hedging within liberalized electricity markets. These limits stems from contractual parties' misalignment of interests regarding price and quantity provisions.⁶

We first analyze the origins of price provisions' misalignment. Spot price fluctuations have opposite effects on retailers and generators profits as any increase in the spot price will affect positively the revenue of the producer to the detriment of the retailer.⁷ As the price risk's profiles of retailers and producers are negatively correlated, long-term fixed price sourcing contracts should, in principle, credibly align long term hedging needs of both parties, and safeguard their economic interests for the best of the contract's performance.⁸ However, in a setting of fixed price contracts, the *ex post* distribution of risks across the parties depends on the duration and magnitude of the periods during which the spot price will be above/below the contractual fixed price, which in the electricity sector is not foreseeable given the proven incapacity of current price forecasting models to capture real electricity price volatility within a very uncertain market environment (Szkuta et al, 1999, Nogales et al, 2002, Lora et al, 2002, Geman, 2005).

As sustainable periods of spot prices below the contract fixed price may induce profitable new entries into the retail market, retailers with significant level of sourcing through fixed-price contracts are exposed to a risk of *price-squeeze*.⁹ Given the low level of entry costs into the retail segment (Ofgem, 2007); low spot prices will represent strategic opportunities for potential new entrants to corner market shares from any retailer locked into high sourcing fixed price contract. Therefore, any fixed price contract gives rise to an opportunity cost

⁶ Or; the main role of long term contract is precisely to align parties' interests to ensure the contract execution's strong compliance (Masten 1996, Brousseau and Glachant, 2002).

⁷ Conversely, any spot price decrease will have opposite effects.

⁸ Moreover by aggregating numerous customer loads, retailers would theoretically be able to lock in the major parts of their expected demand through sourcing contracts which would match the profile and risk features of their downstream portfolio at prices not tied to the volatile spot price.

⁹ For an example of price squeeze, see TXU Europe bankruptcy in 2001 on the UK market (*Power in Europe*, December 2004)

¹⁰ Sourcing contracts indexed to the spot price should theoretically eliminate the aforementioned opportunity cost for retailers. However, in practice, a residual price risk remains due to the smoothing effect embedded in those spot indexed contracts.

for retailers.¹⁰ When contracting, retailers have no guarantee *ex ante* that spot market sourcing will be more costly than contractual sourcing. The effectiveness of any hedging strategy is revealed only *ex post*, when demand is settled on the spot market. A retailer could always diminish its retail prices to match the competitors' prices and protect its market shares but will then face retail margin's reduction. In such setting, another option for retailers would be to renege on their contractual engagements, giving rise to a classical *hold-up* problem that leave generators with stranded long term investments (Klein et al, 1978). Anticipating this risk of opportunism, generators would require a higher contractual premium, making contracts more expensive for retailers.¹¹ The aforementioned *ex ante* uncertainty on a fair contractual risk sharing *ex post* across the parties hinders the incentives for a fixed price long term contract. Absent long term alignment of parties' interests, such contract is not *self enforceable* (Klein, 2000 and 2002).

Provision's misalignments for quantity preferences are also a source of contractual problems and high maladaptation costs (Williamson, 1985). Indeed, retailers prefer to contract on flexible quantities to match their variable load and demand swings whereas generators prefer to secure their revenues through fixed quantities but with contracts that follow their supply availability patterns (i.e. : including technical maintenance and/or fuel shortage clauses). Retailers would not accept such availability clauses unless they have concluded load curtailing contracts downstream (i.e. *interruptible contract*¹²). These quantity and availability clauses' misalignments contribute to a structural lack of matching between capacity and load and to high contractual premia for conceding contractual preferences.¹³

In the absence of any mutually beneficial price/quantity contractual mix to maximize long term revenues of both parties, the limits of

¹⁰ The perceived asset light retailer's counterparty risk of default (which differs from the risk of opportunistic behavior) also contributes to high contractual premia. Electricity retailing is a financial rather than physical activity. As a retailer has no tangible assets (the portfolio of customers is an intangible asset), this represents a source of financial instability for contractual counterparties.

¹¹ Interruptible contracts are equivalent to a *callable forward*. In such contractual framework, the retailer simultaneously (and virtually) sells a *forward* contract to its clients and buy a *call option* from them.

¹³ Wolak (2007) analyses the benefit for a generator to accept a lower output price in exchange of flexible output.

symmetric fixed (or indexed) price contract make asymmetric contracts such as *call options*¹⁴ potentially more suitable for electricity retailers. Indeed, by giving the right to buy electricity at the strike price (Hull, 2005), *call options* enable retailers to mitigate their quantity risk (Boroumand, 2008). With the structural limits of contractual hedging in mind, we now turn to the risk management benefits of a portfolio strategy based upon physical assets.

The merits of upstream integration¹⁵

Three groups of arguments explain the superior performance of physical hedging over contractual hedging in the current market environment. In contrast to contractual hedging, physical hedging restores symmetry between upstream and downstream in terms of risk exposure (Aid et al, 2007). The asymmetric risk exposure between the upstream and downstream segments stems from two specific phenomena inherent to electricity markets. First, retailers define their market shares strategies and corresponding contractual hedging positions under uncertainty at $t=0$, whereas by producing only at $t=1$ ¹⁶, when demand for electricity is revealed, producers face no uncertainty in their generation decisions. Second, due to the specifically high demand/price inelasticity of electricity, the revenues of non integrated retailers depend on both wholesale and retail prices whereas the producers' revenues only depend upon the wholesale price (as demand is inelastic to price, which is not generally the case in other industries). Consequently, retailers are more exposed to market risks. With such asymmetry in risk exposure and therefore in incentives for risk hedging, contractual risk sharing can only be imperfect. Part of the risk remains structurally unhedged for the retailer.

¹⁴ A *call option* is a contract that gives its purchaser the right (but not the obligation) to buy the underlying commodity at a certain price, the strike (noted X in section 5), on or before an agreed date, the maturity of the option (Geman,2005)

¹⁵ We assume that vertical upstream integration is neither constrained by capital raising issues nor scale differences between generation and retail.

¹⁶ Electricity being economically not storable

Price risk is reduced by the flexibility offered by some plants with short start up and ramp up times to respond to unforeseeable demand through their capacity to produce varying quantities of electricity to match load structural swings rather than being constrained through rigid contractual fixed quantities. Even if the integrated firm will continue to face uncorrelated supply and demand, there is no more opposition of interests as for contractual quantity clauses. The option to produce *only if* the spot price is above the plant marginal cost offers an operational flexibility that is not duplicable through standard forward contracts. Electricity plants are then assimilated to “*operational option*”. An integrated retailer can arbitrate between an internal sourcing and transactions on the spot market based upon the relative economic attractiveness of each sourcing mechanism. By internalizing costs and risks (Finon, 2008, Hansmann, 1996) integrated retailers are in position to maximize and balance their overall profit in relation to the risk profiles and margins' levels of their downstream and upstream segments.¹⁷ The exposure to volatile electricity spot price and the risk to be price squeezed do not disappear but are significantly reduced since price fluctuations that disadvantage the downstream segment will benefit to the upstream one (and vice versa). The effect on the total retailer's profit is neutralized.

The combined impacts of reducing the integrated retailer's exposure to spot price risk, enhancing its ability to balance more effectively capacity and load's features prove that physical hedging is a more efficient and sustainable¹⁸ risk management instrument than contracting.¹⁹ Physical hedging is even more relevant for retailers with a strong base of residential clients. Indeed, in contrast to the industrial segment that may be proposed indexed retail contracts or interruptible contracts that transfer some of the risks to the customers, prices paid by domestic customers do not reflect the gap

¹⁷ Analyzing the margins of Centrica UK and E.ON UK provide clear evidence for this type of vertical arbitrage (Wright, 2007).

¹⁸ Electricity contracts are usually for very limited periods (less than 3 years, Anderson et al (2006)). Beyond this duration, non integrated retailers face renegotiations costs and risk of opportunism not shared by integrated ones.

¹⁹ Since our paper compares two mechanisms of vertical relations (ie contracts and upstream vertical integration) on a risk management perspective, costs and rigidity of internal organization are not discussed. Please refer to Joskow, 2005, and Williamson, 1971, 1985 for details on the costs of vertical integration.

between supply and demand. Domestic customers are therefore not sensitive at all to intraday price fluctuations.

Portfolios risk management quantitative analysis

In this section we demonstrate that a retailer cannot reproduce the risk-reducing benefits of physical hedging by pure contractual portfolios. We demonstrate that this is the result of the non-availability of certain types of contracts. A retailer is assumed to have concluded a retail contract (the retail contract is given *ex ante* and is therefore not a portfolio's parameter of choice) with its customers that imply stochastic demand \tilde{V}_t (for $t=1:T$). The demand distribution is known to the retailer and the uncertainty about the actual demand \tilde{V}_t is completely resolved in time t .

To fulfill its retail commitments the retailer can buy electricity on the spot market at the *ex ante* uncertain spot market price P_r .²⁰ The spot market price distribution is known by the retailer. To reduce its risk from buying an uncertain amount of electricity at an uncertain price, the retailer can conclude financial contracts and/or acquire physical generation assets. All contracts (including the retail contract and the physical assets generation volumes) are settled on the spot market that is assumed to be perfectly liquid. Thus, the payoff streams depend on a given number of spot market realizations (one year, i.e., 8760 hours). For example, an annual baseload forward contract implies buying the agreed volume of electricity at the contractual price for 8760 hours.

In Table I five different contracts/assets – namely a retail contract, a forward contract, a power plant, a call option on the spot price and a put option on the spot price²¹ given the spot price – are introduced. If for example, the electricity spot price (P_t) is above the strike price of the options (X), there is a positive payoff of the call option,

²⁰ We ignore balancing markets. This can be justified by the fact that most of the adjustments of retailers take place in the day ahead market.

²¹ A *put option* on the spot price, gives the retailer the right to sell electricity on the spot market at a given price.

Contract	Payoff
Retail Contract	$\pi_{retail} = -\tilde{P}_t \times \tilde{V}_t + E(\tilde{P}_t \times \tilde{V}_t)$
Forward	$\pi_{forward,t} = V_{forward} \times \tilde{P}_t - E(V_{forward} \times \tilde{P}_t)$
Power Plant	$\pi_{plant,t} = V_{plant} \times \max(\tilde{P}_t - mc, 0) - E(V_{plant} \times \max(\tilde{P}_t - mc, 0))$
Call Option on Spot	$\pi_{call,t} = V_{call} \times \max(\tilde{P}_t - X, 0) - E(V_{call} \times \max(\tilde{P}_t - X, 0))$
Put Option on Spot	$\pi_{put,t} = V_{put} \times \max(X - \tilde{P}_t, 0) - E(V_{put} \times \max(X - \tilde{P}_t, 0))$

Table 1

while the payoff of the put option is zero. The payoff of the power plant, depends on the installed capacity of the plant (V_{plant}) and its marginal cost (mc) and only the payoff of the retail contract depends on the stochastic demand (\tilde{V}_t). By subtracting the expected value ($E(\cdot)$) from the gross payoff all contracts/assets are assumed to have zero expected value. That is, we assume that in a perfect market (no market power, no transaction costs, full transparency, etc.) arbitrage would not allow for the existence of systematic profits.²²

To simulate the payoffs certain assumptions on the distribution of the electricity spot price and retail volume have to be made. We rely on real data of the French electricity market from 2006 and 2007. The hourly prices are obtained from the French electricity exchange Powernext and the corresponding loads are obtained from the network operator RTE. Electricity prices depend non-linearly on the total load (see Figure 1). Thus, load and prices are strongly (although not perfectly) correlated (46% in the sample period) and load increases have a stronger impact on prices than load decreases. To obtain realistic simulations we sort the observed price-load combinations by load. Then, the central points (medians) of 3000 windows of 8760 neighboring observations are drawn from a truncated normal distribution.²³ Note that, due to the normal distribution, windows with a median load closer to that of the observed sample are more

²² Without this assumption, the hypothesis concerning the valuation of contracts and assets, that are not in the focus of this paper, would drive the results.

²³ The mean of this distribution is 8760, representing the central point of the 2 years data. The variance of the central points is $8760/4)^2$. The distribution is truncated below 8760/2 and above 17520-8760/2 to fit the data sample.

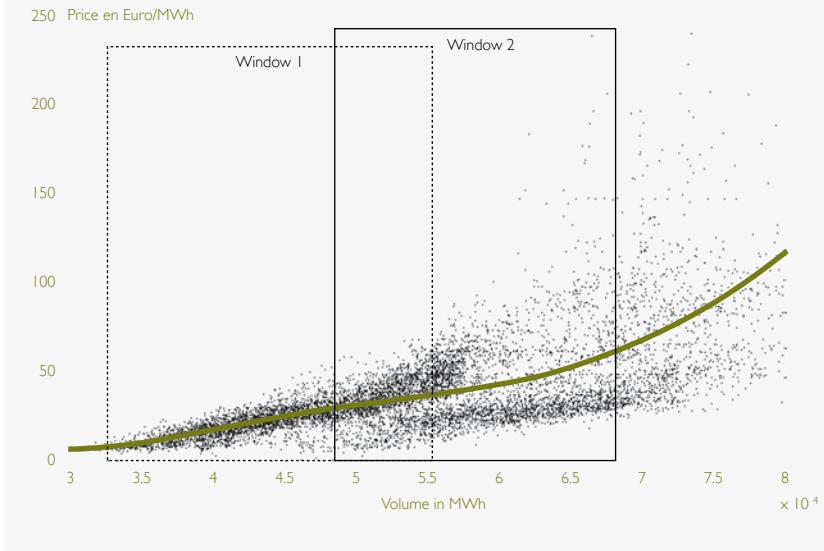


Figure 1

likely than windows with a median very different from that of the real data. Finally, from each of the 3000 windows we draw randomly with replacement 8760 hourly price-load combinations. Consequently, in expectation the median of the observed data (load) is equal to that of the simulated data.²⁴

The marginal generation cost of the power plant is set to the median of the simulated spot prices $mc=39.0$ Euro/MWh, thus representing a peak load power plant. The Strike price of the options is set to the expectation value of the spot price $X=E(P_t)=43.0$ Euro/MWh.²⁵

We can calculate the cumulated annual payoffs of the 8760 hourly

²⁴ Due to the non-normal (joint) distribution of the observed data, the mean of the simulated load is slightly lower (54 instead of 55 GW) than that of the observed loads in 2006-2007. The mean price of the simulated data is slightly lower than that of the observed data (43 instead of 45 Euro/MWh) and the median of the simulated prices is higher (39 instead of 38 Euro/MWh). The variance of the mean (median) price across the 3000 simulations is 29 (20).

²⁵ This is done to make call options and power plants distinguishable as they are equivalent according to Table I if . The intuition of setting the marginal cost to the median price is that thus, the power plant will run exactly 50% of the times. The intuition of setting the strike price to the mean price is that the option is “at the money” in this case.

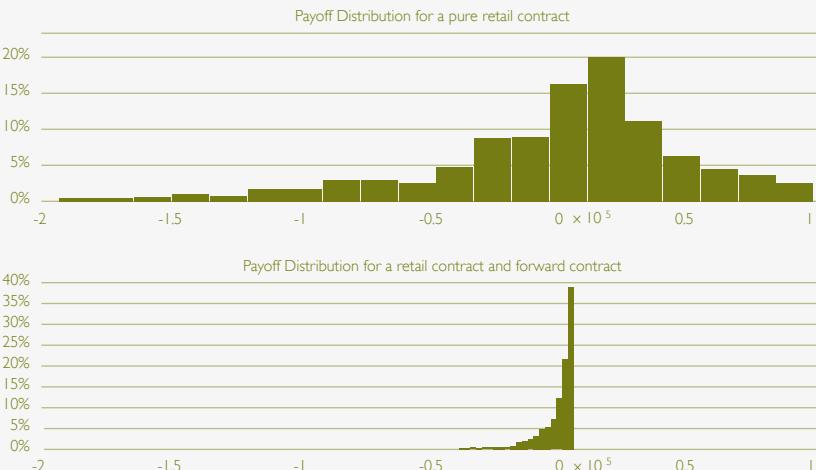
price/volume combinations for all 3000 simulations given the portfolio ($V_{forward}$, V_{plant} , V_{call} , and P_{put}):

$$\begin{aligned}\pi^i = \sum_{t=1}^{8760} & [\pi_{retail,t}(\tilde{P}_t^i, \tilde{V}_t^i)] + [V_{forward} \times \pi_{forward,t}(\tilde{P}_t^i)] + [V_{plant} \times \pi_{plant,t}(\tilde{P}_t^i, mc)] \\ & + [V_{call} \times \pi_{call,t}(\tilde{P}_t^i, X)] + [V_{put} \times \pi_{put,t}(\tilde{P}_t^i, X)]\end{aligned}$$

Thus, π^i is the annual payoff of the i^{th} price and volume simulation given the portfolio defined by $V_{forward}$, V_{plant} , V_{call} , and P_{put} . Employing for example no contracts/assets to reduce the risk, the distribution of the 3000 payoffs from the retail contract stretches from -193,000 Euro in the worst case to 98,000 Euro in the best case. By adding just one forward contract to the portfolio, the risk might be significantly reduced. In the worst case the retailer now loses only -39,000 Euros (see Figure 2). By combining different contracts/assets with the retail contract, the risk could be further reduced. To measure risk, we rely on the Value-at-Risk because it is a good measure of the downside risk of a portfolio that is for example used as preferred criteria for market risk in the Basel II accord.²⁶ The Value-at-Risk for the 95% confidence interval ($VaR(95\%)$) that we use in the remainder of the paper is the one hundred fiftieth lowest of the 3000 payoffs.

Figure 2

Payoff distribution of two portfolios (in Euros)



#	Used assets	Retail				VaR(95%)
1	All contracts		-0.04	0.26	1.24	-0.27
2	without options		0.09	1.33	-	-
3	only options		-	-	1.47	-0.28
4	only forward		0.98	-	-	-
5	only power plant		-	1.46	-	-

Table 2

Using an optimization routine²⁷, the portfolio that produce the lowest *VaR(95%)* can be identified. The objective is to find the portfolio consisting of one 1 MWh baseload retail contract and a linear combination of financial contracts as well as physical assets that reduces the retailers risk. Thus, the factors for the other contracts/assets are also measured in MWh. If the retailer, for example, sold two retail contracts and he would like to hedge this deal with only forward contracts (compare #4 in Table 2), he would have to buy ($2 \times 0,98$ MWh) 1.96 MWh forwards.

The volume of power plant contracts is constrained to be positive, while call option, put option and forward contracts could be both bought and sold at the market (i.e., negative quantities are allowed). In six different scenarios we constrain the volume of certain contract types to zero. Thus, the (non-) substitutability of these contracts for hedging a retailers risk can be assessed.

The *VaR(95%)* of the considered one MWh baseload contract with zero expectation is -97,852 (see Table 5 in the Appendix), i.e., the probability of the retailer to lose more than 100,000 Euro in one year (or more than 10 Euro on average per hour) with this contract is almost five percent. The optimal portfolio if all assets are allowed (#1) produces a *VaR(95%)* of -2,088. The VaR of #1 is thus 98% lower than that of the retail contract without any hedge.²⁸

²⁶ The Value at Risk (VaR) is an aggregated measure of a portfolio total risk. VaR summarizes the expected maximum loss (worst loss) over a target horizon (one year in our paper) within a given confidence interval (generally 95%). Thus, it is measured in monetary units, Euros in our paper. For further details please refer to Ph.Jorion, 2000.

²⁷ We use the „fmincon“ routine in Matlab. As the routine does not necessarily converges for this non-linear problem (especially for the three and four assets case), we rerun the optimization for each case with 100 different randomly drawn starting values. The result of the best run can be considered sufficiently close to the global optimum, as all results tend to be within a fairly narrow range.

Without plants or forwards a $VaR(95\%)$ very close to that of the unconstrained optimal portfolio (#1) can be attained if options are allowed (#3). If options cannot be chosen, the risk management characteristics of #3 can be reproduced without options if power plants and forward contracts are allowed (#2). With only forward contracts allowed (#4), the $VaR(95\%)$ is more than six times bigger than if both, power plants and forward contracts are available portfolio choices (#2). Consequently, if options are no choice for retailers (because for example, nobody is willing to sell them as a counterparty), then power plants – whose payoffs feature option like characteristics – will help retailers to reduce their risk exposure.

If power plants with different marginal costs can be included in the portfolio, the selection decision equals the choice of hedging an underlying with options with different strike prices. If, for example, a low cost technology with marginal cost being equal to the 25 percent percentile of the electricity price and a high cost technology with marginal cost being equal to the 75 percent percentile are introduced, the VaR can be further reduced.

In our example the optimal portfolio (#6 in Table 3) that consists of 0.62 of the normal power plant, 0 of the cheap power plant, 0.59 of the expensive power plant and 0.3 of the forward contract can reduce the $VaR(95\%)$ to 2,112. This implies a slight improvement with respect to the optimal portfolio for the normal power plant and the forward contract ($VaR(95\%) = -2,131$). By allowing only power plants it can be demonstrated that adding a power plant with different payoff characteristics might reduce the VaR of the portfolio. Going for example from #8 ($V_{plant,25}$ and $V_{plant,75}$, $VaR = -2,199$) or #5 ($V_{plant,50}$, $VaR = -2,201$) to #7 (all three power plant types, $VaR = -2,141$) reduces the $VaR(95\%)$ by three percent.

Table 3

Portfolios containing one retail contract and different power plants that maximize the $VaR(95\%)$

#	Used assets	$V_{forward}$	$V_{plant,50}$	$V_{plant,25}$	$V_{plant,75}$	$VaR(95\%)$
6	Forward and 3 plants	0.30	0.62	0.00	0.59	-2,112
7	3 plants		1.17	0.22	0.03	-2,141
8	$V_{plant,25}$ and $V_{plant,75}$			0.76	0.75	-2,199
9	Forward and $V_{plant,75}$	0.50			1.09	-2,183

	Used assets	Retail	$V_{forward}$	V_{plant}	$Var(95\%)$
Constant demand	only forward				0
	only power plant				-4.77
	forward and power plant			0	0
Stochastic demand not correlated to the price	only forward		1.01		-2.56
	only power plant			1.18	-5.80
	forward and power plant		0.81	0.37	-2.05
Demand perfectly correlated to the price	only forward		1.01		-3.66
	only power plant			1.55	-2.89
	forward and power plant		0.46	1.08	-0.81

Table 4

In the above exercise we have shown, that forward contracts are not sufficient to hedge the supply obligations of a retailer. We demonstrated on the example of the French market, that either power plant shares or option contracts on the spot market are necessary to optimally reduce the risk of the portfolio. To understand why this option-like payoff structure is required we proceed with a stylized example. We assume a sinusoid price curve $p(x) = 30 + 15 \times \sin x$ and three different types of retail consumption: constant demand ($q(x) = 40$), stochastic demand ($q(x) = 30 + \varepsilon(x)$) and demand correlated to the price ($q(x) = 40 + 20 \times \sin x$).

The payoffs of a retail contract, a forward contract and a power plant are calculated according to the approach described above.²⁹ Finally, three different types of portfolios are considered: (1) the combination of one retail contract and the optimal number of forward contracts, (2) the combination of one retail contract and the optimal number of power plant shares and (3) the combination of one retail contract, the optimal number of power plant shares and the optimal number of forward contracts.

Assuming demand being constant, a retail contract might be completely hedged by exactly one forward contract (see Table 4 and Figure 5). This is obvious, as the payoffs of the forward contract

²⁸ This level of risk reduction can be approached with very different combinations of assets, as the four assets imply one degree of freedom.

²⁹ As a reminder: the payoff of the assets is computed according to the corresponding formulation in Table 1. Thereby, the marginal cost of the power plant are set at the median spot price and the expectation value of each payoff series is normalized to zero.

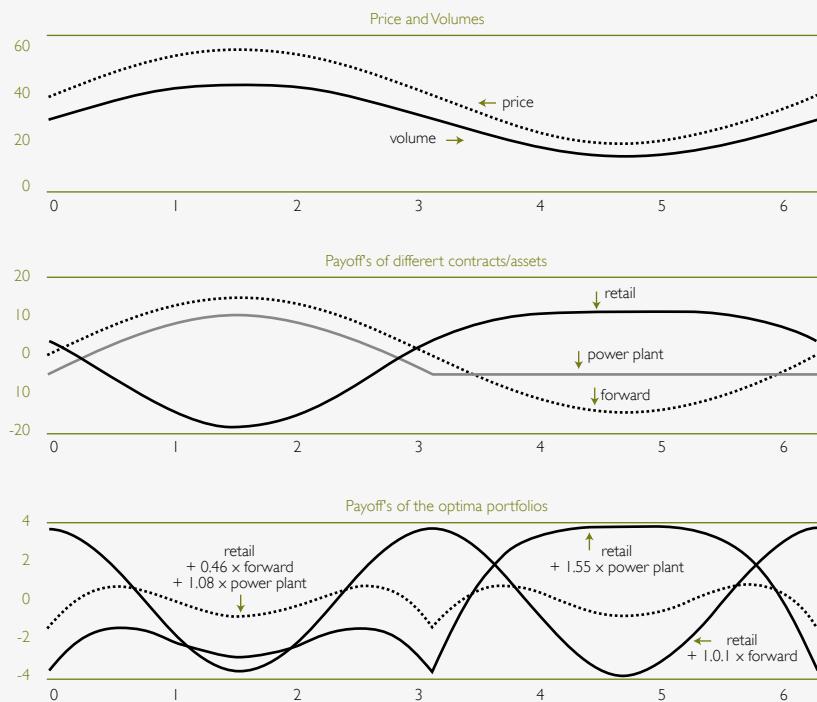
exactly mirror those of the retail contract (see Table 1 and Figure 4). With stochastic demand, a combination of forwards and plant assets forms the optimal portfolio. In cases with non-excessive volatility, forward contracts alone are responsible for the major part of the risk reduction.

When price is assumed to be correlated to demand (Figure 3) again, power plants and forward contracts are needed for optimal risk reduction. In this case power plant shares are responsible for the major part of the risk reduction. Due to the correlation of retail quantities and spot prices retailer's profits have 'flat hills and deep valleys' (see Figure 3).

That is, in periods of high wholesale prices their customers will demand more electricity. Thus, losses in periods where wholesale prices are above retail prices are over proportional. In periods of low wholesale prices the retail customers demand less electricity so that

Figure 3

Stylized examples of the necessity of option-like assets in retail portfolios



a retailer's gain from the positive retail-wholesale price differential is under proportional.

This payoff-structure of retail contracts is almost perfectly mirrored by call options and peak generation assets. Thus, those assets are essential for hedging a retailer's joint price and volume risk. This explains why forward contracts alone are not sufficient for hedging a retail commitment (#4 in Table 2).

