New methods to evaluate key variables in energy markets

José M. Labeaga

Madrid, January the 27th 2016

The problem to address (I)

- Extract information from everywhere to evaluate key variables in energy markets (special focus on demand and particularly price elasticities)
- How can we use the impressive amount of information (web, media, hundred of papers, social networks, etc.) to analyze correlation-causal effects relevant for economic decisions?
- Tools for accessing, storing, extracting, analyzing?
- Methods? Descriptives, graphics, correlation …?

The problem to address (II)

- Big Data is data that exceeds the processing capacity of conventional database systems. The data is too big, moves too fast, or does not fit the structures of our database architectures. To gain value from this data, one must choose alternative ways to process it

- Is it a good idea to storage data (trying to built long time series or long time series of cross-sections) for doing long-run analyses or should these data be analyzed for short-run purposes?
  - Academic vs non-academic analysis

- Is this kind of data useful? Can its use enrich current analyses?

The problem to address (III)

- I do not have closed answers to these questions but it depends
  - On the data
  - On the questions to address
  - On the decisions to inform and their timing
  - ...
- So, my main purpose it to describe some (possibilities of) exercises using this kind of data

Example 1:
Market power, congestion and price-elasticities in the MIBEL market (joint work with R. García-Fierro)

Example 1 (I)

- Set-up of this exercise:
  - Market power in electricity markets
  - Bigerna, Bollino, Polinori (2015), *The Energy Journal* the first empirical exercise to calculate market power with congestion through the estimation of residual demand functions (as in Wolak, 2003 in the *American Economic Review* but modifying it to allow for congestion)
  - The market of our study has very different characteristics to those studied by BBP (Italy) or W (California):
    - Different period
    - Different structure of the energy mix
    - Different structure of producers
    - Different economic structure

Example 1 (II)

- The Iberian Electricity Market (Mercado Ibérico de Electricidad - MIBEL), constitutes a joint initiative from the Governments of Portugal and Spain, and is an important step in the development of an internal electricity market

- In theory, with the materialization of MIBEL, it becomes possible for any consumer in the Iberian zone to acquire electrical energy under a free competition regime, from any producer or retailer that acts in Portugal or Spain

- Data for MIBEL is continuously on the web (but we cannot access continuously to the whole content) so we need tools for storing and processing

Example 1 (III)

- MIBEL's main goals are:
  - To benefit consumers of both countries, through the integration of the respective electric systems
  - To structure the market organization based on the principles of transparency, free competition, objectivity, liquidity, self-financing and self-organization
  - To support the development of the electricity market of both countries, with the existence of a single reference price for the whole of the Iberian Peninsula
  - To allow all the participants free access to the market, under equal conditions of rights and obligations, transparency and objectivity
  - To promote economic efficiency of electrical sector companies, encouraging free competition amongst them

Example 1 (IV)

- MIBEL begun on July the 1st, 2007
- The aim of this exercise is to analyze supply and demand in the day market (not intraday) managed by OMIE (the market operator). We try to study whether there is congestion, market power and whether it affected price-elasticities. In this market there are a number of bids (supply) in day d-1 for the hourly demand during day d
  - Compulsory market for traders with more than 50 MW power. All generating units are required to supply
  - Supply could be simple (only price and quantity) or complex (some conditions added to the supply as minimum revenue, stops, indivisibility of supply, etc. We will only analyze simple bids at a first stage but …
  - Supply and demand consider 24 hours per day (and 25 sections per hour). Price in each section should be increasing for selling and decreasing for purchasing
  - The equilibrium price equals? supply and demand for each hour
  - Price is unique for Spain and Portugal except if there is transport congestion. When transport congestion occurs the process is country specific and we have two different prices at equilibrium

Example 1 (V)

- Summarizing the complexity of our problem:
  - We have a number of (simple) bids (and a number of complex bids) depending on the number of traders for each section of each day for 365(6) days a year for several years (from the 1st July 2007 to the 1st July 2015) (these data can be obtained daily on the web)
  - We have to built aggregated supply (aggregation conditions?)
  - We have to built aggregated demand
  - We have prices at equilibrium (?)
  - We like to have measures of market power, whether congestion affects market power and whether market power affect price-elasticities

Example 1 (VI)

- Summarizing the tasks done:
  - We take data from the day market for all bids at OMIE (http://www.omie.es/inicio)
  - For each bid we have an identification code. This iden allows us to know the trader. There are firms owning several traders (important to know the share of the market for each firm)
  - We take marginal prices and quantities for every hour analyzed
  - All this information allows building aggregated demand for each firm
  - Once we have supply and demand for all firms (and hours) in the market we calculated residual un-contracted demand, price-elasticity and market power

Example 1 (VII)

