FACULTAD DE ECONOMÍA

DEPARTAMENTO DE ECONOMÍA FINANCIERA Y ACTUARIAL

IMPACT OF RENEWABLES IN SPANISH ELECTRICITY MARKETS

PROGRAMA DE DOCTORADO: FINANZAS Y ECONOMÍA CUANTITATIVAS

Tesis presentada por:

Mª Cristina Ballester Chaves

Directora:

Mª Dolores Furió Ortega

Valencia, julio de 2021

A mi padre

Agradecimientos

La presente Tesis Doctoral no habría sido posible sin el apoyo del Ministerio de Educación y Ciencia y la Universidad de Valencia, en el marco de los proyectos ECO2013- 40816-P, PGC2018-093645-B-100 y UV-INV- PRECOMP13-115547. Junto con las valiosas aportaciones de los profesores Hipòlit Torró y Constantino J. García, así como de los participantes de la X Conference Spanish Association of Energy Economics, 2015 y 8th Research Workshop on Energy Markets, 2017.

Mención principal merece el trabajo de Mª Dolores Furió como directora de este proyecto de Tesis, por su extraordinaria y fuerte voluntad, dedicación, esfuerzo y apoyo continuado en el tiempo.   También, me gustaría agradecer la labor docente de todos los profesores del Programa de Doctorado de Finanzas y Economía Cuantitativas, y el apoyo de todos compañeros. En particular, la atención e interés a lo largo del tiempo del coordinador Ángel Pardo, y de los profesores del Departamento de Economía Financiera y Actuarial como Julio Lucía y Francisco Climent; y la amistad y el aliento continuo de compañeros como Beatriz Martínez, José Mª Ballester, Amparo Nagore, Verónica Serrano y Nerea Alarcón.

Debo además agradecer la comprensión y generosidad de mis amigas, y jefas, del trabajo, Teresa García y Reyes Mas, y de mis compañeros Jorge Cano, Mireia Ariza, Manuel Terrádez, Ana Palomares, Diego Serrano, Vicente Castellar, Jordi Ruiz, Elvira Esteve (y todos los demás), gracias por enseñarme tanto, hacer que todo sea más fácil, y estar siempre dispuestos a ayudar.

Y, sobretodo, gracias al apoyo y cariño de mi madre, Mª Luisa, mis hermanas, Elena y Kitty, mis sobrinos, Diego, Elena, Rocío, Elisa, y toda mi familia. Sois lo mejor que me ha pasado en la vida, mi guía y ejemplo para todo. Os admiro y os quiero.

Y en especial, gracias a mi padre, que me llevó de la mano el primer día a la Universidad, y me ha enseñado lo que es tener ilusión con todas las cosas, y con la gente que te encuentras en el camino.

Gracias a todos.

Índice

Introducción

El objetivo de la presente tesis es analizar el impacto de la introducción de las energías renovables en el sistema eléctrico español.

La potencia de generación de electricidad a partir de fuentes de generación renovable ha aumentado considerablemente en los últimos años. El porcentaje del total de potencia instalada en la península ibérica que pertenece a energías renovables en 2007 era el 17%, y en 2013 ascendió al 30%. A continuación, se frenó significativamente la inversión en energías renovables, para retomarse posteriormente a partir de 2018[1]. Este cambio obedece a una mayor concienciación medioambiental en línea con los compromisos internacionales de reducción de emisiones contaminantes asumidos por España, y también al ahorro de costes para las empresas que así consiguen reducir su necesidad de permisos contaminantes. En consecuencia, el *mix* de generación español cambia radicalmente, pasa de estar compuesto principalmente por fuentes de generación convencionales y seguras (en el sentido de la garantía de suministro), a fuentes de generación renovables y de producción intermitente, sobre todo la eólica y la fotovoltaica, que dependen de la existencia de viento y sol (La generación eólica anual pasa de 27.611,65 GWh en 2007, siendo el 10% del total, a 54.713,25 GWh en 2013, el 21% del total generado[2]).

Uno de los temas más discutidos sobre la nueva realidad del mercado eléctrico tras la irrupción de las renovables es la cantidad de recursos destinados al fomento de estas tecnologías de generación. Por un lado, se generaliza la opinión de la supuesta

[1] www.ree.es.
[2] www.ree.es.

insostenibilidad de la remuneración a las energías renovables, que explica la burbuja inversora en España hasta 2013, y su congelación posterior; por otro lado, se establece la necesidad de mantener en el sistema fuentes de generación no-intermitentes para los casos en los que las renovables no están disponibles, las cuales exigen alguna retribución por el mero hecho de servir de *back-up*.
Otro de los temas que están en el centro del debate se relaciona con la mencionada intermitencia de las principales fuentes de generación renovable, la cual se traduce en que su producción es difícilmente predecible, lo que puede aumentar la incertidumbre en el mercado, la volatilidad de los precios y, también, la necesidad de realizar ajustes para garantizar el suministro eléctrico.

En un mercado, como es el de la electricidad, en el que, en la actualidad, el almacenamiento a gran escala no es viable económicamente, es necesario negociar con antelación la entrega, lo cual se lleva a cabo en el mercado diario (con un día de antelación), y posteriormente realizar ajustes para corregir desviaciones y garantizar el suministro eléctrico. El impacto de estas tecnologías no se limita, por tanto, al mercado diario, donde se concentra la mayor liquidez, sino que se extiende, asimismo, y de forma significativa, a los mercados o segmentos de negociación posteriores, los cuales deben ser asimismo objeto de atención por parte de los reguladores para adaptarse a la nueva realidad.

El presente trabajo extiende el estudio del impacto de las renovables sobre el precio en el mercado diario, cuestión que ha sido ampliamente analizada en la literatura, abordando aspectos menos estudiados y complejos, como el efecto del comportamiento estratégico de los participantes y el estudio de dinámicas entre los mercados o segmentos dentro del Sistema Eléctrico. Estas dinámicas pueden ser clave para lograr una mejor integración de las renovables.

El sistema eléctrico español es elegido como ejemplo paradigmático debido al intenso crecimiento observado en las renovables en los últimos años, en especial de energía eólica. Los resultados obtenidos son de utilidad tanto para reguladores como participantes del mercado, profundizando en un mayor conocimiento acerca del funcionamiento del Sistema Eléctrico.

Resumen de los Capítulos

Los capítulos que constituyen el contenido de la Tesis son:

- Capítulo 1: Effects of renewable on the stylized facts of electricity prices.
- Capítulo 2: Impact of wind electricity forecast on bidding strategies.
- Capítulo 3: Analysing the impact of renewables on Spanish electricity final prices using machine learning techniques.

En primer lugar, en el Capítulo 1, se analiza el impacto de las renovables en el precio fijado en la subasta del mercado diario, o precio *spot*, siendo objeto de análisis no solo el comportamiento del precio en niveles (que se espera se reduzca debido a los menores costes marginales de las renovables), sino también sus características principales, como son su elevada volatilidad y la existencia de *spikes (*o saltos inesperados en el precio). Ambas características se observan con frecuencia en los precios eléctricos, pudiendo cobrar aún mayor importancia dada la intermitencia y menor predictibilidad de la producción renovable.

A continuación, en el Capítulo 2, se amplía el estudio extendiendo el análisis a los precios ofrecidos por los agentes en la subasta del mercado diario. Si consideramos que las renovables tienen un impacto relevante en el precio *spot*, entonces es razonable esperar que los participantes tengan en cuenta la previsión de energía renovable (fundamentalmente eólica por su mayor penetración) a la hora de tomar decisiones sobre el precio y la cantidad a ofertar. Es interesante, por tanto, contrastar el comportamiento estratégico de los agentes incorporando sus estimaciones de producción de energía eólica en su matriz de información

Por último, en el Capítulo 3, se completa el estudio mediante el análisis del efecto de las energías renovables en el precio final de la electricidad y en cada uno de sus componentes; componentes que recogen los costes en los que incurre el sistema por la puesta en marcha del resto de mercados y procesos de ajuste que tienen lugar después del mercado diario. En general, se espera que todos estos costes aumenten ante la mayor necesidad de realizar ajustes derivada de la intermitencia y menor predictibilidad de la producción renovable.

Resumen del Capítulo 1:

Effects of renewable on the stylized facts of electricity prices.

Este Capítulo se ha publicado en 2015 bajo el título "Effects of renewable on the stylized facts of electricity prices" en la revista *Renewable and Sustainable Energy Reviews,* revista científica de reconocido prestigio internacional (primer cuartil del © 2019 CLARIVATE ANALYTICS y primer cuartil del Scimago Journal Rank).

Introducción

El precio de la electricidad se caracteriza por la elevada volatilidad y la existencia de *spikes*, o saltos no esperados en el precio ([1]-[6]). Debido a ello, es práctica habitual incorporar procesos discretos de saltos en los modelos de predicción de precios eléctricos, combinados con procesos continuos con la finalidad de poder recoger este comportamiento específico ([7]).

Una idea generalizada en la literatura es que el aumento de las renovables en el mix de generación eléctrica va a provocar la disminución del precio en las subastas del mercado diario. Esto se debe a una cuestión meramente técnica derivada del propio mecanismo de mercado. En la mayoría de países, incluido España, el mercado eléctrico funciona mediante subastas de precio marginal uniforme, que se celebran con antelación al momento de consumo de la energía. El mecanismo más utilizado para asignar la energía en estas subastas es el del orden de mérito; esto es, los agentes del mercado remiten sus ofertas de venta o de compra de electricidad (precio y cantidad) para cada uno de los periodos horarios del día siguiente. Las ofertas se ordenan para construir las curvas agregadas de oferta y de demanda. Y, el precio del mercado diario, o también denominado precio *spot,* es el precio que oferta la última unidad de venta que entra en la casación para satisfacer la demanda, donde las curvas agregadas de demanda y oferta se encuentran, y acaba aplicándose por igual a todos los agentes que participan en el mercado. Las plantas de generación renovable tienen en general costes marginales más bajos que las plantas de generación convencional y pueden ofrecer la energía que producen a precios ofertados inferiores. Por lo tanto, un aumento en las ofertas de renovables debería tener como consecuencia el desplazamiento de la curva de oferta del mercado de tal modo que el precio *spot* podría acabar fijándose en un nivel inferior. Este efecto de las renovables se ha denominado en la literatura *efecto de orden de mérito*. En el caso del mercado español, en [8] se estima, utilizando datos de 2006, que la reducción en el precio *spot* causada por el *efecto de orden de mérito* puede llegar a compensar por el coste en el que se incurre al incentivar la inversión en energía renovable. Por su parte, en [9], utilizando datos horarios, también del mercado español, desde 2005 hasta 2009 se concluye que un incremento marginal de la producción renovable igual a 1GWh puede conllevar una reducción en el precio *spot* de alrededor de 2€/MW.

Sin embargo, hay que tener en cuenta que la producción renovable es intermitente, dado que depende de condiciones climatológicas, de la existencia de viento a una determinada velocidad, o de las horas de sol. De este modo, se considera que las plantas de generación renovable tienden a ser menos predecibles, comparadas con las de energía nuclear o de origen fósil. En este sentido, en la medida en que se recurra *menos* a las fuentes de energía convencional y *más* a las fuentes de energía intermitentes, como las renovables, se podría esperar un aumento en la ocurrencia de saltos en los precios, y/o un incremento de la volatilidad de los precios de la electricidad. No obstante, podría no ser así, ya que al mismo tiempo que aumenta la exposición de los precios a la volatilidad de las renovables, también se reduciría la exposición a la volatilidad del precio de las fuentes de energía convencional.

En este trabajo vamos a profundizar en el análisis del efecto de las renovables en el precio *spot* en el mercado diario español de la electricidad, utilizando un periodo de análisis más amplio que en los trabajos previos (desde 2002 hasta el último año disponible 2013), y completando el estudio con el análisis no sólo del precio en niveles, sino también de sus características principales: volatilidad y ocurrencia de saltos o *spikes*.

Objetivo y Metodología

El objetivo principal de este trabajo es verificar si efectivamente el precio del mercado diario, o *spot*, se ha reducido como consecuencia del incremento de la producción renovable en el mercado; si se ha vuelto más volátil; y, por último, si ocurren con mayor frecuencia los denominados saltos inesperados o *spikes*.

Para ello, los datos utilizados son series temporales de precios *spot* en euros/MWh del mercado diario español con frecuencia horaria desde 2001 hasta 2013 (el último año disponible en el momento de la realización de este análisis). También se utilizan series categóricas que nos indican cuál es la tecnología, o tecnologías, que marcan el precio *spot* cada hora, es decir cuál es la tecnología de generación del participante, o los participantes, en el mercado cuya oferta de precio de venta coincide con el precio resultante de la subasta. Los tipos de tecnología que se consideran en el estudio son los siguientes: renovables (principalmente eólica, pero también solar, cogeneración, biomasa y tratamiento de residuos), térmicas (carbón y fuel-gas), ciclo combinado, nuclear, hidráulica e hidráulica de bombeo. Finalmente, también se utilizan series de volumen de energía casada en la subasta del mercado diario en MWh, con frecuencia horaria de cada uno de los grupos de tecnologías mencionados desde 2008 hasta 2013.

El método de aproximación al problema que se ha seguido ha sido doble: por un lado, se ha realizado un análisis preliminar descriptivo para detectar eventos relevantes en la evolución del precio *spot*; también se ha analizado la evolución de cuál es la tecnología que marca el precio marginal o *spot* en cada momento; y cambios relevantes en el *mix* de generación. Además, este análisis ha resultado ser muy útil para adquirir un valioso conocimiento del mercado. Y, por otro lado, una vez identificados los cambios significativos, se realiza un análisis más profundo mediante herramientas econométricas.

Las herramientas utilizadas son las siguientes:

- Para analizar la relación entre el porcentaje de producción renovable y la tecnología que marca el precio marginal se ha utilizado una regresión lineal. En esta regresión, la variable dependiente es el porcentaje de veces que cada tecnología marca el precio marginal; y la variable independiente es el porcentaje de producción renovable. Las tecnologías consideradas en el estudio son: térmica, ciclo combinado, hidráulica, hidráulica de bombeo y nuclear.
- Para analizar la relación entre el porcentaje de producción renovable y el precio marginal se ha utilizado también una regresión lineal. En este caso la variable dependiente en la regresión es el precio *spot*; y la variable independiente es el porcentaje de producción renovable. Además, aunque el principal foco de atención es el estudio del impacto de las renovables, se amplía el análisis para estudiar también el efecto de otros tipos de tecnologías de generación (térmica, ciclo combinado, hidráulica, hidráulica de bombeo y nuclear).
- Para analizar la relación entre el porcentaje de producción renovable y la volatilidad y la ocurrencia de saltos o *spikes* en el precio, en primer lugar, se ha utilizado el modelo de Cartea y Figueroa (20005) ([7]) para estimar la volatilidad de los precios a lo largo del tiempo e identificar la ocurrencia de saltos. Este modelo es de uso habitual y resulta útil para nuestro propósito ya que incorpora las principales características de los precios eléctricos (reversión a la media, alta volatilidad y ocurrencia de saltos). Una vez realizada la estimación del modelo de precios, se analiza la relación entre la volatilidad de los precios y la volatilidad de la producción renovable mediante un test de correlación de Pearson, repitiendo también el mismo análisis con las otras tecnologías. Y, finalmente, para profundizar en la ocurrencia de saltos, se estima un modelo logístico, en el que la variable dependiente es el logit de la probabilidad de ocurrencia de un salto en el precio, distinguiendo entre: saltos inesperados en el precio hacia arriba (que denominamos saltos positivos), y saltos inesperados en el precio hacia abajo (saltos negativos). Las variables independientes en este modelo de regresión son las series de porcentajes de producción de cada tecnología en diferencias.

Todos los análisis de este capítulo se realizan tres veces, una primera considerando todas las horas de entrega de la energía, una segunda vez considerando solo las horas de entrega de la energía en las que la demanda es elevada, denominadas horas pico (desde las 08:00h hasta las 20:00 h en días laborales), y una tercera vez considerando las horas de entrega de la energía en las que la demanda es baja, denominadas horas valle (desde 00:00h hasta las 08:00h y desde las 20:00h hasta 24:00h en días laborales y las 24h en días no laborales).

Resultados

Los principales resultados del análisis preliminar descriptivo son los siguientes:

- La tecnología de ciclo combinado desde 2006 se convierte en la tecnología que más veces marca el precio marginal en la subasta a lo largo del periodo muestral. Sin embargo, a partir de 2010 es reemplazada por las centrales con tecnología térmica o hidráulica.

- Respecto a la tecnología renovable, aun cuando tiene menos posibilidades de marcar el precio marginal debido a que se sus costes marginales son más bajos, es interesante observar cómo desde 2010 empieza a marcar más veces el precio marginal que en el periodo anterior, en especial en horas en las que la demanda de electricidad es baja (horas valle).

- En cuanto a la evolución de los precios marginales o *spot*, en el análisis preliminar descriptivo de su evolución, no se observa evidencia de que desde 2001 hasta 2013 haya disminuido en nivel significativamente o haya aumentado su desviación típica. Lo que sí se observa es que desde 2010 el coeficiente de asimetría, que antes era positivo, ha pasado a ser negativo, lo que podría tener relación con la caída del precio debido a un aumento en el volumen de ofertas de las renovables. El impacto de las renovables sobre la distribución del precio *spot*, por tanto, podría estar más relacionado con la simetría de la distribución (aparición de precios más bajos con mayor frecuencia), y no tanto con un cambio significativo en su media o en su desviación típica.

- Asimismo, a partir de 2010, el número de *outliers* identificados en la serie de precios *spot* pasa a ser mayor en horas valle que en horas pico, mientras que en el periodo anterior ocurría lo contrario.

- Finalmente, del análisis de la evolución del porcentaje de energía casada en la subasta agrupando la energía por tecnologías de generación se deduce un decrecimiento progresivo de la tecnología de ciclo combinado desde 2008 hasta 2013, al contrario que la producción renovable que presenta un crecimiento continuado.

Los principales resultados de la segunda parte del capítulo, el estudio econométrico, son los siguientes:

- Existe una relación negativa entre el porcentaje de veces que las centrales de generación de ciclo combinado marcan el precio marginal y el porcentaje de renovables (existiendo para otras tecnologías una relación positiva). De este modo, se obtiene evidencia de que la tecnología de ciclo combinado se ha visto en parte desplazada por las renovables en el mercado diario.

- El precio *spot* del mercado diario de electricidad español disminuye cuando hay más generación de origen renovable. Este resultado concuerda con lo observado en la literatura previa, y se debe al propio mecanismo de la subasta en el que las ofertas de generadores con costes variables más bajos (como es el caso de las renovables) son las primeras en resultar casadas (el denominado *efecto de orden de mérito*). Por el contrario, el precio *spot* aumenta cuando hay más generación de las centrales de ciclo combinado y térmicas, resultado que se explica por sus costes marginales más elevados.

- Se detecta la existencia de periodos de intensa aparición de *spikes*, o saltos en el precio, ocurridos entre mayo y junio de 2010 y de 2013, los cuales se producen con mayor intensidad en horas valle, coincidiendo con elevados picos de volatilidad en el precio. Sin embargo, no parecen tener relación con aumentos en la volatilidad de la producción renovable. Por el contrario, coinciden con la aparición de picos de volatilidad en la producción de otras tecnologías, como son la nuclear, la hidráulica y la hidráulica de bombeo.

- El test de correlación confirma que, si excluimos los periodos de intensa volatilidad en el precio, existe una elevada correlación positiva entre la volatilidad de la producción renovable, que es superior a la volatilidad de otras tecnologías debido a su naturaleza intermitente, y la volatilidad del precio spot. Sin embargo, si no excluimos dichos periodos, la correlación baja considerablemente (pasa de +63.4% a +27,9%). Por el contrario, la tecnología que muestra valores de correlación más bajos en todos los casos, incluso negativos, entre su volatilidad y la volatilidad de los precios, es la tecnología de ciclo combinado.

- En cuanto a la ocurrencia de saltos o *spikes* en el precio, los resultados de los modelos logísticos nos indican que un incremento de la producción renovable reduce la probabilidad de que ocurran saltos al alza (positivos) en los precios en horas pico, siendo no relevante su efecto en los precios en horas valle. Por el contrario, el incremento de producción térmica aumenta la probabilidad de que ocurran saltos positivos en el precio en horas valle, al tiempo que reduce la probabilidad de que ocurran saltos negativos en el precio, tanto en horas pico como en horas valle.

Conclusiones

Se confirma para el mercado diario que el incremento de producción renovable reduce el precio *spot* y existe correlación positiva entre la volatilidad de la producción renovable y la volatilidad de los precios. Sin embargo, en el caso de la volatilidad, los resultados obtenidos también ponen de manifiesto que otras tecnologías diferentes a las renovables, como la nuclear y la hidráulica, explican la aparición de la aparición de picos notables de volatilidad. Por otro lado, en contra de lo esperado, el porcentaje de producción renovable ha contribuido a suavizar la ocurrencia de saltos inesperados en el precio (*spikes*) característicos de los precios eléctricos, ya que el incremento de la producción renovable durante las horas pico de demanda ha tenido como resultado la reducción de la probabilidad de que ocurran saltos al alza en el precio.

La distinción entre horas pico y horas valle es sin duda relevante para entender el efecto de la introducción de las energías renovables en el precio del mercado diario, resultando que es en horas valle donde se constata una mayor volatilidad y mayor ocurrencia de *spikes* o saltos en el precio, siendo la producción de las tecnologías térmica y de ciclo combinado las responsables de dicho incremento en la ocurrencia de saltos positivos en el precio (precios más altos).

Los resultados de este capítulo ponen de manifiesto que el impacto de la penetración de renovables en la estructura de generación sobre los precios y el comportamiento de los mismos no se limita a las estrategias de negociación de las centrales de energía renovables que irrumpen en el mercado sino que comporta una acción o reacción del resto de tecnologías, manifestada en su comportamiento asimismo estratégico para adaptarse a las circunstancias del nuevo contexto, en especial aquéllas con capacidad para afectar al precio. Esta cuestión se aborda en el siguiente Capítulo.

Referencias

[1] Schwartz E. The stochastic behavior of commodity prices: valuations and hedging. J Finance 1997;52(3):923–73.

[2] Lucia J, Schwartz E. Electricity prices and power derivatives: evidence from nordic power exchange. In: Review of derivatives research, vol. 5 (1); 2002. p. 5–50.

[3] Deng S. Stochastic models of energy commodity prices and their applications: mean-reversion with jumps and spikes. United States of America: University of California Energy Institute; 2000 (Working paper, PWP-073).

[4] Knittel CR, Roberts MR. An empirical examination of restructured electricity prices. Energy Econ 2005;2005(27):791–817.

[5] Huisman R, Mahieu R. Regime jumps in power prices. Energy Econ 2003;25 (5):425–34.

[6] Karakatsani N, Bunn DW. Modelling the volatility of spot electricity prices, working paper. London: Department of Decision Sciences, London Business School; 2004.

[7] Cartea A, Figueroa G. Pricing in electricity markets: a mean reverting jump diffusion model with seasonality. Appl Math Finance 2005;12(4):313–35.

[8] Sa. nz de Miera G, Del Río P, Vizcaíno I. Analyzing the impact of renewable electricity support schemes on power prices: the case of wind electricity in Spain. Energy Policy 2008; 36:3345–59.

[9] Gelabert L, Labandeira X, Linares P. An ex-post analysis of the effect of renewables and cogeneration on Spanish electricity prices. Energy Econ 2011; 33:S59–65.

Resumen del Capítulo 2:

Impact of wind electricity forecast on bidding strategies.

El Capítulo 2 de la presente tesis doctoral ha sido publicado bajo el título "Impact of wind electricity forecast on bidding strategies" en la revista *Sustainability,* revista científica de reconocido prestigio internacional (segundo cuartil del © 2019 CLARIVATE ANALYTICS y segundo cuartil del Scimago Journal Rank).

Introducción

En la mayoría de los países de nuestro entorno, la electricidad se negocia principalmente con un día de antelación en el mercado diario. Este mercado funciona mediante un mecanismo de subasta de precio marginal uniforme, en el que consumidores y generadores envían sus ofertas de compra y de venta de electricidad, compuestas por precio y volumen, para cada una de las 24 horas del día siguiente. Estas ofertas se ordenan por su precio y el precio resultante de la subasta, o precio *spot*, es aquel que coincide con el precio de venta de la oferta de venta más cara que se necesita para satisfacer el total de la demanda de electricidad para cada hora.

El precio al que una central de generación de electricidad está dispuesta a vender su producción depende de sus costes de producción y estos pueden ser muy diferentes dependiendo de la tecnología utilizada. Las centrales renovables, por ejemplo, tienen costes marginales más bajos y, por tanto, pueden ofrecer su energía a precios menores. Debido a ello, se sitúan en la base de la curva agregada de oferta y suelen ser las primeras en resultar casadas en la subasta del mercado diario. El aumento de la producción de electricidad procedente de fuentes de generación renovable generalmente traerá consigo la reducción del precio marginal o *spot*, ya que el aumento de sus ofertas de venta desplaza la curva de oferta del mercado hacia la derecha y el precio marginal se marca a un nivel inferior, efecto conocido como el *efecto de orden de mérito* de las renovables.

Si el aumento de las ofertas renovables en la subasta del mercado diario tiene el impacto esperado - descenso del precio - y este impacto resulta ser relevante, es lógico esperar que todos los agentes reaccionen incorporando la previsión renovable en sus decisiones de oferta con el fin último de maximizar sus beneficios. En especial, un cambio de estrategia de los agentes que se encuentran con mayor frecuencia entre los que marcan el precio marginal podría tener un impacto significativo en el precio *spot* que no podemos ignorar.

La complejidad de las estrategias de oferta en el mercado eléctrico ha capturado la atención de distintos autores en la literatura ([1]-[5] entre otros). En particular, en [6] se analizan las estrategias de oferta mediante modelos teóricos y se deduce que el *efecto de orden de mérito* de las renovables podría verse incrementado debido a la actuación de participantes con poder de mercado. A su vez, en [7], utilizando datos empíricos del mercado británico, se obtiene que los precios eléctricos podrían ser más elevados, y más volátiles, en escenarios con poder de mercado. En el caso español ([8]), en el periodo temporal que va desde 2002 hasta 2005, se detecta un comportamiento estratégico diferente en las centrales térmicas de fuel dependiendo de su tamaño. Finalmente, también en el caso español, en [9], se identifica un comportamiento estratégico seguido por algunas centrales de generación que consiste en aumentar el precio al que quieren vender en el mercado diario, lo que les dejaría fuera de la casación en la subasta, siendo requerida su participación posteriormente en los servicios de ajuste (en concreto en el proceso de restricciones técnicas) buscando con ello obtener mayores beneficios. Esto tiene lugar en el periodo que va desde julio de 2004 hasta febrero de 2005.

El enfoque de este trabajo es diferente, ya que el interés reside en analizar primero en qué medida los precios de venta ofrecidos por todos los agentes, no sólo los de las centrales de generación térmica, sino también los de las centrales de ciclo combinado, hidráulicas, nucleares y renovables, dependen de la previsión de la producción eólica para el día siguiente; así como estudiar cómo a su vez cada estrategia de oferta impacta en la curva agregada de ofertas del mercado, y por tanto en el precio *spot*.

Objetivo y Metodología

El objetivo de este capítulo es identificar cuáles son los factores de los que dependen los precios ofertados por parte de las centrales de generación de electricidad en la subasta del mercado diario español. En particular, es interesante analizar el efecto que provoca, en su caso, la previsión de la producción eólica para el día siguiente en el nivel ofertado del precio, así como identificar diferencias de comportamiento dependiendo de la tecnología de generación de la central que realiza la oferta, con el fin de dilucidar cómo las diferentes estrategias acaban impactando en el precio marginal resultante de la subasta (el precio *spot*).

Los datos utilizados son series temporales desde 2010 hasta 2013 de unidades de oferta compuestas por precio (euros) y volumen de energía (MWh) presentadas a la subasta del mercado diario para comprar o vender electricidad, tanto las series de ofertas finalmente casadas como las que no resultan finalmente adjudicadas, para cada una de las 24 horas del día siguiente; la previsión de producción eólica para cada hora del día siguiente que Red Eléctrica (REE) publica justo antes de la hora límite de entrega de las ofertas de la subasta. También se incluyen otros datos relevantes para los precios eléctricos como las series de precios del gas natural negociadas en el *National Balance Point* (NBP), las series de futuros sobre el precio de los permisos de emisión de CO2 *European Union Allowances* (EUA) con vencimiento anual y frecuencia diaria o la reserva de agua en España, serie disponible con frecuencia semanal. En total, el conjunto de datos utilizado contiene alrededor de 72 millones de registros.

La metodología utilizada para abordar el análisis consiste en la estimación de modelos de datos de panel en los que la variable dependiente es el precio medio ofrecido en la subasta del mercado *spot* de los agentes que comparten la misma tecnología de generación. Los grupos de tecnología considerados son: ciclo combinado, térmica (que incluye carbón, fuel-gas y fuel-oil), hidroeléctrica, nuclear, y resto de tecnologías de generación, principalmente renovables. Se estiman, por tanto, un total de 6 modelos de datos de panel diferentes, uno para cada tipo de tecnología de generación.

En cada modelo de datos de panel se incluye como sección cruzada la hora de entrega de la electricidad (24 horas). Esto tiene sentido ya que en un mismo momento del tiempo los participantes del mercado deben presentar sus ofertas para cada una de las 24 horas del día siguiente. Cabe esperar, por tanto, que los precios ofrecidos estén correlacionados entre sí, siendo al mismo tiempo diferentes atendiendo a los distintos niveles de demanda de cada hora del día.

Una de las variables explicativas en los modelos de datos de panel es la previsión eólica elaborada por el operador del sistema eléctrico español, Red Eléctrica Española (REE). Esta previsión se pone a disposición de todos los participantes del mercado justo antes de la hora en la que deben enviar sus ofertas para las subastas del mercado diario. Adicionalmente, junto con la previsión eólica, se incluyen otras variables que podrían influir en las decisiones de oferta de los agentes como son: el precio del gas natural negociado en el *hub* NBP; el precio de las emisiones de CO2 (EUA); la reserva de agua en España, que se incluye como porcentaje sobre la capacidad total de los embalses~~;~~ una variable ficticia indicadora de si se trata o no de un día laborable y el precio medio ofrecido del día anterior (lag 1 de la variable dependiente).