Through the presented analysis we provide evidence, that a retailer can hedge the market risks originating from a standard retail contract by either a combination of forwards and options on the spot price or by a combination of forwards and physical assets. In all observed electricity markets, however, liquid derivatives on the spot market are absent (Geman, 2005, Hull, 2005). Thus, the only real choice for a retailer is to hedge its retail obligations through physical assets. These, however, might help to significantly reduce a retailer's risk exposure. In our example the *VaR*(95%) with physical assets decreases by more than 80% compared to a situation where only forward contracts are allowed. Consequently, as long as derivative markets are not sufficiently liquid, retailers will strive to vertically integrate to better hedge their risk exposure. This, on the other hand implies a vicious cycle. The more retailers are vertically integrated the less likely is the development of a liquid contract market, thus forcing non-integrated retailers to leave the market or to move towards physical integration.

Different strategies to exit this vicious cycle might be proposed. By legislative barriers for vertical integration one might boost the demand for certain derivatives. This demand might translate into premiums that make it profitable for banks and generators to provide the demanded derivatives. In the course of time, the number of emitters might increase and drive the premiums down. Another approach would be to reduce the (regulatory) pressure on retail competition allowing for higher margins and thus giving a financial cushion that reduces the risk aversion in the operations. One easily adjusted parameter would, for example, be contract length that is regulated in certain markets. This might allow retailers to maintain potentially risky retail business without vertical integration. But in this case it is unclear whether having a series of oligopolies (retail oligopoly and generation oligopoly) is superior to having an oligopoly of vertically integrated companies.³⁰

Conclusion

Our paper demonstrates that physical hedging, supported to some degree by forward contracting and spot transactions³¹ is an efficient and sustainable approach to risk management in decentralized electricity markets. In contrast to the theoretical premises, financial contracts are imperfect substitutes to vertical integration in the current market environment. The failure of asset-light electricity retailers is indicative of the intrinsic incapacity of this organizational model to manage efficiently the combination of sourcing and market risks.

30 The inelasticity of demand might limit the typical welfare losses from having a series of oligopolies while on the other hand from a customer's perspective the potentially higher prices due to reduced regulation might offset the potential gains of vertical unbundling.

31 Being structurally shorty hedged leaves opportunity for residual contracting options opportunities in volatile electricity markets.

Bibliography

- Aid, R., Chemla, G., Porcher, A., and Touzi, N., (2007), "Hedging and Vertical Integration in Competitive Markets", Paris-Dauphine Working Papers.
- Anderson, A., Edward J., Hu, X., and Winchester, D., (2007), "Forward contracts in electricity markets: The Australian experience." *Energy Policy*, vol.35.
- Bailey, J., (1998), "Intermediation and electronic markets: aggregation and pricing in internet commerce", PhD Thesis, MIT.
- Boroumand, R.H., (2008), "Asset light retailers: the failure of an organizational model", Working Paper Larsen [in French].
- Brousseau, E., and Glachant, J.M., (ed) (2002), *The Economics of Contracts: Theories and Applications*, Cambridge University Press
- Bushnell, J., Mansur, E., Saravia, C., (2007), "Vertical Arrangements, Market Structure, and Competition: an Analysis of Restructured U.S. Electricity Markets," NBER Working Papers, n° 13507.
- Chao, H.P., and Huntington H.G., (ed) (1998), *Designing Competitive Electricity Markets*, Kluwer.
- Chao, H.P., Oren, S., and Wilson, R., (2008), "Reevaluation of Vertical Integration and Unbundling in Restructured Electricity Markets" in
- F.P. Sioshansi (ed.), *Competitive Electricity Markets: Design, Implementation, and Performance*, London, Elsevier.
- Coase, Ronald H., (1960), "The Problem of Social Cost", *Journal of Law and Economics*, vol. 3 (!)
- Coeurderoy, R. and B. Quelin (1997), "Transaction Cost Theory: A Survey on Empirical Studies on Vertical Integration," [in French], *Revue Economie Politique*, vol. 107
- Cornwall Energy and EEE Ltd (2008), *The practices and principles of Electric retail mass markets*, London
- Finon, D., (2008), "Investment Risk Allocation In Restructured Electricity Markets: The Need of Vertical Arrangements" Working Paper Larsen.
- Gehrig, T., (1993), "Intermediation in search markets", *Journal of economics and management strategy*, vol. 2, n°1.
- Geman, H., (2005), *Commodities and Commodity Derivatives. Modeling and Pricing for Agricultural, Metals and Energy*, Wiley Finance editors.
- Hackett, S.C., (1992), "A comparative analysis of merchant and broker", *Journal of Economic Behavior and Organization*, vol. 18(3)
- Hansmann, H., (1996), *The Ownership of Enterprise*, Harvard University Press.
- Hull, J., (2005), *Options, futures, and other derivatives*, 4th Edition, Prentice Hall, Englewood, NJ.
- Hunt, S., (2002), *Making Competition Work in Electricity*, Wiley.
- Hunt, S., and G., Shuttleworth (1997), *Competition and Choice in Electricity*, Wiley.
- Jorion, Ph., (2000), *Value at Risk*, Mc Graw Hill.
- Joskow, P., (1985) "Vertical integration and long-term contracts: the case of coal-burning electric generating plants", *Journal of Law, Economics and Organization* vol. 1(1)
- Klein, B., Crawford, R., and Alchian, A., (1978), "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *Journal of Law and Economics*, vol. 21
- Klein, B., (2000), "The Role of Incomplete Contracts In Self-Enforcing Relationships" *Revue d'Economie Industrielle*, n° 92
- Kühn, K.U., Machado, M., (2004), "Bilateral Market Power and Vertical Integration in the Spanish Electricity Spot Market", CEPR Discussion Papers n°4590.
- Lora, A.T., Santos, J.R., Ramos, J.L., and Exposito A.G., (2002), "Electricity Market Price Forecasting: Neural Networks Versus Weighted-Distance k Nearest Neighbors", *Lecture Notes in Computer Science*, no. 2453.

- Mackay, P., and Moeller, S.B., (2007), "The Maturity Structure of Corporate Risk Management", WP Hong Kong University of Science and Technology.
- Nogales, F., Contreras J., Conejo A.J., and Espinola, R., (2002), "Forecasting Next day electricity prices by time series models," *IEEE Trans. Power Syst.*, vol. 17
- Ofgem, (2007), *Domestic Retail Market Report*. London
- Oxera, (2008), *Energy Supply Market : are they competitive?*, London
- Shelanski, H.A., and PG, Klein (1995), "Empirical research in Transaction Cost Economics: a review and assessment" *Journal of Law, Economics and Organization*, vol. 11(2)
- Szkuta, B.R., Sanabria L.A., and Dillon, T.S., (1999), "Electricity price short-term forecasting using artificial neural networks," *IEEE Trans. Power Syst.*, vol. 14
- Spulber, D.F., (1999), *Market microstructure: intermediaries and the theory of the firm*, Cambridge University Press
- Thomas, S., (2006) "The British Model in Britain: failing slowly" *Energy Policy*, vol 34
- Stoft, S., (2002), *Power System Economics: Designing Markets for electricity*, IEEE Press/Wiley-Interscience.
- Walker, G., and D., Weber (1984), "A Transactions Cost Approach to Make or Buy Decisions," *Administrative Science Quarterly*, vol. 29
- Whinston, M., (2003), "On the Transaction Cost Determinants of Vertical Integration," *Journal of Law, Economics and Organization*, vol. 19
- Whinston, M., and Segal, I., (2003), "Robust Predictions for Bilateral Contracting with Externalities," *Econometrica*, vol. 71
- Williamson, O., (1971), "The Vertical Integration of Production: Market Failure Considerations," *American Economic Review*, vol. 61
- Williamson, O., (1985), *The economic institutions of capitalism: Firms, markets, relational contracting*, Free press (New York and London).
- Working, H., (1949) "The Theory of the Price of Storage", *American Economic Review*, vol. 39.
- Wright, P., (2007) "Competition in Gas and Electricity: Companies Profit, Consumers Day", *Revue de l'Energie*, juin

Appendix

Table 5. Key characteristics of the payoffs of the considered assets

	Variance	Var(95%)	Correlation with retail
Retail contract without hedge	2,384,760,363	-97,852	1
$V_{forward}$	2,195,054,443	-79,179	-0.992
$V_{plant.25}$	1,668,204,401	-64,835	-0.998
$V_{plant.50}$	1,102,804,456	-48,858	-0.999
$V_{plant.75}$	567,871,146	-31,393	-0.993
V_{call}	887,879,490	-42,2	-0.998
V_{put}	332,557,274	-26,045	0.917

Table 6. Characteristics of the observed and simulated data

	observed data	Simulation
mean price	45.1	43.0
median price	37.9	39.0
mean load	54,593	53,949
median load	53,778	53,794
Variance price	1,870	1,056
Variance load	126,726,050	1,548,842

Table 7. Characteristics of the simulated data – Variances

Variance of the median price	20.20
Variance of the mean price	28.60
Variance of the median load	10,825,948
Variance of the mean load	14,512,156

Figure 4. Optimal hedging decision if retail volume is constant

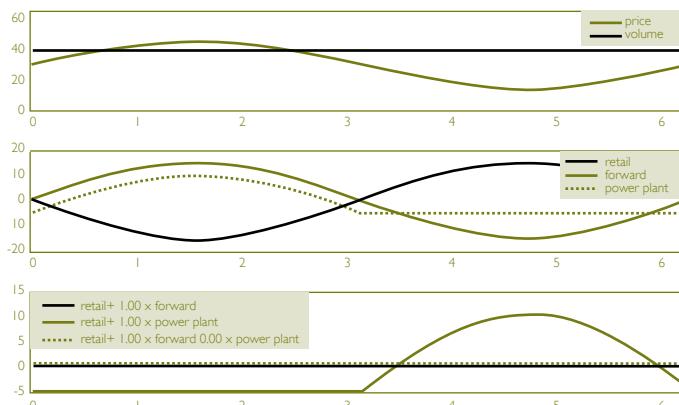
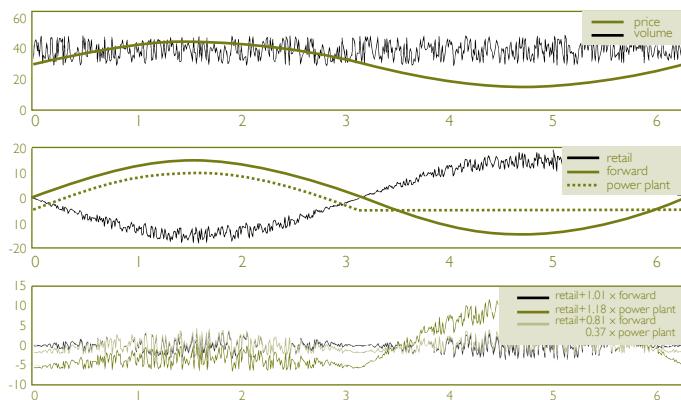


Figure 5. Optimal hedging decision if retail volume is stochastic and uncorrelated to the wholesale price





les politiques de la concurrence et l'organisation des marchés

HHI, an irrelevant indicator without a relevant market

Delphine Perrot-Voisard

LARSEN

Georg Zachmann

LARSEN

For evaluating the concentration effects of horizontal unbundling or mergers in the electricity industry a one-number-indicator is desirable. But, the three most widely applied measures (HHI, PSI and RSI) are per se unable to represent changing sizes of the relevant market due to occasional congestion as they often occur in electricity markets. This paper proposes an adjusted HHI taking into account wider-than-national markets as well as time varying degrees of international competitive pressure. The corresponding adjusted HHI for the French market is 3200 and thus significantly below the HHI of the pure national market (6500).

Introduction

The literature provides several indices of the concentration degree in electricity markets. The most commonly used are the Herfindahl-Hirshmann Index (*HHI*), the Pivotal Supplier Index (*PSI*) and the Residual Supplier Index (*RSI*).¹ Yet, most energy economists would agree that the *HHI* (as well as the *PSI* and the *RSI*) are poor measures of the concentration degree in electricity markets for various reasons: they ignore internal congestion, they ignore vertical integration, they do not take into account the merit order (i.e., the non-linear cost curve of electricity production), etc. Furthermore, much more advanced techniques that cope with certain of those issues exist.² Nonetheless, the desire (of policy makers and lawyers) for a one-number-summary of market concentration has driven a broad usage of those indicators. Especially the *HHI* has been applied and presented in the EU Sector Inquiry, the EU Benchmark reports as well as in certain lawsuits.³ Thus, the economic foundations of key assumptions required to calculate these indicators should be carefully checked.

One factor that is central for all market concentration measures is the definition of the relevant market. Due to the high volatility of demand, the non-storability of electric energy and the technical limitations of the transmission network, the size of the relevant market in the electricity sector might change from hour-to-hour. Because the corresponding *HHI* is as volatile as the relevant market, this concentration indicator would lose much of its appeal. In this paper we propose a pragmatic approach to calculate a one number *HHI* for a certain core market that takes into account the competitive pressure from adjacent markets (with which it forms a relevant market part of the time). In the next section the adjusted *HHI* is introduced. The following section presents the data of our case study (France) and in the fourth section the adjusted *HHI* is calculated for the French market. The last section concludes.

¹ The Herfindahl-Hirshmann index (*HHI*) is defined as follows: , where MS_i is the market share of firm i . It measures the degree of concentration in a given market. The maximum value (monopoly) of the *HHI* is equal to 10000.

² Eg. Oligopolistic models of Cournot or Supply function equilibrium type allow a much deeper understanding of the competitive effects of mergers or divestitures.

³ European Commission (2005, 2006, 2007) and see Vandecast et al. (2006).

The international trade adjusted HHI

A relevant market is defined as the geographic area in which a producer effectively competes with producers of the same product and producers of substitutes.⁴ This theoretical definition is not easily translated into clear market boundaries of real-world electricity markets.

The question 'which regions generators might exercise competitive pressure on a producer?' could be sensibly addressed only by profound modeling. But, building oligopolistic models that replicate all important characteristics of electricity markets has proven difficult. Including highly volatile demand, network characteristics, market mechanisms, forward markets, congestion management mechanisms, etc. in one single model is very challenging.⁵ Consequently, simple initial indicators concerning the size of the relevant market remain desirable.

Thus, a set of pragmatic approaches have been proposed. According to Vandezande et al. (2006) in past cases the EU Commission used certain 'rules of thumb' relying on the occurrence of congestion and price differences as well as the relative size of the interconnectors for determining market definitions.⁶ The Sector Inquiry⁷ proposes that the relevant market might be formed by the domestic market plus the capacity of the transmission lines.⁸ Further 'rules of thumb' have been proposed by other authors.

⁴ Assuming that electricity, at least in the current price region is difficult to substitute in the short- to mid-term we focus on geographic competition relative to product substitutes competition (natural gas for example).

⁵ Using these "structural" models it is possible to deduce the size of the relevant market by estimating whether the strategic choice of a generator is affected by the potential reactions of a competitor or not. In the first case the competitor is part of the relevant market, in the second it is not.

⁶ A wider than national market is an acceptable assumption if: interconnection capacity > 25 percent of national consumption; congestion level < 5 percent; price differences occurring < 10 percent of the time. A wider than national market is not an acceptable assumption if: interconnection capacity < 10 percent of national consumption; congestion level > 25 percent; price differences occurring > 50 percent of the time (see Vandezande et al. (2006)).

⁷ European Commission (2007).

⁸ There is a theoretical difficulty with this approach as transmission lines might either be considered as consumer (export) or producer (import). Thus, a transmission line towards a low generation cost country might increase competition while a connection with a high cost country might curb competitive pressure if interconnector allocation is performed through a competitive mechanism that prohibits withholding.

As has been discussed by Dijkgraaf and Janssen (2008) the relevant electricity market are rather 'flighty'. On the example of the French-Belgian-Dutch trilateral market coupling, they show that certain periods exist, where these three markets form a common 'relevant' market. In presence of congestion, however, the market splits and the relevant market shrinks to a bilateral or unilateral market. Both, extensive oligopolistic modeling and 'rules of thumb' do not take into account the 'flighty' nature of electricity markets. Even if different market sizes for different demand scenarios are calculated, conveying the results to the addressees is difficult. Stating that Polish producer might exercise competitive pressure on prices in Spain in certain hours of the year but that Austrian power plants do not in all cases compete with German ones leaves lawyers uncertain of what to make of these results. Thus, a one-number summary that indicates the competitive situation in the core market, taking into account the occasional coupling with adjacent markets would be desirable. Vandezand et al. (2006) suggest that it is possible to calculate *HHIs* based on two extreme market definitions. Using certain additional indicators (congestion, price differentials) they propose that one might interpolate a definitive result in-between the smallest and the largest market definition.

This method lacks three important qualities for indicators: transparency, reproducibility and simplicity. We therefore propose an alternative approach. The idea is that an adjusted *HHI* for one core market can be calculated taking into account the hourly changing occurrence of congestion at its borders. This concentration indicator is based on weighting the concentration indicators for different market combinations with their respective frequency of being 'relevant'.⁹ Hence, the different competitive effects of export and import congestion are taken into account.

In the case of export congestion, the *HHI* corresponds to the *HHI* in the core market. By way of contrast, in the case of import congestion the competitive pressure of foreign generation is limited to the size of the corresponding interconnector attributed to generators according to their foreign market share (See Table 1).

⁹ See the example in the Appendix.

Doing this analysis for a real-world case thus requires two steps: First the concentration indicator for each feasible grouping of markets (that include our core market) has to be calculated. And second, the frequency at which each market group is relevant (i.e., the biggest joint market) has to be calculated.

This requires defining when a group of markets can be considered as coupled. Markets can be said to be coupled if price (or volume) changes in one market are transferred to the other market. Thus, a market player in one country has to take into account the response on his price/volume decision in the other country. To decide whether two markets are coupled or split, one might either consider physical or financial indicators. If available transmission capacity is not fully used, two markets might be considered coupled.

This physical approach, however, requires full knowledge of the available capacity (including loop flows, reliability margins in transmission capacities etc.). Furthermore, under- or over-utilization of transmission lines with respect to price differentials are very common in Europe. If prices in two interconnected regions are similar, this is a strong indicator that those two markets are coupled.¹⁰ Thus, we assume market coupling if the price difference is below 5%.¹¹ With this definition in hand we can calculate the historic frequency of market coupling for each market combination. Combining these with the corresponding *HHIs* allows computing the synthetic *HHI* for country j (HHI_j).

Table 1

Two Country Example for calculating the adjusted *HHI*

	Generation Capacity of Company A	Generation Capacity of Company B	Generation Capacity of Company C	Size of the relevant market	Frequency of occurrence	HHI
Domestic	90 GW	10 GW	0 GW	100 GW		8200
Foreign	10 GW	0 GW	40 GW	50 GW		6800
Coupled	100 GW	10 GW	40 GW	150 GW	40%	5200
Import Congestion (10 GW line)	92 GW	10 GW	8 GW	110 W	40%	7131
Export Congestion	90 GW	10 GW	0 GW	100 GW	20%	8200
Total						6572

The HHI_j can thus be calculated according to :

$$HHI_j = \frac{1}{T} \sum_{t=1}^T \left(\sum_{k=1}^K \left(100 \times \frac{\sum_{i=1}^n GenCap_{k,i} \times coupling_{i,t}}{\sum_{k=1}^K \sum_{i=1}^n GenCap_{k,i} \times coupling_{i,t}} \right)^2 \right)$$

Where $GenCap_{k,i}$ is the generation capacity of company k in country i and $coupling_{i,t}$ is defined according to :

$$coupling_{i,t} = \begin{cases} 1 & \text{if } -5\% \times price_{j,t} < (price_{i,t} - price_{j,t}) < 5\% \times price_{j,t} \\ 0 & \text{if } price_{i,t} - price_{j,t} \geq 5\% \times price_{j,t} \\ \frac{tc_i}{\sum_{k=1}^K GenCap_{k,i}} & \text{if } price_{i,t} - price_{j,t} \leq -5\% \times price_{j,t} \end{cases}$$

There, tc_i is transmission capacity of country j with country i and $price_{i,t}$ is the price of electricity in country i at time t .

Data

The case study in this paper is focusing on the French and its adjacent markets, Belgium (BE), Italy (IT), Spain (ES), United Kingdom (UK), Germany (DE) as well as the Netherlands (NL).¹² To calculate the HII according to the methodology above the interconnector capacities between the considered markets as well as the ownership structure of generation assets in those markets are required. The former are obtained in the form of net transfer capacities (NTCs) from the UCTE calculations for Summer 2008 (see Appendix).

- ¹⁰ A reservation has to be made: Remaining price differentials might not necessarily indicate splitting of markets but they can also be due to the decentralized price formation in power exchanges and explicit capacity auctions. In fact, penny-sharp arbitrage takes rarely place even in completely integrated markets. Such deviations will abate as market coupling develops.
- ¹¹ The threshold is borrowed from the "Small but Significant and Nontransitory Increase in Price" test but remains somewhat arbitrary. Thus, the corresponding assumption has to be made explicit in a corresponding analysis. In our case study a 10% reduction in the threshold (to 4.5%) increases the HHI by 2.5%.
- ¹² Note that through this approach apart of the Netherlands third countries (e.g., Portugal, Denmark, Austria ...) as well as countries without an hourly spot market (Switzerland) are ignored even though they might form a unique-price-area with France at certain times. Luxemburg is considered as part of the German electricity system with which it is strongly connected.

The ownership of generation capacities is acquired from the Annual reports of the biggest generators (see Appendix). This means that the competition enhancing effects of smaller generators are not fully represented as those are summarized in a single entry ("other generators"). This, however, only results in a very limited bias as the *HHI* is by construction dominated by the market share of the big players. Another point worth mentioning is the way we treated cross-ownership. To keep the calculation as simple as possible, only cross-ownerships among the big players are considered, control stakes ($>50\%$) are taken into account as full ownership while non-control stakes ($<50\%$) are ignored.¹³ For calculating the coupling frequencies we rely on European electricity spot prices from November 2006 (introduction of trilateral market coupling and the Belgian electricity exchange) to August 2008. We collected hourly data from the Italian, Spanish, German, French, Belgium, Dutch and British electricity spot market. We assume that the spot prices obtained in each power exchange reflect the spot price in all the concerned country.

Results

In Table 2 the results for certain market combinations are presented. It is striking, that the French market is almost never (less than 0.5% of the times) completely split from all its neighboring

Table 2

HHIs and coupling frequencies for selected combinations at 5% price difference 11/2006-08/2008

	HHI	Frequency of this being part of the relevant market	Frequency of this being the relevant market
France	6457	100%	0.47%
France - Belgium	5235	89%	13%
France - Germany	2119	26%	2%
France - Italy	2689	7%	0%
France - Spain	2725	16%	1%
France - Belgium - Netherlands	3975	72%	39%
France - Belgium - Nether. - Germany	1694	21%	16%
France - Belgium - Netherlands - Germany - Italy - Spain - UK	825	0%	0%

markets. In fact, in 72% of the cases the trilateral market coupling with Belgium and the Netherlands produces (almost) similar prices in all three countries. Furthermore, also the German (27%), Spanish (16%) and Italian (7%) prices are often (almost) identical to the French price.

From the calculation of the concentration indicators it is obvious that the size of the coupled market significantly influence the *HHI*. While a truly common market (France - Belgium - Netherlands – Germany - Italy - Spain - UK) produces an *HHI* as low as 825; the *HHI* for the “France only” market is at 6457. But both extreme cases were equally unlikely in the observation period. The *HHI* for the French market adjusted for varying international competitive pressure according to the presented approach is 3218. This is noticeably below the traditionally considered ‘France only’ *HHI* of 6457. Nevertheless, the still significant amount of the adjusted *HHI* makes clear that market concentration in France is substantial (>1800) even when taking into account cross-border trade. This is due to the unique position of EdF in the European market, the high concentration in neighboring markets (e.g., Belgium) as well as to the frequent decoupling of certain regions (IT, ES, UK).

But, as the adjusted *HHI* depends on current coupling frequencies, the high French concentration indicators do not necessarily have to persist forever. The coupling frequencies of Dutch and French prices for example increased from 20% before November 2006 to 72% thereafter.¹⁴ If the planned pentagonal market coupling (BE-DE-FR-LU-NL) is introduced and would be as successful as the trilateral-market-coupling (BE-FR-NL) the French *HHI* might for example significantly drop. An equal sized increase of the coupling frequencies between Germany and France (currently 26%) might bring the French *HHI* even below 1800 - the threshold above which a market is generally considered as highly concentrated - as the *HHI* for the French-German-Belgian-Dutch is 1694.

¹³ See Campos and Vega (2002).

¹⁴ This is (at least partially) due to the introduction of a trilateral market coupling between Belgium, France and the Netherlands.

Conclusion

The *HHI* is an inappropriate tool to represent the multi-dimensionality of the competitive situation in electricity markets. But these one-number-summaries still play an important role in the political, regulatory and legislative discussion. Thus, their calculation (and interpretation) should take into account main features of this sector. So far, however, international electricity trade and the corresponding competitive pressure have been neglected in the calculation of the *HHI*. To rectify this situation we propose an *HHI* that is adjusted for time varying international competitive pressure. This concentration indicator is based on weighting the concentration indicators for different market combinations with their respective frequency of forming a unique-price area with the core market. For the French market the *HHI* corresponding to this methodology is 3200 (if assuming no congestion if price difference is below 5% of the French price). This is noticeably below the traditionally considered 'France only' *HHI* of 6500. This significant impact demonstrates the capacity of international electricity trade to reduce market concentration. The proposed methodology has various advantages : it is transparent, reproducible, simple and flexible with respect to the analyzed markets. Furthermore, it could be easily adapted for the use with different concentration indicators (e.g., *RSI* or *PSI*). The main drawback for calculating the adjusted *HHI* is that certain assumptions remain arbitrary to some extent. Especially the threshold for assuming price equalization has to be selected with care, documented well and included in the interpretation.

Appendix

Table 3.
Matrix of biggest Producers : installed production capacities (GW)

	FR	DE	BE	NL	LU	IT	ES	UK
ATEL	0.8							
EDF	92.5		0.4					4.8
EnBW		15.0						
Eon	3.3	26.2		2.0		6.6	3.8	9.8
RWE	0.8	31.7			1.1			10.6
Vattenfall	0.8	15.2						0.09
Poweo	0.3							
Direct Energie	0.3							
GDF/Suez	7.8	0.3	12.9	4.7	0.4		1.2	0.18
EWE (ENR)		0.7						
ESSENT				5.1				
Nuon				4.1				
SPE		1.6						
Cegedel				0.2				
Iberdrola						27		
EDP						3.8		
Gaz Natural						4		
ENEL (incl. Endesa)	1.0				40.4	17.6		
edison					12.0			
ENIIPower					4.5			
Scottish Power						6.35		
SSE						9		
DRAX						3.9		
British Energy						10.7		
Centrica						3.5		
Inter. Power Mitsui						5		
Others*	8.5	35.2	1.2	6.2	0.0	25.6	21	14

* "Others" is the difference of total installed capacity minus the capacity of the biggest generators.
 "Others" are attributed to each country as a different company.

Sources : Companies Annual Reports and OFGEM.