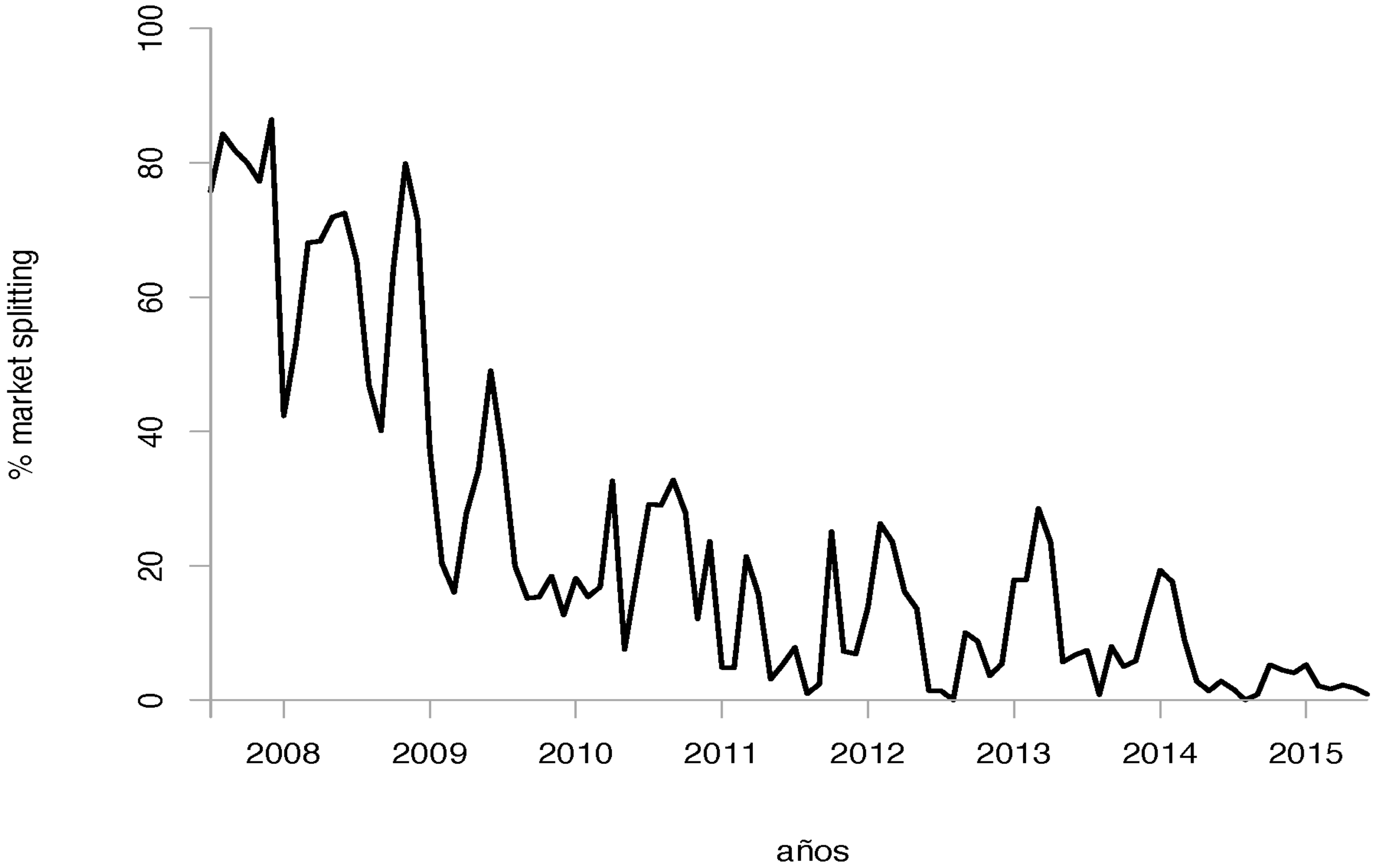

Figure 1. Evolution of percent hours with market splitting (transport congestion)

Example 1 (VIII)

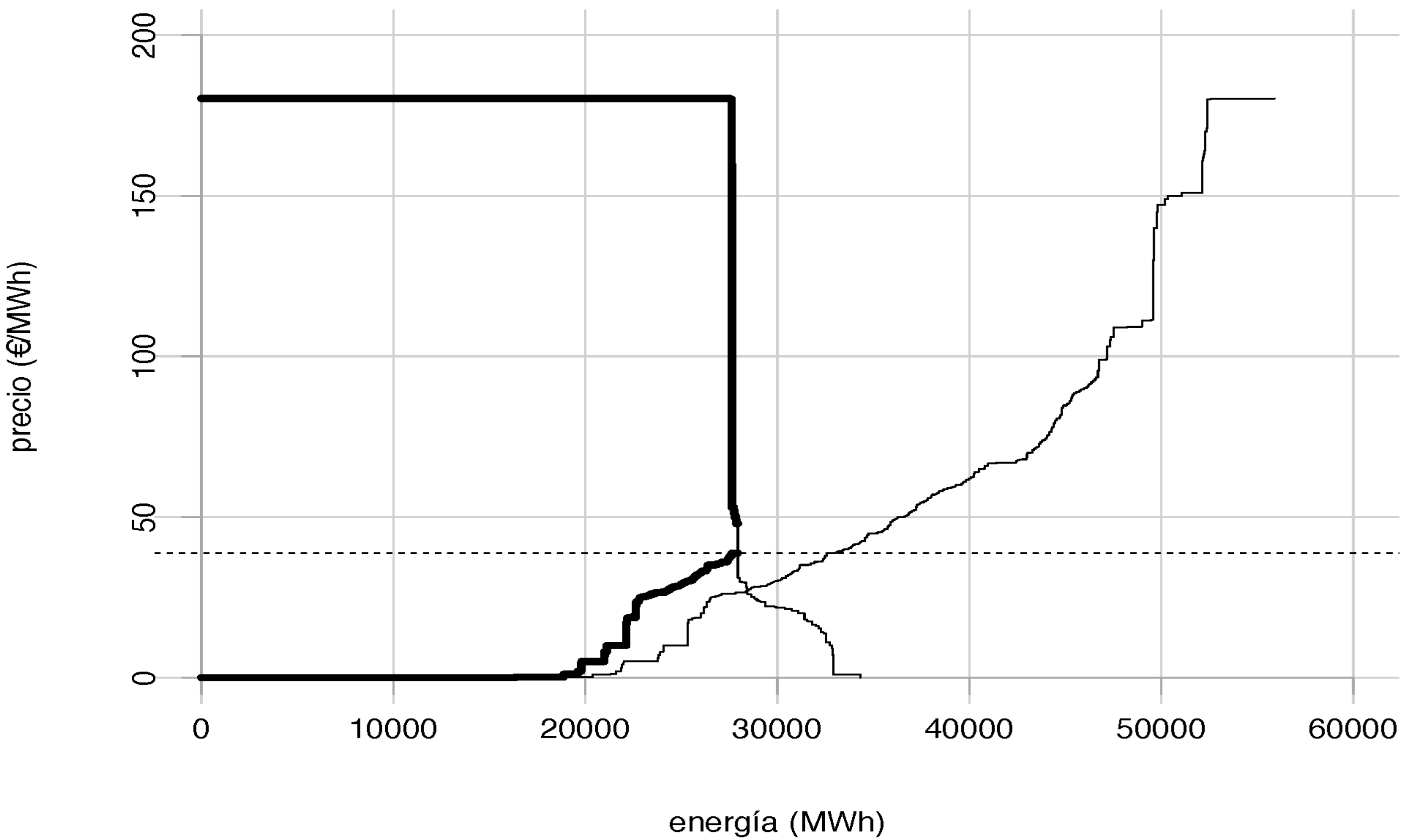

Figure 2. Suppy and demand (hour 10; day 02/07/2007)

Example 1 (IX)