Con anterioridad a la estimación de los precios medios de oferta se realizan test de ajustes para identificar el tipo de modelo de datos de panel adecuado para cada tecnología. Los test aplicados son: el test F, que sirve para contrastar si la sección cruzada, en este caso la hora de entrega, es significativa; El test de Durbin-Wu-Haussman para seleccionar si el modelo de datos de panel más adecuado es el modelo de datos de panel de efectos fijos o de efectos aleatorios; y el test de Maddala-Wu para contrastar la falta de estacionariedad en las series. Los resultados de los test indican que el modelo de datos de panel adecuado para explicar los precios medios ofrecidos por las centrales de ciclo combinado, térmicas e hidráulicas es el modelo de datos de panel con efectos fijos, siendo la sección cruzada la hora de entrega. Por lo tanto, existen características diferentes que dependen de cada hora de entrega que permiten explicar los precios ofrecidos de estas centrales. Por el

contrario, en el caso de las centrales nucleares y renovables, es más adecuado utilizar un modelo de regresión sin sección cruzada.
Adicionalmente, en todos los casos el precio del gas natural (NBP) y el precio de las emisiones de CO2 (EUA) se incluyen en los modelos a estimar en forma de ratio debido a su elevada correlación, con el fin de evitar problemas de multicolinealidad. Para controlar la multicolinealidad, se ha utilizado el indicador VIF (variance inflation factor), obteniéndose que este factor se mantiene en niveles aceptables (inferiores a 2) incluyendo ambas variables como ratio.

Una vez estimados los modelos de datos, y por tanto el efecto de la previsión eólica sobre el precio medio ofrecido por cada tecnología, se completa el análisis con un ejercicio de simulación. En este ejercicio, el precio ofrecido por los participantes del mercado que son generadores es modificado restando a su valor la parte que depende de la previsión eólica, la cual se calcula como la previsión eólica para el día siguiente multiplicada por el coeficiente correspondiente que se ha estimado con el modelo de datos de panel. Con los precios de oferta modificados del modo descrito, se construye una nueva curva de oferta agregada y se simula el procedimiento de casación obteniendo el precio marginal o *spot* resultante para cada hora. Seguidamente, se analizan las diferencias entre este nuevo precio *spot* simulado y el precio *spot* real u observado.

Resultados

Los principales resultados obtenidos en la estimación de los modelos de datos de panel son los siguientes:

- La previsión eólica para el día siguiente es uno de los factores que explican el precio al que los generadores están dispuestos a vender su producción, con la excepción de las centrales nucleares. Este resultado es coherente con la escasa flexibilidad de las centrales nucleares para afrontar paradas en su producción, lo que les otorga un menor margen de maniobra para poder tomar decisiones estratégicas.

- No todas las tecnologías de generación reaccionan de la misma manera ante un posible incremento esperado de la producción eólica para el día siguiente. En particular, destaca el comportamiento de las centrales térmicas y de ciclo combinado, las cuales aumentan sistemáticamente el precio ofertado cuando se espera mayor producción eólica, incrementando por tanto el riesgo de quedar fuera de la casación de la subasta.

- Existen también diferencias de comportamiento que tienen que ver con el nivel de demanda esperado: los precios ofrecidos por las centrales térmicas y centrales de ciclo combinado son más bajos en días laborales, cuando la demanda esperada es alta, y más altos en días festivos, mientras que, para el resto de las centrales, el parámetro asociado a la variable laboral no es significativo.

- El ratio del precio del gas (NBP) y el precio de las emisiones de CO2 (EUA) no es significativo para explicar los precios ofrecidos de las centrales nucleares y renovables. Contrariamente, sí lo son para las centrales de ciclo combinado, hidráulicas y térmicas.

- En general, los agentes reducen sus precios ofertados cuando hay mayor porcentaje de reserva de agua en los embalses, lo cual permitiría una mayor producción hidráulica, excepto las centrales térmicas, que ofrecen precios más altos cuando el nivel de reservas de agua es elevado.

- Finalmente, en cuanto a la hora de entrega de la energía, es relevante destacar cómo las centrales de ciclo combinado y las centrales térmicas ofrecen precios más altos en las primeras ocho horas del día, y más bajos en el resto. Por su parte, las centrales hidráulicas también ofrecen precios más altos en las primeras horas del día; sin embargo, la magnitud del efecto es menor que en el caso de las centrales de ciclo combinado y térmicas.

Cada una de las estrategias de oferta que hemos descrito altera la composición y la forma de la curva agregada de oferta, y, por tanto, puede alterar también el precio *spot*. En la última parte del capítulo, se realiza un ejercicio de simulación para tratar de cuantificar, de forma aproximada, el impacto que tienen las decisiones de precio tomadas por las centrales a partir de la previsión eólica sobre el precio *spot*. Los resultados obtenidos indican que son las centrales térmicas las que tienen un impacto mayor, llegando a incrementar el precio *spot* en un 1.14% en media en el año 2013. Sin embargo, dado que hidráulicas y renovables tienen el comportamiento contrario, el efecto global resultante es algo menor, y al final se estima que el precio *spot* se ve incrementado en un 0.36% en media, debido al efecto de la previsión eólica sobre el comportamiento de oferta de los agentes en la subasta del mercado diario.

Conclusiones

La previsión eólica se ha convertido en un factor clave para explicar los precios ofertados por los agentes del lado de la oferta en la subasta del mercado diario, si bien se detectan notables diferencias según la tecnología de generación. No todos los agentes reaccionan de la misma manera ante un incremento esperado de la producción eólica para el día siguiente. En concreto, cabe destacar el comportamiento estratégico de las centrales térmicas y de ciclo combinado. Así, del análisis realizado se deduce que estas centrales ofertan su electricidad a precios más elevados cuanto mayor es la producción eólica esperada y, por tanto, incurren en un mayor riesgo de quedar fuera de la casación. Este resultado es muy relevante, dado que estas centrales han marcado tradicionalmente el precio marginal de la subasta del mercado diario y son lo suficientemente flexibles como para poder participar en los procesos de ajuste o mercados posteriores, con el consiguiente impacto sobre el precio de todos ellos.

Es importante resaltar que el resto de tecnologías presentan el comportamiento contrario. Su comportamiento estratégico es coherente con el objetivo de garantizar la casación de sus ofertas en las diferentes sesiones de subasta que componen el mercado diario. Es decir, ofrecen precios más bajos cuando la previsión eólica esperada es mayor. De este modo, aunque el comportamiento estratégico de las centrales térmicas y de ciclo combinado impulse el precio hacia arriba, dicho efecto se ve compensado por el comportamiento del resto de las centrales, las cuales ofertan a precios menores en las mencionadas circunstancias de mercado. De este modo, el resultado global es que el precio del mercado diario se ve incrementado, pero en menor medida. Concretamente, se ha estimado que, en media, el precio aumenta alrededor de un 0.36%.

El comportamiento de las centrales térmicas y de ciclo combinado consistente en aumentar el precio ofrecido cuando hay mayor riesgo de quedar fuera de las subastas del mercado diario puede resultar contra-intuitivo. Sin embargo, podría justificarse atendiendo al hecho que el mercado de producción eléctrico no se limita al mercado diario, si bien es el mercado con mayor liquidez, sino que está compuesto por una secuencia de mercados y procesos posteriores hasta llegar a la entrega efectiva de la electricidad. Los resultados obtenidos vienen a confirmar que las decisiones tomadas por los agentes en el mercado diario forman parte de un comportamiento estratégico global que tiene en cuenta su participación no solo en el mercado diario sino también en el resto de mercados y procesos; esto es, en el conjunto del mercado de producción eléctrico. En este sentido, algunos generadores podrían tener el incentivo de desplazar su producción desde el mercado diario hacia otros segmentos de negociación, procesos o mercados posteriores, en los que típicamente se premia la flexibilidad (mercados de balance) y en los que existe una menor competencia, buscando maximizar sus beneficios considerando todos los procesos y segmentos de negociación del mercado considerado como un todo, y no aisladamente cada uno de ellos. Por lo tanto, los mercados y procesos que se celebran tras el mercado diario, cobran especial importancia, sobretodo en el estudio del impacto de las renovables. El estudio del impacto de las renovables en estos mercados y procesos se aborda en el Capítulo 3 de la presente Tesis Doctoral.

Referencias

[1] Weidlich, A.; Veit, D. A critical survey of agent-based wholesale electricity market models. Energy Econ., 2008, 30, 1728–59.

[2] Gountis, V.P.; Bakirtzis, A.G. Bidding strategies for electricity producers in a competitive electricity marketplace. IEEE Trans Power Systems, 2004, 19, pp. 356–65.

[3] Veit, D.J.; Weidlich, A; Krafft, J. An agent-based analysis of the German electricity market with transmission capacity constraints. Energy Policy, 2009, 37, 4132–44.

[4] Jónsson, T.; Pinson, P.; Madsen H. 2010. On the market impact of wind energy forecast. Energy Econ., 2010, 32, 313-320.

[5] Li, G.; Shi, J. Agent-based modelling for trading wind power with uncertainty in the day-ahead wholesale electricity markets of single-sided auctions. Appl. Energy, 2012, 99, 13-22.

[6] Twomey, P.; Neuhoff, K. Wind power and market power in competitive markets, Energy Policy, 2010, 38, 3198-3210.

[7] Green, R.; Vasilakos, N. 2010. Market behaviour with large amounts of intermittent generation. Energy Policy, 2010, 38, 3211-3220.

[8] Ciarreta, A.; Espinosa, M. Supply Function Competition in the Spanish Wholesale Electricity Market. The Energy Journal, 2010, 31, 4, 137-157.

[9] Furió D.; Lucia, J. Congestion management rules and trading strategies in the Spanish electricity market. Energy Econ., 2009, 31, 48-60.

Resumen del Capítulo 3:

Analysing the impact of renewables on Spanish electricity final prices using machine learning techniques.

Introducción

En este capítulo se completa el estudio sobre el impacto de las renovables sobre el mercado de electricidad español analizando su efecto sobre el precio final de la electricidad en España. El precio final de la electricidad depende del precio del mercado diario, o precio *spot*, que es su principal componente, pero también incluye otros componentes que capturan todo lo ocurrido en el resto de mercados que tienen lugar después del mercado diario, y que podrían verse afectados a su vez por el incremento de la producción de origen renovable.

El operador del mercado, OMIE, gestiona el mercado diario, en el que se negocia la energía que se entregará el día siguiente. Este es el mercado que tiene mayor liquidez. Tras el mercado diario, tiene lugar la negociación en el denominado mercado intradiario, cuyo diseño facilita que los mismos agentes que han participado en el mercado diario puedan realizar ajustes a sus posiciones asumidas en este último, en momentos progresivamente más cercanos a la entrega a medida que se van sucediendo las subastas de las diferentes sesiones del mercado intradiario.

El operador del sistema, REE, gestiona los denominados mercados de balance con el objeto de resolver incidencias de índole más técnica, y posibles desviaciones ocurridas posteriormente a la negociación de los agentes. Parte de estos procesos son los denominados mercados de restricciones técnicas que se celebran justo después de cada sesión del mercado diario o mercado intradiario. En estos mercados, se revisa el programa de entrega de electricidad, resultante de la subasta atendiendo a criterios exclusivamente económicos, para determinar si este es viable desde un punto de vista técnico. Otros procesos gestionados por REE son los que tienen que ver con la gestión de las desviaciones respecto al programa de entrega de electricidad, los cuales incluyen el proceso de reserva adicional a subir, banda secundaria, banda terciaria y gestión de desvíos en tiempo real. Adicionalmente, el operador del sistema tiene a su disposición otras dos herramientas para la gestión de los posibles desvíos: los pagos por capacidad y el servicio de interrumpibilidad.

La actuación de estos mercados o segmentos de negociación tiene lugar de forma sucesiva, desde el cierre del mercado diario hasta el momento de la entrega efectiva de la energía, y, lógicamente, todos los costes derivados de los mismos se incorporan al precio final de la electricidad. En este contexto, la intermitencia en la producción de las renovables y la mayor dificultad a la hora de predecir la cantidad de electricidad generada a través de estas fuentes de generación podrían suponer un incremento en los costes asociados a gestionar las desviaciones respecto a los programas de despacho iniciales, y en consecuencia un incremento en el precio final de la electricidad que debe pagar el consumidor.

Autores como [1]-[3] analizan la habilidad de los distintos diseños de mercado para hacer frente al reto que supone la integración de las renovables. En general, se considera que una de las claves para reducir los costes adicionales de ajuste derivados de las renovables son los mercados intradiarios, siendo preferible que los agentes acudan a estos antes que a los otros mercados de ajuste más caros. Sin embargo, no es menos cierto que con la irrupción de las fuentes de generación de origen renovable, los generadores de tipo convencional podrían tener el incentivo económico de trasladar su producción a los mercados de ajuste, típicamente mejor remunerados y con menor competencia. En base a todo ello, entendemos que es de gran interés profundizar en los factores determinantes de la negociación en cada uno de estos procesos o mercados posteriores de negociación para entender mejor su funcionamiento.

En este trabajo se propone abordar el estudio del impacto de las renovables sobre los precios finales de la electricidad a través del estudio de cada uno de sus componentes más allá del precio *spot*, que se corresponden con los costes derivados de todos los procesos intermedios que tienen lugar entre el mercado diario y la entrega efectiva de la electricidad. Para ello, en este capítulo se propone la utilización de técnicas de aprendizaje automático (machine learning), una incorporación reciente a esta rama de la literatura, que hasta donde sabemos, se han aplicado principalmente a la modelización de los precios *spot* ([4], [5]).

Objetivo y Metodología

El principal objetivo de este capítulo es analizar el impacto de la producción renovable en el mercado diario sobre cada uno de los componentes del precio final de la electricidad. Los componentes del precio final objeto de análisis son: (i) el coste de los mercados intradiarios; (ii) el coste de los mercados de seguridad o restricciones técnicas, (iii) el coste del resto de procesos de ajuste agrupados en un mismo epígrafe (mercados de reservas secundaria, terciaria, gestión de desvíos en tiempo real, etc.), (iv) el coste derivado de los pagos por capacidad para remunerar la disponibilidad de determinadas plantas de generación convencionales por si fuera necesario atendiendo al nivel de demanda existente, y, por último, (v) el coste del servicio de interrumpibilidad que se destina a remunerar la retirada de ofertas por parte de grandes consumidores como compensación por reducir su consumo en caso de necesidad.

La metodología aplicada en este tercer Capítulo de la Tesis Doctoral requiere de la elaboración de una exhaustiva base de datos. En total, se consideran 264 variables potencialmente predictoras de cada componente del precio final a lo largo de un amplio periodo de tiempo, desde 2012 hasta 2018. El *set* de variables se genera a partir de información que procede de todos los mercados y procesos que componen el sistema eléctrico, y contiene con frecuencia diaria: el precio, volumen de energía casada, y porcentaje de energía casada por tipos de tecnología de generación en el mercado diario; el índice de precios del Carbón, Argus/McCloskey's Coal Price Index Service (API2), referencia para el carbón que se importa en el Noroeste de Europa,; el precio de futuros del gas natural negociado en el hub de los Paises Bajos denominado Title Transfer Facility (TTF); variables indicativas de diversa información sobre el uso de la interconexión con Francia (número de horas del día con el 100% de utilización en ambos sentidos y el spread de precios *spot* entre España y Francia); variables de calendario cuyo objetivo es captar la estacionalidad a distintos niveles (diaria, mensual y anual); variables retardadas (7 dias) de los componentes del precio final de la electricidad; así como la media de los precios ofertados de venta en el mercado diario y de compra y venta en el mercado intradiario de cada tipo de tecnología de generación.

El ingente número de datos con el que se trabaja, dado el ambicioso número de variables predictoras consideradas, aconseja la utilización de técnicas de machine learning, las cuales están especialmente diseñadas para trabajar con grandes volúmenes de datos. En concreto, se ha profundizado en el estudio y aplicación de la técnica de árboles de regresión para dar respuesta empíricamente a las preguntas e inquietudes planteadas anteriormente.

El procedimiento que se ha seguido es el siguiente:

- En primer lugar, se ha dividido la muestra de datos en dos: El 70% de las observaciones se utiliza para entrenar el algoritmo de aprendizaje automático, y el otro 30% para validarlo.

- En segundo lugar, se procede a realizar los entrenamientos con aprendizaje automático. Dado que a priori se desconoce cuál es la mejor técnica de árboles de regresión para modelizar cada componente del precio final se realizan pruebas con tres técnicas diferentes: Clasification and Regression Tree (CART) y sus versiones más avanzadas: Random Forest y Gradient Boosting. En total, se llevan a cabo 15 entrenamientos.
- Cada entrenamiento se valida utilizando para ello las métricas de error habituales: El error absoluto medio (MAE) y la raíz de la media de la suma de los errores al cuadrado (RMSE).
- A continuación, se mide la importancia de cada una de las 264 variables para explicar cada componente del precio final. Dado que no se trata de un modelo clásico en el que se estiman coeficientes, lo que se utiliza es una medida de importancia de la variable basada en el RMSE. Esta medida es el incremento del RMSE atribuible a la exclusión de dicha variable en la estimación, medida habitual utilizada en este tipo de algoritmos ([6], [7]);
- Finalmente, para examinar la relación de cada variable importante con el coste, se recurre a la generación de gráficos de efecto local acumulado, ALEPlots ([8]). Estos gráficos muestran en el eje de la X los distintos valores de la variable y en el eje de la Y el efecto local que tiene cada valor de la variable sobre el coste estimado por el modelo. Este efecto está centrado sobre la media, de modo que si por ejemplo el efecto local acumulado es + 1.5 para un determinado valor de la variable, esto quiere decir que para dicho valor el coste estimado por el modelo es superior a la media, en concreto se estima que es 1.5 veces superior a la media.

En la interpretación de los resultados se presta atención especial a la importancia que como factor tiene el porcentaje de producción renovable casado en el mercado diario. Sin embargo, también se identifican el resto de potenciales factores determinantes pues la inclusión de tal cantidad de variables constituye una oportunidad única para analizar cómo lo ocurrido en alguno de los mercados o segmentos de negociación del sistema eléctrico afecta a lo que ocurre en los siguientes, lo cual es de gran valor para poder comprender mejor el funcionamiento del sistema eléctrico.

Resultados

De entre las tres técnicas aplicadas, la técnica de árboles de regresión (machine learning) denominada Random Forest es la que permite explicar mejor con un margen de error aceptable (inferior al obtenido con otras técnicas como CART y XGBOOST) los costes imputados al precio final de la electricidad. Asimismo, mediante la utilización de la medida de importancia (el incremento del RMSE), junto con los gráficos del efecto local acumulado (ALEPLOT), se ha podido identificar alrededor de 10 variables que tienen una importancia considerablemente mayor que las demás, así como, lo que es más importante, interpretar el tipo de relación que existe entre cada variable relevante y cada coste. Los principales resultados obtenidos se exponen a continuación:

- Los costes de las restricciones técnicas dependen del porcentaje de renovables casado en el mercado diario, obteniéndose una relación lineal y positiva entre ambas variables: cuanto mayor es el porcentaje de renovables en la casación del mercado diario, mayor es el coste asociado a la gestión de restricciones técnicas. Este resultado es consistente con la idea de que la intermitencia y la consiguiente mayor dificultad de predicción de la electricidad generada a partir de fuentes renovables generarían una mayor necesidad de ajustes, y, por tanto, mayores costes para el sistema. Sin embargo, también se estima un mayor coste derivado de la gestión de restricciones técnicas en periodos de menor demanda, lo que resulta ciertamente más difícil de explicar por razones técnicas, pudiendo corresponderse con comportamiento estratégicos de los agentes.

- Los costes de los otros procesos de ajuste diferentes del proceso de restricciones técnicas (procesos de gestión de reservas y desviaciones sobre el programa) también aumentan cuando el porcentaje de renovables casado en el mercado diario crece. Asimismo, se observa una relación lineal y positiva entre los costes de ajuste y una variable relacionada con el comportamiento estratégico de los agentes: cuando los generadores de las centrales de ciclo combinado aumentan el precio al que están dispuestos a vender su energía en el mercado diario, aumentan los costes estimados de ajuste.

- En lo que respecta al mercado intradiario, los resultados obtenidos son claramente diferentes. El porcentaje de producción renovable casada en el mercado diario no aparece directamente como uno de los factores más relevantes para explicar su coste. Por otra parte, se observa cómo la relación entre la mayor parte de las variables predictoras y el coste a predecir es no lineal y más compleja. Así, el coste del mercado intradiario es decreciente para porcentajes de producción de centrales térmicas en el mercado diario inferiores al 10%, y creciente por encima de este valor. Ocurre lo contrario para el caso del porcentaje de producción de centrales hidráulicas en el mercado diario, que pasa de creciente a decreciente en torno al valor del 9%. Además, se observa una relación más compleja con el precio del mercado diario: precios muy bajos del mercado diario (entre 0 €/MWh y 20 €/MWh) se corresponden con costes del mercado intradiario elevados. A partir de 20 €/MWh la relación es decreciente hasta 45 €/MWh y creciente superado este valor hasta estabilizarse a partir de 60 €/MWh. Adicionalmente, se identifican las siguientes variables como factores relevantes para explicar el coste del mercado intradiario: el precio ofertado por las centrales renovables en el mercado diario y el precio del futuro del gas natural negociado en el hub TTF, siendo el coste del mercado intradiario más elevado en los casos en que los precios ofertados por las centrales renovables se encuentran entre 10 €/MWh y 20 €/MWh (valores intermedios) y cuando el precio del futuro del gas es más bajo (inferior a 20 €).

- Por su parte, los pagos por capacidad dependen principalmente de términos autoregresivos y del nivel de la demanda casada en el mercado diario. Cuando la demanda es elevada, los pagos por capacidad aumentan. Esto es debido a que en su mayor parte se trata de los pagos al servicio de disponibilidad a medio plazo que se calculan en función de la electricidad demandada estimada.

- Por último, en cuanto al coste del servicio de interrumpibilidad, este depende principalmente de términos autoregresivos, del índice de precios del carbón (API2), y de factores de calendario, en especial del año 2018. En este sentido, cabe apuntar que, en 2018, el servicio de interrumpibilidad fue objeto de revisión y reformas para mejorar su competitividad[3] y adaptarlo a la normativa europea[4], siendo el resultado obtenido coherente con este hecho.

Conclusiones

Los resultados obtenidos permiten concluir que el aumento de la producción renovable en el mercado diario tiene un efecto sobre los mercados y procesos que vienen después. En particular, conlleva:

(i) mayores costes derivados de los procesos de restricciones técnicas,
(ii) mayores costes derivados del resto de procesos de gestión de desvíos hasta el momento de la entrega de la energía y
(iii) mayores costes en el mercado intradiario para escenarios en los que el precio del mercado diario es más bajo, lo que contribuye a incrementar el precio final de la electricidad.

Del análisis realizado se obtiene evidencia de la existencia de un comportamiento estratégico por parte de los generadores de electricidad en el mercado español, basado en el diseño de estrategias de negociación que se desarrollan como parte de una estrategia coordinada única tanto en el mercado diario como en el mercado intradiario y demás mercados y procesos de ajuste (gestión técnica y de balance); es decir, los agentes participan en los diferentes mercados considerando el mercado como un todo con el objetivo de maximizar beneficios.

Los factores de los que dependen los componentes del precio final de la electricidad diferentes del precio *spot* están relacionados con el incremento del peso de las renovables en el *mix* de generación, la necesidad de realizar ajustes a las posiciones previamente asumidas en el mercado diario y el comportamiento estratégico de los participantes en el mercado.

[3] Order ETU/1133/2017
[4] Paquete de medidas "Clean Energy for All Europeans", presentado por la Comisión Europea el 30 de noviembre de 2016.

El análisis de los factores determinantes del precio final de la electricidad, así como la interrelación existente entre la negociación en los diferentes mercados y procesos que se celebran o llevan a cabo en el mercado de producción de electricidad han sido poco estudiados en la literatura. Hasta donde sabemos, este es el primer trabajo que aborda este análisis para el mercado español. Detectar y entender las dinámicas existentes en los diferentes segmentos de negociación del mercado es clave para avanzar apropiadamente en las reformas del diseño del mercado, reformas que son necesarias para adaptarse al reto que supone la integración de las renovables en el sistema eléctrico.

Referencias

[1] Weber, C. (2010). "Adequate intraday market design to enable the integration of wind power into the European power systems". Energy Policy 38, 3155-3163.

[2] Borggrefe, F. and Neuhoff, K. (2011). "Balancing and intraday market design: options for wind integration". European Smart Power Market Project. C.P. Initiative.

[3] Henriot, A. (2014) "Market design with centralized wind power management: handling low-predictability in intraday markets". The Energy J. 35, 99-117.

[4] Schnürch, S. and Wagner, A. (2019). "Machine Learning on EPEX Order Books: Insights and Forecasts". arXiv:1906.06248 [stat.AP]

[5] Prasanna, A., Holzhauer, S. & Krebs, F., (2019). Overview of machine learning and data-driven methods in agent-based modeling of energy markets. In: David, K., Geihs, K., Lange, M. & Stumme, G. (Hrsg.), INFORMATIK 2019: 50 Jahre Gesellschaft für Informatik – Informatik für Gesellschaft. Bonn: Gesellschaft fürInformatike.V.. (S. 571-584). DOI: 10.18420/inf2019_73

[6] Hastie, T., Tibshirani, R. and Friedman, J. (2001). "The Elements of Statistical Learning: Data Mining, Inference, and Prediction". Second Edition. Springer Series in Statistics Ed. Springer

[7] Breiman, L. (2001) "Machine Learning" 45: 5. https://doi.org/10.1023/A:1010933404324.

[8] Dan Apley (2018). ALEPlot: Accumulated Local Effects (ALE) Plots and Partial Dependence (PD) Plots. R package version 1.1. https://CRAN.R-project.org/package=ALEPlot

Próximos Pasos

Una primera extensión de este trabajo podría consistir en replicar el análisis aplicando la metodología descrita para estudiar el impacto de las renovables en otros países o áreas de mercado. Los resultados obtenidos podrían ser diferentes atendiendo a diferencias en la regulación o en el *mix* de generación. La comparación de los resultados podría proporcionar lecciones interesantes acerca del funcionamiento y diseño óptimo de los diferentes segmentos de negociación de los mercados eléctricos.

Una segunda extensión sería ampliar el periodo muestral. El periodo que cubre este trabajo comprende los años desde 2002 hasta 2018 (abril), quedando fuera del mismo reformas posteriores llevadas a cabo por el regulador español. Algunas de las reformas más relevantes cuyo impacto podría analizarse en el marco del presente trabajo son: (i) La puesta en marcha de un mercado intradiario basado en un sistema de negociación continua que coexiste con el mercado intradiario organizado en sesiones sucesivas cuya negociación se lleva a cabo mediante el mecanismo de subasta; (ii) La participación de los generadores renovables en los mercados de balance, motivada por la necesidad de incrementar la competencia en dichos mercados y reducir los costes de balance (en especial el mercado de restricciones técnicas); (iii) o la sustitución de los pagos por capacidad por un mercado de capacidad cuyo proyecto de orden para su creación se halla en la fecha de finalización de esta Tesis Doctoral sometido a información pública.

Por último, otra posible extensión del estudio realizado consistiría en analizar el impacto de la generación renovable en la contratación de la electricidad a plazo por parte de los agentes participantes en el mercado, bien en mercados de futuros organizados, bien en mercados OTC o a través de los contratos bilaterales físicos.

Capítulo 1:

Capítulo 1:

Effects of renewable on the stylized facts of electricity prices.

Abstract

Many countries around the world have increased their renewable installed capacity due to a greater awareness of climate concerns. Under this new framework, with renewables being among the main generation sources, the literature warns of a dramatic change in price behaviour. Some of the most commonly claimed effects of having a significant proportion of renewable generating sources in the total electricity production mix include: (i) a systematic decrease in overall wholesale market prices, (ii) a higher occurrence of price jumps, and (iii) a significant increase in price volatility. The goal of the present study is to test whether these changes in price behaviour have actually come about. To do so, we focus on the Spanish day-ahead electricity market as a paradigmatic example. In line with the literature, it is found a statistically negative relationship between the renewable generation share and the day-ahead market marginal prices. As well, we have obtained statistical confirmation of the fact that renewables generation share volatility is transferred to price volatility, though similarly to other generation technologies. Finally, in contrast to the general belief that the introduction of renewable generation would give rise to extreme (positive) prices, according to our results, increases in renewables generation share reduce the probability of upward jumps in prices. The results obtained are of interest for portfolio managers, practitioners and regulators.

1. Introduction

One of the most generalized actions all over the world to deal with climate change has been the promotion of renewable energy sources. Thus, many countries such as Germany, Spain, USA and China have significantly increased their investment in clean energy sources. Particularly, in Spain, the electricity generated by renewables in the day-ahead market was 28% of the total production in 2008 per day, on average, whereas five years later, in 2013, it had reached 58% of total electricity produced per day, on average, followed distantly by the remaining generating technologies (Fig. 1). This substantial change in the Spanish generation mix from conventional generation sources to renewables in a few years' time will likely be expected to have an impact on the price formation process.

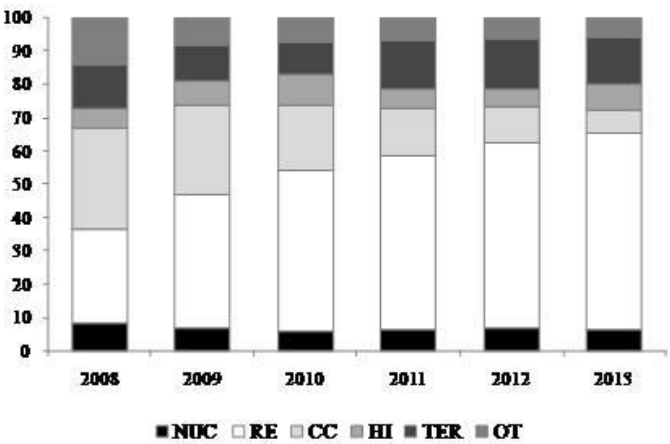

Fig. 1. *Daily average share by technologies in the Spanish day-ahead market Technologies are: nuclear (NUC), renewable (RE), combined cycle (CC), hydraulic (HI), thermal (TER) and others (OT).*

The stylized facts of electricity have been widely pointed out in the literature [1–6]. Due to its well-known intrinsic features, such as its non-storable nature, electricity prices traditionally exhibit high volatility. As well, extreme observations, outliers (atypical values) or jumps normally occur more frequently than with other commodities or financial assets.

The advancement in renewable generation provides social and environmental benefits related to key areas such as rural development, employment or health, as highlighted by

Burgos-Payán et al. [21], and that are not always easy to quantify. In addition, it may also involve changes in the electricity market, with economic impacts.

Several concerns arise when assessing the impact of renewables on the behaviour of electricity prices. These concerns are related to the fact that most of the renewable production is intermittent and somewhat unmanaged. Thus, for instance, wind production depends heavily on the wind speed and direction. In this sense, many voices claim that this intermittent nature of output from renewables will be transferred to electricity prices, with the result of an increase in uncertainty and, hence, in greater price volatility and price risk. The intermittency of renewable generation, when compared to conventional power sources such as nuclear or fossil fuels, which are assumed to be much more secure and reliable, together with the fact that electricity cannot be easily stored, are the main arguments usually given to explain why prices should become even less predictable, and hence more volatile, as long as generation from renewable sources increases. In addition, it is this intermittency that may lead to increases in both the number and magnitude of the so-called price jumps. It should be noted, however, that fuel cost volatility may also be transferred to electricity prices. Therefore, the displacement of conventional power sources by renewable generation may contribute to reduce price volatility[5], instead of increasing it, which is just the opposite effect of the one which is anticipated by those who alert against the use of renewables due to the above-mentioned arguments.