Table 4.
Net Transfer Capacities Values for Summer 2008 (MW)

	From France	To France
Germany	2400	2500
Belgium	2700	1100
Netherlands	2000	1100
Austria	1000	800
Spain	1200	500
Italy	2400	870
UK	2000	2000

Source: ETSO 2008 - non-binding peak values

References

- Borenstein, S., Bushnell, J., Stoft, S., 2000. "The competitive effects of transmission capacity in a deregulated electricity industry". *RAND Journal of Economics* 31(2), 294-325.
- Campos, J., Vega, G., 2002. *Concentration measurement under cross-ownership. The case of the Spanish electricity sector.* Univ. de Palma - documento de trabajo 2002-06.
- Dijkgraaf E., Janssen, M., 2008. *Is there a Belgian Wholesale Electricity Market?* Erasmus Competition and Regulation Institute.
- European Commission, 2005. *Report on progress in creating the internal gas and electricity market (Benchmarking report).* Communication from the commission to the council and the European Parliament COM (2005) 568 final Brussels, 15.11.2005.
- European Commission, 2006. *Report on progress in creating the internal gas and electricity market (Benchmarking report).* Communication from the commission to the council and the European Parliament COM (2006) 841 final Brussels, 10.01.2007.
- European Commission, 2007. *Report on energy sector inquiry (Parts 1,2,3,4)*, DG Competition, Brussels 10.11.2007.
- ETSO, 2008. Online data on the Net Transfer Values, downloaded from [http://www.etso-net.org/NTC_Info/library/e_default.asp].
- Vandezande L., Meeus, L., Delvaux, B., Van Calster, G., Belmans, R., 2006, Evaluation of Economic Merger Control Techniques Applied to the European Electricity Sector. *The Electricity Journal* 19(6).

Legal uncertainty and competition policy in deregulated network industries

The case of long term contracts

Adrien de Hauteclercque

University Paris Sud (ADIS-GRJM) & LARSEN

The case of long-term vertical contracts in the EU electricity markets is a topical example of the difficulties faced by Competition authorities with the liberalization of network industries. Their ambiguous effects on the competitive structure, investment and consumer welfare in the long term made them logically become a priority for antitrust enforcement. However, due to the lack of precedents and the on-going modernization of EC competition law, the legal uncertainty currently perceived in the market place is strong. This article proposes to explore the implications deriving from the strategy implemented by the European Commission to cope with these trade-offs. The article comes up with three conclusions. First, legal uncertainty is largely overstated as both the methodology to analyze these contracts and its implementation principles are clearly emerging. Second, more legal certainty became possible because the coping strategy of the European Commission was to replicate methodologies it had devised in other sectors, especially beer and ice-cream, which upgrades legal certainty but does not guarantee the efficiency of future competition enforcement. Third, this methodology could even go counter the objectives of the European Union in terms of market building and security of supply in electricity.

Introduction

Competition authorities face a considerable challenge with the liberalization of network industries. In fast-evolving market settings, they must fight anti-competitive practices while ensuring a fair degree of legal certainty to market players, without being able to firmly rely on past case law, an intimate knowledge of the market or even economic theory. This is particularly difficult in energy where liberalization opened new doors for sophisticated market abuse while the specifics of the sector are strong and the policy objectives of the European Union often contradictory in practice when enforcing EC Competition law.

The case of long-term vertical contracts (LTC) in the EU electricity markets is a topical example of these difficulties. European markets are still wrapped up with LTC and liberalization has not changed much this traditional sales patterns. Sometimes mere residuals of the former vertically integrated structure,¹ they might now constitute innovative ways to mitigate new uncertainties born from liberalization and facilitate the achievement of other policy objectives such as long-term generation adequacy. While the energy community increasingly doubts the ability of de-integrated markets to ensure an optimal allocation of risks and praises long-term contracting, the European Commission consistently emphasizes the risks of anti-competitive effects inherent in LTC² and made them a priority for antitrust enforcement as the recent proceedings in gas and oil show. In electricity however, the lack of precedents and the on-going modernization of EC competition law cast doubts on the way the Commission will apply competition rules, which has fostered legal uncertainty. This raises serious concerns in a sector where predictability of competition enforcement is crucial to stabilize market players' expectations and hence allow them to sink high fixed-cost investments.

This article proposes to explore how the European Commission and national competition authorities are dealing with one of the most important issues in energy markets since liberalization, the

¹ e.g. stranded power purchase agreements (hereafter PPA) in Hungary or Poland, legacy rights on interconnectors.

² DG Competition report on energy sector inquiry, SEC(2006) 1724, 10 January 2007, 232-244 and 283-294.

long-term supply and transport contracts in the electricity sector. The first section will present the basic competition economics of LTC in electricity and uncover the mechanics of the legal uncertainty currently perceived in the market place. The following section will then depict the methodology which emerged from recent proceedings in gas and oil. The last section will show why this methodology will be applied in electricity and what its consequences could be for efficiency in future competition enforcement.

Competition economics of LTC, legal uncertainty and the modernization of EC competition law

There is a multiplicity of long-term bilateral contracts, all along the electricity supply chain, which complement or replace arm's length market relationships taking place on spot markets. They include fuel supply contracts to power producers, long-term Virtual Power Plants (VPP)³, tolling agreements⁴ and diverse PPA with energy intensive users,⁵ commercial and household customers, or resellers and traders without generation capacities. Long-term reservation agreements on interconnectors must also be considered given their inherent vertical dimension.⁶ If all these LTC lead somehow to varying degree of foreclosure, they also have diverse effects on

³ VPP are a classical remedy in energy which forces dominant firms to make capacity options available for a pre-determined time horizon, which amounts to a virtual divestiture of capacity. See below at Section IV for a discussion.

⁴ e.g. Centrica's contract with Intergen for a 860 MW gas-fired power plant.

⁵ Energy intensive users typically include electricity resellers and some industries where electricity represents an important part of total costs (e.g. steel, chemistry). The latter have often used complex contractual and financial arrangements to pool their electricity purchases. These contracting schemes include for instance risk-sharing in generation between industrial consumers and electricity operators (EPR, Zandvliet, Roselectra), partnership between consumers and generators valuing a secondary fuel (DK6) or consumers cooperative purchasing electricity (Exeltium).

⁶ Long-term horizontal contracts often raise the same sort of issues but are kept out with the scope of this paper. They include for instance joint-marketing, joint-infrastructure development or long-term energy SWAPS. A recent example of the latter is the agreement between EDF and POWEO signed in January 2007. The rationale of the deal was to swap actual base against future peak load capacity. POWEO gains access to EDF nuclear capacities from 2007 to 2021 and gives in return a future access to its CCGT currently in the construction phase, for the same capacity and the same period (160 MW per year over 15 years).

consumer surplus, investment, risk management, entry and spot prices.⁷ From a competition point of view, contract clauses regarding duration, exclusive dealing and use restrictions are particularly relevant in electricity whereas exclusive distribution and destination clauses are more relevant in gas.⁸ Restraints relating to prices such as discrimination, predation or resale price maintenance might also be found but have so far rarely been formally litigated in the context of LTC cases at the Community level. In view of the challenges ahead in terms of security of supply and cheaper prices for the consumers, LTC in electricity need to be carefully regulated by competition authorities. This section briefly shows the pros and cons of LTC in a context of market building and how the modernization of EC competition law as well as the strategy of the European Commission increased the legal uncertainty currently perceived in the market place.

The basic competition economics of LTC in the electricity sector and its ambiguities

Whereas vertical restraints generally tend to be viewed more leniently than horizontal restraints by competition authorities, this is

⁷ EDF-IDEI Report, *Contrat de Long Terme, Concurrence et Efficacité*, December 2007.

⁸ In electricity, exclusive purchase clauses are much more common than exclusive distribution clauses. This is not the case in gas where an important part of Commission enforcement took place in the upstream segment of the industry essentially concerning territorial restrictions. On this, see the following settlements: Commission press releases IP/03/1345 *Gazprom/ENI* of 06.10.2003, IP/05/710 *Gazprom/E.ON Ruhrgas* of 10.06.2005, IP/05/195 *Gazprom/OMV* of 17.02.2005, IP/07/1074 *Sonatrach* of 11.07.2007, IP/02/1869 *NLNG* of 12.12.2002 and IP/02/1084 *GFU* of 17.07.2002. For commentaries see Nyssens, Cultrera and Schnichels, "The Territorial Restrictions Case in the Gas Sector: a State of Play", *I Competition Policy Newsletter* (2004), 48-51; Cultrera, "Les Décisions GDF, la Commission est Formelle: les Clauses de Restrictions Territoriales dans les Contrats de gaz violent l'Article 81", *I Competition Policy Newsletter* (2005), 45-48; Wältare, "Territorial restrictions and profit sharing mechanisms in the gas sector: the Algerian Case", *3 Competition Policy Newsletter* (2007), 19-21, and Talus, "Long-term Gas Agreements and Security of Supply – Between Law and Politics", *32(4) European Law Review* (2007), 535-548. The only important exclusive supply cases in electricity since liberalization have been the remedy concerning CNR in Case M.1853 *EdF/EnBW* (see European Commission press release IP/01/175 of 07.02.2001) and the stranded PPA cases in Hungary, Portugal, Ireland, Italy, United Kingdom, Greece and Poland dealt with under State Aid. CNR, an independent French power producer, was relieved from its long-term exclusive distribution obligation with EDF to foster competition in the French market after the loss of an important potential entrant (EnBV). In the stranded PPA cases, the issue was around the compensation schemes for the operators which suffered from liberalisation due to their long-term commitments or guarantees. On this see Art 24 of Directive 96/92 EC, O.J. 1996, L 27/20; the Commission Communication on the *Methodology for analysing State Aid linked to stranded costs*, available on the website of the European Commission; Hancher, "Energy", in Hancher, Jan Slot and Ottervanger (eds.), *EC State Aids* (Sweet & Maxwell; 3rd ed., 2006), 656-680 and the interesting expert of opinion of Eilmansberger, Jaeger and Thyri, *Compatibility of the Hungarian System of Long-term Capacity and Power Purchase Agreements with EU Energy and Competition Law* (2004).

not the case in energy where the disintegration of vertical market relationships has been one of the key policy items put forward by the European Commission since the first liberalization directive.⁹ However, the current volatility of spot prices, the non-storability of the commodity and the unbundling imposed¹⁰ lead market players to call for more rigid vertical arrangements. The usually concentrated market structure, inelastic demand and high investment costs have also contributed to make LTC, at least as a complement to spot market contracting, a useful contractual structure for individual market players. However, LTC may in parallel hinder market building efforts. The basic rationale for LTC in electricity is hedging price and quantity risks over a certain period of time, so duration is crucial. Besides, LTC in electricity not only define duration but also other features of the transaction such as use restrictions, renegotiation conditions, quantity and price, with some flexibility. It would thus be wrong to solely focus on duration as anti-competitive effects lie as much in other contract clauses and on the competitive structure of the market. Indeed, bilateral contracting does not only enable market players to hedge price and quantity risks, it also expresses and channels their ability to distort competition. In *Gas Natural/Endesa*¹¹ for instance, the Commission¹² explicitly states that the structure of the contract in itself demonstrates the dominant position of Gas Natural. Therefore, as contracting parties take into account *ex ante* the regulatory framework applied to them when devising contracts, competition policy is a way to impact the structuring of competitive behaviors and ultimately to limit abuse of market power.¹³

⁹ Directive 96/92/EC on common rules for the internal market in electricity, O.J. 1996, L 27/20.

¹⁰ See Art 10 and 15 of Directive 2003/54/EC concerning common rules for the internal market in electricity and repealing Directive 96/92/EC, O.J. 2003, L 176/37. See also the proposal of the Commission for furthering vertical unbundling in recital (5) to (15) and Art 8 and 10 of the Proposal for a Directive of the European Parliament and of the Council amending Directive 2003/54/EC concerning common rules for the internal market in electricity 2007/0195 (COD).

¹¹ See XXXth Report on Competition Policy (2000) and European Commission press release IP/00/297 of 27.03.2000.

¹² The 'Commission' or the 'European Commission' will be used interchangeably except if specified.

¹³ Lohmann for instance shows how the modernization of German competition law and the re-inclusion of energy within its ambit led to a natural re-engineering of certain gas contracts in the late 1990's in *The German Path to Natural Gas Liberalisation* (NG14, Oxford Institute for Energy Studies), 92-93.

The main competition concern associated with LTC in electricity is the risk of foreclosure of more efficient players, deemed to equate to a loss of long-term consumer welfare. If a significant part of demand is tied in the long run, a lack of retail outlets may lead to significant output foreclosure at the production level and tied consumers will not be able to benefit from future and potentially more profitable offers by new entrants. LTC may thus constitute a barrier to entry and a negative externality on third parties.¹⁴ Conversely, if the market structure at the producer level is very concentrated, which is usually the case in European energy markets, input foreclosure may occur and prevent entry in retail. LTC by incumbents in electricity became a priority for antitrust enforcement due to these foreclosure effects both in generation and retail.

Much theoretical ambiguity still remains around the impact of long-term supply contracts on spot market deepening¹⁵ a cornerstone of competition policy in European energy markets.¹⁶ Large, liquid and stable spot markets are deemed to facilitate entry in retail and trading, and thus foster competition as well as provide reliable investment signals. If a significant part of electricity flows is contracted on a long-term bilateral basis, spot market development is limited and price volatility increases, which complicates entry and incentivizes market players towards vertical (re)integration or long-term contracting (feed-back effect). However, theoretical arguments¹⁷ have been developed which tend to show that long-term contracting by dominant players incentivize them not to exercise their market power on spot markets as increases in prices would only be profitable on the un-contracted part of their supplies.

LTC effects on welfare are neither clear in the short run nor from a more dynamic, long-term perspective. In the short term, LTC tend to prevent double marginalization problems,¹⁸ which results in both higher profits and lower prices, and facilitate entry when sufficiently

¹⁴ Aghion and Bolton, "Contracts as a Barrier to Entry", *77 American Economic Review* (1987), 388-401.

¹⁵ Bonasina, Creti and Manca, "Imperfectly Competitive Contract Markets for Electricity" (2007), Working Paper.

¹⁶ DG Competition report on energy sector inquiry, see above note 2.

¹⁷ Allaz and Vila, "Cournot Competition, Forward Markets and Efficiency", *59 Journal of Economic Theory* (1993), 1-16.

¹⁸ Onofri, "Electricity market restructuring and energy contracts: a critical note on the EU Commission's NEA Decision", *20 European Journal of Law and Economics* (2005), 71-85.

long.¹⁹ In the longer term, LTC enable contracting parties to sink high investment costs, by-pass the lack of liquidity on spot markets and avoid the transaction costs of repeat business. They also facilitate bank involvement in project financing and might thus be necessary for investments and entry in a capital intensive industry, hence for long-term generation adequacy. LTC may even contribute to approaching optimal diversity in the fuel mix by enabling high fixed-costs investment in nuclear or coal power stations which might not be financed otherwise.²⁰ However, if LTC are generally viewed as facilitating high fixed costs investments, this view needs to be contrasted in electricity as several other criteria are required to reach that effect.

The duration of the contract needs to be long enough and this requires finding counterparties sufficiently capable of predicting their needs to be able to commit. Given the strong uncertainty resulting from the fast-evolving regulatory context and the spot price volatility, such potential contractors can only go slowly upward the learning curve in this sector, which may limit opportunities for long term contracting.²¹ This is particularly visible for producer/retailer relationships as retailers face risk of switching of final consumers which are tied for much shorter periods.²² In case spot prices are lower than the price of the LTC concluded, alternative retailers will be able to propose lower retail prices which will incentivize consumers to switch. The retailer, which contracted for the long term in the previous period, then cannot but align with its rivals and be squeezed.

Overall, competition economics in electricity provides useful insights but theoretical ambiguities remain and the body of empirical work is still limited. In particular, the pattern of entry in generation and the effects of LTC on spot markets are unclear. If LTC may be useful for individual contracting parties, their impact on market building in a deregulated network industry is much harder to regulate than the

¹⁹ Newbery, "Competition, Contracts and Entry in the Electricity Spot Market", 29 *RAND Journal of Economics* (1998), 726-749.

²⁰ Finon and Perez, "Investment Risk Allocation in Restructured Electricity Markets: The Need of Vertical Arrangements", LARSEN Working Paper (2008).

²¹ Vasquez, Rivier and Pérez-Arriaga, "A Market Approach to Long-term Security of Supply", 17(2) *IEEE Transactions on Power Systems* (2002), 349-357.

²² Neuhoff and De Vries, "Insufficient Incentives for Investment in Electricity Generations", 12 *Utilities Policy* (2004), 253-267.

traditional problem of foreclosure in sectors where competition is more mature. As a result, the policy question is more how to approach a workable mix of contract durations than what is the optimal contract length in a deregulated industry, a question far to being settled. This is indeed what antitrust authorities try to do when they impose remedies. A blind enforcement of competition law in electricity may create incentives to further vertical (re)integration and undermine the best contractual allocation of risks among parties, which would go counter the objectives of the European Union in terms of market efficiency and investment in this industry. In the face of such difficulties, disregarding sector specifics and using the usual methods of competition policy maybe tempting. However, this is not the reception of recent insights from the new energy economics which created legal uncertainty but the strategy of the Commission and the evolution of competition tools.

The evolution of the jurisprudence and the modernization of EC competition law

Legal uncertainty and the evolution of the case law. Despite their importance for the success of liberalization, LTC are almost absent in gas and electricity secondary EC law so guidance must be sought in past case law.²³ The first source of legal uncertainty comes from the limited number of decisions on LTC in electricity since the opening-up of markets. Prior to the first liberalization directive, enforcement of competition law did not occur on a regular basis and there were only few instances of LTC cases. Most of them concerned independent power producer supplying the national incumbent on an exclusive basis. Following the Single European Act, the directive on cross-border trade in electricity²⁴ was enacted in 1990 and the Commission started to look at these LTC to limit their duration so that they do not hamper the future opening of markets to competition. The durations

²³ Except in the Directive 2003/55/EC concerning common rules for the internal market in natural gas and repealing Directive 98/30/EC, O.J. 2003, L 176/57 in recital 25: "Long-term contracts will continue to be an important part of the gas supply of Member States and should be maintained as an option for gas supply undertakings in so far as they do not undermine the objective of this directive and are compatible with the Treaty, including competition rules. It is therefore necessary to take them into account in the planning of supply and transportation capacity of gas undertakings."

²⁴ Directive 90/547/EEC on the transit of electricity through transmission grids, O.J. 1990, L 147/37.

were in general limited to 15 years as in *Electricidade de Portugal/Pego*,²⁵ *Isab Energy/Enel*,²⁶ *Sarlux*,²⁷ *Rosen*,²⁸ *REN/Turbogas*,²⁹ *Scottish Nuclear*³⁰ or *Api Energia*³¹ and 25 years in *Transgas/Turbogas*.³²

These cases were characterized by a lack of methodology for the analysis of foreclosure effects, leading to decisions unlikely to be accepted today on the same terms. For instance, the formation of selling and purchasing consortia contracting on a long-term basis was accepted in *Jahrhundertvertrag*³³ to allow the development of local energy sources (coal) for the sake of "national security". However, we can note that the "security of supply" argument as a justification for the long duration was already accepted with some reluctance by the European Commission which refused in *Jahrhundertvertrag* and *Ijsselcentrale*³⁴ to proceed under Art 86(2) and preferred to use Art 81(3).³⁵ In the new context, market players can anticipate that the 15 years duration will probably not be accepted but they cannot get legal uncertainty from this case law.

Since liberalization and apart from *Synergen* and *Gas Natural/Endesa*, relevant cases in this sector have essentially concerned long-term reservation rights on interconnectors signed before liberalization. In fact, most cases concerning LTC in energy have taken place in gas and essentially in the upstream part of the industry (long-term import

²⁵ *Electricidade de Portugal/Pego Project*, notice pursuant to Art 19(3) of Regulation 17/62, O.J. 1993, C265/3 and XXIIIrd Report on Competition Policy (1993),222.

²⁶ *Isab Energy*, notice pursuant to Art 19(3) of Regulation 17/62, O.J. 1996, C 138/3.

²⁷ XXVIth Report on Competition Policy (1996), I34.

²⁸ XXVIth Report on Competition Policy (1996), I34.

²⁹ *REN/Turbogas*, notice pursuant to Art 19(3) of Regulation 17/62, O.J. 1996, C 118/7.

³⁰ See Section IV for more on this case.

³¹ XXVIth Report on Competition Policy (1996), I34.

³² XXVIth Report on Competition Policy (1996), I35.The 25 years duration was justified by an improvement of security of supply due to the development of alternative sources of gas supply.

³³ Case IV/33.151 *Jahrhundertvertrag*, O.J. 1993, L 50/14.

³⁴ Case IV/32.732 *Ijsselcentrale*, O.J. 1991 L 28/32.

³⁵ On the security of supply argument in the pre-liberalization period, see also the position of the ECJ in Case C-72/83 *Campus Oil Limited and Others v. Minister for Industry and Energy and Others*, [1984] ECR 2727, [1984] 3 CMLR 544; Case C-393/92, *Municipality of Almelo and Others v. Energiebedrijf IJsselmij NV*, [1994] ECR I-1477 and Case C-158/94, *Commission v. Italy*, [1997] ECR I-5789.

contracts) where the problems are different due to the geopolitical dimension. In addition, since the early 2000's, DG Competition has publicly and repeatedly voiced strong concerns over the risks of anti-competitive effects inherent in electricity LTC³⁶ whereas long durations had been repeatedly accepted in the former period. The legal uncertainty created by the lack of precedents has been amplified by the split between the current state of Commission thinking and its past decisional practice.

Legal uncertainty and the modernization of EC competition law.

Antitrust is a key policy tools to overcome the current shortcomings of energy liberalization in Europe³⁷ and the new context of EC competition law raises several fundamental problems for legal certainty.

First and from a more procedural point of view, in the old system, legal certainty came from the possibility to notify LTC *ex ante* to the Commission in order to get clearance in case the agreement was not covered by an exemption regulation.³⁸ Since Regulation 1/2003,³⁹ firms and their legal counsels must define the relevant market and self-assess their agreements on that market as well as potential efficiencies pursuant to Art 81(3).⁴⁰ This has increased the regulatory burden on firms and added considerable legal uncertainty, especially in energy.⁴¹ In the former system, considerable uncertainty could already come from the definition of the relevant market.⁴² In liberalizing energy markets, relevant markets are consistently moving with the development of interconnections, which also complicates the self-assessment of efficiencies. In addition, Regulation 1/2003 has

³⁶ DG Competition report on energy sector inquiry, see above note 2.

³⁷ Cameron, *Competition in Energy Markets* (OUP, 2nd ed., 2007), 280.

³⁸ Art 2 and 4(1) Regulation 17: First Regulation implementing Articles 85 and 86 of the Treaty, O.J. 1962, I, 3. However, most notifications were dealt with under comfort letter.

³⁹ Council Regulation 1/2003 on the implementation of the rules on competition laid down in article 81 and 82 of the Treaty, O.J. 2003, L1/I.

⁴⁰ Art 2 Regulation 1/2003.

⁴¹ Even if the Commission may issue in certain cases a guidance letter; see notice on informal guidance relating to novel questions concerning Art 81 and 82 of the Treaty that arise in individual cases, O.J. 2004, C 101/78. However, the length of the exemption procedure and the lack of binding effect hindered the efficiency of the old system.

⁴² Hawk, "System Failure:Vertical Restraints and EC Competition Law", 32 *Common Market Law Review* (1995), 973.

organized decentralization by aligning national competition regimes with EC competition law⁴³ and organizing the shared enforcement of competition rules between the Commission, national competition authorities and national courts, especially as concerns the right to exempt agreements under Art 81(3).⁴⁴ Given the highly political nature of energy markets in most member states, firms fear inconsistent enforcement of competition rules.

Second and more importantly, modernization aimed at implementing a 'more economic' approach based on long-term consumer welfare, which meant gradually shifting from a legal 'form-based' analysis of contracts to a more 'effect-based' approach where the real economic effects of competitive behaviors are more important than the drafting of contracts.⁴⁵ This is expressed in the regular statements of the Commission on the fact that it will take a "case by case" approach to energy cases. Applying a sort of *rule of reason* is already a challenge for competition authorities in most sectors, but applying it in newly liberalized energy markets where the rate of technical change is too low to allow a fast development of competition as in telecommunication might seem intractable in practice and likely to undermine predictability of antitrust enforcement. However, one could argue that the deregulation of network industries rendered the modernization of EC competition policy inevitable. The complexity of competition dynamics in these sectors, as depicted above in the case of LTC, renders a *per se* approach very inefficient. On the other hand, the gains in terms of efficiency must not be offset by the welfare loss from less legal certainty. Applying the 'more economic' approach in deregulated network industries is a good test of the capacity of European competition authorities to fine tune the balancing between predictability and accuracy given a certain level of information. This is also a test of the gains which the society may expect from the modernization of competition law.

⁴³ Art 16 Regulation 1/2003.

⁴⁴ Art 5 and 6 Regulation 1/2003.

⁴⁵ The modernization of EC Competition law has already been widely commented. For a legal account see Ehlermann, 2000, "The Modernization of EC Antitrust Policy: a Legal and Cultural Revolution", 37 *Common Market Law Review* (2000), 537 or Wesseling, *The Modernization of EC Antitrust Law* (Hart Publishing, 2000). For an economic account see Neven, "Competition economics and antitrust in Europe", 21:48 *Economic Policy* (2006), 741-791.

Third, given the highly concentrated market structures in most European electricity markets, LTC are most likely to be caught both under Art 81(1) and Art 82 EC. There is thus a continuum⁴⁶ between Art 81 and 82 EC when enforcing EC competition rules in LTC cases.⁴⁷ Art 81 EC which deals with anti-competitive practices, together with relevant guidelines and notices, does not *a priori* allow or ban LTC, even when they involve dominant firms, unless the agreement contains the so-called 'hard-core' restraints.⁴⁸ It rather provides a framework of analysis to balance anti-competitive aspects and efficiencies according to the contracting parties' market shares and the nature of the restraint involved. Nonetheless, this has not been the case so far with Art 82 EC which tackles abuse of a dominant position. Indeed, the reform of Art 82 EC is recent and still going forward.⁴⁹ This has fostered legal uncertainty since the degree of economic analysis, competition law objectives and methodologies have substantially diverged. Art 82 EC is still based on legal forms and, to a certain extent, protection of competitors, especially when it comes to assessing exclusive dealing clauses.⁵⁰ In particular, pro-competitive and efficiency justifications have been scarcely used under Art 82 EC, which has increased the need to reform in order to achieve

⁴⁶ This expression comes from Lianos, *La Transformation du Droit de la Concurrence par le Recours à l'Analyse Économique* (Sakkroulas, 2007).

⁴⁷ As a general rule an agreement exempted under Art 81(3) EC is unlikely to infringe Art 82 EC, even though the Court has already ruled otherwise, see Case T-51/89, *Tetra Pak Rausing SA v. Commission*, [1990] ECR II-309. For an interesting discussion, see Loewenthal, "The Defense of "Objective Justification" in the Application of Article 82 EC", 28(4) *World Competition* (2005), 461-463. At last, we note that in the case where the LTC has been imposed on the firm by a public authority, the assessment of the agreement with the EC Treaty rules on competition will be pursued under Art 86 EC. See on this the analysis of Eilmansberger, Jaeger and Thiry.