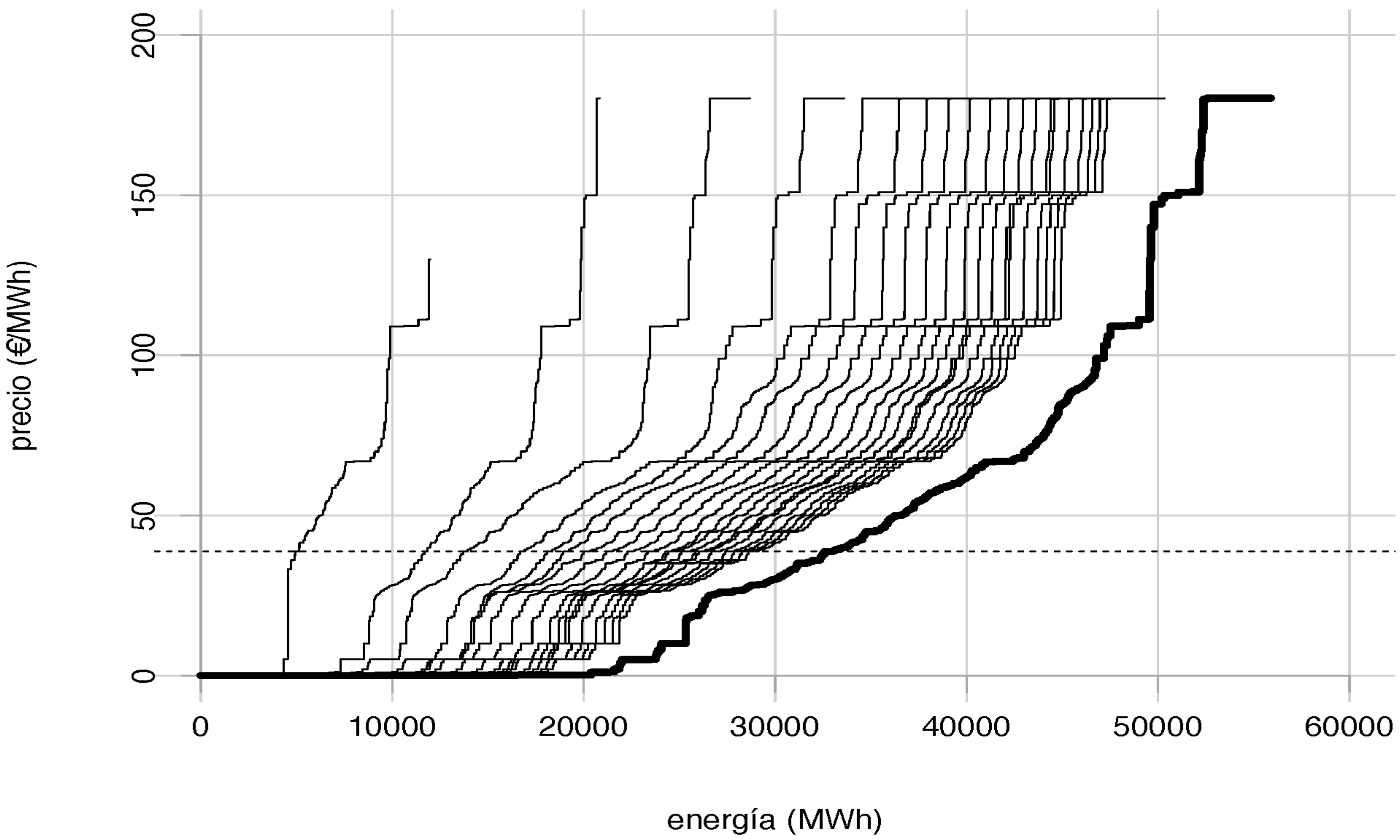

Figure 3. Suppy by trader and sum of supply curves for all traders (hour 10; day 02/07/2007)

Example 1 (X)

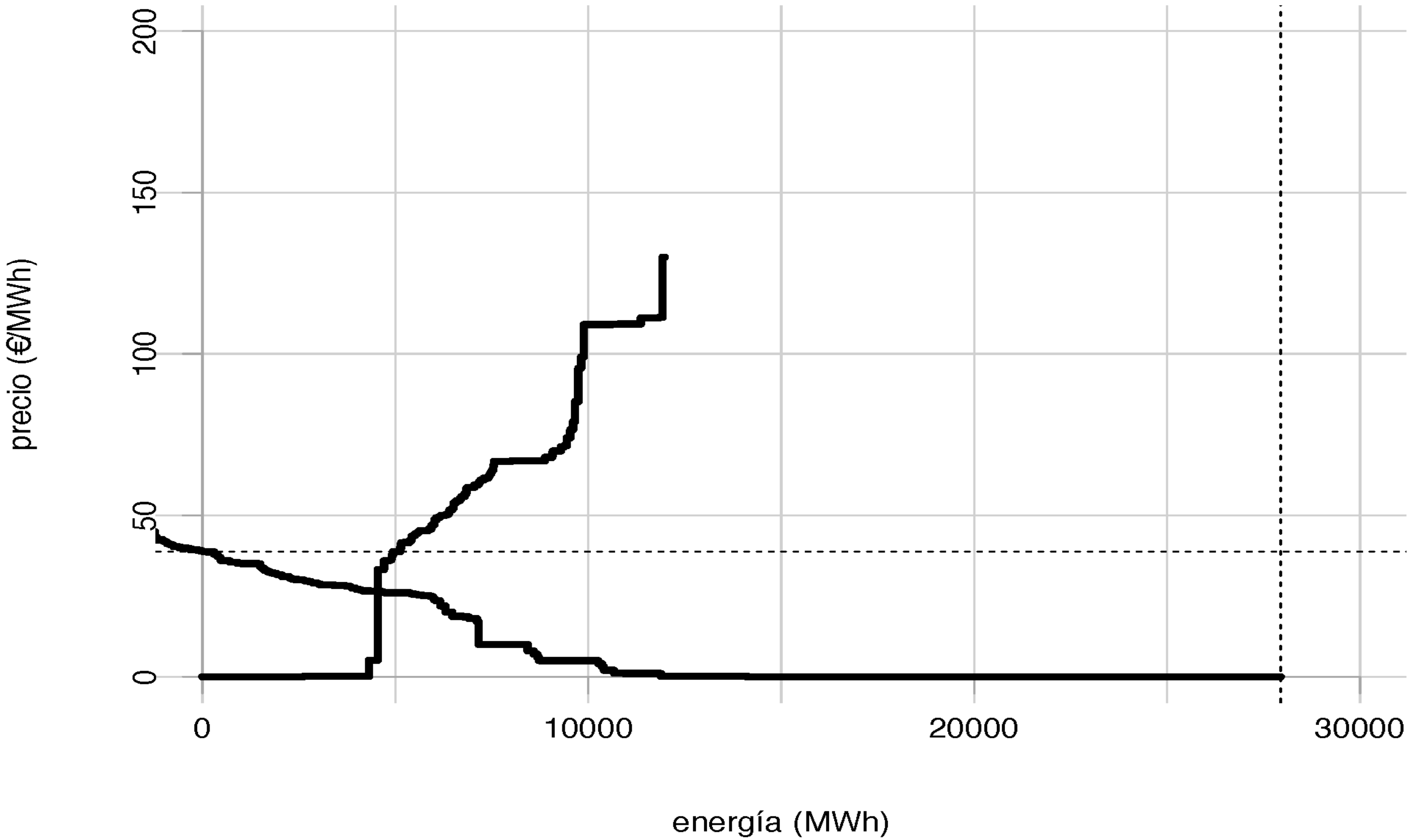

Figure 4. Supply and residual demand of a firm (hour 10; day 02/07/2007)

Example 1 (XI)

- It is necessary to know the residual demand function in order to calculate the price-elasticity of demand (Wolak, 2009), but Figure 4 shows that supply and demand do not cross at the observed price of equilibrium. Why?

- Because of the existence of complex bids. How to do to calculate elasticities with complex bids?