A quite generalized idea related to the impact of the inclusion of renewables as a new generation source in the electricity market is that it will presumably cause a decrease in marginal prices. The reason behind this is that renewable producers can be considered as price takers since they offer very low (close to zero or even zero) prices. Thus, an increase in the amount of these low price offers is expected to shift the supply curve to the right with the result of lower marginal prices. Lower prices for electricity would undoubtedly have positive effects for both consumers and firms, given that the latter use electricity as an input in their manufacturing process. Therefore, a decrease in electricity prices may also contribute to increasing overall productivity.

---

[5] We wish to thank an anonymous reviewer for pointing out this impact on price volatility that may even compensate the previously exposed.

The goal of the present work is to verify whether the above mentioned assumptions have been verified in practice, once the penetration of renewables has been significant enough. Thus, the questions to be answered include:

(i) whether marginal prices may have decreased as a consequence of the entrance of renewables into the system,
(ii) whether marginal prices have become more volatile, and finally,
(iii)whether marginal price jumps occur more frequently than before.

The relationship between renewables and electricity prices has captured the attention of many authors in the literature on energy markets: [7] present an overview of research results on the price effect of renewable production. A common pattern can be observed: in all markets using a merit order dispatch system, generators with lower marginal costs, such as renewable producers, contribute to reducing marginal prices;[8] compare two days with different levels of renewable production but with a similar demand in 2006 in the Spanish case to find that the cost of supporting the development of renewables, initially considered to be very expensive, may have been compensated for by the subsequent decrease in electricity prices; [9] obtain the same conclusions for the German market, in 2006; [10] investigate the economic impact of a large amount of renewables in the Nordic Countries. By employing simulations, they conclude that high penetrations of wind power may push the Nordpool spot market prices down; [11] state that increments in photovoltaic electricity generation lead to lower marginal prices in the Australian electricity market; [12] after studying the effect of weather conditions in the Dutch electricity market (period 2006– 2011), find that an increase in wind speed negatively affects electricity prices; Finally, [13] carry out an ex-post analysis of the effect of renewables and cogeneration in the Spanish electricity market. By using a database of hourly data from 2005 through 2009, their results lead them to conclude that a marginal increase in renewable production of 1 GWh could be associated with a reduction of almost 2€/MW h in marginal electricity prices.

The Spanish case is chosen as a paradigmatic example to embrace the analysis due to the massive introduction of renewables sources into the Spanish generation mix during recent years. To learn about the particular supply-demand situation in the Spanish electricity day-ahead market, Fig. 2 shows the 24-load curves of four typical days of 2013, one

Wednesday a season. Each day has been selected to be Wednesday in order to make them comparable each other. Together with the offered demand, it is represented the offered supply too. Intra-daily seasonality appears to be significant. Thus, the offered supply and the offered demand are generally lower for off-peak hours[6] and, particularly from hour 1 to 8. Similar figures for other two years of the sample period, namely, 2012 and 2010, are presented in Appendix (Figs. A.1 and A.2). As can be observed, in all cases, the offered demand is lower than the offered supply, indicating that there is a permanent situation of excess supply in the market.

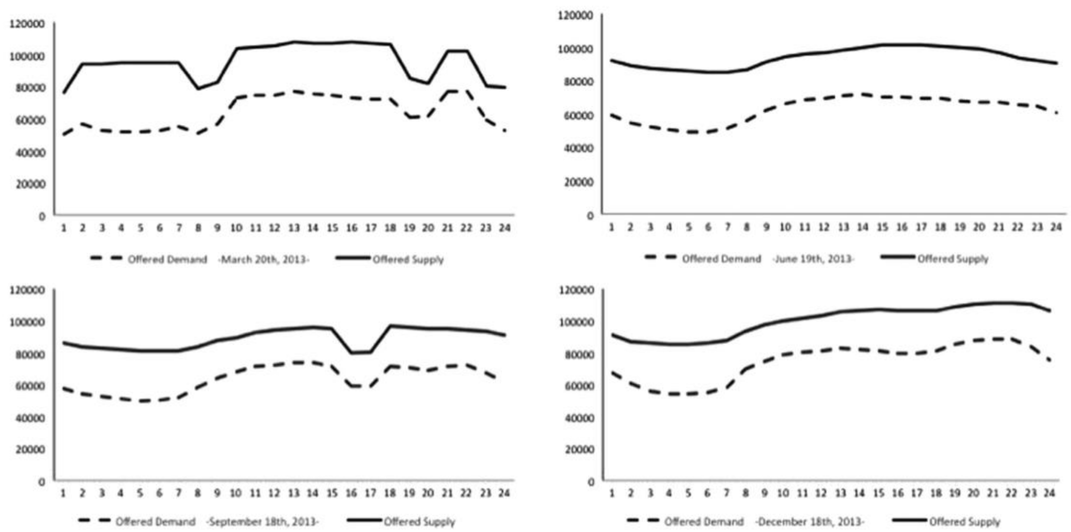

Fig. 2. *24-h load curves of four typical days in 2013 (MW h).*

In addition, the 8760- h load duration curves for the difference between the offered supply and the offered in the last four years of the sample, 2010–2013, is shown in Fig. 3. As can be seen, it is evidenced the excess of supply over demand for every considered year, though such an excess is progressively higher from 2010 to 2013. It is also remarkable the peak observed in 2013. The insight obtained from this graphical analysis is consistent with the fact that the installed capacity is higher in 2013, mainly due to the continuously increasing penetration of renewables into the system.

[6] Peak hours refers to hours from 8:00 h to 20:00 h on business days, while off- peak hours refers to hours from 00:00 h to 8:00 h and from 20:00 h to 24:00 h on working days and the whole day on weekends and holidays.

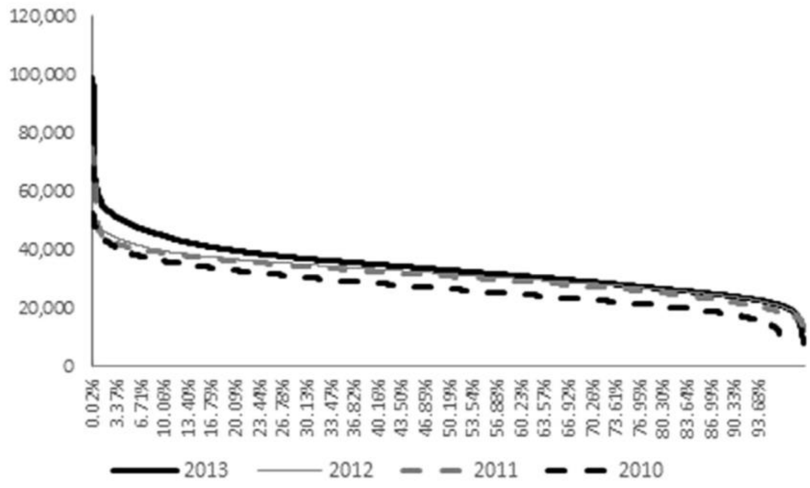

Fig. 3. *8760-load duration curve for the difference between offered supply and offered demand (MW h) for 2010–2013.*

The results of the present study are of interest for both portfolio managers and practitioners, who, being aware of the need to hedge the price variation risk, aim to properly know the true characteristics of price behaviour. In fact, it is the intermittency of renewable generation that is claimed to be responsible for greater price volatility as well as contributing to an increase in the frequency of price jumps. The higher the price volatility, the greater the need to hedge power portfolios in order to minimize the negative effect of adverse price fluctuations.

We extend the previous literature by analyzing the effect of electricity generated by renewable sources on marginal prices, once a sufficiently long enough sample period is available. This period consists of approximately six years of data, since 2008, and may be compared to the earlier years of the whole sample. Besides, the undertaken analysis is more complete than the previously mentioned works, since it does not only cover the effect of renewables on the level of prices but also on price volatility and on the frequency of jumps, taking a two-prong approach. In a first step, a preliminary descriptive analysis is performed, that is certainly helpful to gain overall insights into the research questions addressed by this study and to identify which issues require a more in-depth analysis, which will be carried out in a second step, using econometrical tools.

The rest of the paper is structured as follows. Section 2 describes the data used. Section 3 presents an overview of the changes in the technologies that set marginal prices for the period 2001–2013, the evolution of the day-ahead market marginal price statistics, as well as of variations in the power generation mix within the period of study. Section 4 is devoted to an empirical analysis of the impact of the renewables share on the marginal price, on the number of times each technology sets the marginal price and on the marginal price volatility and jumps. Section 5 summarizes the obtained results and concludes.

2. Data

The dataset used consists of Spanish day-ahead market marginal prices and the generation sources or technologies setting marginal prices with an hourly frequency, from 2001 to 2013. Furthermore, we have employed the amount of electricity produced by technology from 2008 to 2013.This dataset is available at the OMIE webpage[7], where renewables are referred to as special regime. The special regime includes mainly wind[8] but also solar, co-generation, biomass and waste treatment. From now on, we will refer to this group as renewable generation sources, RE, in which hydroelectric plants are not included, and to refer to the remaining technologies, the following nomenclature will be used: TER (thermal: coal and oil-gas), NUC (nuclear), HI (hydroelectric), BG (pumping hydropower) and CC (combined cycle).

Finally, the offered demand and supply hourly volumes sub- mitted to the day-ahead market in of the period covering from 2010 to 2013 have been used to build the corresponding 24-h load curves and the 8760-h load duration curve that are referred to in Section 1.

3. Preliminar Analysis

3.1 Technology setting the marginal prices

In the Spanish day-ahead market, prices and quantities of electricity are determined through a uniform price auction for each delivery hour of the following day. The price

[7] www.omie.es (last accessed April 2014).
[8] In 2013, the 49% of the special regime group comes from wind, whereas the percentages for co-generation, solar, and the remainder included technologies (biomass, waste treatment and mini hydraulic are 29%, 11%, 11% and 11%, respectively (〈www.ree.es〉 , last accessed May 2015).

for each hour is the one paid by market participants whose purchase bids have been accepted after the bid matching process. This price, called the marginal price, equals the price of the last sale bid whose acceptance has been required in order to meet the matched demand [9].Then, it is very relevant to identify the technologies and trading strategies of those plants that set the marginal price, and see whether there have been any changes in the technologies setting the marginal price throughout the considered period, and, specifically, during the period of the sample in which the participation of renewables became significant.

These submitted offers will typically be dependent upon the variable generating costs of the referred technologies but also on the expected offered prices and quantities submitted by the rest of the market participants. Sometimes, more than one offer unit sets the marginal price for a specified hour because they bid at the same price. Indeed, each technology has 24 occasions a day to set the marginal price.

Sale bids from renewable generators are frequently very low. For that reason, a priori they should not be expected to be among the technologies normally setting the marginal price. However, their increasing presence may have altered the supply curve and affect the probability of other generation technologies to deter- mine marginal prices.

Table A.1 in Appendix shows the average percentage of times a day that each technology sets the marginal price from 2001 to 2013. According to it, four different periods can be distinguished:

(i) 2001–2003, in which the main technologies determining the marginal price were HI and TER (approximately 36% on average for base-load prices).
(ii) 2004–2009, a period in which the most remarkable thing is the huge increase in CC setting the marginal price.
(iii)2010–2013, in which the number of times that CC sets the marginal price decreases in favour of other technologies, mainly TER and HI. During this period,

[9] Within the context of electricity, the so-called spot markets are actually day- ahead markets and the marginal prices resulting from the day-ahead market auction are frequently referred to as spot prices.

on average, RE sets marginal prices 9.5% (7.4%, 10.5%) of the time for base-load (peak, off-peak) prices.

When distinguishing between peak and off-peak hours, on the one hand, the leading role of HI for peak hours can be observed. In fact, HI occupies the first place, on average, for all the studied periods except for the period 2005–2009 when it is replaced by CC. The number of times HI sets the marginal price is especially high in 2003, 2010 and 2013. These were very wet years, which allowed reservoirs to reach high water levels (above 60%) [10].On the other hand, it is TER that holds the leader position during 2001– 2005 and 2011–2012, for off-peak hours. For the period 2006– 2010, CC exceeds TER in terms of the number of times it sets the marginal price, whereas in 2010 and 2013 HI becomes the leader.

It is also interesting to see the difference in setting marginal prices by pumping hydropower (BG) between peak and off-peak hours. The number of times the BG technology sets the marginal price reaches 30% in peak hours, whereas this value is much lower for off-peak hours, around 7%.

__3.2__ Descriptives statistics of the day-ahead market marginal prices

Fig. 4 shows the evolution of the Spanish day-ahead market marginal price for the period 2001–2013. Table A.2 in Appendix shows the main descriptive statistics of marginal price by years, distinguishing between peak (Panel B) and off-peak (Panel C) prices.

[10] (www.ree.es)

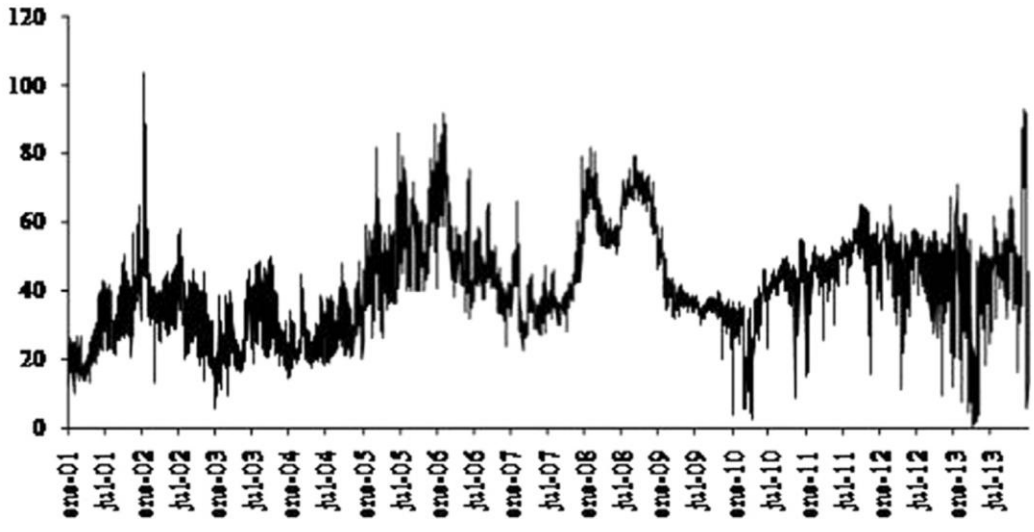

Fig. 4. *Marginal prices in the Spanish day-ahead market.*

The lowest base-load prices on average, around 30 Eur/MW h (Table A.2, Panel A), are those from the early years in the sample, namely 2001, 2003 and 2004. Then, it is in the period 2005–2008 when average marginal prices reach their highest level, around 50 Eur/MW h, and they are particularly high in 2008 (64Eur/MW h). This period coincides with years of drought and low water reservoir levels, as well as with the entrance of combined cycle plants. In the following two years, 2009–2010, prices drop up to 37 Eur/MW h on average. During this period, the drought ends and there is a notable penetration of renewables into the system. However, for 2011–2012, despite the increasing contribution of renewables to electricity production, prices rise again up to levels near 50 Eur/MW h, followed by a slight reduction in prices during 2013. Mean-peak (Table A.2, Panel B) and off-peak (Table A.2, Panel C) prices are shown to follow the same pattern as base-load prices, though, as expected, peak prices are always higher than off- peak prices.

Looking at the standard deviation of daily marginal prices obtained as the daily average of the 24 hourly marginal prices, there is no clear evidence that price volatility has increased as a consequence of renewables for the Spanish case. Only in the last year of the studied sample, 2013, is standard deviation notably higher than in previous years. As can be observed, skewness takes negative values for the later years in the sample, meaning that prices that are below the mean are more frequent than prices exceeding it. This result, a priori, would be consistent with the idea that more renewable production can lead to lower prices. Finally, kurtosis indicates the degree of peakedness of a distribution relative to the normal. According to our results, it seems that the distribution of marginal prices are generally becoming narrower for the last years of the sample.

As mentioned in the introduction section, prices would be expected to become more and more extreme as a consequence of the presence of renewables. A great number of extreme values in the price distribution may be a problem when trying to predict prices. In the literature, before estimating a model, the values that are considered to be too far from the central points of the distribution, normally called outliers, are usually replaced by other more normal values or even ignored. To get some preliminary evidence about the evolution of extreme values throughout the studied sample, Table A.2 also includes the percentage of prices

that could be considered as outliers. In this work, outliers are identified following the procedure described in [14], which consists of defining an outlier as any value which is outside of the interquartile range, i.e. Q3–Q1 (where Q1 and Q3 are, respectively, the first and the third quartiles).

From Table A.2, it can be observed that for the first three years of the sample, 2001–2003, 3.3%, 4.7% and 6.6% of base-load marginal prices can be considered as outliers, according to the described procedure. During the following years there are quite few outliers, except for the approximately 5% of outliers found during 2004– 2005 in peak prices. From 2010 onwards, the number of outliers generally increases, reaching similar levels to the first period. Finally, in 2013, 15% of the observations can be considered as outliers. Furthermore, we note the huge differences found when peak and off-peak hours are analyzed separately. Thus, in contrast to what happened in the first years of the sample (years without renewables), from 2010 onwards, the number of outliers in the time series of off-peak prices becomes much larger than in peak hours.

3.3 Electricity generation by technology type

The weight of RE has significantly grown, increasing from 29% in 2008 to 59% in 2013, on daily average. Since 2011, on some days it has even amounted to 80% of the total production. 78% on average of the energy matched in the Spanish day-ahead market for the period covering 2008–2013 comes mainly from three generating technologies:

renewables (RE), combined cycle (CC) and thermal plants (TER). Hydroelectric (HI) occupies the fourth place, with 7% in 2013.

Table A.3 shows the share of electricity production by generation source, by year, from 2008 through 2013, distinguishing between peak and off-peak prices. It is interesting to observe the continuous growth of RE throughout the sample, which contrasts with the progressive reduction of CC, declining from 30% in 2008 to 7% in 2013. Regarding HI share, it is really quite variable over the years because it strongly depends on annual rainfall and reservoir water levels. Thus, during wet (dry) years, the hydraulic generation actively (hardly) participates in the total production of electricity. Finally, TER share decreases for 2009–2010, though to a lesser degree than CC, to recover a predominant position since 2011. The reason behind this may be found in a new regulation that entered into force in February 2011 (the Royal Decree 134/ 2010), whose aim was to achieve a minimum level of electricity produced by using domestic coal.

In addition, we must take into account that the studied period includes the global financial and economic crisis. In 2009, Spanish GDP growth became negative, 3.8%, and economic activity was considerably reduced, causing a notable decrease in energy demands, -4.7% [11]. The crisis went on during the later years of the sample. Under this context, it should be highlighted that the proportion of renewables keeps growing, during both peak and off-peak hours.

NUC share has reduced slightly overtime, whereas BG share, being more residual (2.5% on average), was increasing during 2008-2012.

4. Empirical Results

The aim of this section is to investigate by using econometric tools, the role that renewable electricity production may have played in the Spanish day-ahead market, particularly, whether:

[11] (www.ree.es)

(i) RE share may have effectively altered the number of times each technology sets the marginal price;
(ii) Marginal prices may have on average decreased as a consequence of the penetration of renewable generation sources into the Spanish electricity system;
(iii) Price volatility may have increased and been explained by RE share volatility; and
(iv) RE share may have made price jumps more frequent.

__4.1__ Renewable share and technology setting marginal price

To study whether the RE share may explain the frequency with which each technology sets marginal price, the following linear regression model is used:

$$m_{e;t} = \alpha_e + \beta_e * RE_t \qquad (1)$$

where $m_{e,t}$ is the percentage of times the technology e sets the marginal price in the day-ahead market on day t and RE_t is the percentage of renewables in the produced electricity in the day- ahead market on day t (RE share).
Estimation results are shown in Table 1. As can be seen, there is a significantly positive relationship between RE share and the percentage of times that TER, HI and BG set the marginal price, whereas such a relationship is statistically negative between RE share and the percentage of times that CC does it. These results are confirmed for base-load, peak and off-peak hours, with the only exception being that for peak hours there is no statistical relation- ship between RE share and the percentage of times that HI sets marginal price. In this way it confirms the idea of RE affecting the probability of other technologies setting the marginal price. Particularly, it can be stated that CC, as technology setting the marginal price, may have been displaced, partially at least, by the irruption of renewables into the system in the Spanish case.

Table 1 Estimates of Model (1). *Ordinary least squares estimates of the univariate model (1). Renewable (RE) share is the independent variable and the dependent variable is the number of times each technology sets marginal price. The considered technologies are: combined cycle (CC), thermal (TER), hydraulic (HI) and pumping hydropower (BG). The Newey-West correction is used to control for heteroscedasticity and serial correlation.*

	CC		TER		HI		BG	
	Value	t-Statistic	Value	t-Statistic	Value	t-Statistic	Value	t-Statistic
Base-load hours (00:00–24:00)								
α	63.7430**	31.96324	14.8768**	11.41163	16.1352**	9.307617	10.22145**	9.579982
β	−0.73435**	−18.97935	0.214637**	7.461359	0.38346**	10.22468	0.112069**	5.179323
Adjusted R^2	0.273002		0.035664		0.101389		0.021506	
Peak hours (08:00 – 20:00, business day)								
α	50.31431**	17.45870	7.376469**	3.927244	35.07231**	13.02311	18.80578**	9.737366
β	−0.553015**	−10.19001	0.258551**	6.055620	0.095845	1.764.239	0.085766*	2.232375
Adjusted R^2	0.120268		0.037509		0.003509			
Off-peak hours (00:00–08:00 and 20:00–24:00, business day, 00:00–24:00 holidays)								
α	73.94679**	33.15469	19.26772**	12.73165	4.286219*	2.406686	3.994430**	4.462314
β	−0.878278	−20.17791	0.185395**	5.750539	0.562738**	13.91629	0.147374**	7.647749
Adjusted R^2	0.298063		0.020864		0.168985			

*Statistical significance at the 1% (5%) level is denoted by ** (*).*

4.2 *Renewable share and marginal price*

As previously indicated, a greater amount of renewable pro- duction is expected to have an impact on the day-ahead price due to the auction mechanism itself. Thus, as renewables generators offer lower prices than most of the other agents in the market, this causes a shift to the right in the supply curve. To examine this issue, the following linear regression model is estimated with RE share as the independent variable:

$$P_t = \alpha * \beta * RE_t \qquad (2)$$

where P_t refers to the marginal price in the day-ahead market on day t and RE_t is the RE share on day t.

As can be seen in Table 2, the beta coefficient is negative, meaning that the marginal price will likely decrease with an increase of RE share, and vice versa, confirming that RE share has the expected effect on the marginal price. In this way, the entry of renewables into the system would have contributed to reducing the price resulting from the day-ahead market auction.

However, as seen in Table A.2 and already commented on in the previous section, despite the fact that RE share has been consider- ably higher for the later years in the sample period, marginal prices on average have not decreased. To shed some light on this issue, the regression model (2) has been newly estimated by substituting RE with each of the other generation sources. According to the results shown in Table 2, similarly to RE share, HI presents a statistically negative relationship with (base-load, peak and off-peak) marginal prices. It is also found a significantly negative relationship between NUC and (base-load and off-peak) prices and between BG and (off-peak) prices.

Regarding the other types of generation sources, a significantly positive relationship is found between TER and CC shares with regards to the marginal price, which is an expected result given that they are technologies with higher generation variable costs. Therefore, to answer the question set out at the beginning of this section, marginal prices get reduced with renewables, which is in line with previous literature ([8,13], among others).

Table 2 Estimates of Model (2). *Ordinary least squares estimates of the univariatemodel (2).The dependent variable is the marginal price, while renewable (RE) production share is included as the explanatory variable. The model is re-estimated by substituting the RE share with the following alternative technologies: combined cycle(CC), thermal (TER), hydraulic(HI), pumping hydropower (BG) and nuclear (NUC).The Newey-West correction is used to control for heteroscedasticity and serial correlation.*

	RE		CC		TER		HI		BG		NUC	
	Value	t-Statistic	Value	t-Statistic	Value	t-Statistic	Value	t-Statistic	Value	t-Statistic	Value	t-Statistic
Base-load hours (00:00–24:00)												
α	70.29**	34.31	38.15**	29.05	31.02**	22.44	57.07**	48.69	47.69**	32.08	51.89**	24.22
β	−0.50**	−13.18	0.47**	7.00	1.21**	14.40	−1.48**	−10.61	−0.52	−1.11	−0.77*	−2.33
Adjusted R^2	0.27		0.14		0.36		0.20		0.00		0.02	
Peak hours (08:00–20:00, business day)												
α	75.61**	27.12	43.49**	24.20	38.21**	20.80	61.50**	34.92	53.01**	26.86	56.75**	18.52
β	−0.51**	−9.54	0.42**	4.84	1.07**	10.01	−1.36**	−5.53	−0.32	−0.69	−0.86	−1.46
Adjusted R^2	0.25		0.11		0.27		0.11		0.00		0.02	
Off-peak hours (00:00–08:00 and 20:00–24:00, business day, 00:00–24:00 holidays)												
α	65.45**	33.69	36.39**	28.48	28.04**	20.87	53.30**	50.75	45.43**	35.58	48.33**	22.27
β	−0.46**	−12.84	0.42**	6.36	1.19**	14.45	−1.39**	−11.64	−1.17*	−2.21	−0.61*	−2.03
Adjusted R^2	0.25		0.11		0.37		0.22		0.00		0.01	

*Statistical significance at the 1% (5%) level is denoted by ** (*).*

4.3 *Renewables share and price volatility*

As mentioned in the Introduction, electricity prices tradition- ally exhibit high volatility. Furthermore, the so-called price jumps are assumed to be relatively frequent. One of the most important concerns about the integration of renewables into the system is that the intermittent nature of these technologies may increase price volatility as well as the number of price jumps, which would end up creating more difficulties when modelling electricity prices, due to greater uncertainty. Then, what needs to be determined is:

(i) whether RE generation may be behind price volatility, and

(ii) whether RE share volatility may contribute to the presence of price jumps.

In order to find this out, the model proposed by [15], which was later applied to the electricity market by Cartea and Figueroa [16], is chosen. This model aims to describe the main features of electricity prices and it is especially interesting for the purposes of this study, as it allows price volatility and jumps to be captured.

The model adapted by Cartea and Figueroa [16] is a stochastic process with mean reversion that includes a discrete jump process (a diffusion model). Under this model, jumps are defined as large price movements at a particular point that break the continuous process followed by the price, and price volatility is calculated day-to-day with a moving window of 30 days. Once estimated, the next step will be to study whether the obtained estimates may have been altered by changes in the electricity production from renewable sources.

4.3.1 *Model definition*

We have (Ω, P, F, {Ft}t c [0, T]) a filtrated and completed probability space with finite time horizon t < ∞. The spot price on time t, 0 <= t <= T, is defined as:

$$P_t = \exp(f(t)+Y(t)) \tag{3}$$

where f(t) is a deterministic function that captures seasonal tendency and Y(t) is a stochastic process whose dynamics are:

$$dY_t = -\alpha Y_t dt+\sigma(t)\, dZ_t + \ln J dq_t \tag{4}$$

Y_t is a diffusion process with jumps and mean reversion of the spot price P_t; σ(t) is the volatility that depends on time; J is the size of the random jump; dZ_t is the increment of standard brownian and dq_t is a Poisson process where l is the intensity or frequency of the process (dq_t is equal to 1 with probability ld_t or it is equal to 0 with probability (1-ld_t)).

J is Lognormal:

$$j \to N(\mu_t,\sigma^2{}_t)$$

$$E(J)=1$$

Their properties are:

$$J=\exp(\phi),\phi\to N(-\sigma^2{}_t/2,\ \sigma^2{}_t) \tag{5}$$

$$E\ [\ln J\ ]=-\sigma^2{}_t/2$$

$$\mathrm{Var} \ln J\ ]=\sigma^2{}_t$$

The steps to estimate the parameters of the model are as follows:

1. Transformation of the price series in log returns. Previously, once the outliers are identified using the method described in Section 3, they are replaced by the average of their neighbouring values[12].

2. Estimation of the long-term trend T_t. The function proposed by [17] is used, a sinusoidal function supplemented by an exponentially weighted moving average (EWMA), being λ=0.975 (value recommended by [17]). Parameters are estimated by nonlinear least-squares, using the Gauss-Newton option on PROC NLIN of SAS.The function is:

$$T_t=a_1+\sin\{2\pi((t/365)+a_2)\}+a_3+a_4\ EWMA^{\lambda}{}_t \tag{6}$$

$$EWMA^{\lambda}{}_t=(1-\lambda)Pt+\ EWMA^{\lambda}{}_{t-1}$$

3. Once the long-term trend defined in the previous step is subtracted, a second seasonal component, S_t, is calculated, which is equal to weekly average.

4. Following [16], the mean reversion is estimated through the following equation:

$$Y_{t+1}-Y_t=\ \alpha Y_t+\ \varepsilon_t \tag{7}$$

where Y_t is the price in logarithms without seasonal components and α is the mean reversion parameter, which is estimated by ordinary least squares.

[12] It should be emphasized that these values are normally excluded and not employed in estimation because they are considered to cause serious distortion. Next, it is crucial to know the number of potential outliers that can be expected within a particular series, and for those series with many outliers, alternative methods are needed, since the removal of them may cause a loss of information which would translate into less efficient estimates.

5. To calculate the price volatility in the model, as it is considered not to remain constant over time, the standard deviation is calculated for a moving window of 30 days, namely, price volatility is calculated day-to-day with a moving window of 30 days.
6. The technique used to identify jumps in our sample is the one used in [16] and [18]. It consists of an iterative algorithm that filters the returns whose absolute value exceeds the standard deviation multiplied by three. The values marked as jumps are replaced by the average of its non-marked neighbours and the procedure goes on until all values in the sample are non-marked values.

Table 3 (Panel A) shows the estimation results of the diffusion model with jumps and mean reversion (4) for the period 2008– 2013, when a non-negligible amount of electricity generation comes from renewable sources, distinguishing between peak and off-peak hours. Firstly, some relevant differences between peak and off-peak hours appear. Thus, the price volatility is higher for off-peak than for peak hours, 0.15 versus 0.9. As well, the frequency of jumps is also notably higher for off-peak hours, whereas the mean reversion is not much lower as indicated by the value of the α coefficient. Secondly, in order to study jumps in detail, the number of jumps is also shown (Panel B), distinguishing between the negative and the positive ones, not only for peak but also for off-peak hours. As can be observed, negative jumps are much more frequent than positive for base-load (62 versus 33), peak (38 versus 16) and off-peak (75 versus 40) hours.