⁴⁸ Hard-core restraints relevant for electricity are market partitioning clauses, use restrictions, minimum resale price maintenance and contractual provisions having similar effects. These restraints contravene the fundamental Treaty objective of market integration and hence will almost never be accepted, which amounts to a *quasi-per se* prohibition.

⁴⁹ See on this the very influential EAGCP report "An Economic Approach to Article 82" (July 2005); the DG Competition discussion paper on the application of Article 82 of the Treaty to exclusionary abuses (December 2005); Lowe, "DG Competition's Review of the Policy on Abuse of Dominance", in Hawk (ed.), *International Antitrust and Policy: Annual Proceedings of the Fordham Corporate Law Institute 2003*, (Juris Publishing, New York, 2004), 163.

⁵⁰ Sher, "The last of the Steam-powered Trains: Modernizing Art 82", *European Competition Law Review* (2004), 244; Jan Slot, "A View from the Mountain: 40 years of developments in EC Competition Law", 41 *Common Market Law Review* (2004), 462; Bellamy and Child, *European Community Law of Competition* (London 2001), 42. Ridyard, "Exclusive Pricing and Price Discrimination Abuses Under Article 82-an Economic Analysis" 23 *European Competition Law Review* (2002), 290-291. For a somehow more contrasting view, see Eilmansberger, "How to Distinguish Good from Bad Competition under Article 82 EC: in Search of Clearer and More Coherent Standards for Anti-Competitive Abuses", 42(1) *Common Market Law Review* (2005), 138.

consistency between Art 81 and 82 as they overlap when it comes to contractual abuses.⁵¹ As a result, the current legal uncertainty regarding LTC in electricity does not only come from the lack of decisions since the opening up of markets or from the legitimate difficulties that the European competition authorities face when regulating the inter-temporal policy trade-offs at stake. It also directly stems from the ongoing evolution of antitrust tools and the flows in their coherence.

However, it is fair to acknowledge that legal uncertainty does not concern all kinds of restraints and all market players. Vertical restraints with market partitioning or use restriction effects have been clearly litigated in energy⁵² and both the Vertical Block Exemption Regulation (VBER)⁵³ and Guidelines on Vertical restraints⁵⁴ (GVR) provide, it is submitted, sufficient guidance for LTC involving non-dominant firms. In fact, legal uncertainty is now concentrated on long-term exclusive supply and purchase obligations involving the former electricity incumbents. Since liberalization, the legal uncertainty arising from the usual lack of precision of Art 81 and 82 EC is aggravated by a lack of consistent, stable and widely accepted methodology for interpretation and application of the EC Treaty rules on competition in a context of market building. This is all the more detrimental to market players given the fast evolutions of both the sector-specific legal framework and the market environment in general. This is thus *in fine* detrimental to final consumers in a sector where the ability to commit in the long-term is crucial to ensure a socially beneficial level of investment. In view of the structural under-investment in generation capacity which the European Union may start to face from the next four or five years onwards, the legal uncertainty currently perceived in the market place becomes a major issue.

⁵¹ See Rousseva, "Modernizing by Eradicating: How the Commission's New Approach to Article 81 EC Dispenses with the Need to Apply Article 82 to Vertical Restraints", *42 Common Market Law Review* (2005), 587-638. Vickers, *Abuse of Market Power*, Speech to the 31st conference of the European Association for Research in Industrial Economics (2004). Kallangher and Sher, "Rebate Revisited: Anti-Competitive Effects and Exclusionary Abuse under Art 82", *European Competition Law Review* (2004). Loewenthal, above note 47.

⁵² For recent cases, see for instance RWE/Transgas in 2006 and 2007 where the Czech Office for the Protection of Competition dealt with problems of market partitioning through destination clauses and discriminatory treatments; and the EUR 208 millions fine imposed by the Bundeskartellamt on seven liquefied gas suppliers (see press release of 19.12.2007).

⁵³ Regulation 2790/1999 on the Application of Art 81(3) to categories of vertical agreement and concerted practices, OJ. 1999, L 336/21.

⁵⁴ Commission Notice on Guidelines on Vertical Restraints, OJ. 2000, C 291/1.

Digging deeper into the new commission methodology

Under a long-term exclusive dealing agreement, the energy company is obliged to meet its entire demand or supply its entire output, or at least a significant part thereof, for the product concerned from or to the dominant firm (usually an incumbent), during an arguably excessive period of time. In the new context, more legal certainty could only come from the clear statement of the relevant facts taken account of by the European Commission and how to interpret these facts in the context of liberalization. Both were missing, it is submitted, until fairly recently. However, for the first time in late 2007, the beginnings of a comprehensive methodology for better analyzing LTC in energy has been sketched out in the *Distrigas* case. This is a clear departure from the pre and even post liberalization period on this issue. The Commission had concerns about liquidity problems on the Belgian wholesale gas market due to LTC concluded by Distrigas with industrial customers and thus opened an Art 82 EC proceeding. The objective of this new methodology was to propose a more integrated framework able to capture the real economic effects of LTC on competition as well as to provide a sound rationale for negotiating remedies. This section shows that, by mixing the *Distrigas* methodology with relevant insights from *Synergen*, *Gas Natural/Endesa*, *Repsol* and *E.ON Ruhrgas*,⁵⁵ the legal uncertainty on the methodology, the relevant facts and the Commission's point of view on the combined relationships of these facts has to a large extent come to an end.⁵⁶ This paper will then show in section IV why this methodology is most likely to be applied in future proceedings across energy sectors, the far reaching consequences in terms of legal certainty and efficiency it will have and the conclusions we may draw on the evolution of antitrust enforcement in the face of the radical uncertainties raised by the liberalization of network industries. When assessing individual cases in energy, the Commission will from now on focus on interactions

⁵⁵ Other relevant cases from national competition authorities will also be used when relevant.

⁵⁶ The remedies will be touched upon in the different paragraphs and discussed in Section IV. The statements of the VBER and GVR confirmed in energy decisions will be highlighted.

among several key elements once the 30% threshold⁵⁷ is exceeded: (i) market characteristics, (ii) competitive position of contracting parties, (iii) the share of the customer's demand tied, (iv) duration, (v) the overall share of the market covered by contracts containing such ties and (vi) efficiencies.⁵⁸

Analysis of market characteristics

Assessing dominance in newly liberalized or emerging markets has been a constant problem in competition policy. Dominance in an emerging market is usually not considered harmful as it often results from an innovative breakthrough and is usually transitory. The Commission is right not to take that path in energy. In the long-term gas supply contract between Spanish incumbents Gas Natural and Endesa, the Commission has concluded that dominance must be assessed even more strictly in highly concentrated liberalizing markets than in more mature settings. Gas Natural being the sole importer and holding more than 90% market share on both free and regulated markets, its dominant position could not be considered transitory.⁵⁹

This conclusion has been confirmed in *Distrigas*.

However, some doubts remain concerning the definition of the relevant product market. The limited development of interconnectors and the restrictions on long-term reservation capacities make LTC between countries unlikely, so the relevant geographic market is likely to remain national for some time. However, recent decisions did not completely settled the question as to whether the wholesale market will be sub-divided according to the different type of customers, namely resellers and big energy users, whose pattern of consumptions is arguably substantially different.

⁵⁷ Art 3 VBER presumes all vertical non-hardcore restraints to be legal so long as the market share threshold of 30% is not exceeded and duration is not indefinite or over 5 years. However, we note that as a general rule exemptions granted under the VBER cannot be pursued when the agreement is between competitors or potential competitors operating at several levels of trade (Art 4). In this case, vertical aspects will be dealt with under the GVR and collusion aspects under the Guidelines on the applicability of Art 81 EC to horizontal cooperation agreements, O.J. 2000, C 118/3.

⁵⁸ This broadly follows the Commission's typology in MEMO/07/40 of 11.10.2007 and in the Sector Enquiry, p.235.

⁵⁹ On this see Fernandez Salas, "Long-term Supply agreements in the context of gas market liberalization: Commission closes investigation of Gas Natural", 2 *Competition Policy Newsletter* (2000), 55-58.

When assessing market characteristics, the Commission will primarily look at future entry in supply and demand, and their real impact on competition. Indeed, in *Synergen*, the Commission stated that the future entry of Viridian which will develop a 340MW power plant would not increase competition intensity due to the “*equilibrium of potential competitive threat*” which should then prevail. The likelihood of new entry in energy depends a great deal on the existence of potential competitors, usually foreign incumbents, especially in markets close geographically. In *Synergen*, the Commission clarified what a potential competitor could be. A potential competitor is usually a firm able to undertake the required investments to enter the relevant market within one year following a small but significant increase in prices. Here the Commission stated that a potential competitor must be judged on the basis of its internal competitive strength: its brand image in the relevant market, ready available capacity in the relevant (gas) market and large financial capacities. The definition of what is a potential competitor is important as the 30% exemption threshold does not apply to vertical restraints between actual or potential competitors. Other barriers to entry such as the level of vertical integration in the market and difficulties in setting up a parallel network of resellers are also key factors to consider (*Repsol*). Forbidding the dominant firm to carry over additional acquisitions of downstream resellers (during two years in *Repsol*) will then be a possible remedy.

The market position of the dominant suppliers

After having analyzed market conditions and their likely evolutions, the focus will be on market shares of the dominant firm and its portfolio of contracts as “*LTC concluded by other suppliers will generally not give rise to concern*.⁶⁰ As in any other sector, the higher the market share of the dominant supplier, the sooner the cumulative market coverage of its LTC will be deemed to create foreclosure. This has to be weighted against the presence of buyer power and whether buyers who represent a substantial part of total market demand on their own are tied for the long term with the dominant supplier (*Gas Natural/Endesa*). In the case of a group of leading suppliers, the Commission will look similarly at the cumulative effects of their LTC but there will be no

⁶⁰ European Commission, MEMO/07/407, 11.10.2007.

need to prove that they lead to tacit collusion (collective dominance) to show that significant foreclosure effects occur.⁶¹ Of course and as recalled in the Sector Enquiry, LTC can be deemed incompatible with Art 81 EC if they result in stabilizing suppliers' market shares over a long period of time and hence lead to collusion. In energy, sole or joint dominance is in most cases self-evident. Except in the old *Almelo* case,⁶² the Commission has not yet dealt with collective dominance in an anti-competitive LTC contract case.

The share of the customer's demand tied under the contract

It is one of the main sources of foreclosure effect. If a customer, all the more if it could have fostered entry for itself,⁶³ must meet all or a big part of its needs with a particular supplier, he does not constitute any longer an available outlet for a potential entrant.⁶⁴ The analysis of the share of the customer's demand tied is closely linked with that of the pattern of consumption. In gas for instance, transaction costs may become too high when negotiating for a small quantity and it may become uneconomic for an alternative supplier to provide less than a certain amount. Competition authorities seem to consider that 20% of a customer demand is the threshold for having incentives to enter into a relationship with a second supplier (implicitly in Art 1(b) VBER, confirmed in *E.ON Ruhrgas*). Some commentators think that the Commission could find foreclosure effects for contracts amounting as low as 50% of a customer demand in case of a network of parallel contracts with the same terms.⁶⁵ In the Commission's view, signing such contract is a way for dominant firms to 'maintain or strengthen

⁶¹ Kjolbye, "Vertical Agreements" in Jones (ed), *EU Energy Law Volume II: EU Competition Law and Energy Markets*, (Claeys and Casteels, 2nd ed., 2007).

⁶² Case C-393/92, *Municipality of Almelo and Others v. Energiebedrijf IJsselmeij NV*, [1994] ECR I-1477.

⁶³ In *Gas Natural/Endesa* (XXXth Report on Competition Policy 2000 and European Commission press release IP/00/297 of 27.03.2000), the Commission reduced the size of the contract from nearly 100% to 75% of Endesa global purchases as Endesa was one of the leading electricity producers in Spain and thus could motivate entry in its own right. In *Thyssengas/STAWAG*, the Bundeskartellamt considered that supplying more than 50% of a major buyer demand on a long-term basis (more than 4 years) could raise antitrust issues (See Lohmann note 13, 95, for the history of the case and Bundeskartellamt press release, November 7 2003).

⁶⁴ We note that if a customer signs several contracts with the same supplier, the Commission will analyze them as one contract to compute the part of the demand tied.

⁶⁵ Schnichels and Nyssens, "Energy" in Faull and Nikpay, *The EC Law of Competition*, (OUP, 2nd ed., 2007).

*their ability to set prices and other conditions on the market*⁶⁶ (confirmed in *Gas Natural/Endesa*). In addition, reduction clauses,⁶⁷ the so-called 'English clauses'⁶⁸ and fidelity rebates granted by dominant firms on remaining volumes are also most likely to infringe competition rules.

The duration of the contracts

The share of the customer's demand tied under the contract has to be analyzed along with its duration. Even if 100% of a customer demand is tied to a particular supplier, foreclosure will not occur if this customer can return to the market every year. However, long-term contracts constitute a barrier to entry when they preclude customers to switch for a more efficient supplier.⁶⁹ As a general rule, the Commission is very suspicious of contracts longer than 5 years and considers that efficiencies generally do not offset foreclosure effects beyond that limit.⁷⁰ It is noteworthy that the Commission

⁶⁶ European Commission, *supra* note 58.

⁶⁷ Reduction clauses allow the buyer to reduce off-take in case the supplier starts reselling in its commercial area. This merely means protecting the buyer's market, which contravenes the fundamental principle of market integration (see *EDF Trading/Wingas, XXXIInd Report on Competition Policy*, 2002, 196). These clauses thus have similar market partitioning effects than exclusive distribution clauses, except that they often entail horizontal restrictions of competition. The Commission will also apply the cumulative effect doctrine to analyze their anti-competitive effects. Clauses of 'right of first refusal' or 'most favored customer' will receive a similar treatment. They remain nonetheless more relevant for gas than electricity. For a complete treatment see Kjolbye, *supra* note 61, 252-270.

⁶⁸ 'English clauses' allow incumbents a right to match the offer of an alternative supplier in case the consumer wants to switch (para.152 GVR). It is worth pointing out that the European Commission has so far never dealt with LTC involving household customers. However, national competition authorities have dealt under Art 82 EC and relevant national provisions with related problems of customer retention strategies by incumbent firms: see *London Electricity* (see the Gas and Electricity Market Authority's Decision under the Competition Act 1998 that London Electricity plc has not infringed the Prohibition Imposed by Section 18(1) of the Act with Regard to a 'Win Back' Offer, 2003), *ENEL Distribuzione* (see Autorita Garante della Concorrenza e del Mercato, press release of 24.10.2007) and *KalibraXE/EDF* (see decision no 07-MC-01 of 25.04.2007, available on the website of the French Competition Council). According to Ofgem, London Electricity abused its dominant position by providing excessive financial incentives to returning customers and subsequently locking them in for a period of 13 months. In *Enel Distribuzione*, the Italian competition authority considered a web of abusive practices of the incumbent and forced the firm to bring its commercial practice in line with competition law principles. In *KalibraXE/EDF*, the French Competition Council did not wait the end of its enquiry on the exclusive dealing clauses of its 2/3 year retail contracts to estimate that the risk of foreclosure was high enough to impose as interim remedy the inclusion of clear termination clauses (appeal is pending). See on this issue: Harker and Waddams Price, "Introducing Competition and Deregulating the British Domestic Markets: a Legal and Economic Discussion", *Journal of Business Law* (2007), 244-268.

⁶⁹ Aghion and Bolton, *supra* note 14.

⁷⁰ Art 5(a) VBER and para.141 GVR.

considers contracts with tacit renewal clauses or no last delivery date as contracts of indefinite duration (confirmed in *E.ON Ruhrgas*) and several contracts signed with the same supplier as one contract. The duration has been and still is an enduring question for competition policy in energy markets. However, recent cases provide, it is submitted, a certain dose of certainty. Duration of contracts accepted by the Commission will mainly depend on the competition position of the counterparty. If the counterparty is an established reseller, duration will not exceed two years as in *Distrigas*. The Bundeskartellamt accepted in *E.ON Ruhrgas* a duration of four years maximum for resellers with more than 50% of overall demand tied, but only two years above 80%.⁷¹ Competition authorities will thus play with the two factors. Where demand requirements are satisfied by several suppliers, the Bundeskartellamt specifies that contracts should distribute the risk of demand fluctuations among suppliers according to the actual supply share provided by each of them so as not to disadvantage the second supplier. In 2005, the Danish Competition Council intervened against a 6 years LTC between the dominant incumbent DONG and the distributors Hovedstadsregionens Naturgas and Naturgas Midt-Nord. It shortened the duration by 2 years to have it terminated by January 2007 and cancelled the exclusive supply clause with a prohibition of such clause in future contracts if they were to renegotiate the agreement. For a new entrant in retail, a duration of 5 years is most likely to be accepted.⁷² One note here that the Commission always thinks in terms of quantities effectively received and not only in terms of contracted quantities. In *Repsol*, a duration of 5 years was accepted for contracts with established resellers (from 25-40 years originally) but the market shares of the dominant firm only reached 30% to 50% in that case, which hardly exceeds the dominance threshold (40%). One also notices a more lenient approach of the Commission towards fuel supply contracts than to producer/reseller contracts. In

⁷¹ Following *E.ON Ruhrgas*, four major gas transmission companies committed in June 2007 to ensure the compatibility of their LTC with EC and German Competition law. See Bundeskartellamt, Decisions B8 - 113/03-6 Bayengas, B8 - 113/03-7 Gas-Union, B8 - 113/03-8 Saar Ferngas of 29 January 2007 and decision B8 - 113/03-15 Vingas.

⁷² We note that in *Direct Energy*, the French Competition Council did not criticize the 5 years duration of the original contract. On *Direct Energy*, see Section IV below and Decision n°07-MC-04 of 28.06.2007 and Decision n°07-D-43 of 10.12.2007, available on the website of the French Competition Council.

Gas Natural/Endesa, the duration has been reduced from 15 to 12 years. This rather long duration, as compared to the 5 years accepted in *Distrigas* for gas supply contracts with power producers and other industrial customers, may be explained by different levels of market opening, the evolution of Commission thinking between 2000 and 2007 and the fact that even dominant firms can claim for some degree of long term security in fuel supply.

The overall share of the market covered by contracts containing such ties

The Commission assesses here the cumulative effect of the parallel network of vertical restraints on market foreclosure. Indeed, LTC can foreclose markets to new entrants only to the extent that a substantial part of market demand is tied for the long term. The fact that a dominant firm be involved in the contract does not change that conclusion from a competition point of view. As a general rule, the Commission considers that a significant cumulative foreclosure effect is unlikely to arise if the total market demand tied does not exceed 30% of global demand. In *E.ON Ruhrgas*, the Bundeskartellamt estimated that the firm contributed significantly to cumulative foreclosure with 75% market shares in its supply area, within a national market where 80% of total demand was tied in the long term. Interestingly, as opposed to the Bundeskartellamt and its rather form-based approach, the Commission in the *Distrigas* case included flexibility parameters for the dominant firm itself. Distrigas was allowed to adjust its portfolio of contracts to its own needs as long as it complies with the duration limitation of 5 years and the 65-70% target. The firm retains by this a fair level of flexibility. Distrigas can thus indifferently have 37.5% of customers supplied under 5 year contracts and 62.5% supplied under one year contracts or 40% supplied under 4 year contracts and 60% supplied under one year contracts. Further flexibility is guaranteed as to protect Distrigas from having to re-open existing long-term gas supply agreements if the volume of gas it supplied decreased. The net effect is that Distrigas can tie under LTC at most 30% of its existing gas supply volumes or 20% of the market, whichever is higher. These commitments will last for a minimum of four years and until Distrigas' market shares decrease below 40% (or another supplier reaches the level of Distrigas market shares minus 20%). Distrigas'

dominance is thus deemed to stop below 40% market shares, which is the traditional threshold for dominance in EC Competition policy.

Efficiencies

The Commission clearly acknowledges that a LTC might be efficiency-enhancing for an individual market player or even for competition in the longer run in case of exemption from Third Party Access for new infrastructure building. However, and as a general rule, the Commission tends to consider that the aggregate effect of those contracts will be detrimental to economic efficiency and consumers from a long-term perspective. In the course of its enforcement practice in recent years, the Commission has made fairly clear what could constitute a pertinent efficiency defense and how it will manage the inter-temporal policy trade-offs raised by LTC.

At first, the Commission has repeatedly accepted the need of LTC for new power plants erection and entry in general.⁷³ In *Distrigas and E.ON Ruhrgas*, restrictions on duration do not apply to new investments in gas fired power plants. In *Synergen*, the Commission accepts both a 15 years gas supply contract with Statoil for 100% of the new power plant needs and a 15 years power purchase agreement for 50% of its production with the electricity incumbent ESB, acknowledging the need of secure dispatch levels to mirror long-term upstream fuel commitments and facilitate project financing. Setting up new power plants is beneficial because it will help ensure long-term generation adequacy and perhaps fuel mix diversity.⁷⁴ Indeed, traditional project finance structures require fuel supply and dispatch contracts lasting longer than 5 years, even in the case of the new hybrid merchant/LTC financial structures. If an investment enables entry, the Commission is highly likely to consider that consumers will receive a fair share of benefits from the vertical restraints, which will fulfill the condition under Art 81(3)b.

The Commission has taken this view to an even greater extent in the field of infrastructure development.⁷⁵ Promoters of new

⁷³ para.44, Commission Guidelines on the application of Article 81(3) of the Treaty O.J. 2004, C 101/97; confirmed in *Synergen* and *Distrigas*.

⁷⁴ Finon and Perez, supra note 20.

⁷⁵ See European Commission press release MEMO/01/76 of 13/03/2001.

interconnectors have been granted either exclusive rights of indefinite duration on the full capacity or a 25 years exemption from third party access (*Viking Cable*,⁷⁶ *Rovigo LNG, Grain 1/2/3, South Hook LNG, Dragon LNG*),⁷⁷ with use-it or lose-it principle though.⁷⁸ Similarly, on the UK-Belgium gas pipeline,⁷⁹ no third party access was required as the Commission judged that the important number of users would allow development of a secondary market. On the other hand, for already existing and amortized interconnectors owned by dominant firms, the Commission deemed long-term capacity reservations to be abuse of a dominant position and required that 100% of capacities be freed up (UK-French submarine interconnector;⁸⁰ Dutch-German interconnector;⁸¹ Norway-Denmark and Denmark-Germany interconnectors following the merger *VEBA/VIAG*⁸²).

However, the mere objective of securing loans might not be sufficient to get exemption as the Commission in other sectors did not always consider it indispensable.⁸³ In future cases, it is likely that energy providers will be required not to prevent their buyers from terminating the exclusive purchase clause and repaying the outstanding part of the loan at any point in time and without payment of penalties.⁸⁴ In addition, the *Synergen* and *Gas Natural/Endesa* decisions clearly showed a different

⁷⁶ *Viking Cable*, notice pursuant to Art 19(3) of Regulation 17, O.J. 2001, C 247/11.

⁷⁷ On *Rovigo LNG, Grain 1/2/3, South Hook LNG, Dragon LNG*, see DG TREN website.

⁷⁸ In Britned, the exemption was limited by the Commission to 10 years as National grid international and Nlink International BV, the two national transmission operators, were deemed to have under-calibrate the capacity of the interconnector to maximize profit.

⁷⁹ *UK/Belgium interconnector*, informal settlement, IP/02/401 of 13.03.2002.

⁸⁰ *UK/France Interconnector*, informal settlement, IP/01/341 of 12.03.2001.

⁸¹ Case C-17/03 *Vereniging voor Energie, Milieu en Water, Amsterdam Power Exchange Spotmarket BV, Eneco NV v Directeur van de Dienst uitvoering en toezicht energie* [2005] ECR I-4983. See Commission Staff Working Paper on the decision C-17/03 of 7 June 2005 of the Court of Justice of the European Communities, SEC (2006) 547, 26 April 2006 and Cameron, *supra* note 37, 343-354. See also Talus and Wälde, "Electricity Interconnectors in EU Law: Energy Security, Long-term Infrastructure Contracts and Competition Law, 32(1) European Law Review (2007), 125-137, where they argue that the ECJ ruling in C-17/03 does not imply, contrary to the European Commission's position, a general ban on long-term contracts on interconnectors, at least for non-dominant undertakings.

⁸² Case M.1673 *VEBA/VIAG*.

⁸³ Nissens and Schnichels, *supra* note 65.

⁸⁴ In case the loan comes from the dominant supplier, it will be considered as an efficiency gain only if it cannot be obtained on the same terms with commercial or investment banks. As a result, to remedy the long duration in *Repsol*, the Commission gave the right to usufruct stations to repay their loan at market value.

treatment according to the market position of the sponsor and contracting parties. If the sponsor is dominant, the duration will be shortened. The reduction of 15 to 12 years in *Gas Natural/Endesa* in 2000 would probably become 5 years today in a more mature market. Similarly, if the off-taker in *Synergen* had not been dominant downstream, the power plant would have probably been allowed to contract 100% of its output over 15 years. It also explains why different remedies will be applied.

One notes here that if long-term generation adequacy is clearly a critical goal of the Commission,⁸⁵ the vague concept of 'security of supply' is approached with much skepticism under competition rules. Prior to liberalization, the idea of ensuring security of supply through fuel mix diversity allowed Member States to secure 20% of the relevant market through LTC between industrial consumers and local producers (*Jahrhundertvertrag*).⁸⁶ Today, only long-term gas import contracts are sure to be accepted on the basis of a 'security of supply' argument,⁸⁷ as long as territorial restrictions are not included.

As shown in the *Synergen* and *Gas Natural/Endesa* cases, long-term contracts often enable the buyer to get cheaper prices. Nonetheless, the parties will have to demonstrate clearly that cost efficiencies are linked with the long duration. Cost savings from coordination will be hard to compute and it will be hard to prove that the consumer benefit outweighs the negative effects of the restriction on competition. In addition, it will be challenging to demonstrate that a sufficient part of the cost efficiency will be passed on to final consumers and that this will outweigh the negative effect of the restriction. In view of the rather large and flexible treatment of Art 81(3)(b) EC by the Commission in energy, a neutral effect on final consumers should be sufficient to pass this test. Once again, cost efficiencies will not be assessed the same way given the market position of the contracting parties. In *Synergen*, the price formula benefits a new power plant and is explicitly

⁸⁵ As evidenced by the new directive on security of supply, Directive 2005/89/EC concerning measures to safeguard security of electricity supply and infrastructure investment, O.J. 2006, L 33/22.

⁸⁶ Hancher, *EU Electricity Law* (Wiley Law Publishing, 1992). See also *Electrabel/Mixed Intercommunal Electricity Distribution Companies* in 1997 where the Commission accepted that a substantial part of the local authorities' electricity requirements still be procured with Electrabel (see XXVIIth Report on Competition Policy, I50 and European Commission press release IP/97/351 of 25.04.1997).