- Simple expression $\varepsilon = -\frac{\Delta q / q}{\Delta p / p}$

Example 1 (XII)

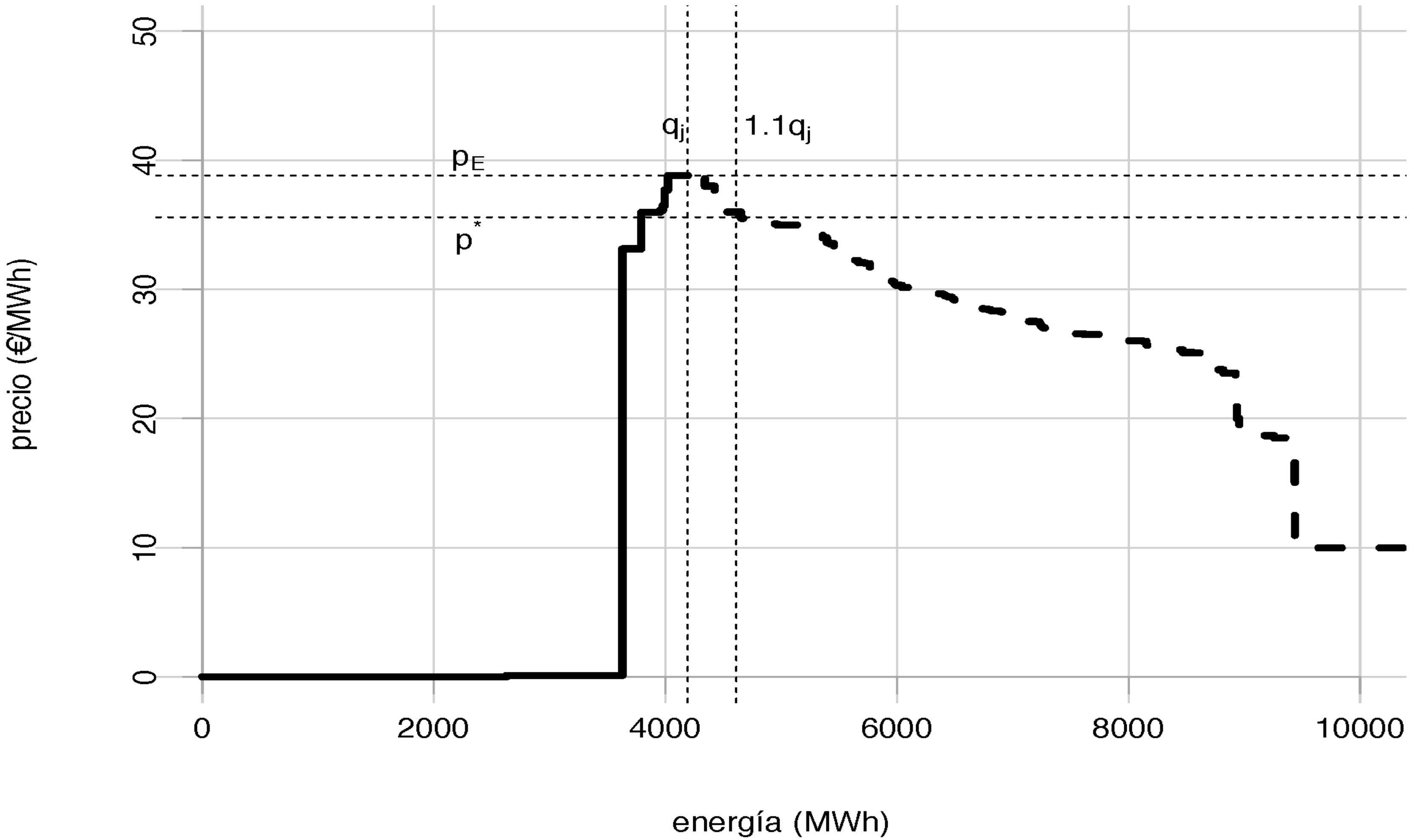

Figure 5. Residual demand at clearing bids (hour 10; day 02/07/2007)

Example 1 (XIII)

- In Figure 5 we can see that there is not residual demand (there is not supply) above the market price. How can we calculate the elasticity?

- We can apply a formula that modifies the proposal by Wolak (2009) to calculate the elasticity (around the value of prices and quantities in Figure 5 –algorithm or approx.-)

- In theory, $$\varepsilon = \frac{RD_j(p^h) - RD_j(p^l)}{p^h - p^l} \frac{p^h - p^l}{RD_j(p^h) - RD_j(p^l)}$$

Example 1 (XIV)

- In practice we can use data observed (see the figure) to solve (an approx to the index of Lerner):

$$aq_j = RD_j(p^*); 1.1q_j = RD_j(p^*)$$

$$\frac{1}{\varepsilon} = \frac{(p^* - p_E)/p_E}{0.10}$$

- Where RD_j is residual demand of firm *j* and q_j is the quantity of firm j that clears the market (sensitivity of the elasticity around the price of equilibrium, i.e., for different values of *a* from 1.1 to 1.01, for instance?)

Example 1 (XV)

- What does happen when there is transport congestion?
  - Total demand in each area is area-specific but in the area with congestion is less than total capacity
  - Supply of rest of generators (used to calculate residual demand) is the sum of supply of the area plus supply of the area where demand is smaller than capacity

$$RD_i^{Esp}(p) = D_T^{Esp} - S_{T-i}^{Esp} + max\{T_{Por\rightarrow Esp}, min[T_{Esp\rightarrow Por}, D_T^{Por} - S_{T-i}^{Por}(p)]\}$$

Example 1 (XVI)