***Table 3. Estimation results of the diffusion model with jumps and mean reversion** (4). Panel A shows the diffusion model estimates, distinguishing between peak and off-peak hours: a1, a2, a3 and a4 are the parameters used to adjust the long-term seasonal component Tt, (7); [σt] is the mean 30 days volatility of the price; α indicates the reversion to the mean, and σj, l are the jump parameters, namely standard deviation and frequency of the jumps, respectively. Panel B, presents detailed information about the number of outliers, total, positive and negative jumps detected in the sample.*

Panel A parameters	$a1$	$a2$	$a3$	$a4$	$[\sigma t]$	α	σj	l
Base-load hours (00:00–24:00)	−0.0412	110.2	−2.0171	2.6769	0.12	0.23064	0.82	15.833
Peak hours (08:00–20:00, business day)	−0.00429	110.2	−2.0519	2.6928	0.09	0.21406	0,43	9
Off-peak hours (00:00–08:00 and 20:00–24:00, business day, 00:00–24:00 holidays)	−0.0450	110.2	−19.715	26.550	0.15	0.18140	0.10305	19,167
Panel B outliers and jumps	Total	2008	2009	2010	2011	2012	2013	
Base-load hours (00:00–24:00)								
Outliers	145	0	3	32	12	30	68	
Jumps	95	0	3	16	8	21	47	
Positive jumps	33	0	2	3	4	6	18	
Negative jumps	62	0	1	13	4	15	29	
Iterations	6							
Peak hours (08:00–20:00, business day)								
Outliers	79	0	3	14	3	16	43	
Jumps	54	11	1	15	2	9	26	
Positive jumps	16	0	1	5	1	3	6	
Negative jumps	38	11	0	10	1	6	20	
Iterations	5							
Off-peak hours (00:00–08:00 and 20:00–24:00, business day, 00:00–24:00 holidays)								
Outliers	128	1	2	26	15	23	61	
Jumps	115	0	7	32	8	21	47	
Positive jumps	40	0	2	15	1	5	17	
Negative jumps	75	0	5	17	7	16	30	
Iterations	11							

Fig. A.3 shows the evolution of jumps throughout the years in the sample. It should be noted the large number of jumps recorded from January to May, in 2010 and in 2013 which contributed to increasing volatility during these two periods (as can be seen in Fig. 5), and which was even more notable for off-peak hours.

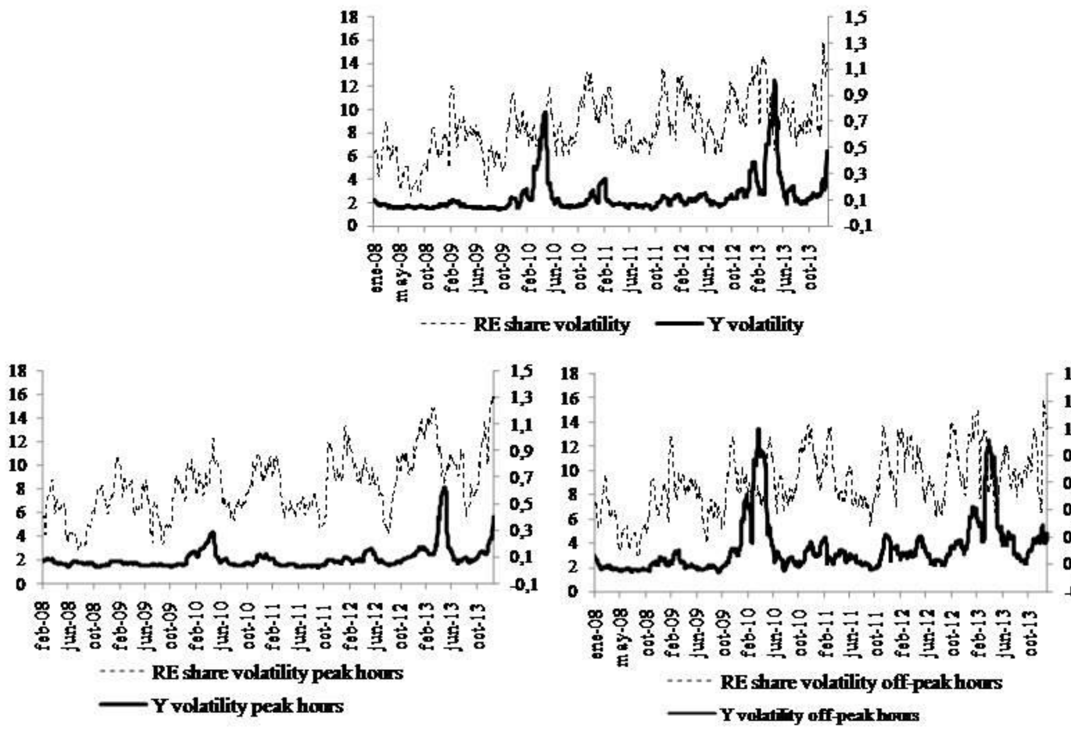

Fig. 5. *RE share volatility and base-load marginal price without seasonal component (Y) volatility.*

In order to measure the relationship between RE share volatility and price volatility, the Pearson test is used. The two detected high volatility periods, i.e., from January to May, 2010, and from January to May, 2013, have been analyzed separately for peak and off-peak hours. Results are shown in Table 4. As can be observed, there is a positive linear relationship between RE share volatility and price volatility for the whole sample, which becomes stronger when excluding the two high-volatility periods mentioned above, when the correlation coefficient reaches 63%, in peak and off-peak hours. Compared to the rest of technologies, this is the highest Pearson test value obtained. Therefore, increases in RE volatility are accompanied by increases in price volatility. This result is consistent with the results of [19] for the English market.

Table 4. Marginal price volatility and production share volatility by technologies*. Pearson test between marginal price (without seasonal component) volatility and the production share volatility for different generation technologies, distinguishing between peak and off-peak prices. Included technologies are: hydraulic (HI), pumping hydropower (BG), nuclear (NUC), combined cycle (CC) and thermal (TER).*

	RE	CC	TER	HI	NUC	BG
Base-load hours (00:00–24:00)						
Total	0.27978**	−0.19049**	0.15136**	0.73057**	0.54517**	0.50646**
Excluded Jan 2010–May 2010 and Jan 2013–May 2013	0.63852**	−0.2125**	0.60925**	0.3417**	0.2564**	0.50315**
Jan 2010–May 2010 and Jan 2013–May 2013	−0.26351**	−0.15735**	−0.16887**	0.68028**	0.67888**	0.72605**
Peak hours (08:00–20:00, business day)						
Total	0.55748**	0.02774	0.30428**	0.67605**	0.48723**	0.45195**
Excluded Jan 2010–May 2010 and Jan 2013–May 2013	0.63255**	−0.03463	0.52936**	0.37192**	0.21579**	0.36904**
Jan 2010–May 2010 and Jan 2013–May 2013	0.01435	0.03957	−0.14151*	0.65415**	0.49694**	0.58671**
Off-peak hours (00:00–08:00 and 20:00–24:00, business day, 00:00–24:00 holidays)						
Total	0.26882**	−0.19898**	0.18314**	0.70779**	0.57265**	0.57316**
Excluded Jan 2010–May 2010 and Jan 2013–May 2013	0.62358**	−0.21330**	0.61021**	0.25600**	0.38560**	0.53368**
Jan 2010–May 2010 and Jan 2013–May 2013	−0.35802**	−0.16690**	−0.19043**	0.67545**	0.66119**	0.71528

*Statistical significance at the 1% (5%) level is denoted by ** (*).*

Nevertheless, it is relevant to mention that significant positive relationships have also been found, above all when excluding the two detected high-volatility periods, between the price volatility and the volatility of the shares of technologies other than RE, such as TER HI, NUC and BG. Furthermore, another very interesting point is the significantly negative relationship obtained between the CC share volatility and the (base-load and off-peak) price volatility.

Focusing on the two periods with the greatest price volatility, namely, from January to May 2010, and from January to May 2013, the correlation is notably higher for HI, NUC and BG for base-load and off-peak hours, and even the sign of the correlation between RE share and price volatility becomes negative. In fact, as is shown in Fig. 6, volatility peaks in HI, NUC and BG generation match marginal price volatility peaks better than RE.

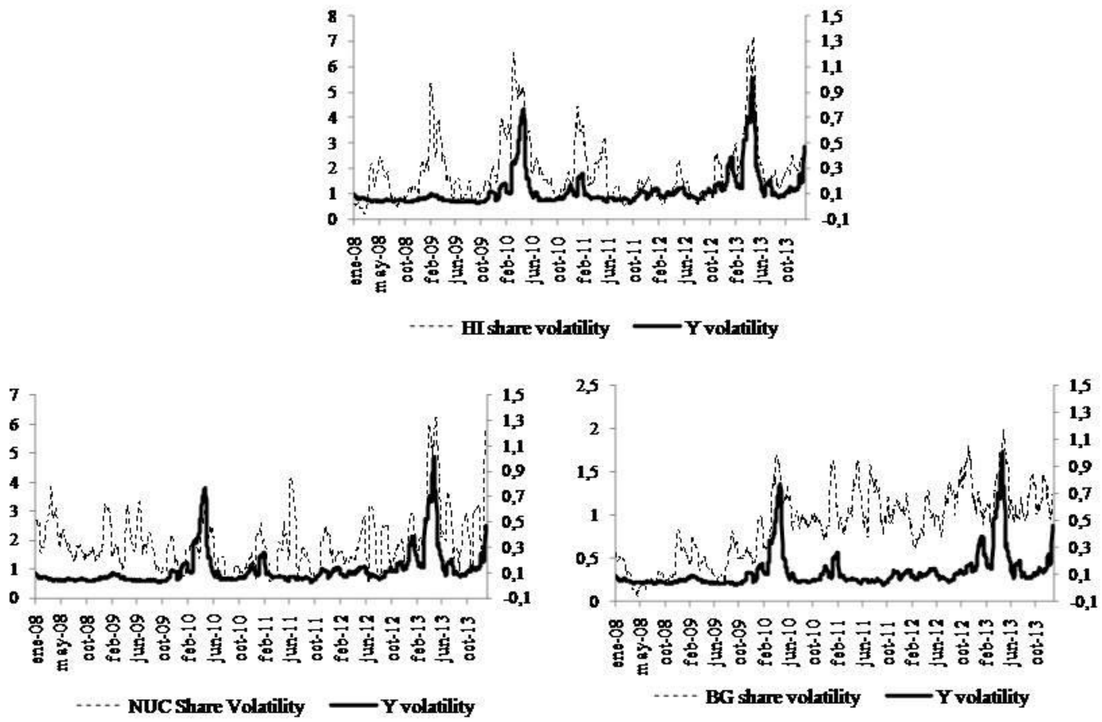

Fig. 6. *Base-load Marginal price without seasonal component (Y) volatility, hydraulic share (HI) volatility and pumping hydropower share (BG) volatility.*

Therefore, the volatility of the electricity produced by the different generation technologies involved in the present study has been transferred to prices, with the only exception of CC, which presents no relationship at all with peak prices volatility and a significantly negative one with (base-load and off-peak) prices volatility.

Finally, in order to find out whether there has been a greater number of price jumps as a consequence of renewables, the jump process must be expressed as a function of the involved technology shares. A model for discrete choice, as that in [20], is adequate for this purpose because the event studied is a discrete event, meaning that it has only two possible outcomes: the jump event occurs or it does not. The model to be estimated is a logistic regression with the different generation technologies as explanatory variables:

$$\mathrm{Logit}(\pi)=\mathrm{Log}(\pi /(1- \pi)=\alpha+\beta'X \tag{8}$$

where π is the probability of a jump (in returns without seasonal component) and X is the matrix of time series in differences of six variables: RE share (d_RE), HI share(d_HI), BG share (d_BG), NUC share (d_NUC), CC share (d_CC) and TER share (d_TER).

With the aim of identifying the technology or technologies that may better explain the jump event, the automated procedure denominated stepwise (backward selection) is used. In the first step, the model does not include any variables. A chi-square test is carried out with each variable (seven variables, including intercept), and the variable that has the strongest relationship with the event enters into the model. In the second step, the exercise is repeated with the rest of variables. Once again, the best variable among those considered is chosen and the model is re-estimated with the two variables. If both variables were significant, they would both remain as candidates. However, if one or both were not significant, then they would be ignored. The process continues as long as there are non-significant variables that may be considered as candidates for entering into the multivariate model.

The estimated results of the logistic regression are shown in Table 5.[13] The HL goodness-of-fit test shows if there is any evidence of a lack of fit in the selected model, and the c-statistic is a measure of association for the variables and the event. A c-statistic equals to 0.50 means that the model is not better than a completely random prediction. However, with a c-statistic equals to 1, then the fit is considered to be perfect.

[13] Table 8 only shows in each case the group of variables that turned out to be significant at the end of the stepwise procedure.

***Table 5. Logistic regression estimates. Model** (8). The dependent variable is the logit function of the probability of a negative jump event in the series of the day-ahead market returns, whereas explanatory variables are the generation share of different technologies, in differences: renewable share (d_RE), hydraulics(d_HI), pumping hydropower(d_HIB), nuclear(d_NUC), combined cycle (d_CC) and thermal (d_TER). The exercise is repeated with positive jumps as a new event and splitting the sample for peak and off-peak hours. The number of regressions are then 6: a, b and c for negative jumps and d, e and f for positive jumps. The coefficient c-statistics is a measure of association, and the LH test is the Hosmer and Lemeshov test of goodness-of-fit. This table only shows the group of variables that, following a stepwise procedure, turns out to be significant in each case.*

Model	Negative jumps						Positive jumps					
	Base-load hours (00:00–24:00) (a)		Peak hours (08:00–20:00, business day) (b)		Off-peak hours (00:00–08:00 and 20:00–24:00, business day, 00:00–24:00 holidays) (c)		Base-load hours (00:00–24:00) (d)		Peak hours (08:00–20:00, business day) (e)		Off-peak hours (00:00–08:00 and 20:00–24:00, business day, 00:00–24:00 holidays) (f)	
Parameter	Value	Wald Chi-square	Value	Wald Chi-square	Value	Wald Chi-square	Value	Wald Chi-square	Value	Wald Chi-square	Value	Wald Chi-square
α	3.616	694.22**	-39.319	413.55**	-3.442	730.46**	-4.329	469.47**	−4.986	213.44**	-4.295	453.84**
d_RE							-0.07	96.19**	-0.109	14.88**		
d_HI												
d_BG					0.287	5.05*						
d_CC											0.078	5.27*
d_TER	-0.097	12.70**	-0.171	34.47**	-0.105	16.78**					0.110	8.03**
d_NUC											-0.147	4.47*
c-Statistic	0.591		0.702		0.642		0.632		0.708		0.716	
HL test	0.001	25.28	0.023	17.80	0.156	11.88	0.002	24.99	0.234	10.46	0.088	13.77
Jumps	62		38		75		33		16		40	
Obs.	2.191		1.523		2.191		2.191		1.523		2.191	

*Statistical significance at the 1% (5%) level is denoted by ** (*).*

Firstly, it should be pointed out that when using the time series of base-load prices (Table 5, models (a) and (d)), there is statistical evidence of lack of fit. Therefore, as estimation results are not valid, they are ignored. The reason can be found in the fact that the sample including the 24 price observations a day is made up of two very different levels of prices. So, price levels that would be considered as a positive (negative) jump under the distribution of off-peak (peak) prices may be considered as normal (meaning that it is not a jump) under the distribution of base-load prices. Then, when the prices for all the 24 h are put together, it turns out to be more difficult for the jumps in prices to be detected. Once the difference between peak and off-peak prices is detected, the picture is more informative.

Thus, starting with negative jumps in prices, it is observed that when TER share is higher than in the previous day, then the probability of a negative jump in the price decreases, for peak and off-peak hours, as indicated by the significantly negative parameter value (-0.17 peak hours, -0.10 off-peak hours). Additionally, the behaviour of BG share is also significant, having a positive effect on the probability of negative jumps, though the fact that this result only applies for off-peak hours is notable.

Regarding positive jumps in prices, a statistically significant relationship is found between increases in RE share and the frequency of price jumps, though only for peak hours. However, the estimated parameter value is negative (-0.10), meaning that an increase in RE share would reduce the probability of positive jumps for peak prices. During off-peak hours, the technologies that would have an impact on the frequency of positive jumps would be: CC and TER, exhibiting positive values for the corresponding estimated parameters (+0.07 and +0.11 respectively), whereas for NUC it is displayed a significantly negative coefficient value (-0.14).

This is quite a striking result since, in contrast to the general belief that the introduction of renewable generation was going to give rise to extreme (positive) prices due to their intermittency and other supposed production planning and/or management problems, our results lead us to conclude just the opposite for the Spanish case. Indeed, it is the probability of a positive jump in peak prices (at the end, higher prices) that is reduced with increases in renewable generation. With regards to off-peak hours, there seems to be no statistical relationship between changes in renewables generation and jumps in prices.

5. Conclusion

The promotion of renewable energy sources in electricity systems has been a priority all over the world to deal with climate change. The advance of renewable technologies has environmental and social benefits, but it also involves economic impacts. The integration of clean energy sources is expected to cause relevant changes in electricity prices. In this work, we focus on the Spanish electricity market to shed some light on this matter.
Together with the evidence obtained regarding the impact of renewables generation on the level and volatility of prices, other results derived from the role of the other involved generation technologies have also been provided. The main conclusions can be summarized as follows.

Firstly, the picture has become much more informative when peak and off-peak hours are analyzed separately, confirming the fact that these price series should each be viewed as different commodities, with different features. Thereby, only when peak and off-peak prices are considered separately, do some changes that may be caused by renewables appear. Thus, for the period from 2002 to 2009, price volatility is higher and jumps are more frequent during peak hours, whereas during the last years of the sample, namely 2010–2013, where renewable generation is much more relevant, the opposite happens.

In line with the literature, there is a statistically negative relationship between the renewable generation share and the day-ahead market marginal prices. In addition, a significant relationship has been found between renewables generation share and the number of times that other technologies such as combined cycle, thermal and hydropower technology sets the marginal price. This relationship is negative only for the combined cycle technology. Therefore, it can be stated that renewables may be responsible for the replacement of CC as the technology setting marginal prices.

As well, we have obtained statistical confirmation of the fact that renewables generation share volatility is transferred to price volatility. However, significant positive relationships between the share volatility of other technologies (such as TER, HI, NUC and BG) and price volatility have been found and are worth being highlighted. Last but not least, this relationship becomes negative for the case of CC share, indicating that increases in this generation technology would contribute to reduce price volatility.

Lastly, in contrast to the general belief that the introduction of renewable generation would give rise to extreme (positive) prices, due to their intermittency and other supposed production planning and/or management problems, according to our results, increases in renewables generation share reduce the probability of upward jumps in peak prices, whereas no significant relation- ship between renewables generation share and jumps in off-peak prices have been found.

The results of this work can help practitioners and regulators understand how the inclusion of renewables into the electricity generation system has actually impacted the level and volatility of day-ahead market prices. One must be conscious of the fact that the intermittency of these sustainable generation technologies may be transferred to subsequent markets such as the intraday market. This issue, together with an analysis of the strategic bidding behaviour by the market participants when considering the transmission of information between the different markets and the information related to the foreseen generation by the different technologies, are left for further research.

Acknowledgements

The authors gratefully acknowledge the financial support of the Spanish Ministry of Education and Science (Project ECO2013- 40816-P) and the University of Valencia (Project UV-INV- PRECOMP13-115547). Thanks are also due to the participants of X Conference Spanish Association of Energy Economics, 2015.

Appendix A

See Tables A.1–A.3 and Figs. A.1–A.3

***Table A. 1. Percentage of times each technology sets marginal price on average**. Included technologies are: hydraulic (HI), thermal (TER), combined cycle (CC), pumping hydropower (BG) and renewable (RE).*

Year	Base-load hours (00:00–24:00)					Peak hours (08:00–20:00, business day)					Off-peak hours (00:00–08:00 and 20:00–24:00, business day, 00:00–24:00 holidays)				
	HI	TER	CC	BG	RE	HI	TER	CC	BG	RE	HI	TER	CC	BG	RE
2001	34.26	41.13	0,00	18.56	0,00	40.51	25.59	0.00	30.27	0.00	31.70	49.25	0.00	11.88	0.00
2002	42.2	37.31	0,00	17.34	0.25	58.92	12.11	0.00	29.40	0.20	34.38	49.97	0.00	10.81	0.27
2003	43.05	34.17	0,00	15.56	1.58	58.07	13.81	0.00	26.21	1.15	34.85	45.35	0.00	9.60	1.79
2004	27.41	34.26	14.57	18.09	1.78	35.49	15.78	12.94	29.18	1.47	23.02	43.75	16.06	11.74	1.92
2005	17.07	38.57	26.47	18.32	2.16	26.32	18.81	22.33	33.40	1.61	12.02	49.12	29.28	9.73	2.45
2006	26.35	32.76	39.33	5.1	4.84	17.83	23.34	29.81	8.27	7.01	31.19	37.11	43.61	3.18	3.65
2007	19.98	27.98	43.47	12.84	9.85	27.31	12.42	44.63	22.79	15.78	15.09	35.95	44.38	7.12	7.21
2008	21.19	22.17	48,00	10.11	3.96	30.70	15.98	43.98	16.08	5.66	14.89	25.33	52.02	6.44	3.14
2009	29.6	17.4	46.56	13.42	4.11	37.37	10.53	41.21	22.08	4.49	25.00	20.83	50.79	8.25	4.18
2010	42.45	15.63	32.64	19.06	9.55	44.98	9.45	23.56	30.28	6.14	40.57	18.24	39.03	12.36	11.13
2011	31.76	32.73	23.86	14.51	7.68	33.89	30.53	22.36	20.52	7.05	30.24	33.84	25.74	10.55	8.17
2012	32.32	30.77	15.69	15.32	12.84	37.81	27.34	12.62	20.88	12.02	28.55	33.32	18.07	11.53	13.35
2013	48.05	31.3	7.89	20.63	7.96	52.30	22.21	4.63	26.80	4.36	45.37	36.37	10.10	16.44	9.74

Table A. 2. Marginal price descriptive statistics. *Descriptive statistics and outliers (in percentage) of the series of the Spanish day-ahead market marginal prices (2001–2013). Outliers have been identified following the procedure proposed by Benth et al. [14].*

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Panel A) Base-load hours (00:00–24:00)													
Mean	30.13	37.40	28.96	27.94	53.68	50.53	39.35	64.43	36.96	37.01	49.92	47.24	44.26
Std Deviation	10.58	12.54	9.22	6.56	12.33	13.60	8.86	7.19	5.58	10.63	4.99	8.84	17.46
Skewness	0.53	1.34	0.26	0.72	0.40	1.02	1.25	−0.07	0.46	−1.08	−1.17	−1.37	−0.44
Kurtosis	−0.15	5.04	−0.97	−0.17	−0.30	0.72	1.97	−0.79	6.10	0.82	4.56	2.60	1.09
% Outliers	3.3%	4.7%	6.6%	0.3%	1.6%	0.5%	0.0%	0.0%	0.3%	5.8%	1.9%	5.2%	15.6%
% Outliers up	2.5%	3.0%	4.9%	0.3%	1.4%	0.3%	0.0%	0.0%	0.0%	3.6%	1.1%	3.0%	7.4%
% Outliers down	0.8%	1.6%	1.6%	0.0%	0.3%	0.3%	0.0%	0.0%	0.3%	2.2%	0.8%	2.2%	8.2%
Panel B) Peak hours (08:00–20:00, business day)													
Mean	37.41	47.07	36.30	33.16	67.04	61.72	46.52	71.12	40.40	42.20	54.51	53.30	51.20
Std Deviation	13.13	15.74	11.67	8.68	16.10	18.32	10.96	8.79	5.97	10.19	6.46	7.59	17.81
Skewness	0.38	1.96	−0.00	0.43	0.35	0.68	1.10	−0.07	0.53	−1.08	0.00	−1.55	−0.62
Kurtosis	−0.23	7.48	−1.23	−0.69	−0.47	−0.29	1.73	−0.75	9.16	1.26	3.38	4.56	1.96
% Outliers	3.5%	2.8%	9.4%	5.5%	4.7%	2.8%	0.0%	0.0%	0.8%	4.3%	0.8%	4.7%	13.8%
%Outliers up	1.6%	1.2%	4.7%	2.4%	2.0%	0.8%	0.0%	0.0%	0.0%	2.8%	0.4%	1.6%	6.3%
% Outliers down	2.0%	1.6%	4.7%	3.1%	2.8%	2.0%	0.0%	0.0%	0.8%	1.6%	0.4%	3.1%	7.5%
Panel C) Off−peak hours (00:00–08:00 and 20:00–24:00, business day, 00:00–24:00 holidays)													
Mean	26.50	32.86	25.48	25.34	46.77	44.61	35.11	60.31	34.83	34.21	47.35	44.33	40.83
Std Deviation	8.65	9.67	7.01	5.07	8.96	11.07	7.78	6.67	5.75	10.51	7.19	8.93	16.79
Skewness	0.44	0.85	0.28	0.86	0.58	1.18	1.30	−0.02	0.05	−0.97	−1.34	−1.19	−0.33
Kurtosis	−0.22	3.10	−0.57	0.36	0.36	1.06	1.60	−0.84	4.38	0.35	3.32	1.66	0.88
% Outliers	1.1%	1.4%	1.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.5%	6.3%	2.7%	4.9%	15.6%
% Outliers up	0.5%	0.5%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	4.1%	1.6%	2.5%	7.1%
% Outliers down	0.5%	0.8%	0.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.5%	2.2%	1.1%	2.5%	8.5%

Table A. 3. Day-ahead market production share of each technology *(daily average percentage). Included technologies are: renewable (RE), combined cycle (CC), hydraulic (HI), thermal share of each technology. (TER), nuclear (NUC) and pumping hydraulic generation (BG).*

Year	Base-load hours (00:00–24:00)						Peak hours (08:00–20:00, business day)						Off-peak hours (00:00–08:00 and 20:00–24:00, business day, 00:00–24:00 holidays)					
	RE	CC	HI	TER	NUC	BG	RE	CC	HI	TER	NUC	BG	RE	CC	HI	TER	NUC	BG
2008	28.49	30.05	6.04	13.27	8.24	0.61	28.27	34.17	6.67	12.12	6.08	0.85	28.41	28.77	5.42	13.73	9.23	0.44
2009	40.35	26.44	7.65	10.57	6.66	1.27	40.03	29.79	7.38	9.50	5.36	1.67	39.71	25.76	7.61	11.19	7.31	1.00
2010	48,00	19.43	9.51	9.7	6.07	2.38	46.86	22.12	8.59	9.50	4.92	3.38	47.93	19.14	9.62	9.91	6.71	1.73
2011	51.93	13.99	6.14	14.8	6.54	2.86	49.95	16.01	6.01	15.66	5.43	3.77	52.55	13.62	6.03	14.47	7.13	2.23
2012	55.32	11.02	5.04	15.27	6.99	2.81	53.36	13.74	4.94	16.10	5.52	3.84	55.45	9.88	5.09	15.53	7.80	2.12
2013	58.93	6.7	7.87	13.95	6.56	2,00	58.98	7.53	7.76	14.99	4.72	2.56	58.27	6.47	7.89	13.95	7.46	1.64

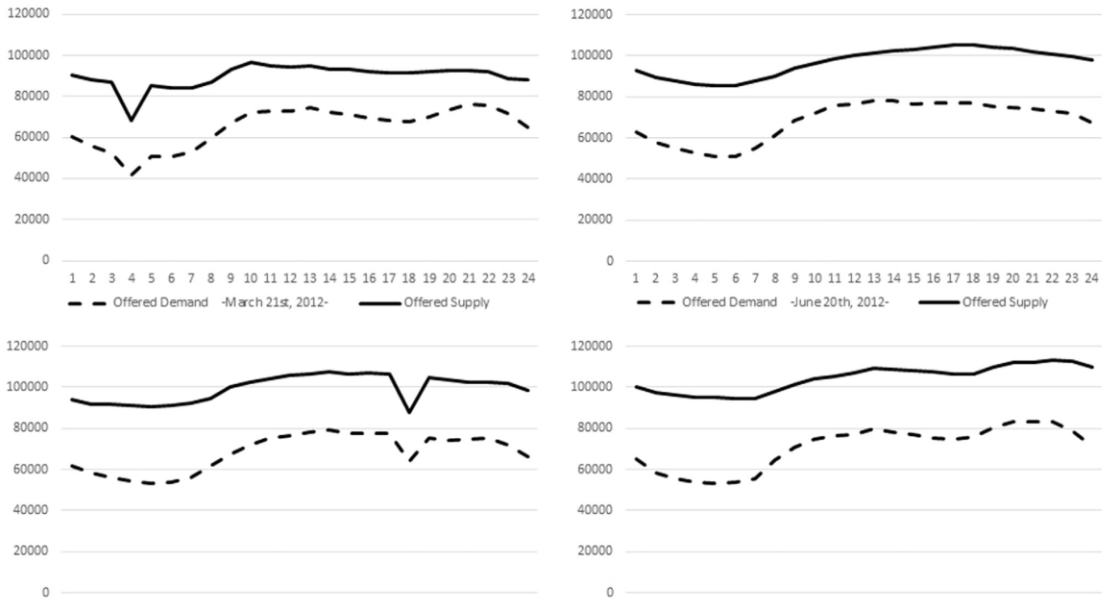

Fig. A. 1. *24-h load curves of four typical days in 2012 (MWh).*

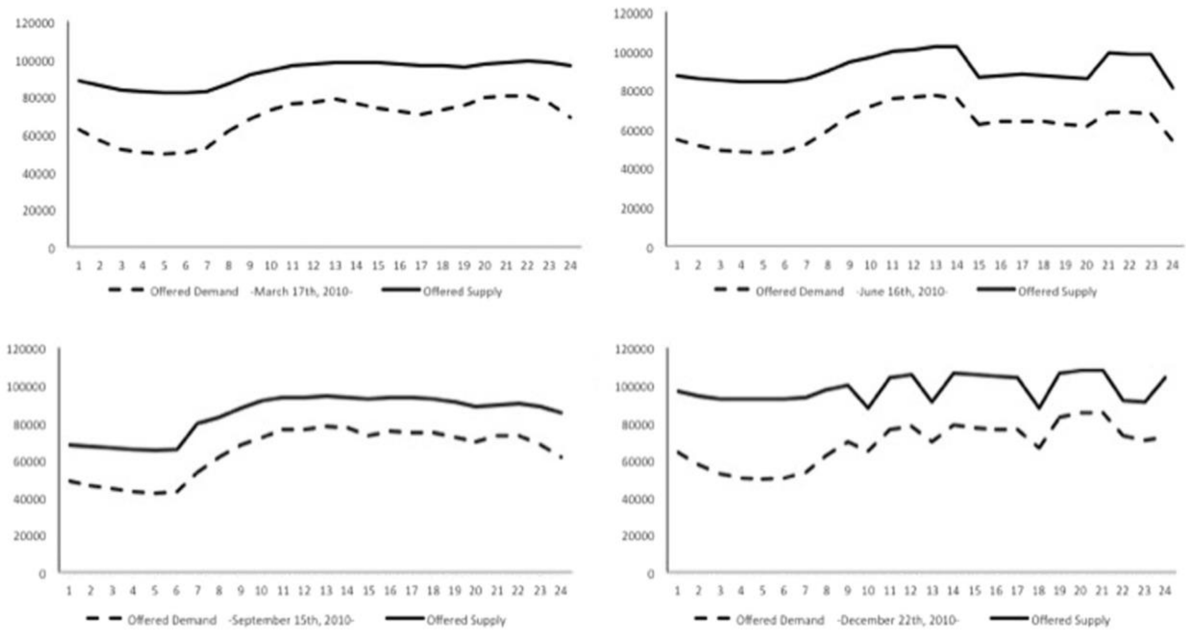

Fig. A. 2. *24-h load curves of four typical days in 2010 (MWh).*

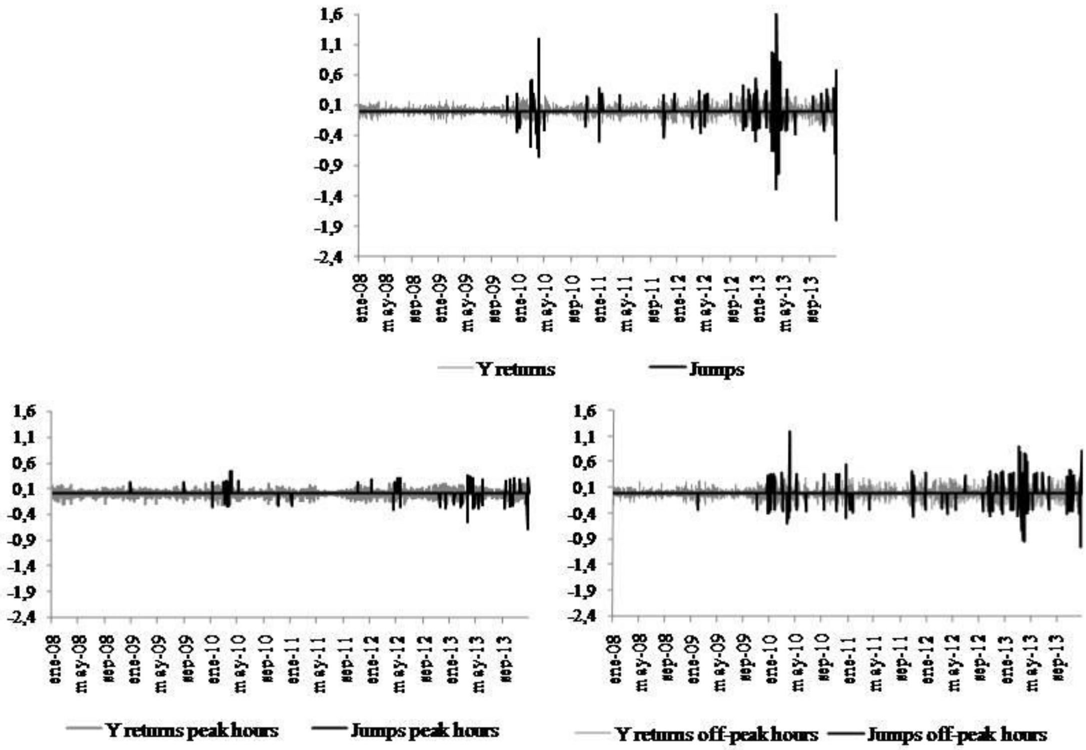

Fig. A. 3. *Price without seasonal component (Y) in returns and Jumps detected (2008–2013).*

References

[1] Schwartz E. The stochastic behavior of commodity prices: valuations and hedging. J Finance 1997;52(3):923–73.