⁸⁷ See Gas Directive, recital 25, quoted above 23.

acknowledged as an efficiency to be counted toward exemption under Art 81(3) EC. To the contrary, the cost efficiency in *Gas Natural/Endesa* is considered to grant an unfair competitive advantage to Endesa and had to be removed. To that extent, *Gas Natural/Endesa* could have been a proceeding based on price discrimination. At last, resale price fixing which does not appear to be a common feature of electricity for the moment, but might become so in the future, are not forbidden *per se* as long as it does not eliminate price competition. Therefore, minimum resale price maintenance will be banned but a maximum price ceiling will be accepted as long as price competition among resellers is economically possible and alignment effects do not occur (Art 4(a) VBER and para.225-228 GVR, confirmed in *Repsol*).

Competition enforcement : legal uncertainty vs. efficiency in the aftermath of *Distrigas*

This paper argues in this section that the methodology depicted above has recently upgraded legal certainty more than market players and commentators tend to usually think and that interesting conclusions can be drawn on the relationship between legal uncertainty and the modernization of EC Competition law. It also argues that the alignment of competition enforcement in energy with enforcement in other sectors expresses the difficulties which the Commission faces when enforcing competition law in deregulated network industries and that it may raise serious concerns for efficiency of future competition enforcement.

The new methodology, legal uncertainty and the modernization of EC competition law

Proposition 1: Legal uncertainty is overstated as the European Commission takes similar views on the optimal mix of contract durations across energy sectors.

The Commission opened in July 2007 two proceedings against EDF and Electrabel for possible breaches of EC Treaty rules on abuse of a dominant position due to their LTC with industrial customers. The Commission argued that '*the cases will take account of the reasoning*

*developed in a competition case concerning Distrigaz and the gas markets in Belgium.*⁸⁸ This is a first indication as to how the Commission will approach LTC in electricity. Beyond the fact that enforcement in gas is logically the best proxy to anticipate future enforcement in electricity, one can see that the few cases we have in electricity since liberalization show that the Commission's view regarding the optimal mix of short and long-term contracts is the same in electricity and in other energy sectors.

Remedies are a laboratory for reform experimentation and DG Competition clarifies its strategy over time. Its main goal when imposing remedies is to improve liquidity in the wholesale market and find a workable mix of contract durations able to accommodate the different market players' needs while limiting foreclosure. As shown above, a maximum duration of about three to four years for contracts with big energy users has been regularly applied in recent oil and gas decisions. When looking at the mix of contracts imposed in remedies in electricity (the so-called Virtual Power Plants), it is interesting to note that the Commission broadly imposed the same durations.

Indeed, in the 2000 *EDF/EnBW* merger proceeding,⁸⁹ EDF was required to auction blocks with durations of three months to three years, amounting overall for one third of eligible consumer demand, during 5 years minimum. In 2001, for the UK-French submarine interconnector, 100% of the capacity had to be freed and auctioned on the basis of 3 years bilateral contracts (1500MW divided in 50MW blocks) and concurrently annual (50MW in 1MW blocks) and daily (150MW in 1MW blocks) auctions. In 2002, in *Synergen*, the original plan submitted to the Commission foresaw that the dominant incumbent ECB would hold 70% of the new 400MW gas-fired power station with the entire plant output to be sold through its retail subsidiary ESBIE, which would clearly reinforce the group's market power on the relevant market. To remedy that situation, the Commission imposed a 600 MW VPP, including 200MW from the Synergen power plant (so for half of eligible consumer demand). The 600MW would have to be sold on the basis of three years bilateral contracts, and subsequently through auctions in case bilateral contracting would not

⁸⁸ See European Commission press release MEMO/07/313 of 26 July 2007.

⁸⁹ Case M.1853 *EDF/EnBW*.

work.VPP or gas release programs are the most important occasions for competition authorities to take on a *quasi-ex ante* regulatory role, which displays with clarity the advancement of their thinking on what a workable mix of contract duration should look like. The durations shown above are similar to what can be found in other energy sectors, which thus tends to show that the Commission takes similar views across energy sectors, at least on durations, and reinforces the idea that the methodology devised in gas will be applied in electricity. Similarly, long durations of 20/25 years were accepted in both sectors for investments in new infrastructure.

Proposition 2: Legal uncertainty is overstated as antitrust enforcement in energy quickly converges with enforcement in other sectors.

In line with past case law, proceedings under Art 82 EC and relevant national provisions, such as *Distrigas* and *E.ON Ruhrgas*, should have entailed a *quasi*-automatic *per se* prohibition. To the opposite, efficiencies and insights from the traditional competition economics of foreclosure such as the analysis of the cumulative market coverage are clearly taken into account in these cases on abuses of a dominant position. We can now safely state that dominant energy firms can implement LTC if the economic and legal contexts allow and that there is no market share threshold beyond which a contractual practice, save hard-core restraints, becomes *per se* illegal.

Most importantly for legal certainty, we can notice that the reasoning and the market share thresholds used under Art 82 in these decisions are clearly similar to the methodology which would have been applied under Art 81 EC and are in line with the VBER and the GVR. Indeed in *Distrigas*, this is primarily the cumulative effect of the network of contracts concluded by the firm which grounds the infringement of Art 82 EC. Historically, the doctrine of cumulative effect on foreclosure has been a cornerstone of the modernization of Art 81 EC and has been regularly endorsed by Community Courts.⁹⁰ It was first treated in the Art 81 EC cases *Brasserie de Haecht* (1967),⁹¹ *Delimitis* (1991)⁹² and

⁹⁰ Verouden, "Vertical Agreements and Article 81(1) EC: The Evolving Role of Economic Analysis", 71 *Antitrust Law Journal* (2003), 525.

⁹¹ Case 23/67, *Brasserie De Haecht v Wilkin*, [1967] ECR 407.

⁹² Case C-234/89, *Stergios Delimitis v Henninger Bräu*, [1991] ECR I-935.

more recently *Langnese-Iglo* (1995),⁹³ *Schöller* (1995),⁹⁴ *Neste* (2000)⁹⁵ or *Van den Bergh Foods* (2003).⁹⁶ This is a well-established tool of competition analysis which *in fine* helps firms themselves analyze if their portfolios of LTC infringe EC Competition law.

One can also take two examples from the field of remedies. The first regards termination rights of existing contracts granted to buyers, which one can find in *E.ON Ruhrgas* and *Distrigas*. In *Distrigas*, existing contracts with energy intensive industries (resellers excluded) enjoyed unilateral termination rights. This is a classical remedy in EC completion policy and the fact that LTC with indefinite durations and clear termination rights are less restrictive than LTC concluded for several years has once again been recalled in the recent Dutch case *Heineken Nederland* and *Neste*.⁹⁷ Second, one can notice that the criteria used to define the duration of commitments becomes in line with what happens in other sectors. Indeed in *Distrigas*, the fact that the commitments will only apply as long as the firm has a market share exceeding 40% and the share of its closest competitors is no more than 20% mirrors a similar approach adopted in the recent *Coca-Cola* case,⁹⁸ where the commitments only applied if the share of Coca-Cola's closest competitor was less than half that of Coca-Cola. At last, traditional market share thresholds under Art 81 EC such as the 30% for automatic exemption defined in the VBER became a benchmark for applying remedies in *Distrigas*. Indeed, the Commission imposed that 65-70% of the firm's customers come back to the market every year as long as Distrigas' market shares still exceed (a very ambitious) 40%, in line with the dominance threshold in other sectors. Another example is the one year duration which renders any exclusive purchase obligation acceptable under Art 81 EC⁹⁹ and has become an explicit target under Art 82 EC as shown above.

⁹³ Case T-7/93, *Langnese-Iglo GmbH v Commission*, [1995] ECR II-1533.

⁹⁴ Case T-9/93, *Scholler Lebensmittel GmbH & Co KG v Commission*, [1995] ECR II-1610.

⁹⁵ Case C-214/99, *Neste Markkinointi Oy v Yötüuli Ky and Others*, [2000] ECR I-11121.

⁹⁶ Case T-65/98, *Van den Bergh Foods Ltd v Commission*, [2003] ECR II-4653, para.82-83.

⁹⁷ In *Kalibraxe/EDF*, the need to clarify termination right is also recalled.

⁹⁸ Case COMP/39,116 . *Coca-Cola*.

⁹⁹ para.141 GVR.

In fact, the policy statements of the VBER and GVR have almost all been confirmed in the course of recent cases in energy. The way to analyze market characteristics and patterns of consumption, the suspicion towards contracts longer than one year and tacit renewal clauses, or even the principles to analyze efficiencies which are in line with the Guidelines on the application of Art 81(3), all show that a unified approach among competition provisions of the EC Treaty is emerging.

The recent decisions analyzed above, and essentially *Distrigas*, show that the reform of Art 82 EC is well under way and it would not be unrealistic to assert that the European Commission is using the deregulation of network industries to complete the modernization of EC Competition law. This trend is confirmed by the opinion of Advocate General Damasco Ruiz-Jarabo Colomer in Case C-468/06 presented the 1st of April 2008, where he encourages the European Court of Justice to clearly state that a *per se* approach is not applied any more in Art 82 cases, even if there is no doubt about the eliminatory intent of the dominant firm, and that anti-competitive effects must be weighed against potential gains for the consumer as is current under Art 81 EC. In the future, energy companies should thus less and less face discrepancies between enforcement under Art 81 and 82 EC, and between enforcement in energy and other sectors, which shows that the legal uncertainty currently perceived in the market place is largely overstated.¹⁰⁰

Proposition 3: The quick alignment of antitrust enforcement in energy also demonstrates the limits of the application of the “more-economic” approach in deregulated network industries

The rationale of the more economic approach in EC competition policy is to better capture industry specifics. Yet, there is no reason to believe that energy at this stage of the liberalization process must be analyzed as the beer or ice-cream sectors, except if energy truly converged with these industries which is not the picture we can find in the Sector Enquiry. True, applying some analytical

100 This effect, which was predicted in the EAGC report and Rey, “On the right test for exclusive dealing”, in Ehlermann and Marquis (eds.), *European Competition Law Annual 2007: A Reformed Approach to Article 82 EC* (2008), is evidenced here.

devices such as the cumulative effect doctrine does bring some relevant insights for competition enforcement in energy. Indeed, seeking consistency in enforcement across sectors is a legitimate goal of competition policy.

However, it seems that the application of the new methodology also expresses a path dependency in competition enforcement and the difficulties the Commission currently faces in energy. When this paper argues that legal certainty has recently been upgraded in electricity, this does not come from a new methodology able to capture real economic effects with a high level of consistency, as purported for instance in the EAGCP report, but from a methodology which the Commission knows, can easily apply and *in fine* be anticipated. The only specificity introduced by the European Commission was the flexibility granted to the dominant firm when applying remedies. This has to be recognized but the Commission could have gone further. Remedies could for instance be gradually decreasing in strength before the dominance threshold, which would be the proof that a proportionality test is really applied under Art 82 EC and that remedies should evolve with a firm going from a "super-dominant"¹⁰¹ to a dominant position.

As a result, the new methodology as applied in recent decisions in energy is closer to an improved *per se* rule than a true unstructured *rule of reason*. Given the limited information at hand on the specific competition dynamics of market building in these sectors and the limited practical insights to be drawn from energy economics, it is not sure that this coping strategy is really sub-optimal from an efficiency point of view. In addition, from a more procedural point of view, we cannot but notice that a limited level of discretion is more suitable to a decentralized application of EC competition law. However, the more-economic approach finds here its limits and providing guidance through cases as the Commission is doing at the moment certainly does not fulfill the objective of legal certainty. Retaining that way some flexibility for future competition enforcement is understandable, but this comes at an unknown cost.

¹⁰¹ As defined by Advocate General Fennelly in Case C-395/96P & C-396/96P, *Compagnie Maritime Belge Transports SA and Others v. Commission*, 2000 ECR II-1365, para. 137; see also *Deutsche Post AG – Interception of cross border mail*, OJ 2002, L331/40.

The new methodology of the European commission in energy : unfinished business ?

This paper aims in this final section to highlight certain theoretical and practical problems which the Commission will face when enforcing competition rules in the near future.

LTC and spot market development: the black box. First, one cannot but notice that the main issue debated by economists when it comes to LTC in electricity is neither taken account of nor debated whatsoever in any decision of the Commission.¹⁰² The impact of LTC on spot market deepening is a complicated issue though. On the one hand, there is a negative effect in the sense that LTC would dry out spot markets, which increases price volatility and incentivizes players to contract bilaterally. This lack of liquidity on spot markets is what the Commission aims to fix. However on the other hand, LTC are a good mitigation device, for certain market structures and types of competition,¹⁰³ when firms may either abuse their market power by strategically withholding capacities or tacitly colluding.¹⁰⁴ These conclusions come from recent economic analyses which depart from the traditional assessment of market structure using e.g HHI index to use instead oligopoly models. As a result, some authors even propose to impose a tax on market players who do not contract long term a substantial part of their supply.¹⁰⁵

The Commission implicitly recognizes the need of different contract durations when imposing remedies, but always for sake of production planning, and not for a smooth development of spot markets. By choosing to tackle the lack of liquidity rather than potential abuses of market power, the Commission seems to take a legitimate course of action as more than 95% of electricity in the EU remains contracted

¹⁰² As Buschnell 2007 states: « *The competitive implications of the ability of firms to trade in transparent forward markets has received considerable attention in the academic literature. Their implications have not had much implication on policy however.* »

¹⁰³ Allaz and Vila 1993, note ?; Mahenc, P., & Salanie, F. (2004). « Softening Competition Through Forward Trading. » *Journal of Economic Theory*, 116, 282–293., Bushnell J., 2007, Oligopoly Equilibria in Electricity Contract Markets. *J Regul Econ* (2007) 32:225–245. Anderson E and X. Hu, Forward contracts and market power in an electricity market, *International Journal of Industrial Organization* 26 (2008) 679–694.

¹⁰⁴ Green, R. J., & Le Coq, C. (2006). The length of contracts and collusion. CSEM Working Paper WP-154, University of California Energy Institute.

¹⁰⁵ Willem B., Market Power Mitigation by Contracts, 2006 Working Paper.

bilaterally. However, the Sector Enquiry¹⁰⁶ acknowledges recurrent problems of abuse of a dominant position on European spot markets. As a result, even if these analysis still tends to rely on fairly strong assumptions, the Commission should not ignore the questions raised by recent advances in economic theory.

Remedies in energy and entry in generation. A second potential problem lies in the way the Commission analyzes market fundamentals and the resulting opportunities for entry in generation. The purpose of the methodology is to better analyze the competitive situation on a given market and then impose remedies to remove barriers to entry. The Commission bets that the long-term benefits of competition will offset the short-term costs incurred by individual players and thus increase social welfare aggregated over several periods of time. The whole methodology must therefore be based on a robust understanding of the pattern of entry. From that perspective, the almost systematic imposition of VPP (or gas release) is, it is submitted, a source of concern.¹⁰⁷

VPP are primarily intended to remedy horizontal concentration at the generation level and increase liquidity on the wholesale market. They force dominant firms to make capacity options available for a pre-determined time horizon, which amounts to a virtual divestiture of capacity. As such, VPP are a way to tackle concentrated market structures in merger and antitrust proceedings when physical asset divestiture is not feasible.¹⁰⁸ VPP are thus hybrid remedies, between structural and behavioral, which should facilitate entry by cancelling the need to invest in generation. In the Commission view indeed, VPP is part of a two-stage strategy where a first wave of entry in retail must create new outlets which will attract entry in production by independent power producer or at least enable resellers to build a sufficiently stable

¹⁰⁶ DG Competition report on energy sector inquiry, *supra* note 2.

¹⁰⁷ VPP have for instance been implemented as remedy in *EDF/EnBW*, *Nuon/Reliant Energy*, *Synergen*, *Direct Energy*. More surprisingly, it has also been imposed by the Bundeskartellamt to RWE in the context of a proceeding concerning abusive pass-on of CO2 certificates to consumers. The auction concerns 46 millions MW of both base-load electricity from lignite power station and pick load electricity from a new hard coal power station. On this, see Bundeskartellamt press release of 27.09.2007.

¹⁰⁸ Competition authorities traditionally favor physical divestitures as it limits subsequent monitoring costs.

customer base to subsequently integrate backward. As any LTC, VPP might have mitigation effects on abuse of market power by dominant firms in the spot market but there are few studies quantifying these effects on firms' strategic bidding and equilibrium prices.¹⁰⁹

As a result, there is to date no convincing evidence of positive effect of VPP on competition.¹¹⁰ This can be explained by the fact that the efficiency of VPP will depend on many factors such as auction design,¹¹¹ contract durations or the investment climate, which have not been systemically analyzed, neither theoretically nor empirically.¹¹² Indeed, the main effect of VPP might well be to deter investment in new capacity, which goes counter the objective of long-term generation adequacy. In balancing the contradictory incentives for entry in retail and production, the length of the VPP is thus important and implementing VPP for periods longer than the period of decision and construction of a new power station does not seem necessary.¹¹³ In addition, the proceedings themselves as well as the monitoring of remedies over many years are not costless. If long-term VPP or gas releases are imposed, or if competition authorities must monitor portfolios of contracts over a long period of time, competition authorities will be durably involved in the day-to-day monitoring of deregulated network industries, taking up a *quasi-ex ante* regulatory role for which they might not be prepared.¹¹⁴

Remaining uncertainties concerning the antitrust treatment of LTC involving energy intensive users. Strong uncertainties remain concerning the different collective buying schemes for energy intensive users¹¹⁵ which might be found at all levels of the supply chain in both

¹⁰⁹ For an attempt, see Boisseleau and Giesbertz, "Assessing Regulatory Measures in Electricity Markets: The Case of VPP in the Netherlands", 29th IAEE International Conference, 2006, Potsdam, Germany, conference proceedings.

¹¹⁰ Boisseleau and Giesbertz, *supra* note 110.

¹¹¹ In most cases, VPP define base and peak load rights with different durations granted through an ascending clock auction.

¹¹² Boisseleau and Giesbertz, *supra* note 110.

¹¹³ As decided by the French Competition Council in *Direct Energy*, see also Léveque, "Le Conseil de la Concurrence au secours des opérateurs alternatives de l'électricité", *Revue Lamy de la Concurrence*, (2008).

¹¹⁴ Leveque 2008, *supra* note 114.

¹¹⁵ See DG Competition report on energy sector inquiry, *supra* note 2, 204-205.

gas and electricity.¹¹⁶ Energy intensive users, who once lobbied for the opening up of markets, pretend now to be squeezed between rising energy costs and the impossibility of passing them on downstream due to the stark international competition. This has pressured certain Member States like Spain and France to act on that question and fight the risk of delocalization.

Prior to liberalization, in *Jahrundertvertrag*, the Commission had cleared joint coal sales and purchasing consortia based on Art 65(2) ECSC Treaty but the exemption reasoning is unlikely to be replicated today. In the new context, these schemes give rise to significant competition problems which are not all answered by applying the new methodology. Given that no Block Exemption Regulation applies and assuming that the *de-minimis* thresholds are exceeded, these schemes will be assessed under Art 81 EC,¹¹⁷ Art 82 EC if a dominant supplier participates, or State Aids if a government or a public firm is involved. To get the exemption under Art 81(3) EC, this is the fulfillment of Art 81(3)(b) which will constitute the biggest hurdle. Indeed, cost efficiencies will not be hard to trace back to the buying scheme and the Commission will probably accept that they will benefit *in fine* final consumers. Under Art 81(3)(b), the agreement in question must not give to the firm "*the opportunity of eliminating competition in respect of a substantial part of the products in question.*" This is where the new methodology brings relevant insights, especially on the cumulative market coverage and the interaction of duration and the percentage of the customer demand tied from a competition perspective. However, it is far from clear what the Commission position is on the likely collusive aspects inherent in such schemes and how it would assess competition effects on markets upstream and downstream of the joint purchasing consortium.

In addition, the opinion issued by the French Competition Council on Exeltium¹¹⁸ in 2005 may cast doubt as to the definition of the relevant market. Indeed, the Council took as relevant only the market for eligible customers who effectively switched. In addition, to assess

¹¹⁶ This could also concern joint buying of technical equipments for construction services, e.g. C4 gas joint venture set up by Fluxys, GDF International and Transco, O.J. 2002, C 166/8.

¹¹⁷ Vertical and horizontal issues will be assessed under the relevant guidelines as well as under the Notice on Art 81(3).

¹¹⁸ See Opinion n°05-A-23 of 5.12.2005 from the French Competition Council.

horizontal competition aspects at the supplier level, it estimated that the supply of 15-20 years LTC by EDF would not distort competition as alternative suppliers with production capacities in France were not willing to commit for such a long period and importers would be most unlikely to compete with EDF given the current restrictions on interconnections. As a result and following the Council reasoning, due to the specific pattern of consumption of energy intensive users, a criteria taken into account in *Distrigas* and *E.ON Ruhrgas*, LTC with such purchasing consortium would constitute a different relevant market where the long duration would not be a problem or at least would be counterbalanced by efficiency gains.

This opinion is all the more striking since it is fairly recent and seems to go counter the late practice of the Commission. The only recent guidance we have concerning duration for LTC with energy intensive users comes from the *Distrigas* case where a 5 years duration is accepted, with termination rights.¹¹⁹ In the context of purchasing consortia, giving unilateral termination rights to the members could be a way to balance benefits from more predictability and cheaper prices against costs from not being able to opt out of the deal to benefit from a better offer. However, the *Distrigas* decision might not be replicable in this context as cost efficiencies arising from these buying schemes might be much more important than for traditional LTC. Indeed in Exeltium, the supplier does not contract directly with buyers but with a highly leveraged project financed special purpose vehicle (debt/equity ratio around 90/10) which enables the buyers to finance off balance-sheet their electricity purchase. In view of this, it is hard to anticipate which direction the Commission will take.

Furthermore, if a dominant firms supplies the consortia below its total costs to drive competitors out of the market, it can be caught by the Commission under Art 82 EC for predatory pricing, or at least for price discrimination. In a case where the supplier is a public firm, as EDF in France in the case of Exeltium, then the Commission could proceed under State Aid if prices are deemed not to reflect market conditions, which will be highly complicated to demonstrate given the limited duration (around 3 years) on European forward markets and the strong uncertainties on prices in the next 15 to 20 years due

¹¹⁹ In *E.ON Ruhrgas*, LTC with energy intensive users are unaffected by the prohibition.

to fuel supply costs, costs of CO₂ emissions and demand evolution. State Aid could also lie in the preferential fiscal regimes granted to intensive users, for instance in Spain (G4 Tariff) and foreseen in France (SOFIBASE). Finally, as pointed out by the French Competition Council, the most intractable problem from a competition point of view could lie in the discriminatory conditions for access to the buying scheme. Additionally, single market concerns exist with these schemes as intra-European competition might be hampered by the cost advantage derived from the different technology portfolios of EU member states, if foreign firms cannot participate due to interconnections restraints or political pressure.

Linked to the problem just described is the position of the Commission regarding the renewed interest for nuclear power in several European countries, especially the UK and recently Italy. In the *Scottish Nuclear* case,¹²⁰ almost 20 years ago, Scottish Nuclear was allowed to sign a 15 years supply contract with Scottish Power and Hydro-Electric. The Commission at that time explicitly recognized the need of long-term dispatch and planning for reaching the scale economies of that technology, even if it hindered price competition between the downstream duopolists.¹²¹ It would be interesting to know whether this jurisprudence is still relevant today and whether such LTC would be acceptable for supply to a single downstream company or consortium, and on what basis.

Price restraints and the location of regulatory powers in energy LTC cases. Apart from resale price maintenance in *Repsol* and to some extent price discrimination in *Gas Natural/Endesa*, the Commission has so far never dealt with price issues in LTC cases. Community and national competition authorities have however some experience in the context of abuse of market power on spot markets¹²² and discriminatory access to essential facilities. This has changed recently with the important *Direct Energy* case dealt with by the French

¹²⁰ Scottish Nuclear decision of 6 July 1991, O.J. 1991, L 178/31.

¹²¹ This LTC was deemed to facilitate transition to a market-based industry in the UK and ultimately to benefit consumers.

¹²² See for recent cases *Viesgo Generacion* where the Spanish Competition Authority found that ENEL had charged abusive prices on the spot market (Tribunal de Defensa de la Competencia, Report 2006) or *Iberdrola Castellon* where the Spanish Competition Authority found that Iberdrola had manipulated wholesale prices.

Competition Authority in late 2007. Direct Energy is a new entrant in French retailing with no production capacity. Due to the overwhelming dominant position of EDF, restraints on interconnectors and a limited spot market, there was no alternative to bilaterally contract with EDF for a substantial part of its base load demand. Direct Energy and EDF thus signed in December 2005 a 5 year contract at a price fairly comparable to that of forward contracts with the longest maturity on Powernext at that time. Following this contract, Direct Energy remained unable to compete with the French incumbent retail subsidiary. It thus decided to sue EDF for an infringement of Art 82 EC mainly on three grounds: price discrimination (between Direct Energy and the EDF retail subsidiary), margin squeeze abuse and discriminatory access to nuclear base load capacities.

The tremendous difficulties of the Council in this case remind the difficulties the competition authorities face as soon as price issues are involved.¹²³ Concerning the margin squeeze aspect of the case, the Council recognized that Direct Energy was squeezed out between the EDF retail tariffs and the high and unstable prices to be found on wholesale markets where EDF was dominant. The Council could only conclude on a presumption of margin squeeze. On the remaining two issues, the Council is much less clear. On the price discrimination aspect, the Council simply stated that it could not conclude due to missing evidence and that further analysis was required as the competitive 'benchmark' price was highly difficult to determine in energy. Even if the Council deemed that access to nuclear capacities was not discriminatory, he acknowledged the need to organize access for alternative suppliers. As a remedy, the Council imposed a VPP to EDF.

The Council thus revived the debate which took place in the telecommunication sector about whether or not to regulate *ex ante* access to upstream products and whether this is an acceptable second best. This also raised the debate about the suitability of competition authorities in dealing with price restraints in general and whether regulatory authorities should not be given jurisdictional

¹²³ Another recent case concerning difficulties to prove anti-competitive pricing in energy, though not in electricity, is *Austrian Airlines/OMV* where the Austrian Federal Competition Authority could not assert with the necessary degree of certainty that OMV was charging excessive prices and thus referred the case to the Cartel Court (Bundeskartellamt, Report 2006-2007).

power on this issue, especially to avoid deterrent effects on entry in generation.¹²⁴ However, since the ECJ upheld the 10th of April 2008 the Commission decision on *Deutsche Telekom*¹²⁵ and confirmed the legality of its methodology, this is unlikely to be the case.

Procedural issues of the new methodology. At last, the end of *ex ante* notification has added a layer of uncertainty and potential problems for efficiency in enforcement in a sector where *ex post* monitoring might not be optimal. Indeed, *ex post* monitoring is likely to be the optimal audit regime only when the competition authority's probability of error is low.¹²⁶ At least relatively to enforcement in unregulated sectors, it is unlikely to be the case in a deregulated network industry.