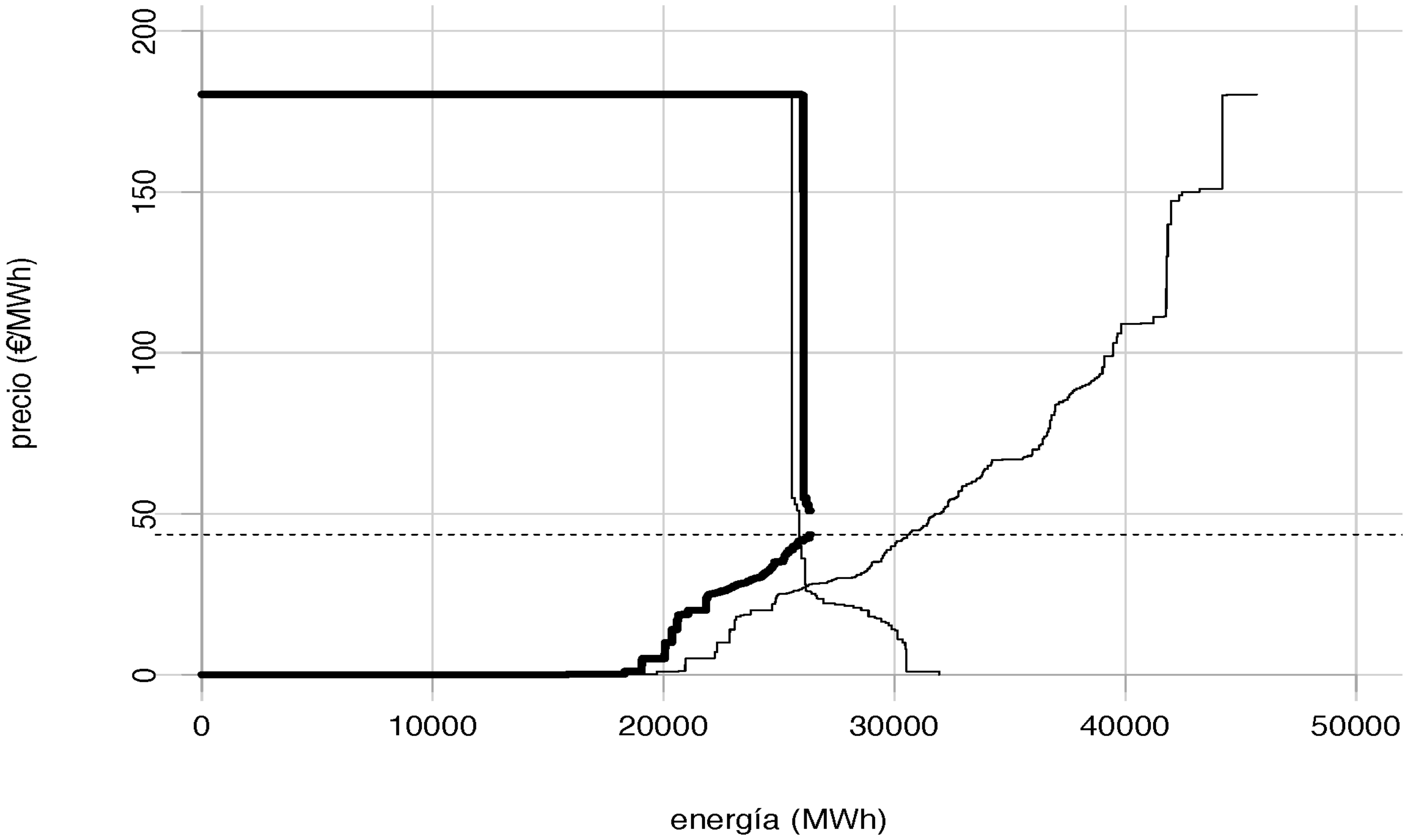

Figure 6. Supply and demand with congestion in Spain (hour 11; day 02/07/2007)

Example 1 (XVII)

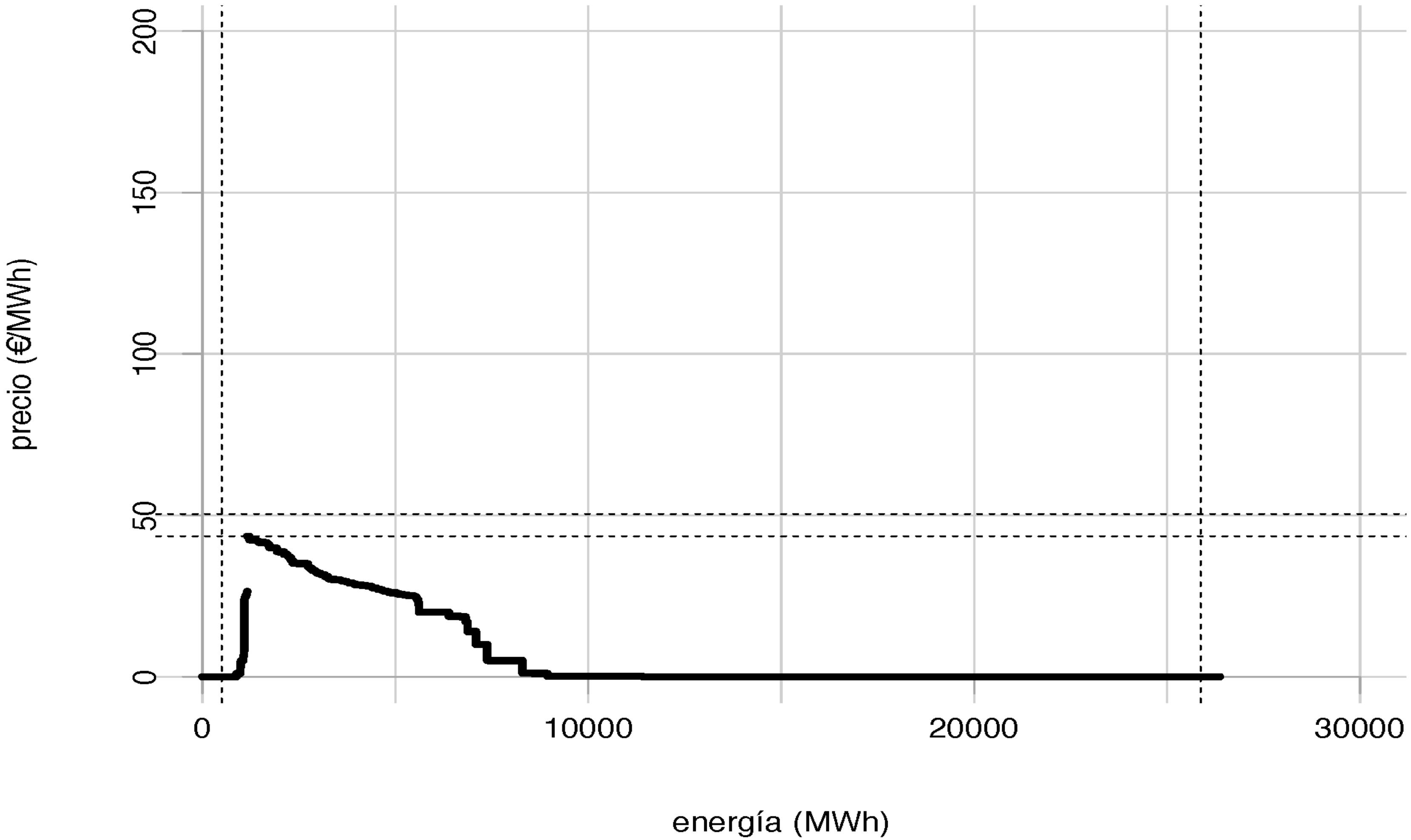

Figure 7. Residual demand with congestion in Spain for exporter (hour 11; day 02/07/2007)

Example 1 (XVIII)