[2] Lucia J, Schwartz E. Electricity prices and power derivatives: evidence from nordic power exchange. In: Review of derivatives research, vol. 5 (1); 2002. p. 5–50.

[3] Deng S. Stochastic models of energy commodity prices and their applications: mean-reversion with jumps and spikes. United States of America: University of California Energy Institute; 2000 (Working paper, PWP-073).

[4] Knittel CR, Roberts MR. An empirical examination of restructured electricity prices. Energy Econ 2005;2005(27):791–817.

[5] Huisman R, Mahieu R. Regime jumps in power prices. Energy Econ 2003;25 (5):425–34.

[6] Karakatsani N, Bunn DW. Modelling the volatility of spot electricity prices, working paper. London: Department of Decision Sciences, London Business School; 2004.

[7] Würzburg K, Labandeira X, Linares P. Renewable generation and electricity prices: taking stock and new evidence for Germany and Austria. Energy Econ 2013;40:S159–71.

[8] Saénz de Miera G, Del Río P, Vizcaíno I. Analyzing the impact of renewable electricity support schemes on power prices: the case of wind electricity in Spain. Energy Policy 2008;36:3345–59.

[9] Sensfuβ F, Ragwitz M, Genoese M. The merit-order effect: a detailed analysis ofthe price effect of renewable electricity generation on spotmarket prices in Germany. Energy Policy 2008;36:3086–94.

[10] Holttinen H. The impact of large scale wind power production on the nordic electricity system. (Dissertation for the title of Doctor of Science in Technol- ogy). Finland: Helsinki University of Technology; 2004.

[11] McConnell D, Hearps P, Eales D, Sandiford M, Dunn R, Wright M, et al. Retrospective modeling of the merit-order effect on wholesale electricity prices from distributed photovoltaic generation in the Australian National Electricity Market. Energy Policy 2013;58:17–27.

[12] Mulder M, Scholtens B. The impact of renewable energy on electricity prices in the Netherlands. Renew Energy 2013;57:94–100.

[13] Gelabert L, Labandeira X, Linares P. An ex-post analysis of the effect of renewables and cogeneration on Spanish electricity prices. Energy Econ 2011;33:S59–65.

[14] Benth E, Benth J S, Koekebakker S. Stochastic modelling of electricity and related markets, 11. Singapore: World Scientific; 2008.

[15] Benth F, Ekeland L, Hauge R, Nielsen B. A note on arbitrage-free pricing of forward contracts in energy markets. Appl Math Finance 2003;10(4):325–36.

[16] Cartea A, Figueroa G. Pricing in electricity markets: a mean reverting jump diffusion model with seasonality. Appl Math Finance 2005;12(4):313–35.

[17] De Jong C. The nature of power spikes: a regime-switch approach. Stud Nonlinear Dyn Econom 2006;10(3):1558–3708.

[18] Benth F, Kiesel R, Nazarova A. A critical empirical study of three electricity spot price models. Energy Econ 2012;34:1589–616.

[19] Green R, Vasilakos N. Marketbehaviour with large amounts of intermittent generation. Energy Policy 2010;38:3211–20.

[20] Hellström J, Lundgren J, Yu H. Why do electricity prices jump? Empirical evidence from the Nordic electricity market Energy Econ 2012;34:1774–81.

[21] Burgos-Payán M, Roldán-Fernández J, Trigo-García A, Bermúdez-Ríos J, Riquelme-Santos J. Costs and benefits of the renewable production of electricity in Spain. Energy Policy 2013;56:259–70.

Capítulo 2:

Capítulo 2:

Impact of wind electricity forecast on bidding strategies.

Abstract

The change in the generation mix from conventional electricity sources to renewables has important implications for bidding behaviour and may have an impact on prices. The main goal of this work is to discover the role played by expected wind production, together with other relevant factors, in explaining the day-ahead market price through a data panel model. The Spanish market, given the huge increase in wind generation observed in the last decade, has been chosen for this study as a paradigmatic example. The results obtained suggest that wind power forecasts are a new key determinant for supply market participants when bidding in the day-ahead market. We also provide a conservative quantification of the effect of such trading strategies on marginal prices at an hourly level for a specific year in the sample. The consequence has been an increase in marginal price to levels higher than what could be expected in a context with notable wind penetration. Therefore, the findings of this work are of interest to practitioners and regulators and support the existence of a wind risk premium embedded in electricity prices to compensate for the uncertainty of wind production.

1. Introduction

Because of deregulation, the price for electricity has come to be determined by competitive bidding by producers and consumers in the wholesale day-ahead market, where an auction system is generally followed. The electricity supply function is discontinuous and increases with the level of demand. The resulting price from the auction, the so-called marginal price, corresponds to the highest price offered by the supply side from those accepted to satisfy demand. The offered prices to sell electricity will, in turn, depend on production costs– and these significantly differ among the generation technologies. Therefore, the generation mix of a specific market area, among other factors, will likely condition the resulting marginal prices and the success of a given market design. Establishing the factors affecting price is crucial for all market participants for obtaining accurate forecasts when planning production and consumption, or when designing hedging strategies to face the price variation risk to which their positions are exposed.

Due to greater climate awareness, the inclusion of renewable production in the electricity system is a goal in most countries. Apart from the promotion of renewable generation, another measure taken to fight climate change has been the creation of carbon emission markets. The mechanism works as follows. At the end of each year, firms must deliver an equivalent number of allowances for their excess emissions. Firms are then provided with a number of emission allowances that depend on their pollution levels (derived from their production). Firms that need to increase their volume of emissions must have–or buy–the corresponding permits in the carbon emissions market. Within this new framework, in addition to input costs, market participants may have internalised the expected future carbon prices and wind production forecasts into their decision-making process when designing their bidding.

In the Spanish case, the development and integration of renewable electricity production in the electricity market has been a target for the regulator over the last decade. Tables 1 and 2 show the annual figures for installed power capacity and electrical energy in Spain per generation technology from 2007 to 2013. These tables show that installed wind power capacity increased around 68% in the mentioned interval, whereas the amount of wind generation grew by 98% and reached 20% of the overall generation during 2013.

Table 1. Installed power capacity (MW) from 2007 to 2013.

Technology	2007	2008	2009	2010	2011	2012	2013	Δ
Hydraulic	17507.3	17555.42	17555.42	17564.63	17571.99	17786.40	17785.98	1.59%
Nuclear	7729.11	7729.11	7729.11	7790.38	7865.99	7865.99	7865.99	1.77%
Coal	11894.79	11897.13	11897.13	11918.11	12158.11	11623.77	11641.23	-2.13%
Fuel + gas	7542.55	7161.10	5994.58	5145.44	3717.33	3428.73	3498.37	-53.62%
Combined Cycle	22390.25	23105.03	24503.01	27146.39	27171.21	27206.47	27206.47	21.51%
Other Hydraulic	1871.49	1981.13	2022.91	2036.94	2042.40	2042.76	2105.70	12.51%
Wind	13667.82	16117.99	18.869,00	19715.31	21174.9	22765.85	23002.3	68.30%
Photovoltaic	636.93	3352.55	3398.1	3838.45	4259.35	4559.53	4667.03	632.74%
Thermal Solar	11.02	60.92	232.22	532.02	998.62	1950.02	2299.52	20766.79%
Thermal Renewable	588.17	634.57	782.12	821.13	887.07	975.41	980.05	66.63%
Thermal non-renewable/cogeneration/ Others	6617.31	6870.29	7076.79	7240.04	7317.65	7280.7	7200.37	8.81%
Overall	90456.74	96465.24	100060.39	103748.84	105164.61	107485.64	108253.01	19.67%

Source: www.ree.es. Last accessed March 2015.

Table 2. Electric energy balance (GWh) from 2007 to 2013.

Technology	2007	2008	2009	2010	2011	2012	2013	Δ
Hydraulic	26351.89	21428.2	23862.23	38652.87	27571.15	19454.73	33970.28	13.01%
Nuclear	55102.47	58973.42	52761.04	61989.95	57731.36	61470.16	56827.39	21.77%
Coal	75027.85	49646.83	37311.24	25478.01	46518.61	57661.6	42397.79	16.24%
Fuel + gas	10784.48	10690.97	10056.01	9552.96	7479.95	7541.49	7002.18	2.68%
Combined cycle	72307.14	95528.68	82239.39	68595.33	55139.86	42510.47	28671.93	10.98%
Generation Consumption	-9634.62	-9256.95	-7999.11	-7572.09	-8128.95	-8511.61	-7053.51	-2.70%
Other hydraulic	4126.5	4639.82	5454.07	6824.32	5295.99	4646.34	7102.2	2.72%
Wind	27611.65	32159.82	38252.83	43545.33	42465.29	48508.34	54713.25	20.96%
Photovoltaic	483.9	2497.96	6072.39	6422.77	7425.12	8202.09	8326.92	3.19%
Thermal solar	7.63	15.38	129.82	691.62	1832.36	3444.13	4441.53	1.70%
Thermal renewable	2588.97	2868.71	3317.34	3332.36	4317.99	4754.77	5074.7	1.94%
Cogeneration/others	23450.43	26721.15	28600.73	30973.32	32318.8	33767.25	32296.38	12.37%
Net generation	288208.29	295913.98	280057.99	288486.75	279967.53	283449.75	273771.03	
Pump consumption	-4432.29	-3802.5	-3794.19	-4457.78	-3214.96	-5022.55	-5957.85	
International Exchange balance	-5750.47	-11039.59	-8086.41	-8332.68	-6090.13	-11199.95	-6732.14	
Overall	278025.54	281071.89	268177.39	275696.29	270662.44	267227.25	261081.04	100.00%

Source: www.ree.es. Last accessed March 2015.

This sustained growth has meant a substantial change in the generation mix from conventional energy sources to renewables, and changes in the input proportions (among them, commodity prices) in electricity production costs. Moreover, the inclusion of new generating technologies in the generation mix may have altered the bidding strategies of generators and this may have an impact on prices.

A number of studies can be found in the literature that analyse the impact on spot prices of increasing renewable electricity production. A common pattern is detected that consists of a decrease in spot prices because of an increase in renewable production. This is due to the auction mechanism that is based on a merit order dispatch system (commonly used in electricity markets). Thus, sellers and buyers, the day before delivery day, submit quantity-price bids to the auction market. These bids are ranked by price and a marginal (or clearing) price is set when the supply aggregate curve matches the demand aggregate curve. Therefore, generators with lower marginal costs, such as renewables, can bid at lower prices – and these bids are normally positioned at the base of the merit-order and so are among the first bids matched in the auction. Therefore, an increase in renewables is expected to change and shift the supply curve in such a way that the spot price could be set at lower levels. This effect has been called in the literature *the merit-order effect*[14] of renewables and has been highlighted in previous studies ([1-8], among others). A reduction in spot prices is welcomed by consumers and regulators. In fact, such a reduction will mean savings for household and industrial consumers, with the well-known implications in productivity gains. Such a reduction should also help to compensate for the economic effort required to finance support for renewables[15].

Agent-based models have been used in the literature to capture the complexity of the bidding strategy in electricity markets. [10] present an overview of the techniques used by researchers to capture the dynamics in electricity markets that focuses on the agent-based models. The authors in [11] propose a model to maximise the benefits of a single generator that includes the expected behaviour of the rest of the participants and some of the characteristics of the electricity markets – such as the existence of congestion in the grid. In [12], an agent-based model is adjusted to the electricity market in Germany and it is found that the reduction in the spot price caused by renewables is higher when there are no transmission capacity constraints. The authors in [13] carry out an interesting theoretical analysis by adding the effect on the spot price of conventional generator's strategies in scenarios considering market power, wind and forward trading. In large wind, conventional generators with market power can follow strategies pushing the price below competitive price, in order to win more pay buck energy, and doing the opposite

[14] For a complete overview of past research on the merit-order effect of renewables see [9].
[15] In Europe, the most commonly adopted renewable support mechanism has been a feed-in-tariffs scheme in which the cost of the project is transferred to customers.

in lower wind. As a consequence, the benefits for conventional and wind generators will be asymmetric. However, forward trading could help to reduce this effect. The model in [13] is empirically tested in [4] by applying it to the British wholesale market, and the predicted asymmetric benefits are observed. The prices received by wind generators are lower than the demand-weighted price. Additionally, it is point out that prices can be higher and more volatile in scenarios with market power. In [14] the relationship between wind power forecasts and spot prices is analyzed in the Western Danish price area of the Nord Pool's Elspot market. In [15] the aim is to find what is the optimal bidding strategy for a wind generator to maximize profits. The wind generation firms can increase their net earnings by improving wind forecasting accuracy. Focused on the Spanish electricity market, [16] test whether the bidding behaviour of large oil-fired thermal generators differed from that of small oil-fired thermal generators from 2002 to 2005.
Our approach is different as we are interested in distinguishing the effects by generation technology. Therefore, we study the impact on prices from bidding strategies by thermal, combined cycle, nuclear, hydroelectric, and renewable generation plants. To do so, we use a panel data model at an hourly level, similar to the panel data used [8] for the Irish single electricity market[16]. The period under study, from 2007 to 2013, is characterised by the installation in Spain of a number of combined cycle and renewable source plants. Finally, an approximate quantification of the effect of wind production forecasts on the day-ahead market price is provided.

The main goal of this study is to identify the factors playing a specific role in the bidding behaviour by generators in the Spanish electricity day-ahead market. We are particularly interested in the role played by expected wind production as a new key determinant in this new context. The Spanish market has been chosen as a paradigmatic example due to the huge increase in wind generation in recent years. Together with expected wind production and based on the Spanish generation mix, we also control for other potential noteworthy factors, such as carbon and natural gas prices, and reservoir levels.

Our findings support the existence of a wind risk premium embedded in electricity prices to compensate for the uncertainty of wind production. It is interesting that it is not wind

[16] In [8] model the spot price using an extension of seemingly unrelated regressions (SUR) for panel data models considering a system of 24 hours, one equation for each hour of the day from 2008 to 2013.

farms who are behind this premium–but the thermal (fuel oil, natural gas, and coal) and combined cycle plants who see how their production is being increasingly replaced by wind power when the wind blows.

This work is organised as follows. Section 2 describes the dataset used to embrace the analysis. Section 3 presents the methodology to detect the key factors explaining generator bidding and the results. Section 4 provides a quantification of the estimated impact of the expected wind production translated into bidding strategies on the marginal price by generation technology. Finally, Section 5 discusses the research results and gives some concluding remarks.

2. Data

The data set covers the period from 1 January 2010 to 31 December 2013 and consists of the following time-series data:

- Quantity-price offers submitted by each generation unit to the day-ahead Spanish electricity market, to sell or buy energy, and by delivery hour (including matched and non-matched offers). This data is available on the website of Iberian Market Operator for Electricity, OMIE (www.omie.es).
- Wind power forecasts released by REE (www.ree.es) on an hourly basis. From all wind power forecasts available on an hourly basis for each delivery day-ahead hour, we carefully selected the last wind power forecasts available just before the deadline for submitting bids to the day-ahead auction markets. The historical series of data was directly received from REE.
- The day-ahead market marginal hourly prices were downloaded from the OMIE website (www.omie.es).
- National balance point natural gas day-ahead prices. The data was obtained from the Thomson Reuters database. Originally quoted in GBp/Therm, the data was transformed into euros/MWh for this study.
- European emission allowances (EUAs) futures prices corresponding to next December maturity with a daily frequency, obtained from the Thomson Reuters database.

- Hydroelectric reservoir data with a weekly frequency, downloaded from the Thomson Reuters database.

In short, the overall data set used in the present study includes 72 million records.

3. Empirical Analysis

This work aims at exploring the main factors affecting bidding by generators, and particularly the role of expected wind electricity production, since wind has recently entered the generation mix in many countries and expected wind production is becoming increasingly important. In areas where wind power has a significant share in the generation portfolio, variations in wind power generation can lead to substantial short-term changes in the overall supply function. Bids made by wind generators are usually among the first matched in the day-ahead auction market. This is due to the market mechanism itself, a merit-order dispatch procedure in which those technologies with lower variable costs (like nuclear, but also wind) can submit bids with lower prices and be among the first to be matched. Thus, if the wind blows, marginal prices are expected to decrease and generation technologies other than wind are likely to be (at least partly) replaced. Therefore, to optimise profits both renewable and non-renewable generators are incentivised to behave strategically when submitting bids to an auction market that will also depend on wind power forecasts.

To embrace this analysis, it was necessary to deal with the data of the whole supply curve of the day-ahead market (all the offered prices of supply side participants) at an hourly level. For generator bidders, the day-ahead market is really made up of 24 auctions, one for each delivery hour. Bids have been grouped by generation technologies to disentangle differing plant-type strategies. According to the classification made on the OMIE webpage, bids from generators have been grouped into the following categories: combined cycle (CC); coal, fuel-gas and fuel-oil thermal plants (CT); hydroelectric (CH); nuclear (CN); and finally, renewable technologies, mainly wind and solar (CR[17]).

[17] This category also includes bids coming from cogeneration and surplus production, but these latter bids are of residual importance because of their relatively scarce associated volume.

A panel data model has been chosen to make the most of the data. Generators submit bids for 24 hour blocks at the same time, but the marginal price is set in a different auction for each delivery hour, and so hourly prices can be considered as separate contracts – but traded at the same time. A panel data model can capture both the unobserved effect (due to the delivery hour) and all the predictive information available at the single moment of bidding.

3.1 *Preliminary Analysis*

Firstly, for the good specification of the model, a check is made as to whether the coefficients representing cross-section-specific characteristics are equal for all cross-sections. To do so, we test for poolability across sections in the panel data model in an F test, with a null hypothesis that assumes homogeneous slope coefficients for all cross-sections (hours). As shown in Table 3 (a) the null is rejected for CT, CC, and CH generation technology plant groups, indicating there are cross-section-specific characteristics that depends on the delivery hour. Additionally, the Durbin-Wu-Haussman test (H test) has been used to differentiate between two options to model the cross-section effect: a fixed-effects model or a random-effects model. Contrarily, the null cannot be rejected for CN and CR – meaning that a pooling model is preferable when analysing these latter groups. The results of the H test confirm the previous result that the fixed-effects model (1) is more suitable than the random-effects model for CT, CC, and CH. Therefore, for these generation technologies, the fixed-effects panel data model is adequate [17] and can be defined as follows:

$$PMO_{t-1,t,i} = \beta^{*}WPF_{t-1,t,i} + \gamma^{*}\,DL_t + \phi^{*}\,NG_{t-1}/EUA_{t-2} + \omega WR_{t-1} + \delta\,PMO_{t-2,t-1,i} + u_{t,i} \tag{1}$$

$$u_{t,i} = v_i + \varepsilon_{t,i}$$

where:

- $PMO_{t-1,t,i}$ is the average supply offered price by the group of generators sharing the same generation technology submitted on a particular day (day t-1) for delivering electricity at the day-ahead (day t) for the hour i.
- $WPF_{t-1,t,i}$ denotes the last available wind power forecast as made public by REE before the deadline for submitting bids to the day-ahead auction market for delivering electricity during hour i on day t, and known at day t-1.

- DL_t is a dummy variable that is equal to 1 if t is a business day and 0 otherwise. It is included in the model to capture the business-day effect on electricity day-ahead prices.
- NG_{t-1}/EUA_{t-2} is the ratio: national balance point natural gas day-ahead prices on day t-1 divided by ICE ECX European emission allowances next December maturity closing futures prices at t-2[18]. Regarding the latter, we use lagged prices because the available closing prices at the closure time of the day-ahead auction market which takes place at t-1 (for delivering electricity at t) are those of the previous trading session, i.e. at t-2, and
- WR_{t-1} is the hydroelectric reservoirs on day t-1.
- The stochastic component, $u_{t,i}$, is a process made up of two components: ν_i, which is assumed to be independent during the days, although it allows for cross-sectional covariance between the hours, and $\varepsilon_{t,i}$, which is the usual homoscedastic component, normally distributed $N(0,\sigma)$. Indeed, it is the specification of a fixed-effects panel model in which the cross-section is the delivery hour i=1,2…24.

As previously stated, the null hypothesis of poolability could not be rejected for CN and CR plant groups, which prevented us from using a panel data model as specified

[18] We firstly considered the inclusion of natural gas and carbon emission price series as separate explanatory variables, but the higher correlation between them prevented us from doing so. We finally opted to use the series of the ratio carbon emission prices/natural gas prices to avoid multicollinearity problems.

in (1) for these generator types. Contrarily, it is more appropriate in this case to use a pooling model for CN and CR plant groups, as follows:

$$PMO_{t-1,t,i} = \alpha + \beta^*WPF_{t-1,t,i} + \gamma^* DL_t + \phi^* NG_{t-1}/EUA_{t-2} + \omega WR_{t-1} + \delta\, PMO_{t-2,t-1,i} + \varepsilon_{t,i} \quad (2)$$

which differs from (1) in the inclusion of an intercept, α, which is the same for all cross-sections substituting the fixed-effects hourly components, v_i, and in that the stochastic component is a process made up of only one component, $\varepsilon_{t,i}$, the usual homoscedastic component, normally distributed $N(0,\sigma)$.

It is well known that regression models for non-stationarity variables give spurious results unless the series are cointegrated. Non-stationarity is at least as serious a problem for panel data sets as it is for aggregate data, since non-stationarity could cause spurious estimates when estimating static panel models, according to [18]. We use the unit root test proposed by [19] to test the stationarity of the dependent variable ($PMO_{t-1,t,i}$). As is displayed in Table 3, the null hypothesis of a lack of stationarity is rejected for all the generation technology plant groups. Furthermore, to control for multicollinearity, the variance inflation factors (VIF) have been obtained (Table 3 (b)). In all cases, VIF is lower than 2.1, a level which is considered acceptable following [17].

Table 3. Panel data specification tests: *(a) presents the results of the F test, and the Durbin-Wu-Haussman and Maddala-Wu tests. The F test (F test in the table) is a poolability test and enables choosing between a fixed-effect panel data model and a pooling model. The null hypothesis is a homogeneous slope coefficient for all cross-sections (hours) indicating that a pooling model is preferred. The Durbin-Wu-Haussman test (H test) is used to select between the existence of random or fixed effects in the panel data model with auxiliary regression and robust covariance estimators. The null hypothesis is that the random-effect estimation is preferred (consistent and more efficient). Finally, the Maddala-Wu test (MW test) is a unit root test for panel data. The null hypothesis is non-stationarity. The tests are performed for all technologies with the exception of the H test, which is only for combined cycle generators, thermal, and hydraulics. The significant codes (Sig.) are: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1; (b) presents the variance inflation factors (VIF) to control for multicolinearity.* $WPF_{t-1,t,i}$ *denotes the wind power forecast made public on day t-1 just before the deadline for submitting bids to the day-ahead market action for delivering electricity during hour i on day t;* DL_t *is a dummy variable that equals 1 if t is a business day, and 0, otherwise;* NG_{t-1} *is the British natural day-ahead trade close price on day t-1;* EUA_{t-2} *refers to the European emission allowances next December maturity futures closing prices on day t-2;* WR_{t-1} *denotes the hydroelectric water reservoirs on day t-1, and* $PMO_{t-2,t-1,i}$ *is the average supply price offered by the group of generators the day-before (day t-2) for delivering electricity on the day-ahead (day t-1) during hour i. combined cycle (CC), thermal (CT), hydraulics (CH), nuclear (CN) and renewable (CR).*

	CC		CT		CH		CN		CR	
(a)	Statistic	Sig.	Statistic	Sig.	Statistic	Sig.	Statistic	Sig.	Statistic	Sig.
F test	6.35	***	18.25	***	3.07	***	0.38		0.75	
H test	539.07	***	3537.80	***	49.42	***				
MW test	104.81	***	1025.70	***	410.33	***	1142.10	***	246.18	***
(b)	VIF		VIF		VIF		VIF		VIF	
$WPF_{t-1,t,i}$	1.096		1.116		1.128		1.085		1.087	
DL_t	1.000		1.002		1.000		1.002		1.001	
NG_{t-1}/EUA_{t-2}	1.640		1.210		1.306		1.105		1.370	
WR_{t-1}	1.448		1.121		1.240		1.077		1.111	
$PMO_{t-2,t-1,i}$	2.064		1.206		1.440		1.039		1.300	
$PMO_{t-2,t-1,i}$	2.064		1.206		1.440		1.039		1.300	

Before estimation, it is interesting to take a brief look at the main descriptive statistics for the variables used in the present study, which are displayed in Table 4. As expected, average offered prices to sell electricity from low variable cost plants, such as CN or CR, are considerably lower than those offered by higher variable cost plants, namely, CC or CT. This is true for the four quantiles of the bid-price distributions.

The range (R) is calculated as the difference between the largest and smallest offered prices and indicates the array of prices at which generators of the same technology have submitted their bids to the auction during the studied sample. This measure of variation gives us an idea about the most actively strategic plant groups, namely, those groups that are flexible enough to adapt their bidding to expected supply and demand levels. Thereby, the highest range is obtained for CC plants (90.11), followed by CT plants (82.52), CH plants (79.27), CN plants (45.30) and CR plants (27.60). However, it is noteworthy that the range measure has the disadvantage that it is based on only two observations and fails to show how the other observations are arranged between them. The interquartile range (IR), calculated as the difference between the third and first quartiles, overcomes this drawback, indicating the spread of the middle 50% of the distribution. Thus, the interquartile range is again the highest for CC plants (23.24), but now followed at a greater distance by CH plants (11.58), CT plants (8.71), CN plants (4.55), and CR plants (3.78). Note that these are still conservative indicators, since we are dealing with average offered prices by generation source and both the range and interquartile range for each generator are expected to be higher.

Another statistic that is frequently used to measure variability in prices is standard deviation. From Table 4, the highest value of standard deviation corresponds to bids from CC plants (15.27), followed by CH plants (10.19), and by CT plants (8.69). The standard deviations of CN and CR plants are remarkably lower (respectively, 3.87 and 4.77). The small standard deviation of the ratio NG/EUA (2.06) needs to be highlighted, as well as the large standard deviation of the wind power forecasts (2792), consistent with the variability of wind production that often makes it difficult to predict.

Table 4. Descriptive summary statistics (2010-2013): *Descriptive summary statistics of the variables included in the analysis: PMOe is the average supply price offered by the group e of generators sharing the same generation technology submitted on a particular day (day t-1) for delivering electricity on the day-ahead (day t) during hour i; WPFt-1,t,i denote the wind power forecast made public just before the deadline to submit bids to the day ahead market on day t-1 for delivering electricity during hour i on day t; NGt-1 is the British natural day-ahead trade close price on day t-1; EUAt-2 refers to the European emission allowances next December maturity futures closing prices on day t-2 (EUAt-1); WRt-1 denotes the hydroelectric water reservoirs on day t-1. Combined cycle generators (CC), thermal (CT), hydraulics (CH), nuclear (CN) and renewable (CR). [Min=minimun; Max=maximum; median; mean; 1st Qu=25% quantile; 3rd Qu.=75% quantile; Sd=standard deviation; R=range; IR=interquartile range]*

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max	Sd	R	IR
PMOe=CC	30.03	49.59	66.19	63.35	72.83	120.14	15.27	90.11	23.24
PMOe=CT	42.05	53.99	58.35	59.20	62.70	124.57	8.68	82.52	8.71
PMOe=CH	21.74	48.70	55.55	54.12	60.28	101.01	10.19	79.27	11.58
PMOe=CN	00.00	18.20	20.22	20.62	22.75	45.30	3.87	45.30	4.55
PMOe=CR	05.62	19.40	21.51	20.31	23.18	33.22	4.77	27.60	3.78
WPFt-1,t,i	508.00	3124.00	4776.00	5289.00	6984.00	16264.00	2791.99	15756.00	3860.00
NGt-1/EUAt-2	0.73	1.35	2.85	3.12	4.64	10.11	2.06	9.38	3.29
WRt-1	47.30	65.10	77.70	73.03	81.60	89.20	11.91	41.90	16.50

Table 5 displays the average prices offered by generation technology and by hour and distinguishing between business and non-business days. Interestingly, the average offered prices by CC and CT are remarkably higher in the early hours of the day, more specifically, in the first eight hours, regardless of whether it is a business or a non-business day.