Conclusion

Due to their ambiguous effects on competition, investment and welfare, long-term vertical contracts will remain a key issue of competition enforcement in the EU electricity markets for many years to come. Even if some uncertainty remains, this paper has showed that a clear methodology emerged from recent decisions in other energy industries and that this methodology is most likely to be applied in electricity. The strong legal uncertainty currently perceived in the market place is largely overstated but the European Commission would be well advised to clearly and publicly state its strategy given the negative externalities it may create in energy markets and *in fine* the society as a whole.

The analysis of the parallel development of EC competition law and energy markets liberalization raised interesting insights on antitrust

¹²⁴ For a discussion in the context of the telecommunication sector, see Geradin and O'Donoghue, "The Concurrent Application of Competition Law and Regulation: the Case of Margin Squeeze Abuses in the Telecommunications sector", 1(2) *Journal of Competition Law and Economics* (2005), at 355-425. Fletcher and Jardine discuss the pros and cons of price regulation by competition authorities for exploitative conduct in Fletcher and Jardine, "Towards an Appropriate Policy for Excessive Pricing", in Ehlermann and Marquis (eds.), *supra* note 100.

¹²⁵ Case T-271/03, *Deutsch Telecom v Commission*, 10 April 2008, not yet published.

¹²⁶ Loss et al., "European competition policy modernization: From Notifications to legal exception", 52 *European Economic Review* (2008), 77-98.

enforcement in deregulated network industries. In face of a radically new context, the European Commission largely disregarded sector specifics and simply replicated competition analysis it had devised in other sectors, especially beer and ice-cream. This cannot only be explained by the implicit antitrust objective of fostering consistency across sectors of the economy but rather expresses the difficulties competition authorities tend to face with the liberalization of network industries. In electricity, the pattern of entry in generation and the impact of LTC on spot markets remain little known and this is precisely where the new methodology implemented by the Commission will display its most important shortcomings. While the Commission states in the explanatory memorandum attached to the Third Energy Package presented in January 2008 that it will provide soon guidance on LTC, this paper shows that many dimensions might not be addressed.

Signaux-prix et équilibre de long-terme

Reconsidérer l'organisation des marchés électriques

Dominique Finon

CIRED & LARSEN

Christophe Defeuilley

LARSEN

Frédéric Marty

(CNRS, Université Nice Sophia-Antipolis), OFCE & LARSEN

Les industries électriques ont été libéralisées en référence au modèle d'équilibre de marché. Dans ce modèle, les agents se font concurrence aux différents étages de la filière sur des marchés de gros et de détail où les prix (actuels et en espérance) suffisent à articuler équilibre de court terme et équilibre de long terme. L'accent est mis sur la structure de marché dé-intégrée verticalement, qui est supposée permettre l'optimalité des décisions d'agents décentralisés à partir de prix de court terme et de marchés supposés complets. Ce modèle du marché décentralisé présente plusieurs défauts. D'une part le marché ne permet pas d'assurer un équilibre physique en pointe et en toute circonstance. D'autre part le marché ne permet pas d'orienter correctement les investissements pour aboutir à un parc optimal de production. Il y a 'défaillances de marché' en ce qui concerne la question centrale des investissements en production. Il faut dès lors accepter que les intervenants sur les marchés puissent utiliser d'autres modalités de coordination (contrats de long-terme, intégration) pour corriger ces défauts et assurer une gestion de risques que le marché ne prend pas en charge. Ce qui nécessite un aménagement du cadre d'organisation du secteur électrique allant vers une reconnaissance de l'utilité de cette coordination hors-marché. La concurrence imparfaite pourrait être vue non plus comme le symptôme d'une réforme inachevée mais comme le palliatif à des défaillances de marché.

Introduction

Les industries électriques ont été libéralisées en référence au modèle d'équilibre de marché qui ont inspiré et inspirent encore réformateurs et régulateurs. Dans ce modèle, les agents se font concurrence aux différents étages de la filière sur des marchés de gros et de détail où les prix (actuels et en espérance) suffisent à articuler équilibre de court terme et équilibre de long terme. L'accent est mis sur la structure de marché dé-intégrée verticalement, qui est supposée permettre l'optimalité des décisions d'investissement d'agents décentralisés à partir de prix de court terme et de marchés supposés complets. Toutefois la complexité de l'industrie met à mal ce schéma. Produisant et fournissant un produit non-stockable sous obligation de fourniture, l'industrie électrique présente la particularité que l'on doit recourir à un éventail de technologies de production de structures de coût très différentes pour satisfaire des demandes variables d'une heure à l'autre. Du fait du caractère non stockable de l'électricité qui oblige à organiser le marché physique en marchés horaires, de la non-transmission des prix horaires de gros dans les prix finaux, de l'inélasticité des fonctions de demande horaire et de la présence de biens publics (la sécurité de fourniture), le modèle du marché décentralisé présente plusieurs défauts. D'une part le marché ne permet pas d'assurer un équilibre physique en pointe et en toute circonstance. D'autre part le marché ne permet pas d'orienter correctement les investissements pour aboutir à un parc optimal de production conduisant à des prix de marché les moins élevés en moyenne annuelle. Nous sommes confrontés à des 'défaillances de marché'¹.

L'objet de cet article est de montrer que la réforme du secteur électrique, fondée sur une grille de lecture où l'articulation court-terme / long-terme est effectuée par le prix de marché ne conduit pas à un résultat efficient en termes d'investissement. Des défaillances de marché provoquent des distorsions dans les choix : sous-investissement en pointe, biais dans les choix techniques. Il faut dès lors accepter que les intervenants sur les marchés puissent utiliser d'autres modalités de

¹ On considère qu'il y a une défaillance du marché lorsque le fonctionnement spontané du marché engendre une situation qui n'est pas optimale. Notons que nous ne traitons pas ici de la question des investissements concernant les réseaux électriques (transport et distribution), tout aussi capitalistiques que les investissements en production, dans la mesure où ils relèvent d'un autre contexte institutionnel, hors marché (celui de la régulation en monopole naturel).

coordination (contrats de long-terme, intégration) pour corriger ces défauts. Ce qui nécessite un aménagement du cadre d'organisation du secteur électrique allant vers une reconnaissance de l'utilité de cette coordination hors-marché. La concurrence imparfaite pourrait être vue non plus comme le symptôme d'une réforme inachevée mais comme le palliatif à des défaillances de marché.

Notre propos se structure en quatre étapes. Dans un premier temps les éléments de base de la littérature 'standard' de l'investissement en incertain, qui sert de socle théorique à cette articulation court-terme / long-terme, sont exposés. Puis on présente le modèle des marchés électriques décentralisés sur lequel s'adossent les réformes en mettant en lumière les questions relatives à l'investissement en production laissées en suspens, et qui peuvent conduire, sur le long-terme, à des choix socialement inefficaces. Enfin, deux solutions d'organisation industrielle susceptibles d'y remédier, les contrats long-terme et l'intégration verticale et horizontale, sont analysées.

Signaux de prix et décision d'investissement en environnement incertain

Les réformes dans le secteur électrique accordent un rôle central aux coordinations marchandes et au prix comme signal pour orienter les décisions et allouer les ressources, aussi bien à court-terme qu'à long-terme. Les signaux-prix de court-terme sont réputés guider la prise de décision en matière d'investissement, même dans un environnement incertain.

Comme dans la fiction du marché parfait de la théorie de la tarification optimale au coût marginal de long terme, le marché 'concret' est supposé reproduire de façon parfaite les prix optimaux assurant l'articulation court terme-long terme en garantissant un niveau d'investissement et une répartition entre technologies socialement optimale pour suivre le développement des demandes horaires. Les prix de ces marchés horaires sont certes très variables, mais on va supposer qu'il existe une relation inter-temporelle des prix de court terme susceptible d'aboutir à des choix socialement efficaces sur

longue période, notamment grâce à l'existence d'instruments financiers de couverture du risque (l'hypothèse des marchés complets d'Arrow).

Dans le schéma de la tarification marginaliste du monopole de service public, l'entreprise pouvait planifier et optimiser le développement de son parc d'équipements pour définir ses prix tout en évitant la difficulté posée par la variabilité de ses coûts marginaux de court-terme. Ceci est idéalement assuré par la correspondance entre courbes de coûts marginaux de court terme (CMCT) et courbes de coûts marginaux de long terme (CMLT), sa courbe enveloppe, lorsque le parc est adapté (Boiteux M., 1956). Le prix au CMLT permettait en quelque sorte le lissage des CMCT tout en assurant le recouvrement des coûts d'investissement. C'est cette liaison inter-temporelle que reprend implicitement le modèle théorique du marché décentralisé dans un contexte où l'on prendrait en compte les effets de l'incertitude, de l'irréversibilité et des anticipations des acteurs en concurrence.

Mais cette relation inter-temporelle dans le modèle théorique n'est respectée qu'au prix de la mobilisation d'hypothèses fortes sur la nature de l'incertitude et sur les anticipations rationnelles des agents en concurrence, qui éloignent les modèles des situations réellement rencontrées sur les marchés. Lorsqu'elles sont relâchées, on pourrait imaginer que l'utilisation d'instruments financiers de couverture du risque pourrait être la solution à notre problème d'articulation court terme-long terme. Mais dans les faits elle n'est pas d'un grand secours pour les investisseurs s'engageant dans des décisions de long-terme et lourdes en capital. Nous allons donc d'abord remettre à plat les hypothèses implicites au modèle d'équilibre d'investissement en univers incertain avant d'en identifier les limites.

La théorie standard. La théorie 'standard' assigne aux prix un rôle central dans la coordination inter-temporelle en articulant sans distorsion les décisions de court terme et de long terme. Elle confère un rôle informationnel aux prix en raisonnant en avenir probabilisable et dans un univers quasi-statique en ce qui concerne les choix d'équipement et l'innovation. Les technologies sont supposées spontanément émerger d'un panier de technologies à des dates prévisibles et des effets d'apprentissage connus en correspondance avec l'évolution attendue des prix relatifs des facteurs de production.

Dans ce cadre, la décision d'investissement, lorsqu'elle s'applique à des choix irréversibles en situation d'information croissante, doit tenir compte des valeurs d'option qu'il peut y avoir à reporter certaines décisions (Arrow K. et Fischer A., 1974 ; Henry C., 1974). L'information étant croissante, l'apparition de nouvelles informations susceptibles de modifier la décision ou d'en faire varier les bénéfices espérés, doit être valorisée et c'est l'objet de cette quasi-option de le faire. Elle mesure donc la valeur accordée par les agents à la flexibilité dans la prise de décision.

Dans un secteur très capitaliste comme celui de l'électricité, les décisions portent sur des investissements totalement ou partiellement irréversibles. Les dépenses consenties sont irrécupérables dans la mesure où l'investisseur ne pourra pas recouvrer tous ses coûts s'il choisit de se retirer lorsque les conditions de marché deviennent défavorables (celles-ci affectant le revenu futur escompté pour tout nouvel acheteur). Compte tenu de l'importance des coûts fixes et de contraintes techniques (particulièrement fortes dans le cas de la plupart des moyens de production), il ne pourra plus mener à bien la construction de l'ouvrage, puis décider de ne pas l'utiliser, sans que cela conduise à de lourdes pertes.

Toutefois, il bénéficie d'une certaine liberté d'action dans le timing de son investissement : il peut le retarder, voire le différer s'il considère qu'il peut être avantageux pour lui d'attendre afin de disposer d'une meilleure information sur les évolutions à attendre des principaux paramètres qui impactent la rentabilité de son projet. C'est notamment le cas lorsque les conditions de marché s'avèrent particulièrement incertaines. L'opérateur dispose d'une option d'investissement, qu'il exerce lorsqu'il déclenche sa décision d'investir. Il abandonne alors la possibilité d'attendre et de gagner une meilleure information pour réduire l'incertitude entourant la rentabilité de son projet. Donc, la valeur de l'option perdue (non utilisée) représente un coût d'opportunité qui doit être considéré lors de la prise de décision².

Il en découle que s'ils anticipent une amélioration de l'information, les investisseurs auront une préférence pour la flexibilité et pourront choisir d'attendre avant de s'engager dans des décisions irréversibles (Dixit A. et Pindyck R., 1994 ; Caballero R., 1991). La prise en compte de la valeur de cette option rend plus contraignante les conditions

² Voir C. Chaton (2001) pour une application des décisions d'investissement sous incertitude au secteur électrique.

de rentabilisation d'un investissement totalement ou partiellement irréversible³. Parce que la décision d'attendre avant d'investir affecte la valeur du projet, calculée selon la formule classique de la VAN (Valeur Actualisée Nette) espérée. Il est alors nécessaire d'aller au-delà de ce calcul en prenant en compte la valeur du report d'investissement, c'est-à-dire la valeur du gain d'informations et de la réduction de l'incertitude. Cela affectera le calcul des flux actualisés de ressources et de dépenses liées au projet, calcul qui sert de socle à l'établissement de la VAN (Pindyck R., 1991).

Mais, en tout état de cause, la décision d'investir dépendra toujours d'une estimation des flux actualisés de recettes et de dépenses sur la durée de vie de l'investissement. Cette estimation s'appuie sur une représentation probabiliste de l'incertitude, qui est généralement de la forme d'une marche aléatoire. C'est-à-dire que le prix actuel évoluera vers un prix futur choisi au hasard.

Si le prix actuel est élevé, et compte tenu d'une distribution gaussienne des probabilités, la meilleure prévision possible dans une marche aléatoire pour l'ensemble des prix futurs conduit à retenir un niveau élevé. Autrement dit, si le prix actuel est élevé, la marche aléatoire se déplace vers le haut et aboutit à une espérance de prix plus élevée dans le futur (Dixit A., 1992).

Ce qui signifie que les prix actuel et futurs sont liés entre eux, le prix actuel jouant comme un signal susceptible de déclencher l'investissement lorsqu'il se traduit par des espérances de flux actualisés de recettes favorables. Cette séquence inter-temporelle de prix ne fonctionne en espérance que si l'on considère qu'aucun choc (variation inattendue du niveau de demande, choc durable de prix des combustibles, irruption d'une nouvelle technologie, etc.) ne viendra la perturber tout au long de la durée de vie de l'investissement. C'est à ce prix qu'une firme, soumise au passage du temps, pourra définir le timing et le niveau optimal de son investissement

Le raisonnement est étendu au niveau d'un secteur tout entier placé en concurrence pure et parfaite. On aboutit au même résultat que dans le cas d'une firme unique. Les décisions de toutes les firmes convergent vers un point unique d'équilibre optimal tant en ce qui concerne le timing de l'investissement que son niveau. Le processus

³ Et inversement tend à valoriser davantage un projet flexible, à fort degré de réversibilité.

s'effectue de la manière suivante : les firmes en place, identiques et preneuses de prix, envisagent d'investir lorsque le prix sera suffisant élevé pour justifier le déclenchement de l'investissement. Car leurs espérances de profit seront positives. Cela attire de nouveaux entrants, qui envisagent aussi d'investir, avec comme conséquence de réduire les espérances de profit de l'ensemble du secteur. Un processus itératif d'anticipations croisées s'engage entre toutes les firmes, qui débouche sur un niveau d'investissement optimal à partir d'une convergence de vue sur le prix.

Cet équilibre d'ensemble est possible parce que les firmes sont supposées avoir des anticipations rationnelles. Chaque firme est en mesure d'utiliser toute l'information disponible et d'effectuer des prévisions correctes (en espérance) non seulement pour elle-même, mais pour le reste de la population de firmes (décider pour soi revient ici à deviner ce que les autres anticipent). Ce qui signifie qu'elles partagent une connaissance commune forte : l'ensemble de la population de firmes réagit de manière identique à l'information qui est à sa disposition et effectue le même raisonnement. Ce qui explique que l'on puisse aboutir à des anticipations convergentes (Guesnerie R., 1992). En résumé, la littérature relative à l'investissement sous incertitude accorde une place centrale aux signaux-prix dans la coordination et la prise de décision. L'information apportée par les prix actuels permet de faire la meilleure estimation possible des flux de revenus futurs, compte tenu de l'information disponible. Les hypothèses retenues (une représentation probabiliste de l'avenir et des anticipations rationnelles de la part des agents) permettent d'aboutir à un résultat optimal en terme de choix d'investissement ⁴.

Limites. Ce résultat est mis à mal lorsqu'on relâche les trois hypothèses relatives à la représentation de l'incertitude, aux anticipations des acteurs et à l'existence de marchés complets :

a) Une représentation de l'incertitude qui laisse la place à la survenance de chocs, de frictions et de transformation des conditions

⁴ Ces hypothèses, très restrictives, expliquent certainement pourquoi, de l'aveu même de leurs concepteurs, ces modèles rencontrent tant de difficulté à être validés empiriquement. " I believe that the main reason the investment literature does so poorly empirically is that adjustment frictions are vastly more complex than we have modeled. Of course this is true of almost everything we do in economics, but it seems to be worse for investment" (Caballero R., 2000. Voir également Carruth A. et al., 1998).

d'offre et de demande. Par exemple, la prise en compte de l'innovation, comme facteur de déstabilisation permanente des équilibres de marchés et des fonctions de production conduit à considérer la concurrence comme un processus de création de ressources dans le cadre de processus d'ajustements hors équilibre. Dans cette perspective, les prix ne jouent plus un rôle central de coordination car l'innovation et les restructurations qu'elle implique aggravent l'incertitude et l'hétérogénéité, brouillant ainsi les signaux-prix et leur signification. Ce faisant, l'innovation introduit une rupture dans la complémentarité inter-temporelle des marchés. Les investisseurs devront donc baser leurs décisions sur d'autres éléments que sur le seul signal-prix de court terme. Celui-ci peut même être déficient et décourager l'innovation si les conditions de marché sont trop incertaines et les prix trop volatiles. Des mécanismes complémentaires de coordination peuvent alors être développés par les acteurs en présence pour pallier ces déficiences et réduire l'incertitude qui entoure leur prise de décision : communication d'informations ou rigidification des mouvements de prix. Dans ce contexte le rôle des prix n'est pas d'envoyer un signal mais de créer les conditions pour éviter des déséquilibres de marché excessifs qui mettent en question les possibilités d'innovation (Amendola M. et Gaffard J-L., 1998 ; Amendola M. *et al.*, 2004).

b) Des acteurs qui ne se comportent pas comme le suppose la théorie des anticipations rationnelles. Les firmes en concurrence peuvent, et c'est une hypothèse assez réaliste, ne pas bénéficier des mêmes informations, ne pas utiliser le même raisonnement pour établir leurs prévisions. Dès lors, des anticipations divergentes peuvent voir le jour; des comportements d'aversion au risque se développer; ou des phénomènes de mimétisme apparaître, créant les conditions de périodes de sur-investissement ou de sous-investissement. Par exemple, si une firme observe que personne n'investit alors qu'elle envisage de le faire, elle peut en inférer que ses propres estimations de rentabilité sont basées sur des hypothèses trop optimistes. Elle sera incitée à les réviser à la baisse, considérant que si les autres firmes n'investissent pas c'est parce qu'elles détiennent des informations privées de nature différente. Si chacun se comporte de la même manière, une période de sous-investissement s'en suivra (Stiglitz J., 1989). Les anticipations rationnelles garantissent que les acteurs s'engagent dans un processus d'itération qui convergera vers un point d'équilibre. Lorsque cette hy-

pothèse est relâchée, c'est-à-dire lorsqu'on introduit des asymétries d'information et/ou de l'hétérogénéité dans les raisonnements, les anticipations des acteurs peuvent diverger; rien ne garantissant que le processus de décision qui va alors se mettre en place conduira à un choix collectif optimal. Dans les cas où les agents ne se comportent pas selon l'hypothèse des anticipations rationnelles et où l'incertitude peut conduire à des retournements brusques et inattendus des conditions de marché, les décisions d'investissement peuvent s'éloigner de l'efficience de long-terme. Des comportements de mimétisme ou de myopie de la part des investisseurs, provoqués par une aversion au risque ou des anticipations non convergentes peuvent produire des biais dans les décisions d'investissement.

c) Des instruments financiers de couverture du risque qui ne permettent pas de restaurer la relation inter-temporelle des prix. Malgré les deux problèmes précédents, la relation inter-temporelle des prix pourrait être, en théorie, restaurée par l'intermédiaire des marchés financiers. Ils sont susceptibles de proposer des instruments de couverture (contrats à terme, options) permettant à l'investisseur de maîtriser son risque en lui garantissant des revenus sur longue période. Dans ce cas, les intermédiaires financiers prendraient à leur charge le risque induit par l'établissement de contrats à prix fixes portant sur de longues durées. Bénéficiant de tels contrats, les investisseurs pourraient s'engager dans des décisions d'investissement en ayant l'assurance d'obtenir un flux de ressources garanties (si par exemple ils optent pour un contrat de vente à terme de leur production future).

Mais ce n'est pas le cas : les marchés à terme, qui permettraient en théorie de réduire l'exposition des investisseurs aux fluctuations de prix, courrent sur des horizons de temps limités au regard des durées de mise en place et de recouvrement des coûts d'investissement. Pas au-delà de trois ans pour les produits réellement échangés, donc liquides, sur les marchés électriques ⁵. Les producteurs, les commer-

5 Les durées peuvent être plus longues pour d'autres industries qui engagent des investissements très capitalistiques (par exemple les commodités), mais ne recouvrent pas, là encore, la durée de vie des équipements. Il existe néanmoins une base théorique à l'utilisation de *future de court terme* pour se couvrir des évolutions de prix à échéance de plusieurs années. Différents modèles proposent des solutions théoriques pour évaluer la rentabilité future d'un investissement du type gisement ou mine, et pour construire un portefeuille de couverture établi à partir d'une succession de contrats à terme de courte durée. Cependant, force est de constater que les agents répugnent à les mettre en application. Ce qui suggère que les modèles théoriques disponibles ne permettent pas (ou pas encore) de fonder une stratégie de couverture adéquate des risques sur le long-terme (Defeuilley C. et Meunier G., 2008).

cialisateurs et les *traders* sont susceptibles de s'engager dans ce genre de transaction dans la mesure où, pour ces échéances, un certain nombre d'éléments fondamentaux affectant l'offre et la demande (donc le niveau des prix) sont, sinon connus, du moins prédictibles avec une marge d'erreur acceptable. Au-delà de ces horizons courts et au regard des délais d'investissement, les acteurs ne peuvent se fonder que sur leurs propres anticipations pour déterminer ce que seront les prix de marché à l'échéance du contrat et établir leur position en conséquence.

La transcription du paradigme de l'équilibre concurrentiel de long-terme dans le secteur électrique

Le modèle de référence des industries électriques auquel se réfèrent les régulateurs nationaux et la Commission européenne est celui du marché décentralisé dont le fonctionnement pleinement concurrentiel est supposé assurer l'efficacité de court terme et le développement optimal des capacités à long-terme. Ce modèle laisse une question en suspens, celle des investissements en production, qui est marquée par deux défaillances de marché (insuffisance des investissements de pointe pour assurer la sécurité de court terme en tant que bien collectif, biais dans les choix de filières techniques), qui peuvent conduire, sur le long-terme, à des choix socialement inefficaces.

Modèle de référence et application dans l'Union européenne

Pour se conformer à ce modèle, les réformes jouent à la fois sur les structures industrielles et les modalités de coordination. En ce qui concerne les structures, le réformateur décide de la séparation des réseaux de transport-distribution des activités concurrentielles (production, ventes) pour assurer la non-discrimination entre l'ancienne *utility* verticalement intégrée et les entrants. La régulation cherche aussi à réduire au minimum les liens verticaux entre la production et la fourniture, notamment en limitant l'intégration verticale et les arrangements contractuels de long terme, de manière à abaisser les barrières à l'entrée et à stimuler l'arrivée de nouveaux acteurs.

En ce qui concerne les modalités de coordination, les réformes s'appuient sur des travaux de référence qui montrent les possibilités d'organiser les diverses coordinations de court terme et de long terme de l'industrie électrique par des marchés sur l'énergie, les droits d'accès au réseau et les services auxiliaires (équilibrage, réserves, etc.), complétés par des marchés de contrats financiers à chaque niveau (Hunt S., 2002 ; Bohn R. *et al.*, 1988 ; Joskow P. et Schmalensee R., 1984 ; Chao U.P. et Peck S., 1998 ; Wilson R., 2002). Un marché organisé à plusieurs étages est créé pour les ventes horaires d'énergie et les services offerts au gestionnaire du système, complétés par des marchés de couverture de risque. Des règles d'accès précises et transparentes au réseau de transport et aux interconnexions entre systèmes assurent les échanges de façon non discriminatoire.

Le gestionnaire de réseau peut s'appuyer sur des mécanismes de marché pour les ajustements du système (marché de réserves opérationnelles, mécanisme d'équilibrage pour des services en temps réel) et faire payer les droits d'accès aux congestions du réseau par des prix établis par des mécanismes de marchés (enchères) qui seraient censés être incitatifs aux investissements localisés du gestionnaire de réseau. L'intégration des marchés entre pays par la suppression des goulots d'étranglement physique et des barrières réglementaires, et l'harmonisation des règles de marché amplifient la concurrence sur les anciennes aires géographiques des monopoles régionaux ou nationaux.

La mise en oeuvre de ce modèle repose sur une croyance forte dans l'efficacité des marchés de produits financiers qui permettraient aux acteurs d'assurer leur couverture de risque de court terme, mais également de protéger leurs investissements, ce qui s'inscrit bien dans la représentation de l'équilibre de marché avec liaison inter-temporelle et spatiale des prix.

De ce modèle résulte une cohérence entre trois dimensions critiques :

- La première concerne le modèle réglementaire déduit de ce paradigme du marché décentralisé, modèle qui recouvre la séparation du réseau et de la gestion du système d'avec les activités concurrentielles (production et fourniture) ainsi que la limitation d'une intégration verticale production / fourniture et la dé-intégration horizontale de ces deux activités. Dans ce cadre, la dé-intégration des structures industrielles et l'introduction de règles de marché à tous les étages de

la filière constituent un tout cohérent. Toute distance prise vis-à-vis de ce modèle dans une seule de ses dimensions conduit nécessairement à des inefficiences et des dysfonctionnements (Joskow P., 2007).

- La deuxième dimension porte sur le type de marché proposé, qui s'appuie largement sur un modèle d'économie financiarisée. Le *business model* de référence dans la production est celui de firmes spécialisées, producteur pur d'un côté, fournisseur pur de l'autre. Le producteur pur devrait même être un *pure player* spécialisé sur une seule technologie de production pour présenter un profil de risque suffisamment homogène, qui puisse être maîtrisé par les investisseurs financiers dans le cadre de financement structuré en *project finance*. De son côté le fournisseur d'électricité devait pouvoir jouer un rôle d'intermédiaire (mise en correspondance des portefeuilles de contrats de vente et contrat d'achat) sans posséder d'actifs physiques et approvisionner ses clients en faisant appel au marché pour les différents horizons temporels considérés (des contrats à terme à la livraison en temps réel) (Boroumand R. et Zachmann G., 2009).
- La troisième dimension concerne la réglementation et les politiques de la concurrence. Il s'agit de limiter l'intégration verticale ainsi que les contrats de long terme entre les activités concurrentielles de production et de commercialisation pour réduire les risques de forclusion aux différents niveaux de marché. De même pour les ventes aux gros consommateurs. En parallèle il s'agit aussi de limiter les concentrations excessives et les pratiques d'ententes et de coordination dans le cadre de structures oligopolistiques éloignées du modèle de référence, ce qui explique l'encouragement des régulateurs à la déconcentration horizontale par le biais de ventes d'actifs de production et de commercialisation, que ces derniers soient physiques ou virtuels (sous forme de droits d'accès contractualisés aux productions des équipements du producteur historique, les *Virtual Power Plants*).