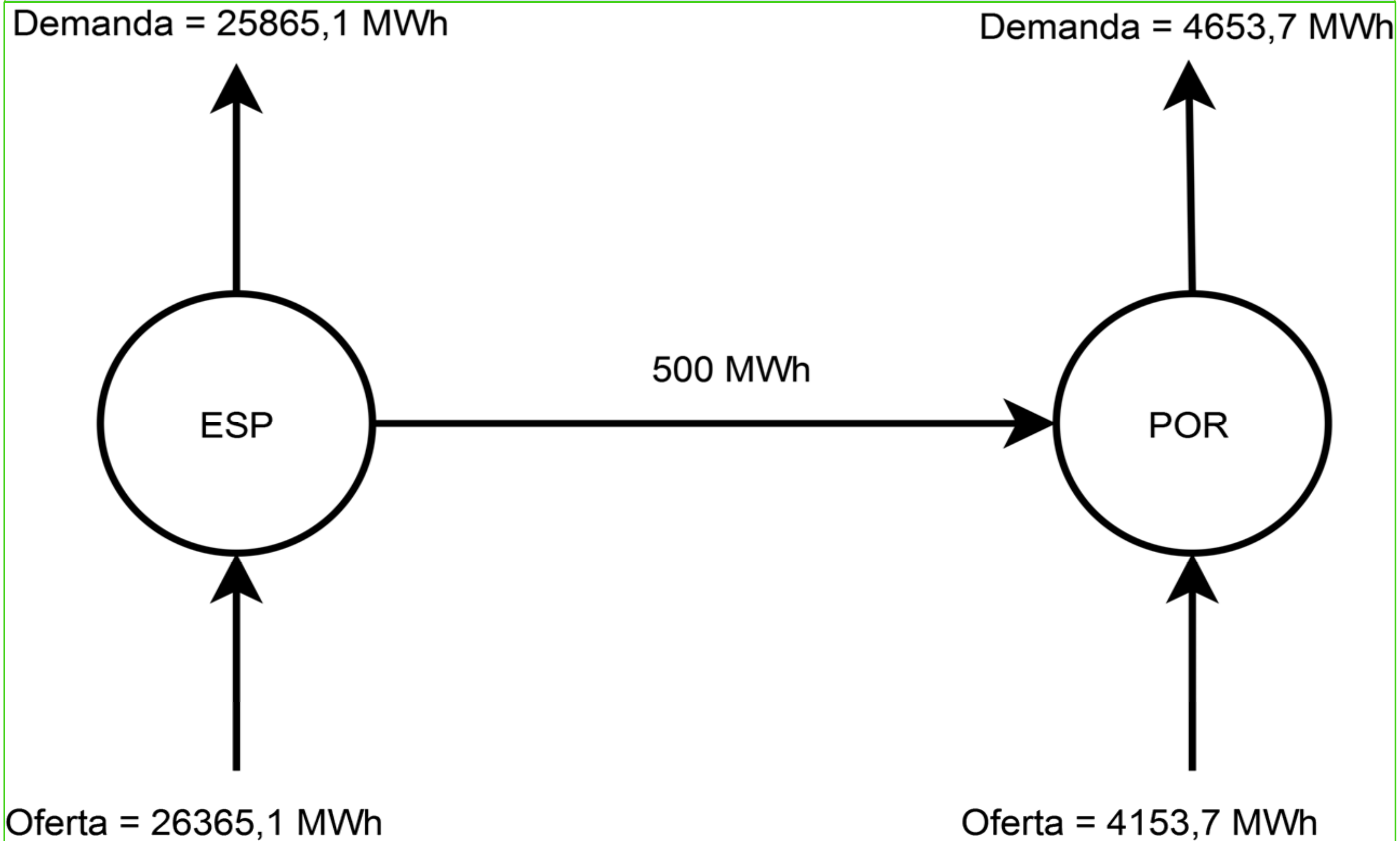

Figure 8. Scheme of the market functioning with congestion

Example 1 (XIX)

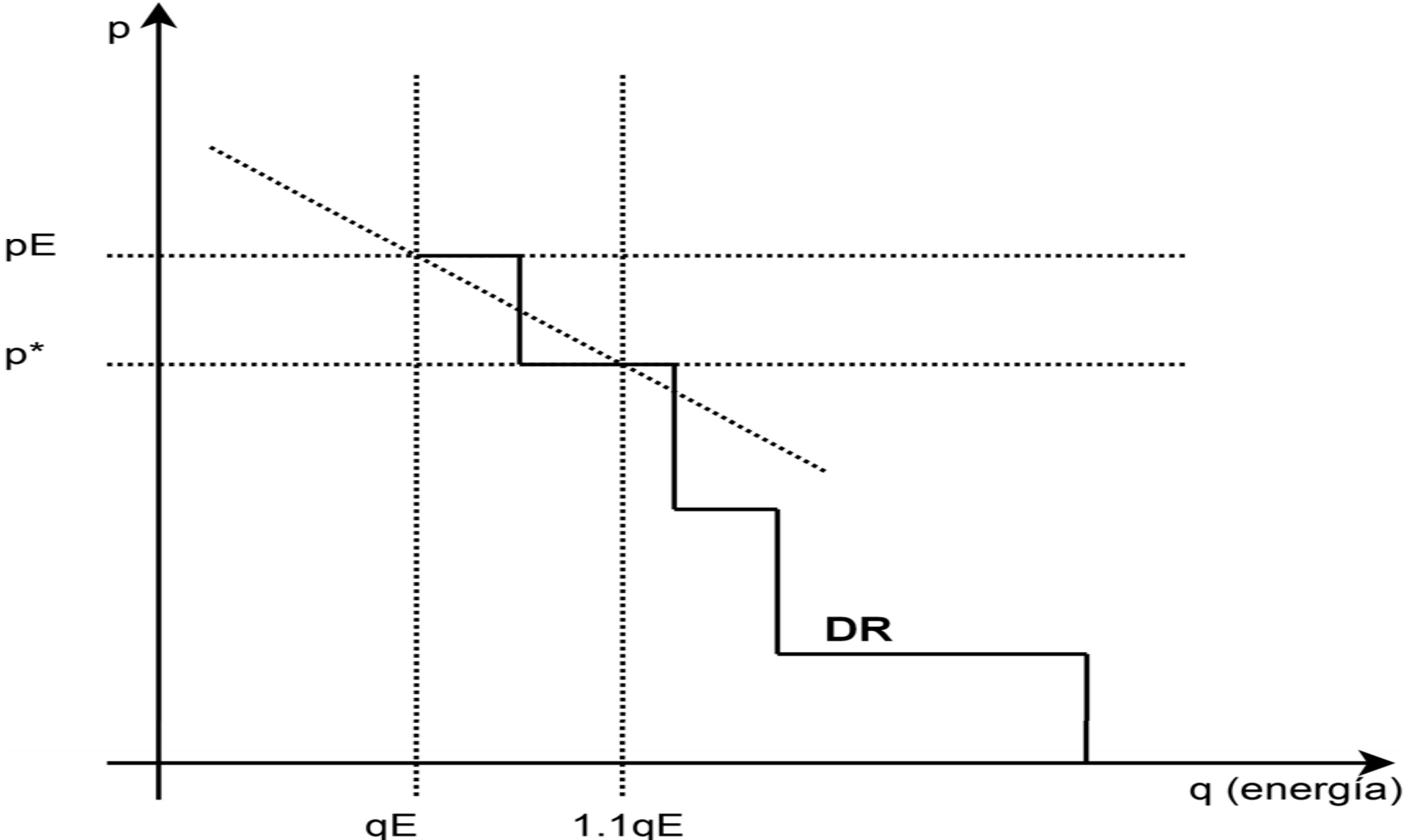

Figure 9. Price-elasticity of demand for electricity in the MIBEL market

Example 1 (XX)