Table 5. Average supply price offered by technologies, hours, and business/non-business days*.*
Combined cycle (CC), thermal (CT), hydraulics (CH), nuclear (CN) and renewable (CR).

	Overall					Business day					Non-business day				
Hour	CC	CT	CH	CN	CR	CC	CT	CH	CN	CR	CC	CT	CH	CN	CR
1	69.53	67.36	54.56	20.24	23.67	68.87	66.98	54.78	20.24	23.62	71.05	68.24	54.04	20.26	23.80
2	73.12	66.81	54.85	20.62	21.78	72.60	66.62	55.03	20.64	21.76	74.31	67.26	54.44	20.58	21.82
3	72.50	67.07	55.25	20.57	21.70	72.21	67.11	55.36	20.60	21.66	73.17	66.99	55.01	20.51	21.79
4	72.95	66.60	55.39	20.51	21.74	72.60	66.60	55.47	20.53	21.71	73.75	66.60	55.21	20.44	21.81
5	71.47	64.63	55.42	20.42	21.79	71.18	64.60	55.50	20.45	21.76	72.14	64.71	55.24	20.34	21.86
6	71.08	63.13	55.62	20.40	22.16	70.70	62.89	55.80	20.42	22.17	71.93	63.69	55.21	20.34	22.14
7	69.57	59.99	55.78	20.37	22.21	68.83	59.52	56.06	20.43	22.28	71.24	61.06	55.14	20.24	22.05
8	65.34	56.83	55.22	20.36	20.51	64.30	56.53	55.39	20.45	20.56	67.70	57.51	54.83	20.16	20.40
9	60.22	56.52	54.63	20.43	19.68	59.31	56.27	54.62	20.53	19.67	62.28	57.09	54.68	20.21	19.70
10	59.72	56.05	53.90	20.56	19.39	58.92	55.82	53.74	20.66	19.40	61.54	56.58	54.26	20.36	19.37
11	59.57	55.65	53.48	20.63	19.36	58.95	55.41	53.26	20.71	19.37	60.96	56.19	53.98	20.46	19.34
12	59.51	54.95	53.28	20.62	19.29	58.93	54.74	53.03	20.70	19.28	60.85	55.43	53.84	20.46	19.31
13	59.49	55.81	53.25	20.70	19.16	58.92	55.57	53.06	20.80	19.15	60.78	56.34	53.69	20.48	19.17
14	59.49	57.89	53.35	20.74	19.04	58.93	57.61	53.23	20.84	19.03	60.77	58.52	53.62	20.52	19.07
15	59.59	57.95	53.51	20.81	18.94	59.06	57.67	53.41	20.93	18.89	60.80	58.58	53.73	20.53	19.04
16	59.60	57.98	53.61	20.86	18.91	59.05	57.68	53.50	20.96	18.86	60.84	58.66	53.84	20.62	19.02
17	59.56	58.12	53.76	20.87	18.95	59.01	57.76	53.65	20.97	18.93	60.81	58.95	54.01	20.65	19.00
18	59.50	58.16	53.87	20.84	19.09	58.96	57.86	53.78	20.93	19.08	60.72	58.85	54.07	20.61	19.10
19	59.34	58.33	53.82	20.86	19.35	58.87	58.06	53.81	20.94	19.36	60.40	58.95	53.85	20.66	19.33
20	59.35	58.40	53.71	20.81	19.64	58.90	58.18	53.79	20.90	19.65	60.36	58.90	53.51	20.60	19.62
21	59.34	58.27	53.20	20.74	19.99	58.92	58.01	53.35	20.80	19.96	60.30	58.85	52.86	20.59	20.05
22	59.33	57.70	52.97	20.76	20.34	58.90	57.49	53.16	20.81	20.28	60.32	58.15	52.52	20.65	20.47
23	59.74	58.18	53.02	20.69	20.55	59.34	57.97	53.22	20.75	20.48	60.65	58.63	52.57	20.56	20.70
24	61.51	57.92	53.53	20.48	20.38	61.25	57.71	53.82	20.58	20.27	62.09	58.41	52.89	20.27	20.62

To see whether those differences are statistically significant, a test for equality of means between the block of the first eight hours of the day and the remaining block of hours is conducted. Table 6 presents the results in three panels, distinguishing between: the overall sample (a); the sample only including business-day observations (b); and the sample including only non-business-day observations (c). As can be observed, the average offered prices for the first eight hours of the day are significantly different from those for the remaining hours for all the considered generation technologies. From (a), it is noticeable that the average offered price for the first eight hours is remarkably higher than the average offered prices for the block of the remaining hours in the cases of CT and CC. This difference is just slightly (but significantly) higher in the case of CH and CR and lower in the case of CN. These results remain the same when moving to (b) (business days) and (c) (non-business days), with the only exception being CR (that does not offer a significantly different price for delivery hours on non-business days).

Table 6. Equality of mean test*. A test for equality of means in the bid prices for the block of the first eight hours and the block of the remaining hours was conducted: (a) presents the results for the overall sample, distinguishing by generation technology.; (b) and (c) presents the results focusing on the business (non-business) days. Combined cycle (CC), thermal (CT), hydraulics (CH), nuclear (CN) and renewable (CR). The sample mean estimates in samples are calculated and t-statistics are shown. The significant codes (Sig.) are: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.*

	CC		CT		CH		CN		CR	
(a)	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**
Mean	70.7	59.6	64.0	57.4	55.3	53.5	20.4	20.7	22.0	19.5
t-statistic	61.0		56.6		14.5		-6.2		44.8	
Sig.	***		***		***		***		***	
(b)										
[Business days]	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**
Mean estimates	*70.1*	*59.1*	*63.8*	*57.1*	*55.4*	*53.5*	*20.4*	*20.8*	*21.9*	*19.5*
t-statistic	*51*		*47.7*		*13.5*		*-6.3*		*37.3*	
Sig.	***		***		***		***		***	
(c)										
Non-business days	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**	**[1-8h]**	**[9-24h]**
Mean	*71.9*	*60.9*	*64.5*	*57.9*	*54.8*	*53.6*	*20.3*	*20.5*	*21.9*	*19.5*
t-statistic	*33.6*		*30.6*		*5.9*		*-1.8*		*24.7*	
Sig	***		***		***		.		***	

3.2 Estimation Results

The panel data estimation results are presented in Table 7 (for CC, CT, and CH generation technologies) and in Table 8 (for CN and CR generation technologies).

The standard errors have been obtained following [20], since they are robust to serial correlation over time and specifically convenient when cross-sectional dependence is present, according to [21][19]. The obtained R-squared is above 90% for CC, CH, and CR – and above 60% for all considered generation technologies except CN (54%).

[19] To reinforce the robustness of the results, following [22], the standard errors have also been calculated using a robust covariance matrix that controls for heteroskedasticity and serial (cross-sectional) correlation for fixed-effect models. Moreover, a wild cluster bootstrapped t-statistics estimation for cluster-robust standard errors in fixed-effect models was conducted ([23]) that provides asymptotic refinement when the number of clusters is fewer than 30. A double-clustering robust covariance matrix estimation for panel models was also conducted. All these approaches led us to the same results as those presented in the main text and so they hold up with remarkable consistency.

Table 7. Panel data fixed-effects estimation. *PMOt-1,t,i = β*WPFt-1,t,i+ γ* DLt + φ* NGt-1/EUAt-2+ωWRt-1+ δ PMOt-2,t-1,i + ut,i ; ut,i =νi + εt,i (1). Fixed-effects panel linear models estimation with nonparametric robust covariance matrix estimators with cross-sectional and serial correlation ([20]). PMOt-1,t,i is the average supply price offered by the group of generators sharing the same generation technology submitted on a particular day (day t-1) for delivering electricity on the day-ahead (day t) during hour i; WPFt-1,t,i denotes the wind power forecast made public on day t-1 just before the deadline for submitting bids to the day-ahead market action for delivering electricity during hour i on day t; DLt is a dummy variable that equals 1 if t is a business day, and 0, otherwise; NGt-1 is the British natural day-ahead trade close price on day t-1; EUAt-2 refers to the European emission allowances next December maturity futures closing prices on day t-2; WRt-1 denotes the hydroelectric water reservoirs on day t-1, and PMOt-2,t-1,i is the average supply price offered by the group of generators the day before (day t-2) for delivering electricity on the day-ahead (day t-1) during hour i. The stochastic component, ut,i, is a process made up of two components: νi, which is assumed to be independent over the days but allows for cross-sectional covariance between the hours and εt,i, which is the usual homoscedastic component, normally distributed N(0, σ). There are three panel data fixed-effect models, one for each generation technology group. Combined cycle generators (CC), thermal (CT), hydraulics (CH). The significant codes (Sig.) are: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.*

	CC			CT			CH		
Parameter	**Value**	**t-statistic**	**Sig.**	**Value**	**t-statistic**	**Sig.**	**Value**	**t-statistic**	**Sig.**
β	0.00014	8.00	***	0.00034	8.09	***	-0.00016	-5.99	***
γ	-1.33477	-11.93	***	-1.16985	-8.85	***	0.14415	1.30	
φ	0.09844	3.35	***	-0.34550	-5.64	***	0.14921	3.91	***
ω	-0.01561	-3.50	***	0.04181	5.31	***	-0.01644	-3.94	***
δ	0.95877	195.46	***	0.76171	36.94	***	0.94246	83.82	***
V_1	3.87672	6.68	***	13.05886	10.36	***	4.67539	5.54	***
V_2	4.03756	6.78	***	12.96304	10.35	***	4.67676	5.54	***
V_3	4.02245	6.79	***	13.0547	10.37	***	4.6857	5.53	***
V_4	4.06191	6.84	***	12.97138	10.36	***	4.67988	5.52	***
V_5	4.01183	6.83	***	12.53579	10.29	***	4.66785	5.51	***
V_6	4.00611	6.84	***	12.20608	10.23	***	4.66998	5.51	***
V_7	3.95008	6.82	***	11.47737	10.07	***	4.67175	5.51	***
V_8	3.77923	6.72	***	10.72787	9.88	***	4.63409	5.50	***
V_9	3.57346	6.60	***	10.65872	9.86	***	4.60065	5.50	***
V_{10}	3.55247	6.59	***	10.54579	9.83	***	4.56004	5.50	***
V_{11}	3.53845	6.57	***	10.43079	9.79	***	4.54366	5.51	***
V_{12}	3.52077	6.53	***	10.22657	9.72	***	4.55033	5.52	***
V_{13}	3.50304	6.50	***	10.39007	9.76	***	4.56618	5.53	***
V_{14}	3.48723	6.47	***	10.84854	9.88	***	4.59104	5.54	***
V_{15}	3.47714	6.44	***	10.82819	9.87	***	4.61774	5.55	***
V_{16}	3.46436	6.41	***	10.80613	9.86	***	4.63886	5.56	***
V_{17}	3.45504	6.40	***	10.82202	9.86	***	4.65649	5.56	***
V_{18}	3.44887	6.39	***	10.82314	9.86	***	4.66564	5.56	***
V_{19}	3.43959	6.38	***	10.85556	9.87	***	4.6661	5.56	***
V_{20}	3.44031	6.38	***	10.87333	9.87	***	4.66328	5.57	***

V_{21}	3.44165	6.38	***	10.84575	9.87	***	4.63425	5.57	***
V_{22}	3.44362	6.39	***	10.7137	9.83	***	4.61469	5.56	***
V_{23}	3.46485	6.41	***	10.84151	9.87	***	4.61226	5.56	***
V_{24}	3.54811	6.48	***	10.80414	9.86	***	4.63071	5.56	***
R-squared	0.95756			0.68220			0.93162		
Adj.R-squared	0.95753			0.68194			0.93157		

Table 8. Panel data pooling estimation**. $PMO_{t-1,t,i} = \alpha + \beta^* WPF_{t-1,t,i} + \gamma^* DL_t + \phi^* NG_{t-1}/EUA_{t-2} + \omega WRt\text{-}1 + \delta PMO_{t-2,t-1,i} + \varepsilon_{t,i}$ (2). Pooling panel linear models with nonparametric robust covariance matrix estimators for panel models with cross-sectional and serial correlation ([20]). $PMO_{t-1,t,i}$ is the average supply price offered by the group of generators sharing the same generation technology submitted on a particular day (day t-1) for delivering electricity on the day-ahead (day t) during hour i; $WPF_{t-1,t,i}$ denotes the wind power forecast made public on day t-1 just before the deadline for submitting bids to the day-ahead market action for delivering electricity during hour i on day t; DL_t is a dummy variable that equals 1 if t is a business day, and 0, otherwise; NG_{t-1} is the British natural day-ahead trade close price on day t-1; EUA_{t-2} refers to the European emission allowances next December maturity futures closing prices on day t-2; WR_{t-1} denotes the hydroelectric water reservoirs on day t-1, and $PMO_{t-2,t-1},I$ is the average supply price offered by the group of generators the day-before (day t-2) for delivering electricity on the day-ahead (day t-1) during hour i. The stochastic component $\varepsilon_{t,i}$, which is the usual homoscedastic component, normally distributed $N(0,\sigma)$. Nuclear generators (CN) and renewables (CR). The significant codes (Sig.) are: 0 '' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.*

	CN			**CR**		
Parameter	**Value**	**t-statistic**	**Sig.**	**Value**	**t-statistic**	**Sig.**
α	6.90959	6.18	***	2.55307	4.49	***
β	-0.00002	-0.97		-0.00008	-7.68	***
γ	0.03202	0.24		-0.08613	-1.54	
ϕ	0.03684	0.90		-0.05055	-1.60	
ω	-0.01684	-2.66	**	-0.00819	-2.41	*
δ	0.72399	17.89	***	0.93633	80.35	***
R-squared	0.54326			0.91421		
Adj.R-squared	0.54320			0.91420		

Wind production forecasts (β)

The t-test for the significance of the coefficient that accompanies the WPF series (β) indicates that the estimated β parameter value is statistically significant for all generation technologies, implying that wind production forecasts have become relevant for all the supply market participants, with the only exception being CN plants which are shown to be indifferent with regards to expected wind production. Nuclear plants have the clear incentive to continuously generate electricity due to the high costs of stopping production.

It is worth emphasising that the series of WPF employed in the present study corresponds to the last hourly series made public by REE just before the deadline for submitting bids for the day-ahead market – since this is the most informative series. The underlying idea is that generators are expected to consider the wind power forecasts when designing their trading strategies and consider the most updated predictions just before that deadline. An a priori expected result is a negative value of the β parameter, given that the expected marginal price is supposed to decrease with increased wind production. Thereby, a logical reaction from bidders would consist in offering lower prices so as not to become unmatched, obviously without exceeding their own production costs. However, according to our results, the CT and CC plants would have been generally offering their production at higher prices when expecting increases in wind production, perhaps trying to compensate for the presumably lower income resulting from a less-likely required thermal and combined cycle production. The increase in the resulting marginal prices may be viewed as a wind premium for the risk of generating less electricity than usual, but at the same time, these generators would incur a risk of not being matched and so being dropped from the day-ahead auction.

Business-day dummy (γ) and effects by hour (vi)

From Table 7 and Table 8, the parameter value associated with the business dummy variable, γ, is significantly negative for CC and CT plants, whereas it is not statistically different from zero (at the 5% level) for CN, CH, and CR plants. The positive value of this parameter for CN plants implies that they would offer their production at higher prices for delivery hours on business days, when marginal prices are typically higher due to increased demand. However, the opposite holds for CC and CT plants. Therefore, these latter appear to bid at lower prices for business hours.

As mentioned before, the data panel model for CC, CT, and CH generation technologies enables differentiation for the specific effects of each hour. Thus, the results are subsequently enriched with the ve,i parameter value for each generation technology group e and hour i. Table 7 shows how it is found that CC and CT plants submit their bids to the market at higher prices during the early hours of the day (from the first to the sixth hour) when electricity demand levels are lowest. It is of note that during these low demand periods, electricity is usually generated by plants with the lowest marginal costs because bids from the remaining plants are normally too high to determine the marginal price and they drop out of the auction. The differences in the hourly bids from CH plants are, on the contrary, very small. Regarding CN and CR plants, as previously explained, the pooling model is shown to be more suitable, which leads us to conclude that there is no evidence of differences in strategic bidding behaviour between hours from CN and CR plants.

Natural gas / CO2 price ratio (ϕ)

The next step is to look at the effect of natural gas prices and carbon prices on the supply bids to the day-ahead market. Following previous literature, to avoid multicollinearity problems caused by correlation between explanatory variables, the ratio natural gas prices/carbon prices is chosen instead of considering these two series of prices separately, i.e. as individual independent variables. The natural gas and carbon price series selected are the corresponding international price benchmarks. The former corresponds to national balance point natural gas day-ahead prices on day t-1, whereas the carbon price series is the ICE ECX European emission allowances next December maturity closing futures prices at t-2. Regarding this latter price series, we use lagged prices because at the closure time of the day-ahead auction market which takes place at t-1 (for delivering electricity at t), the available closing prices are those of the previous trading session, i.e. at t-2.

Coming back to Tables 7 and 8, there is statistical evidence that the so-defined ratio does have an impact on the prices offered by CC, CT, and CH, according to the estimated ϕ parameter value, which is statistically different from zero. The resulting sign of this coefficient also offers interesting insights. Thus, it is positive for the CC and CH, meaning that the offered prices by these plant groups would be increasing with natural gas prices and/or generally decreasing with carbon emission allowance prices. Given that CC plants use natural gas as fuel to generate electricity, they are negatively affected by increases in natural gas prices. The plants are then expected to incorporate this information into their bids as an extra cost. As this type of plant is usually among those that set the marginal price, whenever they need to generate electricity to satisfy the overall demand an increase in marginal price would be expected. However, under these market circumstances, CH plants, being reasonably sure that their bids will still be lower than those from CC plants, may behave strategically and submit higher bids than usual and seek to profit from higher marginal prices set by themselves.

In contrast, the sign of the estimated ϕ parameter is negative for CT, indicating that they would offer higher prices when expecting lower natural gas prices and/or higher carbon emission allowance prices. This result is consistent with the fact that thermal plants, being the most pollutant technology, may have internalised the cost of paying for the carbon emission allowances into their bids and submit higher bid prices when expecting higher carbon prices.

Hydroelectric reservoir levels (ω)

Hydroelectric plants enjoy an important advantage since they can easily adjust their production to the amount needed. Of course, the electricity they can produce depends on annual rainfall, and more specifically, on the water reservoir levels. As only weekly data was available, each datum is repeated for seven daily periods. The way in which larger reservoirs can impact on prices is very similar to that of increased wind. In fact, more reservoirs would imply more capacity to produce electricity and more supply. In times of water reservoir excess, hydroelectric generators, with very low variable costs, can bid into the auction market at lower prices. The obtained results are consistent with that idea, with the only exception being the coefficient for the thermal generation plants. In particular, the estimated ω parameter value is significantly negative for all the generation technology plant groups, meaning that bid prices will decrease with hydroelectric water reservoirs, except for the CT plants for which the estimated ω value is shown to be significantly positive.

Lagged dependent variable (δ)

Finally, the parameter δ of the lagged dependent variable is strongly significant for all technologies. In other words, the bid price is strongly influenced by the same bid price submitted the day before for the same hour.

Summarising the results, firstly, it is evidenced that expected wind power production has become a new price determinant, since it has been shown to impact on generator bidding strategies –with the only exception of CN. Secondly, because of variations in the wind production forecasts, natural gas prices, carbon prices, and hydroelectric reservoirs, the CH, CN, and CR plants generally behave as expected and consistently submit bids with the aim of being matched in the day-ahead auction market.

Nevertheless, according to our results, CC and CT plants offer their production at higher prices when there is a larger supply of low-cost electricity, i.e. assuming a higher risk of dropping out of the day-ahead auction. It has been shown that CC and CT plants submit bids at higher prices: (i) when wind production forecast is larger; (ii) on non-business days; and (iii) from hour 1 to hour 7 – namely coinciding with low demand levels. CT plants also bid at higher prices when there is more water in the hydroelectric reservoirs.

4. Quantifying the impact of wind power forecasts on the marginal price

As shown in the previous section, supply bidders react to wind production forecasting in a different manner depending on the generation technology. Our aim here is to quantify the effect of this result on the marginal price level. To do so, we simulate the day-ahead hourly marginal prices by intersecting the actual (aggregate) demand curve and a fictitious (aggregate) supply curve built as follows:

$$\hat{P}_{a\epsilon e,t-1,t,i} = P_{a\epsilon e,t-1,t,i} - \beta_e WPF_{t-1,t,I} \qquad (3)$$

where $\hat{P}_{a\epsilon e,t-1,t,i}$ denotes the modified offered price; $P_{a\epsilon e,t-1,t,i}$ is the actual offered price submitted on dayt-1 by the market participant a of the generation technology group e for the hour i of the delivery day t; β_e is the estimated parameter obtained for the generation technology group e (displayed in Tables 7 and 8), and $WPF_{t-1,t,i}$ is the wind power forecast, known at day t-1, just before the deadline for submitting bids to the day-ahead auction market for hour i of day-ahead t.

The simulation exercise covers the whole of 2013 and consists of intersecting the modified aggregate supply curve and the actual demand to obtain the simulated marginal price, isolating the impact of the wind power forecast on the bidding behaviour of each generation technology, and ultimately, on marginal prices. A total of 24 x 5 marginal prices were simulated. Differences between actual and simulated prices are displayed in Table 9.

Table 9. Impact of the WPF on the marginal price*. Displays the difference between the actual marginal price and the simulated spot price, distinguishing by generation technology group. The daily difference is shown together with the difference per hour for the whole year 2013. Combined cycle generators (CC), thermal (CT), hydroelectric (CH), and renewable (CR).*

	CC	CT	CH	CR	TOTAL
Hour/daily average price	**0.06**	**0.50**	**-0.39**	**-0.22**	**0.16**
1	0.05	0.48	-0.45	-0.22	0.14
2	0.03	0.48	-0.38	-0.29	0.18
3	-0.17	0.26	-0.54	-0.50	-0.13
4	-0.18	0.31	-0.49	-0.50	-0.02
5	-0.10	0.34	-0.45	-0.46	0.04
6	-0.03	0.44	-0.38	-0.35	0.16
7	0.03	0.59	-0.25	-0.27	0.33
8	0.08	0.55	-0.33	-0.18	0.28
9	0.07	0.55	-0.31	-0.22	0.25
10	0.06	0.43	-0.45	-0.15	0.10
11	0.12	0.48	-0.40	-0.13	0.15
12	0.12	0.47	-0.40	-0.16	0.14
13	0.13	0.49	-0.42	-0.14	0.13
14	0.12	0.54	-0.35	-0.15	0.20
15	0.11	0.57	-0.36	-0.19	0.22
16	0.14	0.63	-0.33	-0.17	0.31
17	0.13	0.68	-0.33	-0.16	0.35
18	0.14	0.61	-0.32	-0.17	0.26
19	0.12	0.63	-0.39	-0.17	0.23
20	0.11	0.49	-0.41	-0.14	0.06
21	0.12	0.48	-0.43	-0.13	0.00
22	0.13	0.49	-0.43	-0.12	0.03
23	0.09	0.53	-0.40	-0.13	0.10
24	0.15	0.63	-0.28	-0.15	0.39

The average actual daily marginal price during 2013 was 44.05 euro/MWh, whereas the average simulated marginal price after removing the estimated effect of the wind power forecast on the offered prices by the thermal generation group according to the formula (3), amounts to 43.55 euro/MWh.

Therefore, the overall daily effect for the year 2013 that may be attributable to the thermal generation bids may be quantified, on average, at a minimum increase of 0.5 euros/MWh. Distinguishing between hours, the average increase in the marginal price oscillates between 0.68 euros/MWh in hour 17and 0.26 euros/MWh in hour 3 (Table 9).

The overall daily effect for combined cycle plants also means an increase in prices of 0.06 euros/MWh (a maximum of 0.15 euros/MWh in the hour 24 and a minimum of -0.18 euros/MWh in hour 4).

As the beta value for the remaining generation technology plant groups is statistically negative, the effect of their reaction to wind production forecasts leads to a decrease in the estimated marginal prices. On average, the decrease in the marginal price attributable to CH and CR generation groups bidding when reacting to the wind production forecasts, would have respectively been of -0.39 euros/MWh and -0.22euros/MWh.
In short, the overall effect of the generator bidding strategies linked to the wind production forecasts for the simulated sample period (year 2013) was an increase in average marginal prices. Hence, the impact of such strategies is shown to be large enough to overwhelm the well-known merit order effect of renewables.

5. Conclusion and concluding remarks

The purpose of this work is to analyse the way in which the bidding strategies by generators have been conditioned by expected wind production, among other key variables, in the Spanish electricity day-ahead market (which has experienced a continuously increasing proportion of wind power in the electricity generation mix).

To summarise the results: expected wind production is a new price determinant and is shown as relevant for the considered supply side participants when submitting their bids to the day-ahead auction market. Nuclear generators are the only exception – as they cannot afford to stop production and so lack the flexibility to maximise profits by bidding strategically.

The average prices offered by CC and CT plants for a given hour have proved to be systematically higher when expected demand levels are low for the studied sample(when delivery is taking place on a non-business day and for the first seven delivery hours of business and non-business days). These results can make sense for generation plants with low variable costs and enough flexibility (such as CR or CH plants) since it is in situations of low demand when they tend to submit less aggressive bids and profit from higher prices if they are successful in setting the marginal price. Anticipating that during these low-demand hours, the electricity produced by CN, CH, and CR generation plants will likely be sufficient to meet demand, the bids submitted by generators can push the marginal price upward. However, it is a priori difficult to establish why CC and CT, with much higher variable costs, bid higher prices when expected demand levels remain low, so incurring the risk of dropping from the auction.

According to the obtained results, the case of CT generation plants deserves a special mention, since they seem to submit bids at higher prices when: (i) the ratio natural gas/carbon price is expected to be lower, which is a logical result given that expected higher carbon prices may imply higher costs precisely for these generation plants; but also when (ii) wind production forecast is greater; and when (iii) there is more water in the hydroelectric reservoirs. Additionally, as mentioned above, their offered prices are higher, on average, for non-business days and from hour 1 to hour 7.

These results seem to lead to counterintuitive conclusions; however, it does not need to be so. As is generally known, the spot electricity market is made up of several sequential trading markets. Market participants submit their bids to buy or sell electricity for each of the 24 hours of the following day through the day ahead market, which is usually the most liquid market and whose price serves as the benchmark for forward contracts. However, given the nature of electricity, the result of the day-ahead 24 auctions must also be feasible from a technical point of view. In the Spanish case, it is the system operator (R.E.E.) who takes the responsibility for validating the technical viability of the day-ahead auction results, as well as for guaranteeing an annually fixed share of domestic coal for producing electricity to reduce external dependence within the supply security constraint regulation process. It is possible for market participants to rectify their previously open positions in the intraday market, which is a balancing market structured into six new consecutive auction markets. In addition, the system operator also manages several additional regulated markets to solve real-time deviations.

Therefore, the possibility of trading in sequential markets with different prices and/or the possibility of being required to produce electricity (to solve technical constraints, to assure a predetermined level of production by using domestic coal, or to guarantee system security in exchange for prices different from the resulting marginal price in the day-ahead market)may lead generators to coordinate bidding in the day-ahead, and subsequently, balancing regulated markets or processes.

Moreover, if these latter prices are systematically higher than the day-ahead market price then generators may prefer to hold back capacity in the day-ahead market to facilitate subsequent offerings in the next sequential markets, or other processes in which they can participate. The fact that market participants may consider the outcome of the sequence of markets, and not each market in isolation, was addressed by [24] for the Californian market. Also, for the Spanish case, [25] pointed out that some generation plants could have been submitting sale orders in the day-ahead market at high prices that would not be matched – and so that they would finally be required to produce electricity to solve congestion. This approach would have been more profitable according to the regulations in force during the period in question (from July 2004 to February 2005). [26] investigated the potential of coordinated bidding in the spot and balancing markets and concluded that significant profits could be made from such a coordination for the Nord Pool.
Concerning the results of the simulation exercise to quantify the effect of wind power forecasts on the day-ahead marginal price, the overall effect for the year 2013 was shown to be negative for electricity consumers, since the aggregated impact of the bidding strategies carried out by the CH and CR generation plant groups (offering lower prices for low-demand-level delivery periods) may have pushed marginal prices down (consistent with the merit-order effect of renewables) but did not overwhelm the increase in marginal prices produced by the effect of the bidding strategies implemented by the CT and CC generation plant groups.

The findings obtained in the present work are of interest to practitioners and regulators, given that they shed light on how the inclusion of renewable generation in the electricity market has altered the trading strategies of the supply market participants in the day-ahead market.

The strong presence of renewable generation in many power markets entails changes in power system planning, operating, and monitoring. These changes need to be considered and adapted to the operational processes. Conventional technologies such as thermal or combined cycle plants have been displaced by wind or solar generation that is characterised by flexibility and significantly lower variable costs of production. In addition, non-flexible conventional generation plants incur high operating and maintenance costs when starting and shutting down. In those cases, where storage capabilities such as hydro resources are insufficient, conventional generation plants may be useful inproviding operating reserves as a backup generation to manage the intermittency of renewable generation, or for guaranteeing system security by resolving output forecasting errors in the renewable generation models used for planning the day-ahead plant schedules.

The European Commission (2013) pointed out that the economic impact of balancing costs needs to be considered in well-designed renewable support schemes. For as long as renewable generation continues to be intermittent and wind output forecasts for periods other than very short-term are insufficiently accurate, then three(related)crucial issues will be: (i) assessing and revising the design of balancing rules considering the particularities of all the generation technologies involved;(ii) determining the amount of operating reserves needed to keep the power system functioning securely; and (iii) revising the way these backup reserves should be remunerated and providing the appropriate incentives as market signals to incumbent and new entrants that may lead to more efficient operations in the whole electricity market.

Our results suggest that some conventional generators react to the entrance of renewable generating sources by behaving in a somewhat strategic manner that contributes to pushing up marginal prices, and thereby creating a wind risk premium that should be considered by regulators when thinking about changes in regulation or market design to adapt to the new market situation. Furthermore, as the need for balancing power is expected to increase with the growth of fluctuating renewable production, an analysis of coordinated bidding in the day-ahead and subsequent markets, as well as in the provision of balancing services for the Spanish case, would also be of interest. This is left for further research.

Acknowledgements

The authors wish to thank Hipòlit Torró, Constantino J. García, and the participants at the 6th Research Workshop on Energy Markets for their valuable comments, as well as the financial support of the Spanish Ministry of Education and Science (project ECO2013-40816-P)

References

[1] Holttinen, H. The Impact of Large Scale Wind Power Production on the Nordic Electricity System. Helsinki University of Technology, 2004. Dissertation for the title of Doctor of Science in Technology.