Ce modèle décentralisé sert de ligne directrice à l'Union européenne, qui de manière directe, via le jeu des directives successives, introduit les règles de marché aux différents niveaux de la filière et, de manière indirecte, via les politiques de la concurrence, cherche à avoir un impact sur l'évolution des structures industrielles. L'Union européenne n'ayant pas les pouvoirs institutionnels de provoquer une remise à plat des structures industrielles existantes, elle a dû accepter que les

réformes décidées par les Etats membres n'atteignent pas la cohérence souhaitée en la matière. La politique antitrust de la Commission européenne et des autorités nationales est devenue l'instrument privilégié pour réduire, autant que faire se peut, les écarts constatés entre les structures industrielles nationales et l'idéal type du modèle décentralisé ⁶.

Deux défaillances de marché

La question des investissements en production est marquée par les deux défaillances de marché déjà signalées. D'une part le marché laissé à lui-même ne permet pas d'assurer un équilibre physique en pointe et en toute circonstance (ou tout au moins avec une très faible probabilité de défaillance). D'autre part le marché ne permet pas d'orienter correctement les investissements pour aboutir à un parc optimal de production.

L'insuffisance des investissements de pointe. Les prix qui se forment sur le marché pendant l'extrême pointe et les situations exceptionnelles peuvent monter à des niveaux très élevés (plus de 300 fois le niveau moyen ordinaire) en cas de tension sur les capacités totales et les marges de réserve. Cette montée des prix aux extrêmes dans certaines situations aléatoires (pointe de demande, indisponibilité des équipements) est rendue possible par l'inélasticité-prix de la demande horaire en temps réel ; elle est aussi liée à la rigidité des fonctions d'offre en extrême pointe du fait de la non-stockabilité de l'électricité (voir figure 1). Ce surplus tiré de revenus aléatoires pendant quelques heures chaque année reflète une rente de rareté quand les marges de réserve du système électrique diminuent en dessous du niveau de prudence défini par les ingénieurs et avalisé par le régulateur. C'est l'anticipation de ce surplus qui serait supposé déclencher l'investissement en pointe et garantit l'optimalité des capacités en place pour assurer une marge de réserve suffisante (Joskow P., Tirole J., 2007).

⁶ De nombreuses décisions l'illustrent : accord entre la DG Comp et les groupes allemands E.ON et RWE pour un désengagement partiel sur leur marché historique (électricité pour le premier gaz pour le second) en échange d'une extinction des poursuites engagées à leur encontre. Contreparties négociées avec les autorités européennes et nationales par plusieurs grands groupes suite à des opérations de rachat ou de fusion conduisant à une diminution des parts de marché des opérateurs concernés sur leur marché historique respectif, au bénéfice de nouveaux entrants (EDF en France lors du rachat de EnBW en 2001, Suez en Belgique lors de la fusion avec GDF en 2008, etc.).

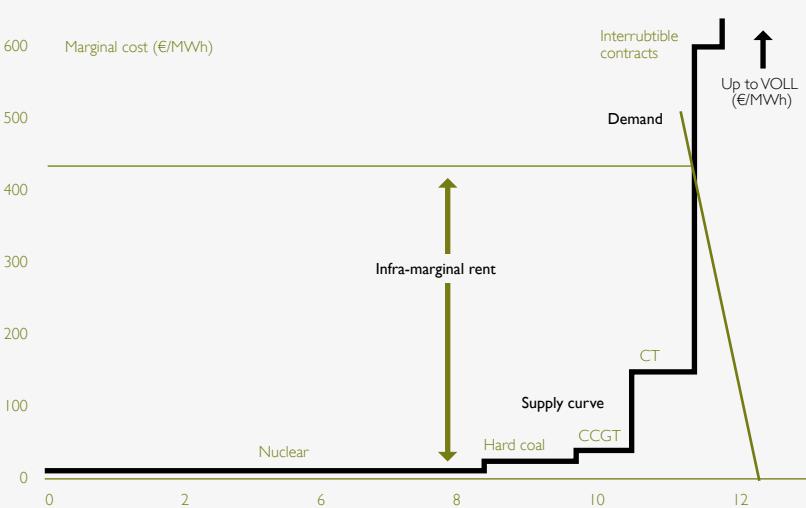
Pour ce faire cette montée des prix aux extrêmes est nécessaire, car elle permet d'inciter à investir dans les équipements de pointe flexibles, peu capitalistiques, mais ne bénéficiant de revenus que pendant le petit nombre d'heures des pics de prix extrêmes.

La question des capacités en pointe est critique pour deux raisons : d'une part les producteurs ont une forte aversion au risque face au caractère très aléatoire des revenus tirés de ce type d'équipements pour rentabiliser leur investissement. D'autre part, le signalement d'une rareté se fait par l'exercice de pouvoir de marché (Stoft S., 2002). En effet l'inélasticité-prix permet à n'importe quel producteur, même petit, d'influencer le prix de court terme pendant les périodes de tension sur les capacités. Cette situation a deux effets : en premier lieu cela accroît le caractère de l'incertitude pour les investisseurs en concurrence car ils ne peuvent anticiper jusqu'où et sur quelle durée les prix vont monter ; en second lieu, lorsque les prix vont s'établir à 5-10000€/MWh alors que le coût du moyen de production marginal appelé par le marché est de 300€/MWh se pose un problème d'acceptabilité sociale.

Rente de rareté et pouvoir de marché étant inextricablement liés, il s'ensuit que les régulateurs sont enclins à mettre un plafond sur les prix horaires lors de ces épisodes de tension sur les capacités, ce qui transforme la défaillance de marché en imperfection règlementaire.

Figure 1

Hourly market equilibrium



La question d'investir en capacités en pointe sous ce régime de *price cap* devient critique car on supprime les revenus que donneraient les marchés électriques horaires sans contrainte sur les prix en pointe. Mais même quand le régulateur ne met pas de plafonds de prix et laisse les prix monter aux extrêmes, les recettes des nouveaux équipements de pointe sont trop aléatoires pour couvrir les coûts fixes et assurer une rentabilité normale de l'investissement. C'est le problème du *missing money* démontré dans les calculs de P. Cramton et S. Stoft (2006) et P. Joskow (2007). La raison provient d'une imperfection réglementaire, à savoir le manque de garde-fous qui empêcheraient les interventions prématurées du gestionnaire de réseau pour assurer la stabilité du système en période de tension sur les capacités. Celui-ci tend à intervenir de façon prématuée par des mesures techniques pour éviter la défaillance du système, ce qui conduit à réduire les revenus tirés des marchés de l'énergie et des réserves pendant les périodes exceptionnelles.

Que ce soit à cause des imperfections de marché ou de la réglementation, le problème de l'adéquation de capacité reste entier parce qu'il est difficile de faire s'exprimer les dispositions à payer l'électricité et la garantie d'être fournie dans toutes les situations par les acteurs en présence, l'adéquation de capacité (donc la sécurité de fourniture à court terme) ayant les caractéristiques d'un bien public ⁷.

Différents instruments sont envisagés pour ajouter un revenu au prix de l'électricité sur le marché de gros pour rémunérer les capacités et leur disponibilité pendant les périodes de tension (De Vries L., 2007 ; De Vries et L. Neuhoff K., 2004 ; Vasquez C. *et al.*, 2002 ; Stoft S., 2002 ; Cramton P. et Stoft S., 2006 ; Finon D. et Pignon V., 2008 ; Marty F., 2007). Ces instruments cherchent à garantir un niveau d'installation suffisant (assurer un niveau adéquat de capacité à l'horizon d'installation de nouveaux équipements) et visent à inciter à la disponibilité maximale des équipements existants en proposant des revenus supplémentaires (c'est la dimension du court terme). Les trois 'mécanismes de capacité' (réserve stratégique, mécanisme de paiement de capacité, obligation de capacité associée à un mécanisme d'échange

⁷ L'adéquation de la capacité d'ensemble des parcs des concurrents est un bien non rival (chacun des producteurs profite de la garantie assurée par l'ensemble des capacités des autres), et non excluible (il n'y a pas moyen de séparer les besoins individuels dans des relations marchandes bilatérales).

de droits) utilisés au départ des réformes présentent des limitations qui ont été bien identifiées et qui en ralentissent le déploiement sur les marchés électriques.

Le premier dispositif, connu sous le nom de *réserve stratégique*, consiste à autoriser le gestionnaire du réseau de transport (GRT) à disposer de réserves contractuelles de long terme ou de ses propres équipements de réserve. Il suscite la crainte de voir le GRT faire appel de façon discrétionnaire à ces réserves et fausser les prix de marché de l'énergie pendant les périodes de pics de prix. Dans les faits il reste utilisé dans les pays à dominante de production hydraulique.

Le second dispositif connu sous le nom de *paiement de capacité* rémunère toutes les capacités⁸. Cette coordination décentralisée par le prix de marché pose un problème car si elle donne une rémunération garantie, il n'assure pas que soit atteint le niveau de capacité et de marge de réserve souhaité par le régulateur pour faire face à toutes les situations. Il a été en général abandonné. Le troisième mécanisme, connu sous le nom de *marché de capacité* agit en fait d'abord par les quantités ; il consiste à imposer aux fournisseurs en concurrence la détention de droits de capacité (équipements en propre et contrats de long terme de fourniture de gros) en relation avec la demande maximale qui leur est adressée en pointe. Il inclut un marché secondaire pour la flexibilité du dispositif. Cet instrument est utilisé dans les marchés régionaux nord-américains (PJM, New England, Californie notamment) mais ne donne pas les résultats escomptés pour deux raisons combinées : il repose sur les décisions des acteurs décentralisés et il était mal conçu (Cramton P. et Stoft S., 2006). En particulier il n'y avait pas de correspondance de temporalités entre les rémunérations offertes par ce mécanisme (droits définis sur le court terme) et le besoin de revenus stables pour les nouveaux équipements.⁹

L'analyse des limites pratiques de ces instruments a conduit à proposer un autre mécanisme. C'est un mécanisme centralisé d'attribution de contrats de capacité de long terme par enchères pour un mon-

⁸ Elle se réfère soit au coût marginal de défaillance estimé par le régulateur dans une logique d'optimum collectif (comme c'était le cas dans la première conception de l'architecture de marchés britannique), ou bien au coût de l'unité de pointe amortie sur quelques années (comme en Espagne).

⁹ En plus du défaut de durée du droits, il donnait aux investisseurs des rémunérations peu incitatives car très aléatoires (variant brutalement entre zéro et le niveau prévu de pénalité prévue en cas de non-disponibilité au moment de l'appel par le gestionnaire de réseau).

tant de capacité correspondant à la demande maximale de puissance anticipée rehaussée de la marge de réserve (Vasquez C. *et al.*, 2002 ; Cramton P. et Stoft S., 2006). Ce mécanisme est connu sous le nom *marché de contrat de capacité* (forward capacity mechanism). Il assure une rémunération de la capacité par des contrats longs pour tous les équipements, et en particulier pour les équipements neufs en conformité avec les délais de construction et le temps de recouvrement de l'investissement. Le design de ce mécanisme est efficace parce que les contrats mis aux enchères assurent une marge de réserve au système (précisons qu'une pénalité suffisamment dissuasive est adjointe au dispositif pour inciter à l'investissement une fois le contrat passé). Les expériences récentes sur le marché du New England et de New York ont montré l'efficacité de ce nouveau mécanisme, qui est aussi utilisé avec succès dans quelques pays d'Amérique latine (Colombie, Pérou, Brésil). Mais il reste très complexe¹⁰.

Les pays qui n'ont pas mis en œuvre des mécanismes cités s'en remettent en général à leurs opérateurs historiques pour investir dans des moyens de pointe. L'économie politique de ces marchés veut alors que la menace réglementaire sur ces grands opérateurs historiques (recours à des enchères pour des contrats de réserve stratégique associé à des équipements nouveaux de pointe, ou bien menace implicite de dé-investissement pour rechercher plus de concurrence propice aux entrées) suffise à les inciter à investir, même à perte¹¹.

Biais dans les choix des filières techniques. Le second défaut de marché concerne le choix des technologies de production. Dans le modèle du monopole électrique vertical réglementé, les équipements électriques pouvaient être choisis de façon planifiée. Les risques d'investissement étaient limités car les coûts étaient transférés sur les consommateurs par la réglementation en *cost-plus* par l'intermédiaire des tarifs. Dans le modèle de marché décentralisé, les logiques sont inversées. Les prix résultant de la concurrence de court terme sont

¹⁰ On doit en effet préciser qu'un marché secondaire accompagne ce dispositif pour des échanges de plus court terme.

¹¹ On ajoutera qu'il existe des procédures informationnelles régulières de type programmation à long terme par le gestionnaire de système qui conduisent à une convergence (programmation en fonction des prévisions de croissance de demande et des annonces des producteurs) qui amènent à éclairer collectivement les besoins de nouvelles capacités et pointent les opportunités d'investissement en équipement de pointe.

sans lien avec les coûts complets, et les producteurs investissent sans garantie que les prix de marché en principe alignés sur les coûts marginaux de court terme rencontrent les coûts marginaux de long terme (et donc assurent le recouvrement des coûts fixes de leurs nouveaux équipements). Ce qui crée, comme on va le voir, un biais dans leurs choix d'investissement.

La régulation de l'industrie électrique en monopole présente l'avantage de pouvoir se reposer sur la fiction d'un équilibre dynamique de long terme, le monopole règlementé étant en mesure d'optimiser les investissements de production et de pratiquer une tarification répercutant les coûts marginaux de développement. Ce sont les coûts de production qui prévalent par anticipation une fois que la capacité productive a été déterminée à un horizon donné par la firme optimisante sur une longue période de temps (Allais M., 1943). Ils sont alignés sur les coûts complets de l'équipement marginal selon les périodes horaires de l'année, notamment pour les tarifs de pointe qui rencontrent facilement l'exigence de recouvrement des coûts fixes des équipements de pointe (Boiteux M., 1949). Le coût marginal de long terme (CMLT) doit prendre en compte le service de la dette de la firme car l'endettement permet à la firme de lisser les déséquilibres temporaires entre coûts et revenus. La référence au CMLT permet une stabilité du prix alors que les coûts marginaux de court terme sont beaucoup plus volatiles. Ce fonctionnement est en cohérence avec une planification du développement des équipements qui soit adaptée à l'évolution de la demande horaire et saisonnière (Boiteux M., 1956). L'autre réglementation plus rudimentaire et largement répandue, celle des tarifs en *cost-plus* alignés sur les coûts comptables, permet aussi intrinsèquement une limitation des risques d'investissement, en supprimant le risque-prix et le risque-volume et en assurant un recouvrement des coûts.

Le modèle du marché décentralisé repose sur la même fiction d'un marché en concurrence parfaite et en information parfaite (Oren S., 2003). Le marché est supposé reproduire de façon parfaite les prix optimaux assurant l'articulation court terme-long terme. Le fonctionnement des marchés horaires sur lesquels l'ensemble du système repose devrait garantir un niveau d'investissement et une répartition entre technologies socialement optimale pour suivre le

développement des demandes horaires. Les seuls signaux de prix horaires permettraient d'obtenir sur un marché donné une structure de parc optimale minimisant les coûts de long terme en couvrant les besoins horaires et en garantissant la fiabilité de fourniture. Les niveaux de prix horaires alignés sur les coûts variables de la centrale qui est appelée la dernière par le marché sont réputés être à même de couvrir en moyenne annuelle l'ensemble des coûts fixes des différentes technologies de pointe et de base, sur la base de ces hypothèses d'information parfaite et de concurrence parfaite en incluant une rémunération du capital compatible avec le régime de marché. L'anticipation probabilisée du surplus des équipements infra-marginaux sur les marchés horaires permettrait aux opérateurs de déclencher les investissements idoines pour garantir à la fois l'adéquation de capacités et l'optimalité du mix d'équipements. En cas de changement de l'environnement institutionnel et économique du secteur (choc d'intégration des marchés, choc de prix de combustibles, internalisation du coût du carbone), l'investisseur pourra décider d'investir dans d'autres technologies après avoir identifié les rentes de rareté résultant de ces chocs qui éloignent de l'état d'équilibre basé sur un parc optimal, formé par l'addition des parcs de l'ensemble des producteurs en concurrence.

La nature même des marchés électriques invalide cette représentation d'équilibre dynamique. Les prix se forment sur des marchés horaires marqués par des inélasticités conjointes de l'offre et de la demande en situation de capacité tendue. Ils sont supposés guider les choix de long terme des agents en concurrence selon le même miracle d'équilibre de marché auquel se référerait le planificateur marginaliste. Si l'on peut très bien admettre que ce dernier ne prenait pas en compte la nature des mécanismes d'information et les risques mis en jeu par le fonctionnement d'un marché électrique qui pour lui n'était que fictif, il n'en est plus de même quand il s'agit de marchés réels. Or, en régime de marché, pour avoir un équilibre de long terme on doit, comme déjà dit, faire l'hypothèse forte que les prix de court terme intègrent les anticipations de long terme sur la relation demande-offre horaire, les limites de capacité des producteurs en concurrence avec un portefeuille de différentes technologies sur les différents marchés horaires, les tendances d'évolution des prix des intrants (les combustibles dans le cas de l'électricité)

et du coût des technologies. La multitude d'incertitudes n'assure en rien cette correspondance qui serait miraculeuse, même en espérance mathématique par rapport à un scénario moyen d'évolution de la demande et d'aléas de court terme (apport hydraulique, fonctionnement des équipements). La volatilité des prix de court terme et le peu de lisibilité des tendances de long terme brouillent les anticipations de revenus. Des prix de court terme volatils qui sont alignés sur les coûts marginaux de la dernière technologie appelée sur les marchés horaires successifs ne sont pas révélateurs de l'état des fondamentaux et des raretés de capacité sur les équipements de base et de semi-base, car rien ne garantit en tendance leur alignement sur les coûts marginaux de long terme.

Le peu de lisibilité des tendances de long terme dissuade les agents à investir dans les équipements très capitalistiques à coût irrécupérable élevé et les conduit à retarder leurs décisions, selon le résultat classique de Dixit A. et Pindyck R. (1994) pour un choix avec une seule technologie. Ou bien, comme les agents ont le choix entre des technologies à structure de coût différente dans l'industrie électrique, cela les incite à choisir la technologie la moins capitalistique, mais aussi présentant une bonne corrélation entre la partie de ses coûts variables et les prix de court terme. Le choix d'investissement n'est plus seulement une question de minimisation de coût, mais aussi une question de gestion de risque, sachant que les risques varient largement entre technologies (voir tableau I).

De ces profils de risque différents, il s'ensuit que l'investissement en technique à gaz en cycle combiné (CCGT) qui est moins capitaliste, plus divisible, avec des temps d'installation courts, et flexible en exploitation a spontanément les faveurs des investisseurs (Finon D. et Roques F., 2008). La structure de coûts de cette technologie avec coût d'investissement faible et coûts variables élevés (combustible) tend à en faire la technologie marginale sur les marchés horaires sur une grande partie de l'année quand le prix du gaz dépasse un certain niveau. Comme elle détermine le prix sur le marché horaire, les investisseurs dans cette technologie bénéficient pour gérer leur risque de la bonne corrélation du prix du combustible (la majorité du prix de revient pour eux) avec le prix horaire de l'électricité. La qualité de faible intensité capitaliste se combine donc avec une bonne corrélation entre

la partie variable de ses coûts et les prix horaires pour faciliter la gestion du risque. Ce n'est pas le profil des investissements dans des techniques très capitalistiques à faible coût variable (nucléaire, hydraulique, charbon) qui risquent, une fois l'équipement investi et mis en exploitation, de voir leurs revenus horaires tombés en dessous des coûts fixes en cas de retournement du marché. Un tel profil de risque tend à rebouter les investisseurs financiers, bien que la technologie présente un avantage symétrique en termes de couverture de risque, et en espérance mathématique, des coûts complets plus faibles, toutes choses égales par ailleurs.

En univers libéralisé, les investisseurs ont donc une préférence pour les CCGT même pour des productions en base. Cette préférence se renforce à mesure que d'autres investisseurs font le même choix. Leurs choix parallèles contribuent à faire des CCGT la technologie marginale sur les marchés horaires sur la grande majorité de l'année. Par effet d'auto renforcement ceci tend à faire des mouvements de prix du gaz le *facteur principal* de changement du prix de l'électricité à court terme, réduisant ainsi le risque encouru à investir dans cette technologie.

Ces choix assurent certes un équilibre de marché pour toutes les heures de l'année, mais ils ne convergent pas avec ce que serait l'optimum technologique de long terme en environnement de marché. Rien ne dit qu'à long terme le choix d'investir dans les CCGT au détriment d'autres technologies soit le meilleur choix possible. Cela dépend très largement du prix des combustibles fossiles (pétrole,

Tableau 1

Comparaison des caractéristiques des technologies de production électrique vis-à-vis du risque d'investissement

Technologie	Turbine gaz (100 MW)	CCGT (400-600 MW)	Charbon (2x700 MW)	Nucléaire (1500 MW)	Eolien (200 MW)
Coût d'investissement	20 millions	100-200 millions	700-1000 millions	2-3 milliards	300 millions
Durée de construction	Très courte	Courte	Longue	Longue	Moyenne
Risque lié à la construction	Faible	Faible	Forte	Forte	Moyenne
Part du combustible dans le coût	Très élevé	Elevé	Moyen	Faible	Nul
Coût du CO ₂	Moyen	Moyen	Fort	Nul	Nul
Risque lié au prix du combustible	Fort	Fort	Moyen	Faible	Nul
Risque lié au prix de marché	Fort	Faible	Moyen	Fort	Fort

Coûts d'investissements exprimés en Euros.

Source : d'après IEA, *Investment in power sector, 2007*

gaz) et du prix du carbone, deux composantes majeures des prix de l'électricité produite à partir des CCGT. En faisant l'hypothèse, plausible, que ces deux composantes auront tendance à augmenter dans les prochaines décennies à un rythme plus élevé que les autres composantes de coûts des autres technologies disponibles, on peut imaginer que ce choix s'avère inefficace. Mais, du fait, d'une part, de la durée de vie des équipements en question, et d'autre part, de l'effet auto renforçant évoqué précédemment, il existe un risque que, bien que plus dispendieuse que les autres technologies, les investisseurs continuent à investir dans les CCGT et s'engagent à long-terme sur une trajectoire potentiellement inefficace, construisant un mix de production qui s'éloigne de l'optimalité.

On est bien ici dans une séquence de choix d'investissement dans l'incertain qui s'éloigne de l'optimum de long-terme évoqué dans la théorie. Les écarts constatés sont liés aux comportements avares au risque des investisseurs, qui, faute de pouvoir décider dans un environnement incertain probabiliste et faute de pouvoir faire converger leurs prévisions à l'aide des anticipations rationnelles, ne peuvent s'en remettre au prix actuel pour calculer leurs espérances de gains futurs ¹². Et ce n'est pas l'utilisation des instruments financiers de couverture du risque qui pourra totalement restaurer la relation inter-temporelle des prix. L'incomplétude des marchés est donc un des éléments qui incitent les investisseurs à choisir les technologies de production les moins exposées, en l'occurrence les CCGT.

Nous avons donc identifié deux défaillances de marché, liées toutes deux à une imparfaite complémentarité inter-temporelle des marchés, qui peuvent conduire à des choix d'investissement socialement inefficaces : manque d'investissements dans les capacités de pointe, préférence pour des filières techniques peu capitalistiques pour les équipements de base mais qui peuvent s'avérer plus coûteuses sur le long-terme. Au final, le simple jeu du marché donne une orientation inefficace (dimensionnement, composition) des parcs de production électrique.

¹² On ne traite ici que des biais (aversion au risque) dans les choix dans un panier de technologies. On aurait pu aussi étendre l'analyse aux mouvements cycliques de sur/sous-investissements, qui sont une autre illustration de la difficulté à dégager un équilibre de long-terme. Voir P.Joskow (2006), pour une analyse de l'épisode de sur-investissement en capacité de production CCGT qu'a connu les USA au début des années 2000.

Des solutions : contrats de long-terme et intégration

Comment répondre à ces défaillances de marché et pallier les incomplétudes du modèle de marché décentralisé en matière d'investissement en production ? Nous avons vu que, sur les marchés décentralisés, au-delà de quelques années, les investisseurs ne pouvaient plus compter sur le signal-prix pour guider leurs choix. Ils s'en remettent à leurs anticipations des états futurs des marchés pour asséoir leur décision. Ils cherchent alors à circonscrire les risques pris aux travers de ces engagements en améliorant leurs informations disponibles et en organisant un partage des risques. Deux voies de partage des risques du producteur sont ici explorées : soit avec leurs clients (gros fournisseurs ou consommateurs) dans le cadre de contrat de long-terme, soit, avec d'autres catégories d'actifs dans le cadre d'une intégration verticale et horizontale.

Les contrats de long-terme (CLT)

Les gros acheteurs, intermédiaires ou consommateurs électro-intensifs auraient des intérêts convergents avec les investisseurs en production pour se couvrir contre le risque de prix de marché par des contrats de vente à prix fixe. Mais dans la réalité une telle convergence d'intérêts n'opère pas spontanément. Le risque d'opportunisme des acheteurs engagés dans un CLT est la source de l'aversion des producteurs à contracter ex ante à long terme à prix fixes pour investir dans des équipements capitalistiques (De Vries L., Neuhoff K., 2004). Il est alors nécessaire, pour crédibiliser les engagements des gros acheteurs, que certaines conditions soient réunies au préalable. A ces conditions, les contrats de long-terme peuvent constituer des instruments efficaces de couverture des risques facilitant l'engagement d'équipements capitalistiques en production.