- Next, we analyze unilateral market power (of traders –each one and averages-) using the Lerner index
  - Lerner index is the inverse of the residual demand elasticity
  - Since price-elasticity is higher (in absolute value) without congestion than with congestion, the index is smaller without congestion (as expected) at any hour analyzed (we have analyzed 70,128 hours (16,453 with congestion). Figures are average of 0.32 up to 2010 with congestion and 0.27 without
- Unilateral market power vanished as long as congestion is now (2015) not important (demand has become less sensitivity to prices)

Example 1 (XXI)

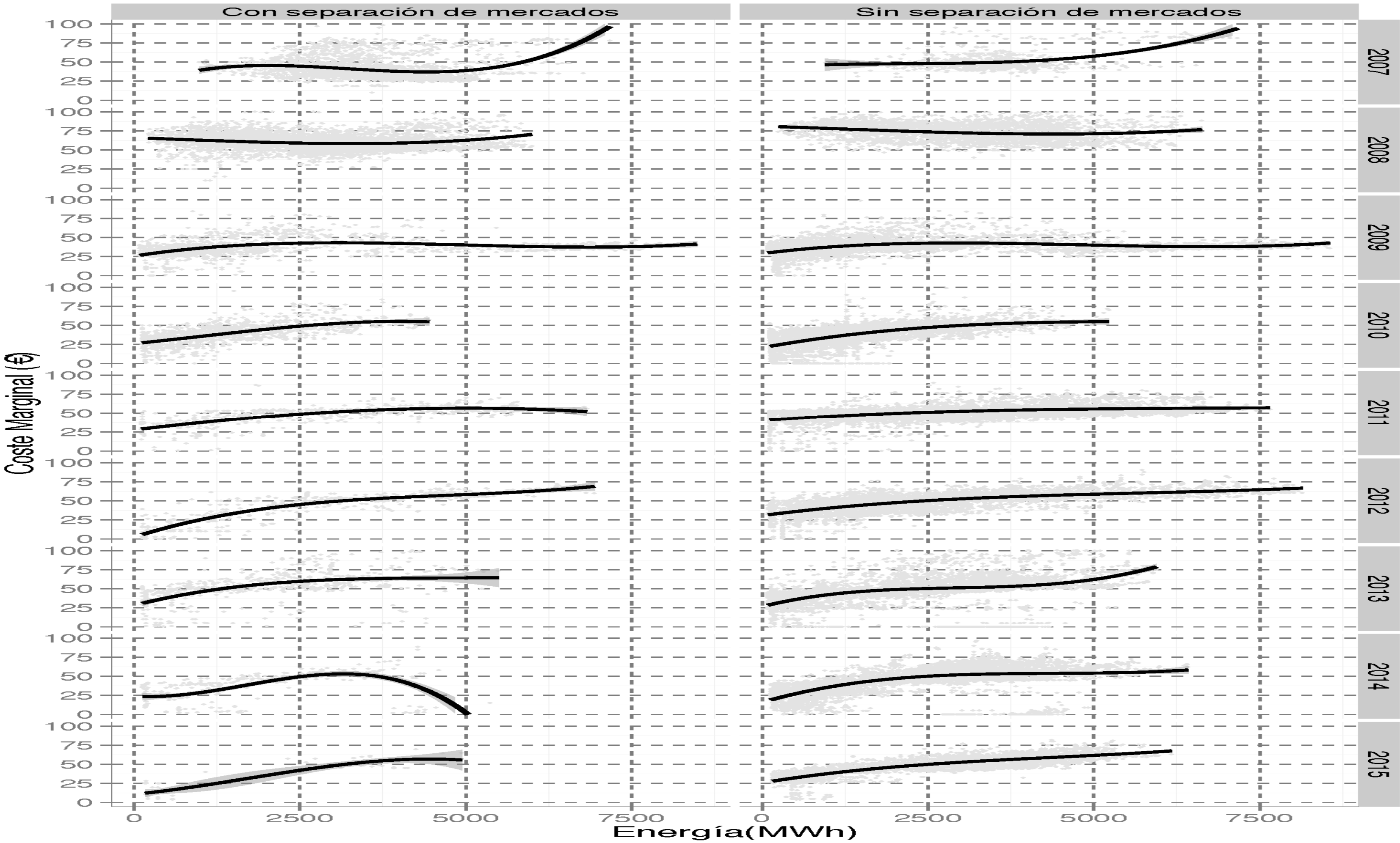

Figure 10. Average marginal cost function of all suppliers using the Lerner index function (cubic function)

Example 1 (XXII)

- Our final aim:
  - To estimate price-elasticities (and market power) in an agent-based microsimulation model to be able to:
    - Simulate tax changes (considering strategic behavior of bidders)
    - Simulate policy measures (regulation with strategic behavior)
    - Estimate market power in the black box (using the complex bids)
    - Etc.
  - Preliminary results:
    - Market power is not an issue (except through the black box)

Example 2:
Using GDELT data to evaluate the confidence on the Spanish government energy policy (joint work with D. Bodas-Sagi)

Example 2 (I)

- Set-up of this exercise:
  - How do sentiment indicators affect energy prices and energy demand?
  - I do not know of any paper analyzing this issue in energy markets
  - I do not know of any economic paper using this database
  - Everything in the presentation constitutes “ad hoc” ideas
  - I will only give correlations

Example 2 (II)

- We use Global Database on Events Language and Tone (GDELT)

  - The GDELT Project is supported by Google Ideas and it shares real-time information and metadata with the world. This codified metadata (but not the text of the articles, which we will use later on) is then released as an open data stream, updated every 15 minutes, providing a multilingual annotated index of the information

  - It includes broadcast, print, and (all) online news sources. The project shares a database with trillions data points. Although, data is available as downloadable CSV files, few users have the storing capacity and processing power to download terabytes of data, and effectively query and analyze it. Google's BigQuery platform provides a way to interact with this huge information source

  - GDELT is a clear example of Big Data, while Google's BigQuery is an example of Infrastructure As a Service (IaaS) technology

Example 2 (III)

- GDELT …
  - GDELT maintains the GDELT Event Database, and the GDELT Global Knowledge Graph (GKG). The GKG begins April 1, 2013 and "… attempts to connect every person, organization, location, count, theme, news source, and event across the planet into a single massive network that captures what's happening around the world, what its context is and who's involved, and how the world is feeling about it, every single day"
  - The data files use Conflict and Mediation Event Observations and we can identify each event (extracted from the original information), which allows rich contextualization of them. We can also extract information (words) on messages, news, etc. , related to a theme or subject

Example 2 (IV)