[2] Sensfuβ, F.; Ragwitz, M.; Genoese, M. The merit-order effect: a detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany. Energy Policy, 2008, 36, 3086–3094.

[3] Saenz de Miera, G.; Del Río, P.; Vizcaíno, I. Analyzing the impact of renewable electricity support schemes on power prices: The case of wind electricity in Spain. Energy Policy, 2008, 36, 3345-3359.

[4] Green, R.; Vasilakos, N. 2010. Market behaviour with large amounts of intermittent generation. Energy Policy, 2010, 38, 3211-3220.

[5] Gelabert, L.; Labandeira, X.; Linares, P. An ex-post analysis of the effect of renewables and cogeneration on Spanish electricity prices. Energy Econ. , 2011, 33, S59-S65.

[6] Gil, H.A.; Gomez-Quiles C.; Riquelme, J. Large-scale wind power integration and wholesale electricity trading benefits: estimation via ex post approach. Energy Policy, 2012, 41, 849-859.

[7] Ballester, C.; Furió, D. Effects of Renewables on the Stylized Facts of Electricity Prices. Renew. Sustainable Energy Rev., 2015, 52, 1596-1609.

[8] Di Cosmo, V.; Malaguzzi L. Wind, Storage, Interconnection and the cost of Electricity Generation. Fondazione Eni Enrico Mattei and Economic and Social Research Institute, 2017.

[9] Würzburg, K.; Labandeira, X.; Linares, P. Renewable generation and electricity prices: Taking stock and new evidence for Germany and Austria. Energy Econ., 2013, 40, S159-S171.

[10] Weidlich, A.; Veit, D. A critical survey of agent-based wholesale electricity market models. Energy Econ., 2008, 30, 1728–59.

[11] Gountis, V.P.; Bakirtzis, A.G. Bidding strategies for electricity producers in a competitive electricity marketplace. IEEE Trans Power Systems, 2004, 19, pp. 356–65.

[12] Veit, D.J.; Weidlich, A; Krafft, J. An agent-based analysis of the German electricity market with transmission capacity constraints. Energy Policy, 2009, 37, 4132–44.

[13] Twomey, P.; Neuhoff, K. Wind power and market power in competitive markets, Energy Policy, 2010, 38, 3198-3210.

[14] Jónsson, T.; Pinson, P.; Madsen H. 2010. On the market impact of wind energy forecast. Energy Econ., 2010, 32, 313-320.

[15] Li, G.; Shi, J. Agent-based modelling for trading wind power with uncertainty in the day-ahead wholesale electricity markets of single-sided auctions. Appl. Energy, 2012, 99, 13-22.

[16] Ciarreta, A.; Espinosa, M. Supply Function Competition in the Spanish Wholesale Electricity Market. The Energy Journal, 2010, 31, 4, 137-157.

[17] Green, W.H. Econometric Analysis. Fifth Edition. 2002, Prentice Hall

[18] Hall, S.; Urga G. (2000) New Developments in the Analysis of Panel Data Sets. The Current State of Business Disciplines. Business Economics, 2000, 2, 537-564.

[19] Maddala GS; Wu S. A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. Oxford Bulletin of Economics and Statistics, 1999, 61 (Supplement 1), 631–652.

[20] Driscol J.; Kraay, A. Consistent Covariance Matrix Estimation With Spatially Dependent Panel Data. The Review of Economics and Statistics, 1998, 80- 4, 549-560.

[21] Cameron, A.; Miller, D. A practioner's Guide to Cluster-Robust Inference. Human Resources 2015, 317-372, doi:10.3368/jhr.50.2.317.

[22] Arellano, M. Computing Robust Standard Errors for Within-groups Estimators. Oxford Bulletin of Economics and Statistics, 1987, 49, 4.

[23] Cameron, A.; Gelbach, J.; Miller, D. Bootstrap-Based Improvements for Inference with Clustered Errors. The Review of Economics and Statistics, 2008, 90- 3, 414-427.

[24] Borenstein, S.; Bushnell, J.A.; Knittle, C.; Wolfram, C. Inefficiencies and market power in financial arbitrage: a study of California's electricity markets. Center for Study of Energy Markets. University of California Energy Institute. Working Paper, 2004, 138.

[25] Furió D.; Lucia, J. Congestion management rules and trading strategies in the Spanish electricity market. Energy Econ., 2009, 31, 48-60.

[26] Boomsma, T.K.; Juul, N.; Fleten, S-E. Bidding in sequential electricity markets: The Nordic case. European Journal of Operational Research, 2014, 238, 797-809.

Capítulo 3:

Capítulo 3:

Analysing the impact of Renewables on Spanish Electricity Markets using Machine Learning Techniques

Abstract

After the introduction of wholesale electricity markets to promote competition in the electricity liberalisation process undertaken by many countries in recent decades, the other great challenge to address is the integration of the expanding renewable generation into electricity systems. Of course, the entry of renewable power into the mix has significant effects on prices. While much of the previous literature has focused on the impact of renewable energy on day-ahead market prices, our aim is to extend the analysis to electricity final prices. To do so, we apply machine learning techniques that allow us to uncover their main drivers by analysing an exhaustive dataset of variables. Factors influencing the components of electricity final prices, other than the day-ahead market price, are much less studied, though they are key to gain insight into the dynamics between the interrelated trading segments and a-priori technical processes included in the wholesale electricity market. We expect our results will be of interest to both practitioners and regulators, as they will provide a better understanding of the functioning of the market and have implications in the restructuring of the market towards a more sustainable and competitive electricity system.

1. Introduction

After the electricity liberalisation process undertaken by many countries in recent decades, which in most cases entailed, among other measures, the establishment of wholesale electricity markets to promote competition, the other great challenge to address is undoubtedly the integration of renewable generation with the aim of achieving a sustainable electricity system as a base to support the transition to a low-emissions economy.

Supporting the increased use of clean power (mainly wind and solar) is a key part of energy policy all over the world. As a result, structural changes to electricity systems are occurring. One of these changes undoubtedly is the impact of high penetration of renewable sources on electricity prices.

While most of the previous papers focus on the impact of renewables on day-ahead market prices (the so-called spot prices), we extend the analysis to study the impact on final electricity prices, which include not only the day-ahead price but also the costs incurred in the subsequent processes until the real time delivery of electricity, by identifying the main determinants of each of their cost components.

Given the day-ahead system-marginal-price auction and the lower generation cost of renewable sources, such as wind and PV-solar, renewable generators can bid at very low prices or even bid in at zero, participating as price takers in the day-ahead market. A large enough number of low-price bids from renewable players can shift the supply-offer curve in such a way that the resulting auction price is set at a lower level. As a result, more renewable production is expected to translate into lower resulting prices. This is the so-called *merit order* effect of renewables, which refers to the reduction in day-ahead market prices due to the introduction of renewables into the electricity system and is well documented in the literature [1] in the Nord Pool, [2] and [3] in the German market, [4] in the Australian market, and [5] or [6] in the Spanish market, among others.[20]

[20]For a complete overview of past research on the merit-order effect of renewables see [7], and recently in the Iberian Market see [8].

However, the intermittency of the main renewable energy sources, together with the non-storability (at a large scale) of electricity, may entail higher market balancing needs and costs. As electricity is a non-storable commodity, its delivery must be scheduled in advance. Later, adjustments are normally needed to find a solution for unexpected deviations from that scheduled. Thereby, the wholesale electricity market is usually composed of a series of interrelated markets: the day-ahead market, which is the main market for physical delivery; the intraday market for more short-term adjustments; and, finally, balancing markets to handle the remaining deviations and other technical issues. The variability and limited predictability of renewable generation, especially wind (which depends on wind velocity and direction), could increase the need for balancing in order to ensure the electricity supply at the delivery moment. Therefore, more renewable production could lead to more balancing needs, and hence more balancing costs, which could in the end drive up final electricity prices. In this regard, [9] find a significant positive difference between real-time and day-ahead market prices, particularly for wind electricity for the Italian market.

As a consequence of the increasing share of renewable energy sources (RES), the design of electricity markets is currently being revisited A number of papers have examined the ability of different market designs to respond to a significant rise in the introduction of RES. [10] studies the behaviour of total physical adjustment capacity to show that only with small wind capacity, total adjustments are dominated by conventional (non-renewable) energy sources. Furthermore, the work highlights the relevance of the flexibility and liquidity in intraday markets to adjust previously assumed positions, recommending the Spanish intraday market design as the most effective way to manage the adjustments. The attractiveness of intraday markets as a mechanism to set balancing prices, avoiding or (at least) reducing the need for more expensive procedures, has captured the attention of other research papers. [11] compare different electricity market designs to determine whether they provide adequate flexibility to deal with intermittency, finding that the finite auction design of the Spanish intraday market has a better performance than the alternative continuous trading system, though the number of trading sessions remains an open question. [12] evaluates the benefits for wind producers to trade in intraday markets and concludes that participating in intraday markets could be a non-

optimal choice for market participants under some circumstances, which may generate additional costs for the system.
Logically, the growth in renewable power generation reduces the need for production from conventional sources and, as previously mentioned, due to the merit-order effect, also reduces day-ahead prices. Thus, conventional-source power plants tend to produce less and, simultaneously, the price received for their production is generally lower than it used to be before the growth in the penetration of RES. In this new context, the survival of these power plants may be in jeopardy. In the Spanish market, those conventional-source facilities that are flexible enough have been receiving capacity payments to compensate them for their lower revenues (and ultimately to prevent them from closing) in exchange for acting as a backup to renewables when needed. This mechanism entails an additional cost that may in the end contribute to pushing up final prices.

Furthermore, the adaptation of these conventional power plants to the new context should also be considered. Thus, strategic bidding behaviour consisting in avoiding the day-ahead market in order to force the sale of the production through other "*a priori* technical" market segments (because these generally end up being more profitable) has previously been documented in the literature ([13], [14], [15] or [16]). [13] identify some generating plants submitting their sale orders to the Spanish day-ahead market at anomalously higher prices during the period July 2004 to February 2005, in order to remain out of the auction results and be able to produce their electricity to solve congestions, which was remunerated at higher prices. [14] also detect strategic bidding behaviour by some thermal power plants using domestic coal under a particular regulation (Royal Decree 134/2010) which was in force from 2010 to 2015. This procedure provided priority dispatch to domestic coal production in the day-ahead market and required a minimum coal share in the Spanish generation mix. To avoid being matched in the day-ahead market auction and instead engage in later balancing processes with a remuneration higher than the day-ahead marginal price, some thermal plants engineered their participation in these latter processes to obtain higher remuneration. Additionally, [15] provide evidence of counter-intuitive behaviour by thermal generators, showing that at times of higher wind power forecasts, thermal generators systematically increase their offer prices, incurring the risk of not being matched in the day-ahead market auction. Therefore, under some circumstances, thermal agents appear to be incentivised to withdraw production from the day-ahead market to operate in subsequent markets and procedures that begin after it closes, i.e. intraday and/or balancing markets and processes. [16] estimate more gains for a flexible plant (hydropower) with coordinated bids considering both markets, the day-ahead and the subsequent balancing market, in the Nordic Market. The key lies in being able to anticipate balancing market opportunities before the day-ahead market closure, which requires an accurate forecast of balancing volumes and balancing pricing, still a difficult task, but potentially a very profitable one. The consequences, as the authors point out, could lead to a supply curve transformation, generally with higher prices at the upper end of the bid curve.

The main goal of this paper is to study the impact of renewable sources on electricity final prices by means of an exhaustive analysis of each of their components. These price components, other than the day-ahead market price, mainly respond to intermediate processes between that market and the real-time delivery of electricity, designed to guarantee continuous supply and system reliability. The Spanish market is chosen as a paradigmatic example due to its high level of renewables as well as the particular design of its intraday market, which has been underlined in previous literature, as indicated above. Complementarily, the analysis made allows us to disentangle the main drivers of electricity final prices, taking a step beyond the study of electricity day-ahead market prices.

This paper is also a great opportunity to explore the dynamics between different trading segments in the wholesale electricity market through the analysis of final electricity cost components. To do so, a complete dataset with a huge volume of predictor variables (264 variables) is generated. To deal with such an exhaustive list of variables, we use machine learning techniques and, particularly, regression trees, since they allow us to examine the interactions between markets in an attempt to clarify whether what happens in one market has an effect on the subsequent ones, while handling a large amount of data, most of them highly correlated, not necessarily linearly. The extensive dataset generated obliges us to opt for these models to the detriment of other classical parametric models that impose assumptions that our series do not meet. Machine Learning Algorithms (MLA) are a quite recent addition to this branch of literature. [17] evaluate the use of the machine learning techniques (random forest and neural networks) to forecast spot electricity prices, obtaining competitive results using demand and supply curves as the input of the models.

[18] summarise the advances of MLA in agent-based modelling of energy markets. The most common subject to forecast is electricity market prices, but load or renewable generation are also predicted in recent studies.

The rest of the paper is structured as follows. Section 2 briefly describes the Spanish electricity system. Section 3 presents the cost components included in the final electricity price. Section 4 lists the dataset used. Section 5 is devoted to present and discuss the results after identifying the final electricity price components using Machine Learning Techniques. Finally, section 6 summarises the main results and concludes.

2. The Spanish Electricity Market

The Electricity Sector Law 54/1997, of 27 November 1997, marks the beginning of the electricity industry liberalisation process in Spain. Later, in 2007, the Spanish and the Portuguese electricity systems were integrated into a common market area, the Iberian Market. This paper focuses on the Spanish market area, and thus the data used refers exclusively to the Spanish area.

The management of the liberalised market in the Spanish System is divided into two areas: economic management, which is assigned to the Market Operator, (OMIE)[21], and technical and transport network management, which is assigned to the System Operator, (REE).[22] The wholesale electricity market is composed of: (i) day-ahead market, (ii) intraday market and (iii) balancing markets.[23]

The day-ahead market is a daily uniform price auction managed by the Market Operator in which the participants submit their bids to purchase or sell electricity for the 24 hours of the following day. The resulting price for each specific hour is determined by the point at which the supply and demand curves meet, according to a marginal pricing system. The intraday market allows decisions to be corrected on the day-ahead market using more

[21] OMIE-POLO ESPAÑOL, S.A.
[22] Red Eléctrica Española S.A.
[23] Additionally, physical and financial bilateral contracts are traded, but they are beyond of the scope of this study.

updated and accurate forecasts. It is composed of six consecutive auction sessions, each of them involving several scheduled periods closer to the delivery date.[24]

In a first step, the resulting marginal price and traded volume of the auctions in both the day-ahead and the intraday markets are obtained based on merely economic criteria. Subsequently, it is necessary to ensure that they are also technically feasible. The System Operator is the entity responsible for their validation from a technical perspective, through the so-called management of the system's technical constraints. Thus, network capacity is analysed to determine whether it is sufficient to accommodate demand, given that the electricity flows from the generation plants to the consumption points under conditions that should be sufficiently reliable. As a consequence, the day-ahead and intraday market auction results are just preliminary and can be altered.

In addition, there are other adjustment processes or balancing markets to ensure the operation of the system that the Spanish System Operator is responsible for: (i) the Additional Upward reserve power market mechanism, whose purpose is to provide the system with the estimated necessary level of upward power reserve; (ii) the Secondary Control Band, designed to maintain the generation-demand balance by correcting deviations in temporary action horizons ranging from 20 seconds to 15 minutes; (iii) Tertiary Control, to resolve the deviations between generation and consumption and the restoration of the secondary control band reserve used; and (iv) real-time deviation management processes.

Lastly, there are still two management tools that the System Operator uses to manage imbalances: capacity payments and the interruptibility service. Capacity payments are paid by final consumers to ensure the availability of sufficient generation capacity to meet the demand for electricity at any time. The increasing share of renewables in the electricity system motivated their inclusion as an extra cost, due to the intermittent nature of some renewable sources, specifically wind, and to the reduction in generation by conventional power plants, which are being progressively replaced by renewable sources and may not be able to cover costs, to the point of being forced to close. Given the above, some flexible conventional power plants started to be remunerated by the system through

[24] The details of the opening and closing times of each session of the intraday market can be consulted on the Spanish Market Operator website (www.omie.es). *Last accessed: March 2019*

the capacity payments mechanism, for merely being available, thereby serving as a backup to renewable electricity sources. The interruptibility service, on the other hand, in force since 2015, is provided by some authorised large power consumers by reducing their consumption (when required by the System Operator) to maintain the balance between generation and demand during periods when demand exceeds supply.

Consequently, final electricity prices include several costs other than the day-ahead market price (its main component). This deserves a closer look in order to disentangle the effect of renewables, among other potentially key variables, on this final electricity price, by means of analysing their effects on each of its components.

To summarise, the components of the final electricity price are the following: (i) the day-ahead market price, (ii) the cost resulting from solving technical constraints, (iii) the intraday market price, (iv) the cost arising from the processes related to ancillary services and deviation management, (v) the capacity payments and (vi) the cost stemming from the interruptibility service.

3. Data

The data used are the price series of the components of the Spanish liberalised market final electricity prices, at an hourly frequency, from April 2012 to April 2018.[25] They are all expressed in €/MWh and are publicly available on the Spanish National Commission on Markets and Competition website.[26] In particular, each component refers to: (i) the average hourly price series of the day-ahead market (DM, from now on); (ii) the average hourly price series of the intraday market component (IM), which captures the net effect on the final price of the six sessions of the intraday market; (iii) the average hourly net effect on the final price of the procedure to solve technical constraints (TTCC), which includes the costs incurred to handle technical constraints after the day-ahead market auction, after each of the intraday markets auction, and in the real-time market; (iv) the

[25] The Spanish liberalised market involves 89% of the total generated energy.
[26] www.cnmc.es (Final Price Free Market (*COM). File: PFMHORAS_COM_yyyymmdd_yyyymmdd). Last accessed: April 2019

average hourly cost resulting from ancillary services and deviation management (SO); (v) the average hourly cost related to capacity payments (CP) and, finally, (vi) the average hourly cost associated with the interruptibility service (IS).

Additionally, we also use the hourly series of bids (price and amount of power) individually submitted by market participants in order to buy or sell energy, distinguishing between matched and non-matched bids, both in the day-ahead and in the first session of the intraday market, since this session is the one in which most of the intraday market liquidity is concentrated ([19]) and for the same sample period. The entire supply and demand curves can be found on the OMIE website.[27] Other energy price series from April 2012 to April 2018 are also included in the analysis: (i) the Dutch TTF (Title Transfer Facility) futures prices[28], which have become the natural gas benchmark in Europe ([20]); (ii) the API 2 index for the coal price[29] and (iii) the European Emission Allowances (EUA) futures prices.[30] Finally, data from the interconnexion France-Spain are also included: the percentage of hours with 100% use, two series (both sides), and the day-ahead spread (Spanish day-ahead market price minus French day-ahead market price).[31]

4. Machine Learning Estimation

To start, each of the components of the electricity final price, other than the day-ahead market price, is chosen as a target of each machine learning model: IM (the net cost of the intraday markets), TTCC (the cost of the system technical constraints together with the cost of the transitory promotion of domestic coal), SO (the cost of the rest of the balancing processes managed by the System Operator), CP (capacity payments) and IS (the cost of the interruptibility service).

A total of 264 variables are used as predictors for each one of the models (Table 1). The variables can be grouped as follows: (i) variables generated from the day-ahead market

[27] www.omie.es: (file:curva_pibc_yyyymm.zip, curva_pbc_yyymm.zip and pdbc_stota_yyyymmdd).
[28] Source: Thomson Reuters database
[29] Source: Thomson Reuters database
[30] These series are available at www.investing.com.
[31] These series are available on the IESOE (Electricity Interconnection in South–Western Europe) webpage (www.iesoe.eu).

data: the day-ahead market price; the total energy matched in the day-ahead market; and the mean offer price to sell electricity grouped by technology: combined-cycle plants (denoted by CC), thermal (CT), hydroelectric (CH), nuclear (CN) and renewables (CR), mainly wind and solar[32], as well as the share of the electricity matched in the day-ahead market auction over the amount of electricity offered to be sold, also aggregated by generation technology. In this latter case, we add a group of offers that are not generators (OO), but mainly commercial units, self-production or pumping units, among others; (ii) variables generated from the first session of the intraday market data: mean offer prices to sell electricity; mean offer prices to purchase electricity; mean offer prices to sell or purchase electricity, grouped by generation technologies; (iii) balancing costs and regulated payments, which are the previously mentioned components of the electricity final price (TTCC, SO, IM, CP and IS); (iv) other commodity prices that are expected to be determinants of electricity prices, such as natural gas prices (TTF), coal prices (API2) and carbon prices (EUA); (v) data from the interconnexion France-Spain; and, finally, (v) calendar variables to control for seasonality (day of the week, monthly and yearly).

Table 1. Predictors for the models

Source	Variables	Denoted by:
Day-ahead market	The spot price	DM_Price
	Total power matched	DM_E
	Mean offer price to sell by combined-cycle power plants	DM_OPSELL_CH
	Mean offer price to sell by hydroelectric power plants	DM_OPSELL_CH
	Mean offer price to sell by nuclear plants	DM_OPSELL_CN
	Mean offer price to sell by renewable plants	DM_OPSELL_CR
	Mean offer price to sell by thermal plants	DM_OPSELL_CT
	Mean offer price to sell by no generator agents	DM_OPSELL_OO
	Share of power sold by combined-cycle generators	DM_PCT_CC
	Share of power sold by hydroelectric plants	DM_PCT_CH
	Share of power sold by nuclear generators	DM_PCT_CN
	Share of power sold by renewable generators	DM_PCT_CR
	Share of power sold by thermal generators	DM_PCT_CT
	Share of power sold by no generators	DM_PCT_OO

[32] This category also includes bids coming from cogeneration and surplus production, but these latter bids are really of minimal importance because of their relatively limited associated volume.

Spread	The difference between Spanish Spot Price and French Spot Price	SpreadESFR
Intraday Market Session 1	Mean offer price to purchase Mean offer price to purchase by combined- cycle plants Mean offer price to purchase by hydroelectric plants Mean offer price to purchase by nuclear plants Mean offer price to purchase by renewable plants Mean offer price to purchase by thermal plants Mean offer price to sell Mean offer price to sell by combined-cycle plants Mean offer price to sell by hydroelectric plants Mean offer price to sell by nuclear plants Mean offer price to sell by renewable plants Mean offer price to sell by thermal plants	IM_OPPURCHASE_ IM_OPPURCHASE_CC IM_OPPURCHASE_CH IM_OPPURCHASE_CN IM_OPPURCHASE_CR IM_OPPURCHASE_CT IM_OPSELL IM_OPSELL_CC IM_OPSELL_CH IM_OPSELL_CN IM_OPSELL_CR IM_OPSELL_CT
Final price cost components	The capacity payment cost component The interruptibility service cost component The system operator processes cost component The technical constraints cost component The intraday market cost component	CP IS SO TTCC IM
Interconnection capacity	Percentage of hours with 100% use Spain -> France Percentage of hours with 100% use France -> Spain	Phu_ESFR Phu_FRES
Commodity prices and Calendar variables	Month of the year: 11 dummies (January excluded) Year: 5 dummies (Year 2011 excluded) Day of week: 6 dummies (Sunday excluded) TTF Dutch gas prices Carbon Prices API2 coal Prices	February…, December 2012…, 2018 Monday,..., Saturday TTF carbon API2

It should be highlighted that the variables from the day-ahead market can be predictors for the same day of the target, whereas the rest of variables are lagged. Moreover, all variables in the dataset are lagged seven days in order to test the relevance of autoregressive effects.

4.1 Methodology

We use regression trees, a machine learning algorithm, to estimate each of the five models associated with each cost component of the final price. The algorithm is based on a recursive partition of the feature space represented by a tree growing. The starting point is a root node, which is the space containing all observations. The space is divided into regions and the target is modelled in a simple way, for instance, as the mean of each region. The split-point that allows the space to be divided into regions is the one with the best fit (the one with the lowest estimation error), namely, the one that shows different separation conditions (for example, day-ahead price above 50MWh). Each split-point drives to a new node (or sub-region), called a leaf, and then new branches are derived from these until a stop criterion is applied (usually, the sub-region minimum size or the maximum number of split-points).

There are different versions of the algorithm. The most basic version is known as CART. The root node is split into two leaf nodes considering the following criteria: taking a set of predictors $\{X_1, X_2, ...X_p\}$, the goal is to select one of them, Xj, and the split-point c to obtain two sub-regions: $R1=\{X|X_j<c\}$ and $R2=\{X|X_j>=c\}$ in such a way that the following measure is minimised:

$$RSS_T=RSS_1+RSS_2=\sum_{i:xi\in R1}(y_i-\hat{}y_{R1})^2 + \sum_{xii\in R2}(y_i-\hat{}y_{R2})^2 \quad (1)$$

where y_i denotes the target observed for the region R_i; $\hat{}y_{Ri}$ is the estimated target (the mean) for the region Ri; and RSS_i refers to the residual sum of squares for the region R_i. In this way, the partition that minimises the total residual sum of squares is chosen.

It should be noted that this is a non-parametric procedure, which has interesting advantages since it allows us to handle non-normal data or multicollinearity. In addition, it is also robust even if there are outliers or missing values. Nevertheless, this first version of the algorithm does exhibit some drawbacks that should be pointed out, such as less accuracy in prediction compared to other techniques, a high variance in the outcomes, and a tendency to overfit. To overcome these drawbacks, improved algorithms have been introduced. Among them, the most popular are Random Forest and Boosting Methods.

The Random Forest technique reduces the variance by estimating more trees and using bootstrap. The procedure is simple. First, different samples with different sets of predictors are generated with bootstrap. Second, a regression tree is fitted in each of the samples. Finally, the mean of the predictions using all the trees is the definite prediction of the target.

The Boosting technique increases the accuracy of the result by fitting a chain of multiple trees, taking as starting point in each step the residues of each previous tree. The procedure involves first selecting a loss function to be optimised, for example, the Root Mean Squared Error (RMSE). Then, one tree regression is fitted, followed by a sequence of tree regressions that are also fitted to explain the errors of the previous tree model. Finally, a gradient descent procedure is employed to add the outputs, minimising the loss function.

A priori, it is not known which one will be the best method for our dataset. Therefore, we apply the three methods: the most basic version (CART) and the other two more advanced ones (Random Forest and Boosted Regression Trees) to test which one explains and predicts the targets better. The dataset used to estimate the models include 591,360 items (2,240 observations for each of the 264 variables covering the period from 1 April 2012 to 19 April 2018), except for the case of the interruptibility service cost component (IS) model, which is shorter, with 315,480 items (1,195 observations for each of the 264 variables covering the period from 8 January 2015 to 19 April 2018).

The procedure consists of several steps. Firstly, the sample is randomly split into two samples: a training sample, which contains 70% of the total sample, devoted to training the algorithm; and a test sample, which uses the remaining 30% of the sample to evaluate the predictive power of the model. Secondly, each target is estimated using each of the three versions of the tree-based methods: CART, Random Forest and Boosting. Thirdly, the performance is evaluated, both in the training and the test sample, to select the most suitable version for our dataset. To test the market performance, the mean absolute error (MAE) and the root mean squared error (RMSE) are used. The next step is to extract the relevant variables, based on the returns of the RMSE. Lastly, the relationship between the relevant variables and the target is explored in order to obtain the marginal effect of each variable on the outcome of each of the models.[33]

[33]The software used are the R packages rpart, Random Forest and xgboost. This software also provides feature importance measures. [21], [22], [23], [24]

4.2 Parameters

The CART method needs an additional process, the so-called pruning method, to prevent overfitting. The aim here is to reduce the number of branches, eliminating those that are not contributing to predicting and may be causing overfitting. To identify the optimal size for a tree, a tree pruning method using cross validation (CV) is used, following [25]. Through this method we, first, leave out one observation for training and, second, use the resulting model to predict the observation that has been left out. However, a full leave-one-out cross-validation is more costly computationally, so it is better to work with k-fold-cross-validation and cost-complexity function in order to reduce the number of fits required. A cost-complexity function for trees is CC (tree) = $\sum RSS_i + \lambda$, which is the sum of squared residuals of all terminal nodes plus λ, where parameter λ is the number of terminal nodes. In practice, the parameter CP (cost complexity) is used, computed as CP $=\lambda/RSS$, with RSS being the sum of squared residuals in a tree with no branches. Finally, the pruning strategy consists of growing a large tree and then pruning it back, considering the smallest sub-tree with a CV error within one standard error of the minimum.[34]

To apply the Random Forest technique, the following parameters are used. First, the number of trees should not be set too small to ensure that every input row gets predicted at least a few times to obtain more stable outcomes. In this study, the number of selected trees is 500. In the splitting process, the variables are selected randomly to prevent overfitting. The default number of variables to work with in regression trees is p/3 in each step, where p is the number of predictors (264 in this case). Next, for the boosted trees method, the RMSE is selected as the loss function. The level of the parameter *eta*, the learning rate control, must be decided because a low value for this parameter prevents overfitting, but a very low value would imply lower computing. In this study, the level is chosen by trial and error to be 0.05. To avoid overfitting, the subsample ratio is set at 0.5, which means that just half of the sample is used for growing trees, and the value of the column sample by tree is 0.85, implying that 85% of the variables are considered when building each tree.

---

[34]The rest of the parameters used are: CP = 0.01 in the initial tree without pruning; MinSplit=20, the minimum number of observations in a sub-region to be split; MinBucket=20/3 (default value); xval=10, the number of cross-validations; and maxdepht=30, the maximum number of levels in a tree.

__4.3__ Results and Discussion

The metrics of the performance obtained for each target and version are shown in Table 2.

Table 2. Performance Metrics. *MAE (Mean Absolute Error) and RMSE (Root Mean Squared Error)*

Target	Algorithm	MAE		RMSE	
		Training	Test	Training	Test
IM	CART	0.0693	0.0773	0.1092	0.1197
	RAND.FOREST	0.0259	0.0739	0.0424	0.1130
	XGBOOST	0.0378	0.0749	0.0510	0.1140
TTCC	CART	0.4859	0.6501	0.6503	0.9289
	RAND.FOREST	0.1897	0.4984	0.2782	0.7225
	XGBOOST	0.2085	0.4706	0.2817	0.7030
SO	CART	0.4336	0.4605	0.6514	0.7420
	RAND.FOREST	0.1447	0.3741	0.2590	0.6576
	XGBOOST	0.1841	0.3708	0.2663	0.6471
CP	CART	0.1852	0.2426	0.3618	0.5273
	RAND.FOREST	0.0741	0.2071	0.1520	0.4509
	XGBOOST	0.0841	0.1898	0.1394	0.4351
IS	CART	0.0100	0.0102	0.0230	0.0239
	RAND.FOREST	0.0043	0.0105	0.0122	0.0259
	XGBOOST	0.0110	0.0136	0.0177	0.0240

From Table 2, it is confirmed that the CART method, the simplest one, has the worst performance, since it provides the highest error metrics (MAE and RMSE) for all the targets. In addition, in spite of the precautions taken in the estimation process, there is some evidence of overfitting. The error metrics, in general, are a bit higher for the test sample than for the training sample, which is indicative of overfitting problems. As we are interested in identifying the relevant factors that would be able to explain the targets, a good performance in the training sample without too much overfitting may be considered sufficient for our purpose. It can be observed that the Random Forest method overall meets the requirements and so this is the method used to carry out the analysis. It should also be highlighted that the best performance achieved corresponds to the interruptibility service and the intraday market, given that their error metrics are the lowest (<0.045).