Il existe deux types principaux de contractants possibles avec un producteur d'électricité : les fournisseurs et les industriels électro-intensifs. Un fournisseur d'électricité est un intermédiaire cherchant à équilibrer en temps réel ses portefeuilles d'approvisionnement et de vente. L'équilibrage en continu est complexifié par l'absence de contrôle sur l'échange physique d'électricité vers ses clients qui

consomment à tout moment ce qu'ils veulent dans les limites de puissance définies dans leur contrat. L'équilibrage contractuel de son côté est complexifié par les différences en termes de durée et de prix entre ses contrats d'approvisionnement d'un côté et ses contrats de vente de l'autre ¹³. Le portefeuille de ventes est constitué de divers types de contrats de durées courtes (de six mois à deux/trois ans maximum) : contrats à prix plats, contrats à prix révisable à intervalles réguliers en fonction de l'évolution des coûts de fourniture, contrats indexés sur le prix du marché de gros, etc. Une règle imposée par les régulateurs pour activer la concurrence de détail instaure une asymétrie dans la relation contractuelle, car elle permet aux acheteurs de sortir à échéance courte, ce qui amplifie le risque volume du fournisseur. Gérer conjointement leur portefeuille de contrats d'achat et celui des contrats de vente est complexe car les fournisseurs vendent la plupart de leur électricité dans le cadre de contrats à prix plats à moyen terme. Ils ne peuvent s'engager dans des transactions longues en amont à prix fixes, car ils risquent d'être pris en tenaille de prix entre leurs engagements à prix ferme en aval et leurs achats en amont lorsque le prix de marché de gros se retourne à la baisse et devient inférieur au prix de leur contrat d'achat. Le *price squeeze* se traduit pour eux par un risque de pertes de part de marché au bénéfice d'entrants achetant l'électricité de gros à prix bas après le retournement du marché, sauf à baisser leurs prix de détail en dessous du prix de leurs contrats d'achat à prix fixes. Engagés dans de tels contrats, ils sont incités à les remettre en question, à les renégocier ou à en sortir.

Du côté des consommateurs électro-intensifs, les difficultés de contracter à long terme avec un producteur sont assez semblables à celles des fournisseurs. L'électro-intensif opère généralement dans des secteurs dont le marché est globalisé. Il est exposé à la concurrence d'industriels dans des pays où ils peuvent bénéficier de meilleures conditions d'achat. Il est aussi exposé au risque de retournement du marché de ses produits, ce qui peut l'inciter à demander la renégociation de son contrat à prix fixes. Le comportement opportuniste

¹³ A la différence d'un grossiste dans d'autres filières industrielles, un fournisseur d'électricité est limité dans sa capacité à ajouter de la valeur au produit électricité. Son apport essentiel tient à sa connaissance du marché et à sa capacité d'agrégation, même s'il peut chercher à offrir des services pour se différencier des concurrents.

dont peut faire preuve l'électro-intensif en cas de retournement du marché de son output est similaire à celui des fournisseurs en cas de *price squeeze*. Prévoir dans le contrat une clause d'indexation du prix contractuel de l'électricité sur le prix mondial de leur output n'est pas forcément acceptable pour le producteur qui a investi dans un équipement de production à coût irréversible élevé ¹⁴.

Devant ce désalignement des intérêts des parties, il convient d'identifier ce qui limite l'opportunisme des parties et crédibilise les engagements de long terme. Dans le cas des fournisseurs, la condition qui permet des engagements dans des contrats longs sur des bases crédibles est principalement l'existence d'une large base de *sticky consumers*.

Pour un fournisseur, cela permet d'accroître son pouvoir de négociation dans le processus de rigidification des prix et des quantités lors de la négociation d'un contrat ex-ante associé à un nouvel équipement avec un nouveau producteur, et d'articuler la gestion de ses risques sur ses approvisionnements et sur ses ventes. Cette base de consommateurs, peu volatile, garantit leur capacité à faire face à leurs engagements contractuels dans la vente de détail. Elle permet de transférer une majeure partie du risque d'investissement sur les clients fidélisés, comme le soulignent R. Green (2004) et P. Joskow (2007). Une autre solution serait le maintien de monopole réduit à la fourniture aux petits consommateurs avec mise aux enchères de la concession par le régulateur ou les autorités concédantes, comme c'est le cas dans certaines juridictions américaines (Ohio) (Littlechild S., 2008).

Dans les deux cas, si l'on avance dans le sens de l'efficience de long terme en permettant l'engagement crédible du fournisseur, la contrepartie est que l'on renforce le besoin de surveillance des marchés de détail ou de régulation des concessions. Il s'agit en particulier de dissuader l'exercice du pouvoir de marché par le fournisseur dominant dans un cas, et de créer une réglementation incitative du monopole de la fourniture de détail dans l'autre.

En ce qui concerne les consommateurs électro-intensifs, le besoin de garantie dans les engagements de long terme conduit le plus souvent

¹⁴ D'autres contrats long terme, avec clauses de flexibilité sur les prix et les quantités, sont possibles entre une entité productrice et de gros acheteurs dans le modèle décentralisé, mais ils ne résoudront pas le problème de divergences d'intérêts des parties et de partage de risque.

à des montages en partenariat avec le producteur. Dans ce type de projet, le ou les consommateurs¹⁵ (électrométallurgiste, électrochimiste, papetiers) sont co-propriétaires de l'équipement de production d'électricité. De grands équipements en CCGT à gaz produisant chaleur et électricité ont ainsi été installés dans le cadre de tels montages, la propriété commune limitant le risque de 'hold-up' (risque d'opportunisme) de part et d'autre. D'autres paramètres de ces projets renforcent la crédibilité des engagements. En premier lieu l'utilisation de sous-produits des procédés de l'industriel (gaz de haut fourneau par exemple pour des aciéries) pour alimenter la centrale, de même que la production jointe de chaleur de process, renforcent la spécificité de l'investissement à la transaction. En deuxième lieu l'achat de combustible et la gestion du risque associée se font en commun, qui se renforce le partage des risques entre partenaires. En troisième lieu la capacité en puissance électrique peut être surdimensionnée par rapport aux besoins en électricité de l'industriel, pour que l'énergéticien trouve dans le projet une rentabilité supplémentaire (possibilité de vendre le surplus sur le marché), qui faciliterait son engagement¹⁶.

Venons en maintenant à la seconde voie utilisée par les acteurs pour circonscrire leurs risques, l'intégration verticale et horizontale.

L'intégration verticale et horizontale

En tant qu'arrangement institutionnel, l'intégration verticale partielle ou totale permet d'internaliser un ensemble de coûts et de risques au sein de la firme (Hansmann H., 1988). Elle limite l'exposition aux prix, le changement de prix qui désavantage une partie de la firme pouvant avantager une autre partie. Elle économise ensuite sur les coûts de la contractualisation (à la fois les coûts de mise en place et les coûts ex post dus aux imperfections de la contractualisation) en dispensant des besoins de contrôle nécessaires à la réalisation du

¹⁵ Ils peuvent au préalable constituer des arrangements horizontaux de type consortium d'achat avant de s'engager avec un producteur qui construit l'équipement.

¹⁶ Les électro-intensifs peuvent également constituer une coopérative de production qui investit dans des équipements de technologies dont les coûts de production sont indépendants des coûts de combustible fossiles et des biens environnementaux associés. Une telle coopérative qui s'était développée en Finlande dans les années 70-80 en regroupant des papetiers pour investir dans deux centrales nucléaires a décidé en 2005 de commander un nouveau réacteur de 1700 MW pour 3 milliards de € en ayant l'assurance d'écouler son électricité à prix coûtant aux industriels membres de la coopérative et aux fournisseurs municipaux qui sont entrés entre temps dans la coopérative (Finon D. et Roques F., 2008).

contrat. Le coût de contracter dans un cadre de marché pour un acheteur inclut, outre ces coûts classiques de contractualisation, les coûts induits par un déséquilibre possible du pouvoir de négociation en sa défaveur. L'asymétrie d'informations entre les deux parties au contrat au détriment de l'acheteur ouvre la possibilité pour la partie la mieux informée de négocier de meilleurs termes. De ce point de vue, l'intégration verticale présente l'avantage classique souligné par O. Hart (1995) de supprimer l'incertitude des négociations ex post. Détaillons les intérêts respectifs des parties à la transaction

Les producteurs purs d'électricité ont donc intérêt à l'intégration verticale qui efface le risque de 'hold-up' de la part de la partie la moins engagée dans la transaction, qui dans notre cas, est l'acheteur (industriel, fournisseur). Ils ont aussi intérêt à descendre en aval pour garantir leurs débouchés et répartir leur risque sur la chaîne d'activités. S'allier à, ou racheter, des fournisseurs en place qui détiennent une base de clientèle leur permettront de bénéficier de cette base sur laquelle ils peuvent faire passer une partie de leur risque d'investissement.

Mais c'est du côté des fournisseurs que l'incitation à s'intégrer verticalement est la plus forte. Comme leur activité d'intermédiation ajoute peu de valeur ajoutée par rapport à la même activité dans d'autres secteurs (banque, télécoms, grande distribution), la faiblesse de leur rôle économique dans ce secteur réduit leurs possibilités de s'affranchir, même partiellement, de leurs conditions d'approvisionnement pour déterminer leur politique de prix à destination des clients du marché de masse. Les fournisseurs ont donc intérêt à organiser une couverture « physique » d'une grande partie de leurs besoins en investissant dans des moyens de production leur permettant de limiter leur exposition aux fluctuations des prix du marché de gros (Boroumand R., Zachmann G., 2009). Un des enjeux importants pour les fournisseurs concerne la couverture du risque pendant les périodes de pics de prix sur le marché de gros, périodes pendant lesquelles ils sont exposés à des variations importantes. La préférence des fournisseurs pour éviter les pics de prix et gérer leur risque les conduit à favoriser la détention d'actifs physiques, en particulier des équipements de pointe flexibles et, par extension, de s'équiper en équipement flexible de semi-base qui couvre ces besoins de pointe. C'est une façon pour eux de se prémunir contre les possibilités d'exercice de pouvoir de marché de la part des producteurs en période de tension sur les capacités.

Une confirmation empirique du développement de l'intégration verticale comme stratégie préférentielle de couverture des risques est apportée par l'observation de l'intégration verticale au Royaume uni, en Australie, en Nouvelle Zélande et aux USA dans les juridictions où les régulateurs avaient imposé aux anciennes *utilities* de se séparer de tout ou partie de leurs unités de production ¹⁷.

L'évolution des structures de marché y témoigne d'un éloignement progressif du modèle décentralisé, où le principe de dé-intégration verticale et horizontale était supposé faciliter les entrées en production et en fourniture. Ce mouvement d'intégration verticale peut être interprété comme une réponse structurelle à l'incapacité des acteurs à gérer leurs risques en faisant appel aux marchés. Placés dans un environnement incertain et confrontés à des variations de prix qui peuvent être de grande amplitude, les acteurs du marché de l'électricité supportent des risques importants qui ne sont pas entièrement couverts par les instruments de marché qui sont à leur disposition (marché à terme, options). Comme déjà signalé plus haut, les marchés à terme (*futures, forward*) ne sont réellement échangés que sur des périodes relativement courtes (jusqu'à trois ans sur les bourses électriques européennes) et les options restent très peu utilisées dans le secteur électrique (marché peu liquide, risques de contrepartie). En d'autres termes, on peut dire que les instruments marchands de couverture du risque ne sont que des substituts imparfaits à la détention d'actifs physiques.

On observera ensuite que le renouveau des besoins d'investissement en production a joué un rôle dans ce processus de ré-intégration. Les réformes ont été définies au départ dans un contexte de surcapacité dans beaucoup de juridictions, ce qui a permis au schéma de marchés décentralisé et de coordination par les prix de court terme de trouver une validation. Le besoin d'arrangements verticaux ne s'est manifesté qu'après l'effacement des surcapacités et l'entrée dans un nouveau cycle d'investissement. On observera également que là où les structures verticales de départ ont été peu réformées comme c'est le cas dans la majorité des pays de l'Union européenne, les entreprises verticalement intégrées sont celles qui portent l'essentiel de la diversification des investissements vers les technologies les

¹⁷ Voir, pour les Etats-Unis, Michaels R. (2006) ; Chao H. et al. (2008) ; Mansour E. (2007) ; pour le Royaume Uni Newbery D. (2005) ; Thomas S. (2006) ; pour la Nouvelle Zélande, Evans L. et Meade R. (2005).

plus capitalistiques (charbon, nucléaire)¹⁸ sur les marchés européens. Cette capacité à s'engager dans des investissements à longue durée d'immobilisations de capitaux révèle aussi l'avantage de la taille et de la diversification géographique qui l'accompagne. Celles-ci permettent une meilleure diversification des risques, l'accès à un financement 'corporate', un taux d'emprunt limité avec prime de risque réduite et, dans certains cas, la détention de larges compétences d'ingénierie pour le contrôle des coûts d'installation de grands équipements de production. Le choix des gouvernements de ne pas réformer radicalement les structures de marché reposait sur la conviction que des firmes intégrées verticalement et diversifiées horizontalement étaient une garantie pour que les capacités puissent être développées de façon adaptée aux besoins et s'inscrivent dans un panier de technologies préservant la sécurité énergétique, comme cela a été le cas en Allemagne, en Belgique, en Espagne et en France. L'avenir devrait probablement confirmer le bien-fondé de ce prémissse, notamment quand il faudra développer à grande échelle des équipements de production non émetteurs en CO₂.

Efficience de long-terme versus imperfections de marché

Pour résumer, le besoin d'investissements diversifiés en production et la gestion de risque des fournisseurs ont co-déterminé le mouvement d'évolution des structures industrielles, qui s'organisent autour d'une structure stabilisée d'oligopole vertical, marquée par un petit nombre de firmes intégrées verticalement, avec une frange concurrentielle. Mais un nouveau problème se pose, celui de l'équilibre à trouver entre les gains d'efficience de long terme permis par les arrangements verticaux d'un côté, et le risque de pouvoir de marché et les barrières à l'entrée que créent ces arrangements de l'autre côté. Formulée autrement, la question est de savoir jusqu'à quel degré l'intégration verticale et l'expansion des contrats de long terme peuvent être tolérées, degré au-delà duquel les échanges de court terme seraient trop limités pour que se maintienne un marché spot suffisamment liquide et se perpétuent les pressions concurrentielles à l'efficacité productive.

¹⁸ Voir la place des investissements en centrales charbon et nucléaire des sept principales entreprises historiques EDF, ENEL, Endesa, E.ON, Iberdrola, RWE, Vattenfall par rapport aux autres entreprises sur les marchés européens, in *Power in Europe, Investment Survey Supplement*, March 2008.

On notera d'abord que les contrats de long terme à prix fixés ou indépendants du prix spot ne sont pas une source d'aggravation de pouvoir de marché. Les opérateurs impliqués dans des ventes de long terme pour une grande partie de leur production ont moins d'incitations à spéculer sur le spot par des retraits de capacité (ce que montre des travaux théoriques généraux ou appliqués à l'électricité, voir Allaz B. et Vila J., 1993). Certains en viennent même à recommander que les grands producteurs soient contraints de vendre une partie donnée de leur production en contrat de long terme à prix fixés. (Willems B., et De Cortes E., 2008). Par ailleurs dans une perspective de politique de la concurrence, des structures de marchés imparfaites sont tolérables tant qu'il n'y a pas de *monopolization* au sens américain, c'est-à-dire d'acquisition ou de pérennisation de la position dominante par l'intermédiaire de pratiques limitant la concurrence et pouvant dissuader les entrées. Simplement la stabilisation des marchés ne doit pas se faire en confortant les positions acquises. Il faut maintenir les possibilités d'entrée et la crédibilité des menaces pour discipliner les entreprises en place.

Conclusion

Dans le modèle de référence des réformes, la coordination des échanges se fait de manière décentralisée sur une place de marché, la couverture de risque s'effectue grâce à des contrats à terme et à des contrats financiers d'options. Ceci évite le développement d'autres mécanismes de coordination et de couverture de risque (contrats long-terme et intégration verticale) et réduit ainsi les problèmes de forclusion et de pouvoir de marché. Dans cette perspective, le maintien de structures industrielles intégrées, des relations encadrées par des engagements de long-terme gêneraient le développement souhaitable de la concurrence et l'efficacité des coordinations par les prix. Ce serait donc le signe d'une réforme qui n'a pas été menée à son terme ou d'une trop grande tolérance de la part des autorités de la concurrence.

Ce point de vue mérite d'être mis en question. Un marché électrique n'est jamais à l'équilibre dans une perspective de long terme du fait de l'incapacité des signaux de court terme à indiquer les besoins de

capacité de long terme en niveau et structure d'équipements. Pour l'investisseur; il y a une difficulté générale à se référer aux prix du marché horaire et à prédire les conditions d'offre et de demande au-delà d'un petit nombre d'années. Pour préserver le processus de marché dans la durée et éviter les défaillances de marché qui se traduisent par un sous-capacité en extrême pointe et une concentration des investissements sur la technique la meilleure en gestion de risque, mais dont rien ne garantit la supériorité à long-terme, il est nécessaire de tolérer des imperfections de marché. Il s'agit, en termes plus généraux, d'accepter une divergence dans les structures industrielles vis-à-vis du cadre de la concurrence pure et parfaite afin de faciliter les coordinations de long terme entre entités opérant sur le marché. Cette tolérance conduit à un éloignement vis-à-vis des prescriptions traditionnelles en matière de politiques de concurrence.

Références

- Allais M., (1943), *Traité d'économie pure*, Paris, Imprimerie nationale.
- Allaz B., Vila J. (1993). "Cournot competition, futures markets and efficiency", *Journal of Economic Theory*, vol. 59, n°1
- Amendola M., Gaffard J.-L., Marty F. (2004), "Prices, costs and investment : why sticky prices are necessary. An illustration with the case of regulated-deregulated Industries", *10th Schumpeter Society International Conference*, Milan, juin.
- Amendola M., Gaffard J.L. (1998), *Out of Equilibrium*, Clarendon Press, Oxford
- Arrow K., Fisher A. (1974), "Environmental Preservation, Uncertainty, and Irreversibility", *Quarterly Journal of Economics*, vol. 88, n°1
- Bohn E., Caramanis M., Schwepp F., Tabors R. (1988), *Spot Pricing of Electricity*, Kluwer Academic Publishers
- Boiteux M. (1949), "De la tarification des demandes en pointe : application de la théorie de la vente au coût marginal", *Revue Générale d'Électricité*, vol. 58, n° 8
- Boiteux M. (1956), "Sur la gestion des monopoles publics astreints à l'équilibre budgétaire", *Econometrica*, January.
- Boroumand R., Zachmann G. (2009), "Risk management and vertical arrangements in liberalized electricity market: the retailer's perspective", working paper n°19, LARSEN, Fontenay aux Roses, www.gis-larsen.org
- Caballero R. (1991), "On the sign of the investment-uncertainty relationship", *American Economic Review*, vol. 81, n°1
- Caballero R. (2000), "Aggregate Investment: Lessons from the Previous Millennium", AEA Session. In Memoriam: Robert Eisner
- Carruth A., Dickerson A. and Henley A. (2000), "What do we know about investment under uncertainty?", *Journal of Economic Surveys*, vol. 14, n°2
- Chao H.-P., Peck S. (1998), "Reliability management in competitive electricity markets", *Journal of Regulatory Economics*, vol. 3
- Chao H., Oren S., Wilson R. (2008), "Restructured Electricity Markets: Reevaluation of Vertical Integration and Unbundling", in Sioshansi F. (dir.), *Competitive electricity markets. design, implementation and performance*, Elsevier, London
- Chaton C. (2001), "Décisions d'investissement et de démantèlement sous incertitude : une application au secteur électrique", *Economie et Prévision*, n°149
- Cramton P., Stoft S. (2006), *The Convergence of Market Designs for Adequate Generating Capacity, with Special Attention to the CAISO's Resource Adequacy Problem*, White Paper for the Electricity Oversight Board, March
- Defeuilley C., Meunier G. (2008), "Gestion du risque et intégration verticale dans l'électricité", Working paper n°12, LARSEN, Fontenay aux Roses
- De Vries L. (2007), "Generation adequacy : helping the market to do its job", *Utilities Policy*, vol. 15, n°1
- De Vries L., Neuhoff K. (2004), "Insufficient incentives for investment in electricity generation", *Cambridge Working Papers in Economics*, CWPE 0428, University of Cambridge, Department of Applied Economics
- Dixit A. (1992), "Investment and Hysteresis", *Journal of Economic Perspectives*, vol. 6, n°1
- Evans L., Meade R. (2005), *Alternating Currents or Counter-Revolution? Contemporary Electricity Reform in New Zealand*, Victoria University Press, Wellington
- Finon D., Roques F. (2008), "Contractual and financing arrangements for nuclear investment in liberalized markets : which efficient combination?", *Competition and Regulation in Network industries*, vol.9, n°3

- Finon D., Pignon V. (dir) (2008), "Special issue on capacity mechanisms in imperfect electricity markets", *Utilities Policy*, vol. 16, n°3
- Guesnerie R. (1992), "Est-il rationnel d'avoir des anticipations rationnelles", *L'actualité économique*, vol. 68, n°4
- Green R. (2004), "Retail competition and electricity contracts", CMI working papers 33, Cambridge University, Cambridge
- Hansmann H. (1988), "Ownership of the firm", *Journal of Law, Economics and Organization*, vol.4, n°2
- Hart O. (1995), *Firms, Contracts, and Financial Structure*. Clarendon Lectures in Economics, Oxford University Press, Oxford
- Henry C. (1974), "Investment decisions under uncertainty : the irreversibility effect", *American Economic Review*, vol. 64
- Hunt S. (2002), *Making competition work in electricity markets*, Wiley, London
- Joskow P. (2007), "Market electricity markets and investment in new generating capacity", in Helm D. (dir) *The New Energy Paradigm*, Oxford University Press, Oxford
- Joskow P., Schmalensee R. (1984), *Markets for Power*, Cambridge, MIT Press
- Joskow P., Tirole J. (2007), "Reliability and Competitive Electricity Markets", *Rand Journal of Economics*, vol. 38, n°1
- Littlechild S. (2008), "Municipal aggregation and retail competition in the Ohio energy sector", *Journal of Regulatory Economics*, vol.34, n°2
- Mansur E. (2007), "Upstream Competition and Vertical Integration in Electricity Markets", *Journal of Law and Economics*, vol. 50
- Marty F. (2007), "La sécurité de l'approvisionnement électrique : quels enjeux pour la régulation?", *La Revue de l'OFCEN*, n° 101, avril
- Michaels R. (2006), "Vertical Integration and the Restructuring of the US Electricity Industry", *CATO Institute Policy Analysis Series*, 572
- Newbery D. (2005), "Electricity liberalisation in Britain: the quest for a satisfactory wholesale market design", *Energy Journal*, Special Issue on European Electricity Liberalisation
- Oren S. (2003), *Ensuring generation adequacy in competitive electricity markets*, Univ. of California Energy institute, Working Paper UCEI, EPE-007
- Pindyck R. (1991), "Irreversibility, Uncertainty, and Investment", *Journal of Economic Literature*, vol. 29, n°3
- Pindyck R., Dixit A. (1994), *Investment under uncertainty*, Princeton University Press, Princeton
- Stiglitz J. (1989), "Incentives, Information and Organizational Design," *Empirica*, vol. 16, n°1
- Stoft S. (2002), *Power System Economics*, Wiley Interscience
- Thomas S. (2006), "The British model in Britain : failing slowly", *Energy Policy*, vol 34, n°5
- Vazquez C., Rivier M., Pérez-Arriaga I. (2002), "A Market approach to long-term security of supply", *IEEE Transactions on Power Systems*, vol. 17
- Willems B., De Corte E. (2008). "Market power mitigation by regulating contract portfolio risk", *Energy Policy*, vol. 36, n°10
- Wilson R. (2002), "Architecture of power markets", *Econometrica*, vol. 70

Working papers 2006-2009 du Gis Larsen

2006

La gestion du risque d'une commodité non-stockable

*Christophe Defeuilley,
Guy Meunier*

Electricité et sécurité de fourniture de long-terme

Dominique Finon, Virginie Pignon

Les conditions de l'investissement nucléaire en environnement de marché

Dominique Finon

Alternative market designs in EU electricity balancing

*Jean-Michel Glachant,
Marcelo Saguan*

2007

La concurrence de détail dans l'électricité

Christophe Defeuilley

Sub optimal investment in oligopolistic electricity markets

Dominique Finon, Guy Meunier

La formation des prix sur un marché à dominante nucléaire

Elliot Romano

2008

La hausse inéluctable des prix de l'électricité en France

Dominique Finon, Jean-Michel Glachant

The social efficiency of long-term capacity reserve mechanisms

Dominique Finon, Guy Meunier, Virginie Pignon

Les politiques de la concurrence sur les

marchés électriques en Europe

Ute Dubois

Gestion du risque et intégration verticale dans l'électricité

*Christophe Defeuilley,
Guy Meunier*

Investment risk allocation in restructured electricity markets

The need of vertical arrangements

Dominique Finon, Yannick Perez

Financing arrangements and industrial organisation for new nuclear build in electricity markets

Dominique Finon, Fabien Roques

Legal uncertainty and competition policy in deregulated network industries. The case of long term contracts

Adrien de Hauteclocque

Le modèle d'activité du fournisseur pur d'électricité est-il viable ?

Raphael Boroumand

Generation adequacy and transmission interconnection in regional electricity markets

Mauricio Cepeda, Marcelo Saguan, Virginie Pignon

La dynamique de la concurrence avec coûts de changement. British Gas (1997-2007).

*Christophe Defeuilley,
Mathieu Mollard*

Capacity choice, technology mix and market power

Guy Meunier

2009

Accompagner l'extinction des tarifs réglementés

*Christophe Defeuilley,
Dominique Finon*

Photovoltaïque, l'inadéquation du tarif d'achat

Dominique Finon

HHI, an irrelevant indicator without a relevant market

*Delphine Perrot-Voisard,
Georg Zachmann*

Retailers'risk management and vertical arrangements in electricity markets

*Raphaël H.Boroumand,
Georg Zachmann*

Large-scale wind power in European electricity markets : time for revisiting support schemes and market designs ?

Céline Hiroux, Marcelo Saguan

Catching the maximum market value of the electricity storage

Xian He, Georg Zachmann

Signaux-prix et équilibre de long-terme. Reconsidérer l'organisation des marchés électriques

*Dominique Finon,
Christophe Defeuilley,
Frédéric Marty*

Comparison of long-term contracts and vertical integration in decentralised electricity markets

Richard Meade, Seini O'Connor



larsen

Laboratoire
d'Analyse
économique
des Réseaux et
des Systèmes
Energétiques

Larsen est un Groupement d'Intérêt Scientifique (GIS) créé en 2006, dont la vocation est de participer aux recherches réalisées dans le domaine de l'analyse économique des réseaux et des systèmes énergétiques, et de contribuer à la réflexion scientifique sur la libéralisation des industries électriques et gazières. Sa mission consiste à stimuler les recherches sur des objectifs scientifiques propres et favoriser l'émergence de nouveaux projets. Elle consiste également à diffuser les connaissances scientifiques et la promotion des meilleurs travaux nationaux ou internationaux par l'organisation de séminaires et de colloques sur les grandes questions scientifiques, théoriques ou appliquées. Larsen rassemble des chercheurs et doctorants en provenance du CIRED (Centre International de Recherche sur l'Environnement et le Développement), unité mixte de recherche de l'EHESS et du CNRS, en provenance des équipes de recherche éco, d'EDF Recherche et Développement et enfin, du GRJM (Groupe Réseaux Jean Monnet), une équipe du laboratoire ADIS de l'Université Paris Sud.