Once the model has been estimated using the Random Forest method, we have 264 variables to be analysed. A usual way to proceed with machine learning models is to use an importance measure, which allows us to identify which variables are the most relevant ones. We use the increments in percentage of the root mean squared error, which is a commonly provided measure by the Random Forest statistical packages.[35] The method to measure the importance variable is based on the idea that the accuracy of a model will fall drastically if an important variable is altered, and it would not fall, or would fall by a lesser degree, if this variable was unimportant (see [26] and [27]).

Accordingly, we proceed as follows. The values of each variable are permutated at random. Using an out-of-bag portion of data (a portion of data saved which is not used to train the tree model)[36], the prediction accuracy is calculated before and after value permutation to check whether there are changes in the accuracy of the estimation. The differences are averaged and normalised by the standard error. If the standard error is equal to 0 for a variable, the division is not done (the measure is almost always equal to 0 in that case). Finally, the increments in percentage of the root mean squared error are computed and the variables are ranked.

Table 3 shows the ranking for each model. The measure of the importance variable (the increment in the percentage of the root mean squared error, denoted by %) notably decreases in the first positions of the ranking. From a visual inspection of the graphical representation of this measure for each model[37], we limit the number of relevant variables to ten, since overall, this appears to be sufficient to capture the most relevant factors in the models.

[35] We use the function varimp from package RandomForest in [23].
[36] The importance variable estimations have been calculated according to [26] using the out-of-bag method. Before each tree is constructed, the training set is divided with bootstrapping into two samples; one is used to construct the tree and the other, the out-of-bag portion, is saved internally to estimate importance variable measures. The tree is run with the out-of-bag examples twice, once with the values of the variables intact and once with the values of the variables permutated at random. The differences in accuracy obtained are used to obtain the measure of the importance variable.
[37] Figures are available upon request by the authors.

Table 3. Importance variable measure. *The increment in the percentage of the root mean squared of each variable (%).*

TTCC (Technical Constraints)		SO (System Operator Processes)		IM (Intraday Cost)		CP (Capacity Payments)		IS (Interruptibility Service)	
variable	%	variable	%	variable	%	variable	%	variable	%
TTCC_lag1	45.14	SO_lag1	24.35	IM_Price_lag1	17.99	CP_lag7	40.49	IS_lag1	21.46
DM_E	22.58	DM_Price	13.38	DM_PCT_CT	10.43	CP_lag1	25.26	2018	17.44
DM_Price	19.44	DM_PCT_CR	11.63	IM_Price_lag7	9.09	August	24.64	IS_lag2	14.43
DM_PCT_CC	18.19	DM_PCT_CT	10.28	DM_Price	8.67	March	19.2	April	13.21
TTCC_lag7	17.96	SO_lag6	8.97	DM_PCT_CH	8.04	CP_lag6	16.69	September	12.46
DM_PCT_CT	17.56	CP_lag5	8.19	TTF_lag4	7.97	Saturday	16.53	IS_lag3	11.84
TTCC_lag2	15.05	SO_lag7	7.97	TTF_lag3	7.54	CP_lag3	14.63	November	10.52
DM_PCT_CR	13.47	DM_PCT_CC	7.54	SO_lag7	7.39	CP_lag4	14.25	July	9.53
Sunday	11.23	DM_E	7.03	TTF_lag2	7.23	DM_E	13.76	API2_lag3	9.14
DM_E_lag7	10.36	DM_OPSELL_CC	6.81	DM_OPSELL_CR	6.86	CP_lag2	13.4	API2_lag6	8.8

Additionally, the analysis is completed with the accumulated local effects plots (ALE plots) of each relevant variable (Figs 1 to 5). These plots are made to see the mean effect of the variable at a certain value compared to the average prediction of the data. In the abscissa axis we see the values of the variable, whereas in the ordinate axis we see the estimated local effect following the ALE method.[38] The estimated local effect is centred. Thus, for example, a negative (positive) ALE estimation value equal to -2 (+2) in the ordinate axis at x=30 in the plot indicates that the predicted value will be lower (higher) than two times the average of the dependent variable. The magnitude of the estimated values allows us to rank the variables in importance. Therefore, by plotting the estimated local effects the relationship between each predictor and the target can be seen.

4.3.1 Technical constraints cost

Fig. 1 shows the ALE plots for the technical constraints cost component corresponding to its top ten relevant significant factors, which are: the amount of electricity matched in the day-ahead market (DM_E); the marginal price in the day-ahead market (DM_Price); the share of power sold in the day-ahead market by combined-cycle plants (DM_PCT_CC), thermal plants (DM_PCT_CT) and renewable plants (DM_PCT_CR);

[38] We use the R package ALEPlot. The ALE method is recommended for explaining machine learning models when predictors are correlated between them [27].

the technical constraints cost lagged one, two and seven daily periods; and, finally, the day of week (Sunday) and the amount of electricity matched in the day-ahead market lagged seven days (DM_E_lag_7). As can be observed, the technical constraints cost depends positively on the share of renewables in the day-ahead market, meaning that increases in the share of renewable generation are followed by increases in the technical constraints cost. In contrast, it depends negatively on the share of power generated by combined-cycle and thermal plants, indicating that when there are higher levels of the share of power produced by these plants, there are fewer costs for solving technical constraints.

This result is consistent with the idea that the intermittency and more limited predictability of renewable production may provoke balancing needs and, as a result, balancing costs. In addition, technical constraints costs seem to be higher when the levels of the marginal price, as well as the amount of power sold in the day-ahead market, are lower. Higher technical restriction costs when demand levels are low are hardly justifiable due to technical reasons, but they may be explained by strategic bidding behaviour by market participants. Note that when demand is low, nuclear and renewable plants may produce enough power to satisfy much of the demand, displacing thermal and combined-cycle plants, whose variable costs are considerably higher. Under these circumstances, these plants may try to participate in the subsequent balancing processes, which have been proved to provide higher incomes than the day-ahead market ([13], [14], [15]).

Additionally, the 7-day lagged amount of electricity sold in the day-ahead market is also relevant to explain technical constraints costs, as well as autoregressive terms of order 1, 2, and 7, capturing daily seasonality effects.

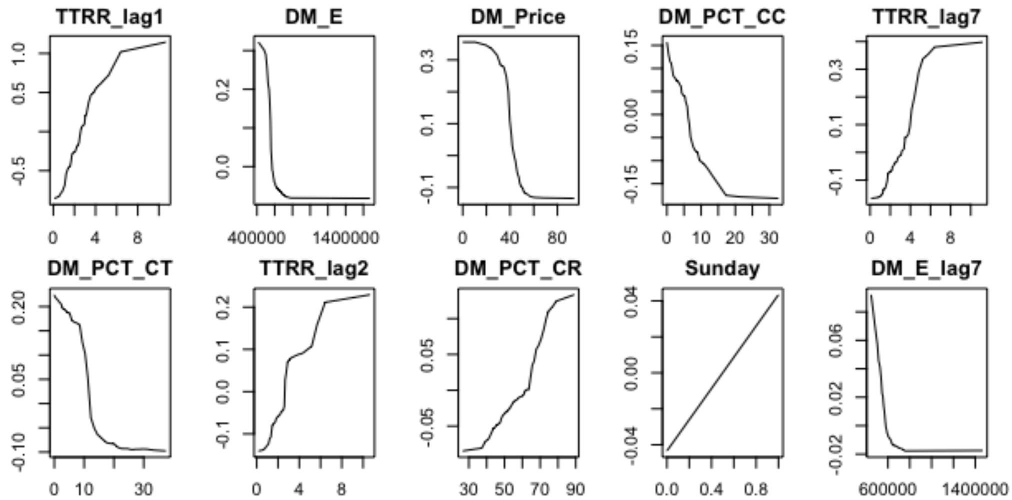

Fig. 1. *Ale Plots TTCC Model*

4.3.2 *System operator processes cost*

Fig. 2 shows that the System Operator processes cost is negatively related to the share of thermal and combined-cycle and positively related to the share of renewable power in the day-ahead market. Thus, these costs are greater with higher shares of renewable generation. As opposite to conventional generation such as thermal and combined-cycle plants, higher renewable production involves a greater need for the management of deviations. On the other hand, the higher the marginal price and/or the amount of electricity sold in the day-ahead market auction, the lower the System Operator processes cost is. Note that it is in scenarios of high demand (and thus high marginal prices) that thermal and combined-cycle plants are needed to generate power in the day-ahead market. In addition, there appears to be a positive relationship between the System Operator processes cost and the capacity payment cost of previous days (CP_lag 5) and also a positive relationship between this cost component and the average offer price submitted to the day-ahead market auction by combined-cycle plants (DM_OPSELL_CC). The amount of operating reserves needed to keep the power system functioning securely and efficiently is a critical issue in power system operation with a large volume of intermittent production ([29]). The higher the renewable generation, the higher the spare capacity that the System Operator will require to ensure system reliability. Such a balancing service is provided by (flexible) plants like combined cycle plants. From the results, it can be deducted that the bidding behavior of combined cycle plants in the day-ahead market has a strong connection with the System Operator processes cost, since the average offer price submitted to the day-ahead market by those plants arises as one of the drivers of this cost. Indeed, while submitting higher offer prices to the day-ahead market, combine cycle plants would drop out of the auction and retain available capacity to participate in subsequent reserve capacity mechanisms. It is important to note that these plants may have the incentive to bid in this way, since the strategic reserve capacity provides them with higher income than the day-ahead market. Finally, autoregressive terms of order 1, 6 and 7 are also relevant to explain the current level of the System Operator processes cost.

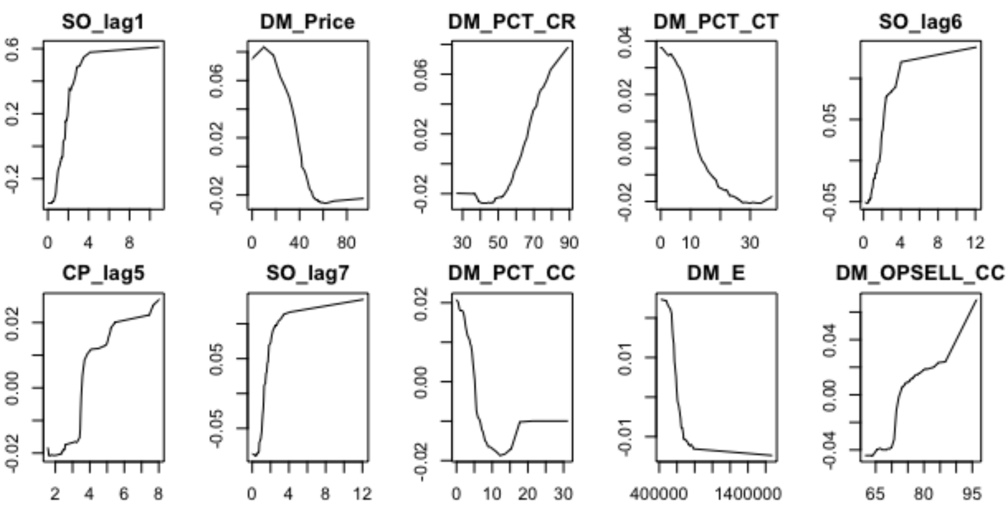

Fig. 2 *Ale Plots SO Model*

4.3.3 *Intraday market cost*

The results of the intraday market model are quite different to those of the TTCC and SO models and are displayed in Fig. 3. Some non-linear relationships between the intraday market cost and the main variables that arise as its best predictors have captured our attention. On the one hand, we note the negative relationship between the intraday market cost and the share of thermal power in the day-ahead market (DM_PCT_CT)) for very low levels of thermal power[39] in the generation mix of the latter, which turns strongly positive for levels of thermal power higher than 10%. On the other hand, the relationship between the intraday market cost and the share of hydroelectric power in the day-ahead

[39] Levels under the first quartile.

market (DM_PCT_CH) is positive for values of the share of hydroelectric power in the day-ahead market from 0% to 9%[40], whereas from that level on the relationship switches to negative.

Thus, the intraday market cost would decrease (increase) with the share of thermal power in the day-ahead market, whenever this share was lower (higher) than 10%; and would increase (decrease) with the share of hydroelectric power in the day-ahead market, for values of this share lower (higher) than 9%.

Furthermore, the day-ahead market marginal price (DM_Price) contributes to explaining the intraday market cost and, again, several intervals can be distinguished: this cost is decreasing for day-ahead market marginal prices between 0 €/MWh and 45 €/MWh, increasing for prices between 45 €/MWh and 60 €/MWh and quite stable for prices higher than 60 €/MWh.

Another remarkable result that is directly linked to bidding strategies is the relevance of the average offer price submitted to the day-ahead market auction by renewable plants (DM_OPSELL_CR). As shown in Figure 3, the relationship between the intraday market cost and the average offer price for selling electricity by renewable plants appears to be a bit more complex. In fact, it starts off as a decreasing relationship between them for very low levels of this average offer price, turns into increasing for average offer prices around 10 €/MWh and it is mainly decreasing for average offer prices above 20 €/MWh.

In addition, natural gas futures prices (TTF lags of 2nd, 3rd and 4th order) are also relevant to explain the intraday market cost component. Two "states" are observed here, the intraday market costs are higher for lower values of the lagged TTF prices (lags of 2nd, 3rd and 4th order), and lower, for higher values of the lagged TTF prices (lags of 2nd, 3rd and 4th order). The transition point is around 20 €/MWh. Finally, the list of predictors is completed with some autoregressive effects (lags of 1st and 7th order) and the 7-lagged System Operator processes cost (highlighting the relationship between the intraday markets and the balancing markets).

[40] Approximately in half of the cases during the sample period.

A number of interesting results come out of the empirical analysis regarding the intraday market. To summarize, except on days with a low percentage of thermal generation in the day-ahead auction market; generally, the greater the weight of thermal generation in the day-ahead market, the higher the intraday cost. The increasing need for thermal generation is usually an indication of strong demand and/or a decrease in available generation capacity from other generation sources, which would lead to high prices in both the day-ahead and the intraday markets. This is in line with the result that the intraday cost is increasing with the day-ahead marginal price when this latter remains between 45 €/MWh and 60 €/MWh, namely between approximately the average price and the 88th percentile during the sample period.

On the other hand, the flexibility of hydroelectric plants allows their managers to bid strategically in the day-ahead and intraday markets to maximize profits. Thereby, as long as there is no shortage of water reservoirs; in particular, when the share of hydroelectric power in the day-ahead market exceeds the average value, it is found a negative relationship between such a share and the intraday market cost, which could be explained by a shift of hydroelectric plants' generation from the day-ahead market to the intraday market where to bid at higher prices with the aim of raising the intraday market auction's marginal price.

Interestingly, the intraday market cost appears to be critically linked to the bid prices submitted by renewable generation plants to the day-ahead market for selling their electricity and not to the amount of renewable generation. The offered prices at which the relationship between both variables changes its sign, i.e. 10 €/MWh and 20 €/MWh, respectively correspond to the 52th and the 83th percentile of the distribution. Therefore, the intraday market cost would increase for mean and somewhat high day-ahead market submitted prices by renewables generations and decrease for low and very high levels of those offer prices.

Finally, natural gas prices and some autoregressive effects complete the list of relevant variables that help to explain the intraday market cost.

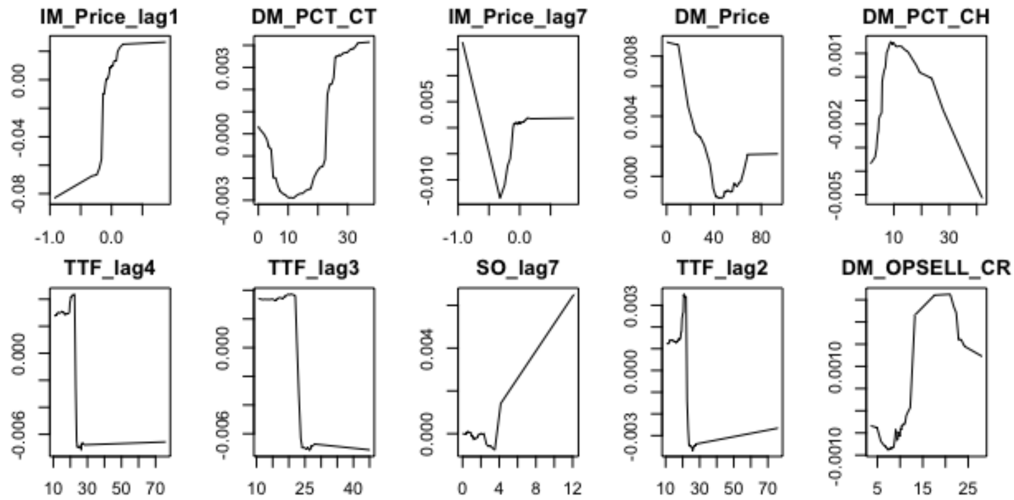

Fig. 3. *Ale Plots IM Model*

4.3.4 Capacity payment cost

The relationships between the capacity payment cost and the variables that arise as its main determinants are displayed in Fig. 4.

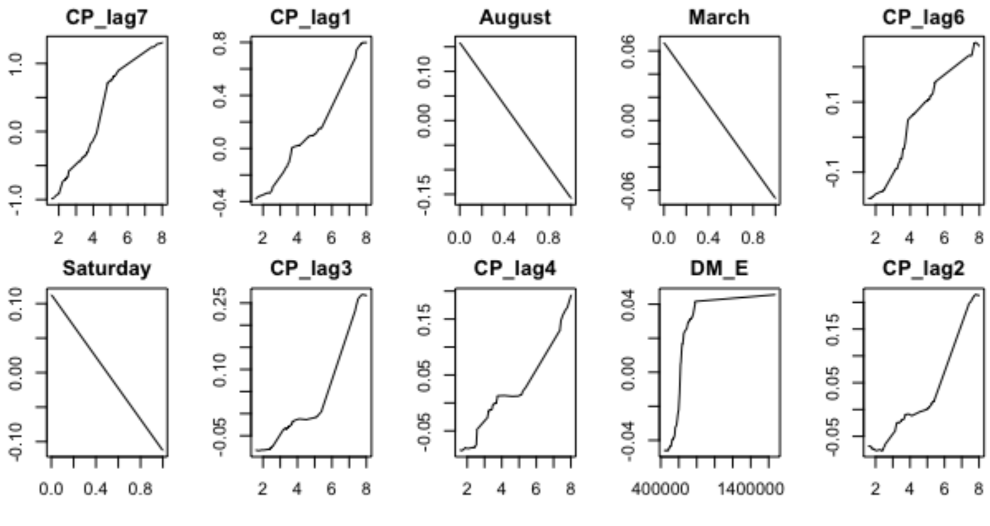

Fig. 4 *Ale Plots CP Model*

As can be seen, capacity payments in previous days are shown to be followed by current capacity payments (autoregressive effects with lags of 1st, 2nd, 3rd, 4th, 6th and 7th order). Additionally, capacity payments are higher with higher levels of electricity sold in the day-ahead market. In fact, they decrease in periods of low demand, such as during the months of March (mild temperatures, no need for cooling or heating) and August (typically a vacation period with lower electricity demand from factories) or, at a daily level, on Saturdays. This result is logical since capacity payments include the medium-term availability service that is calculated by the System Operator based on demand levels. Given that the higher the demand, the greater the risk of a lack of supply, capacity payments are expected to rise (decrease) in periods of high (low) demand. As mentioned earlier, capacity payments were conceived as payments to be received by conventional source plants with the aim of ensuring long-term capacity investments. Motivated by the increasing levels of renewable generation in the electricity system, which were reducing the share of conventional generation, a new regulation (Minister Order ITC/3127/2011 [30]) was introduced to implement the medium-term availability service, whose objective was to guarantee sufficient available generation capacity to meet the demand for electricity at any time. As a result, some flexible conventional plants started to be remunerated by the system through the capacity payments mechanism, for merely being available, thereby serving as a backup to renewable electricity sources. It should be noted that the amount of money paid by consumers through this mechanism has exceeded in a number of years the amount actually received by generators. This excess, far from being reduced to exactly match the amount actually due as capacity payments, has been used to reduce the deficit of the electricity system. Therefore, starting from 2014 in particular, the capacity payment component of final prices could have been much less than it actually was. As stated by the Spanish Commission for Markets and Competition[41], it would have been recommendable that each component of the final price reflects the costs for which it was created, in the interest of transparency and the transmission of appropriate price signals.

4.3.5 *Interruptibility service cost*

Regarding the interruptibility service cost (Fig. 5), this variable itself, but lagged 1, 2 and 3 periods (days), helps to explain it, together with some calendar effects as follows. According to our results, this cost appears to be higher during the months of April and September, while it is lower during the months of July and November and the year 2018. In that year, the System Operator, REE, introduced changes to the procedure with the aim of achieving greater efficiency in the application of the service (Order ETU/1133/2017),

[41] Comisión Nacional de los Mercados y la Competencia (www.cnmc.es)

as well as adapting to EU regulations (the legislative package called "Clean Energy for All Europeans" presented by the European Commission on 30 November 2016), which emphasises reaching a more competitive allocation procedure. Among the changes introduced, we highlight the reduction of the maximum period of time between the notice that the System Operator issues to service providers and the effective start of the execution option that involves the reduction of power made available to the system and the readjustment of the price which serves as a reference to the variable remuneration corresponding to the effective provision of the service. This change in regulation has resulted in a reduction in the cost of the service of 0.2 €/MWh from 2018 on. Finally, coal prices lagged 3 and 6 periods complete the list of predictors.

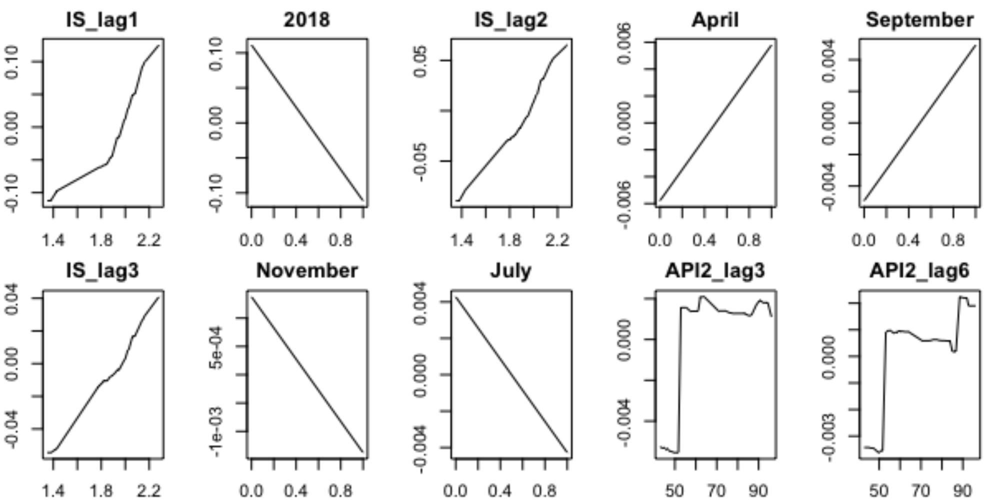

***Fig. 5** Ale Plots IS Model*

5. Conclusion

This paper mainly aimed to investigate the impact of renewable generation sources on final electricity prices, particularly, in the costs incurred in the subsequent processes until the real time delivery of electricity, designed to guarantee continuous supply and system reliability.

Additionally, the analysis made has allowed us to disentangle the main drivers of electricity final prices, which have turned out to be related to the integration of renewable generation, balancing needs, strategic bidding behaviour by market participants and changes to the regulations.

We also find that the increasing share of renewable power in the day-ahead market auction arises as one of the main determinants explaining rises in those technical processes and services that are necessary for the secure and reliable operation of power systems; in particular, higher renewable generation in the day-ahead market auction involves (i) higher costs derived from the technical constraint resolution process, (ii) a greater need for the management of deviations by the System Operator which translates also into greater costs and, finally, (iii) higher prices in the intraday market compared to the day-ahead market (when day-ahead prices are not too low), which contribute to push final prices up. It is remarkable that the net effect of these opposing forces for the Spanish electricity market during the studied period has been a reduction in final prices. Thus, the added system integration costs from the increasing penetration of renewables haven´t been enough to compensate for the drops in day-ahead market auction prices due to the merit-order effect that has been previously highlighted in literature.

The obtained results shed light on the overall impact of renewable generation on electricity prices, providing new evidence of the fact that market participants strategically plan their bidding behaviour considering the market as a whole, trying to maximize their profits as a result of their global participation in the day-ahead market, the intraday market and the rest of the balancing and deviation management processes.

The factors driving the components of electricity final prices, other than the day-ahead market price, are much less studied, but they are key in order to gain further insight into the dynamics between the interrelated trading segments and the technical processes involved in the wholesale electricity market that should be considered when assessing changes in the design of the market towards a more sustainable and competitive electricity system.

Acknowledgements

This work was supported by the Spanish Ministry of Science, Innovation and Universities [project number PGC2018-093645-B-100].

References

[1] Holttinen, H. (2004). "The Impact of Large Scale Wind Power Production on the Nordic Electricity System". Dissertation for the title of Doctor of Science in Technology, Helsinki University of Technology.

[2] Sensfuβ F., Ragwitz, M. and Genoese, M. (2008). "The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany". Energy Policy 36, 3086–3094.

[3] Martin de Lagarde, C. and Lantz, F. (2018) "How renewable production depresses electricity prices: Evidence from the German market". Energy Policy 117, 263–277.

[4] McConnell, D., Hearps, P., Eales, D., Sandford, M., Dunn, R., Wright, M. and Bateman, L. (2013). "Retrospective modelling of the merit-order effect on wholesale electricity prices from distributed photovoltaic generation in the Australian National Electricity Market". Energy Policy, 58, 17-27.

[5] Saénz de Miera, G., Del Río, P. and Vizcaíno, I. (2008). "Analysing the impact of renewable electricity support schemes on power prices: The case of wind electricity in Spain", Energy Policy 36, 3345-3359.

[6] Ballester, C. and Furió, D. (2015). "Effects of renewables on the stylized facts of electricity prices". Renew. and Sustainable Energy Reviews, 52, 1596-1609.

[7] Würzburg, K., Labandeira, X. and Linares, P. (2013). "Renewable generation and electricity prices: Taking stock and new evidence for Germany and Austria". Energy Economics 40, S159-S171.

[8] Carvalho, N. and Pereira, P. (2019). "The 'merit order effect' of wind and solar power. Volatility and determinants". Renewable and Sustainable Energy Reviews, 102, 54-62.

[9] Gianfreda, A., Parisio, L. and Pelagatti, M (2018). "A review of balancing costs in Italy before and after RES introduction". Renew. and Sustainable Energy Reviews Volume 91, pages 549-563.

[10] Weber, C. (2010). "Adequate intraday market design to enable the integration of wind power into the European power systems". Energy Policy 38, 3155-3163.

[11] Borggrefe, F. and Neuhoff, K. (2011). "Balancing and intraday market design: options for wind integration". European Smart Power Market Project. C.P. Initiative.

[12] Henriot, A. (2014) "Market design with centralized wind power management: handling low-predictability in intraday markets". The Energy J. 35, 99-117.

[13] Furió, D. and J.J. Lucia. (2009). "Congestion management rules and trading strategies in the Spanish electricity market". Energy Economics, 31, 48-60.

[14] Chaves-Avila, J. and Fernandes, C. (2015). "The Spanish intraday market design: A successful solution to balance renewable generation?" Renew. Energy 74, 422-432.

[15] Ballester, C. and Furió, D. (2017). "Impact of Wind Electricity Forecasts on Bidding Strategies". Sustainability 2017, 9, 1318.

[16] Klæboe, G., Braathen, J., Eriksrud, A. and Fleten, S. (2019). "Day-Ahead Market Bidding Taking the Balancing Power Market Into Account". Available at SSRN: https://ssrn.com/abstract=3434318 or http://dx.doi.org/10.2139/ssrn.3434318

[17] Schnürch, S. and Wagner, A. (2019). "Machine Learning on EPEX Order Books: Insights and Forecasts". arXiv:1906.06248 [stat.AP]

[18] Prasanna, A., Holzhauer, S. & Krebs, F., (2019). Overview of machine learning and data-driven methods in agent-based modeling of energy markets. In: David, K., Geihs, K., Lange, M. & Stumme, G. (Hrsg.), INFORMATIK 2019: 50 Jahre Gesellschaft für Informatik – Informatik für Gesellschaft. Bonn: Gesellschaft fürInformatike.V.. (S. 571-584). DOI: 10.18420/inf2019_73

[19] Furió, M., Lucia, J., Meneu, V.(2011) "The Spanish electricity intraday market : prices and liquidity risk".European economic and political developments. - New York, NY : Nova Science Publ., ISBN 1-61761-581-1. - 2011, p. 85-102

[20] Chuliá, H., Furió, M. and Uribe, JM. (2019). "Volatility spillovers in Energy Markets". The Energy J. Volume 40, 3, 173-198.

[21] R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

[22] Terry Therneau and Beth Atkinson (2018). rpart: Recursive Partitioning and Regression Trees. R package version 4.1-1. URL https://CRAN.R-project.org/package=rpart

[23] A. Liaw and M. Wiener (2002). Classification and Regression by randomForest. R News 2(3), 18--22. URL https://CRAN.R-project.org/doc/Rnews/

[24] Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, Tianyi Zhou, Mu Li, Junyuan Xie, Min Lin, Yifeng Geng and Yutian Li (2018). xgboost: Extreme Gradient Boosting. R package version 0.71.1. https://CRAN.R-project.org/package=xgboost

[25] Hastie, T., Tibshirani, R. and Friedman, J. (2001). "The Elements of Statistical Learning: Data Mining, Inference, and Prediction". Second Edition. Springer Series in Statistics Ed. Springer

[26] Breiman, L. (2001) "Machine Learning" 45: 5. https://doi.org/10.1023/A:1010933404324.

[27] Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., and Zeileis, A. (2008). "Conditional variable importance for random forests". BMC Bioinformatics, 9, 307. http://doi.org/10.1186/1471-2105-9-307.

[28] Dan Apley (2018). ALEPlot: Accumulated Local Effects (ALE) Plots and Partial Dependence (PD) Plots. R package version 1.1. https://CRAN.R-project.org/package=ALEPlot

[29] Pérez-Arriaga, I.J, and Batllé, C. (2012). “Impacts of Intermittent Renewables on Electricity Generation System Operation”. Economics of Energy & Environmental Policy, 1, 2.

[30] Order ITC/3127/2011, of 17 November, regulating the power availability service of capacity payments and modifying the investment incentive referred to in annex III of Order ITC/2794/2007, of 27 September, revising electricity tariffs from 1 October 2007 BOE-A-2011-18064