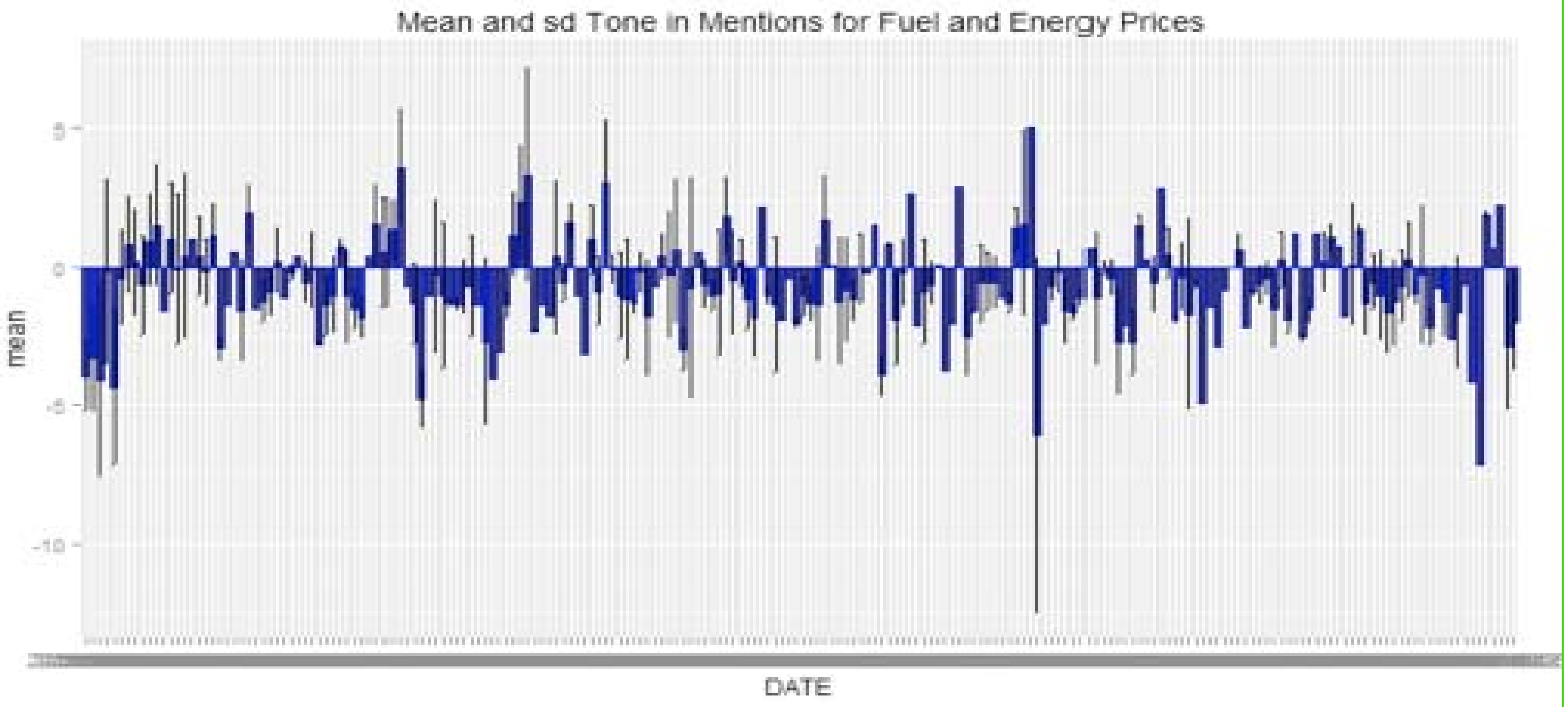

Figure 1. All mentions to fuel prices in Spain (February, the 18th 2015 to October, the 25th 2015)

Example 2 (V)

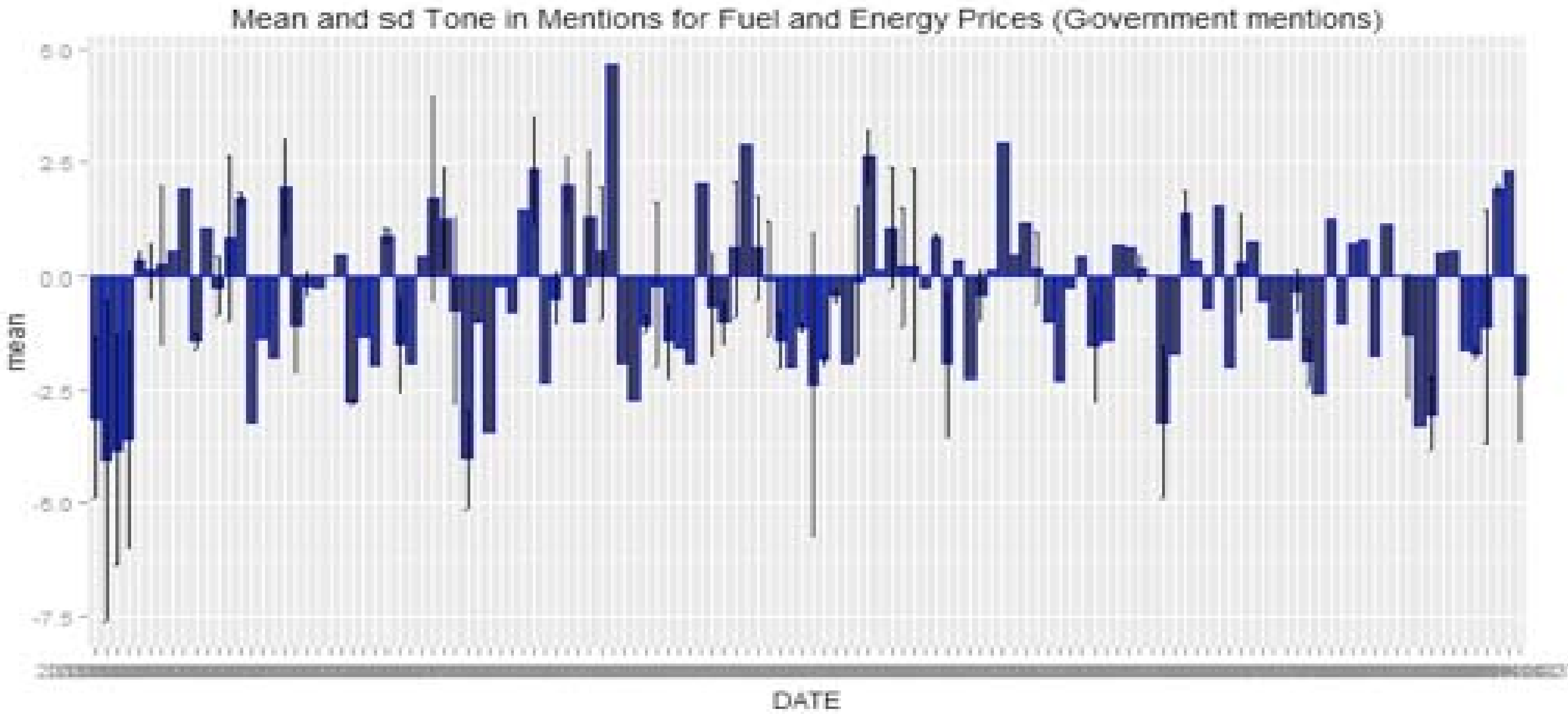

Figure 2. All mentions to fuel prices in Spain (February, the 18th 2015 to October, the 25th 2015) filtering words (government mentions)

Example 2 (VI)

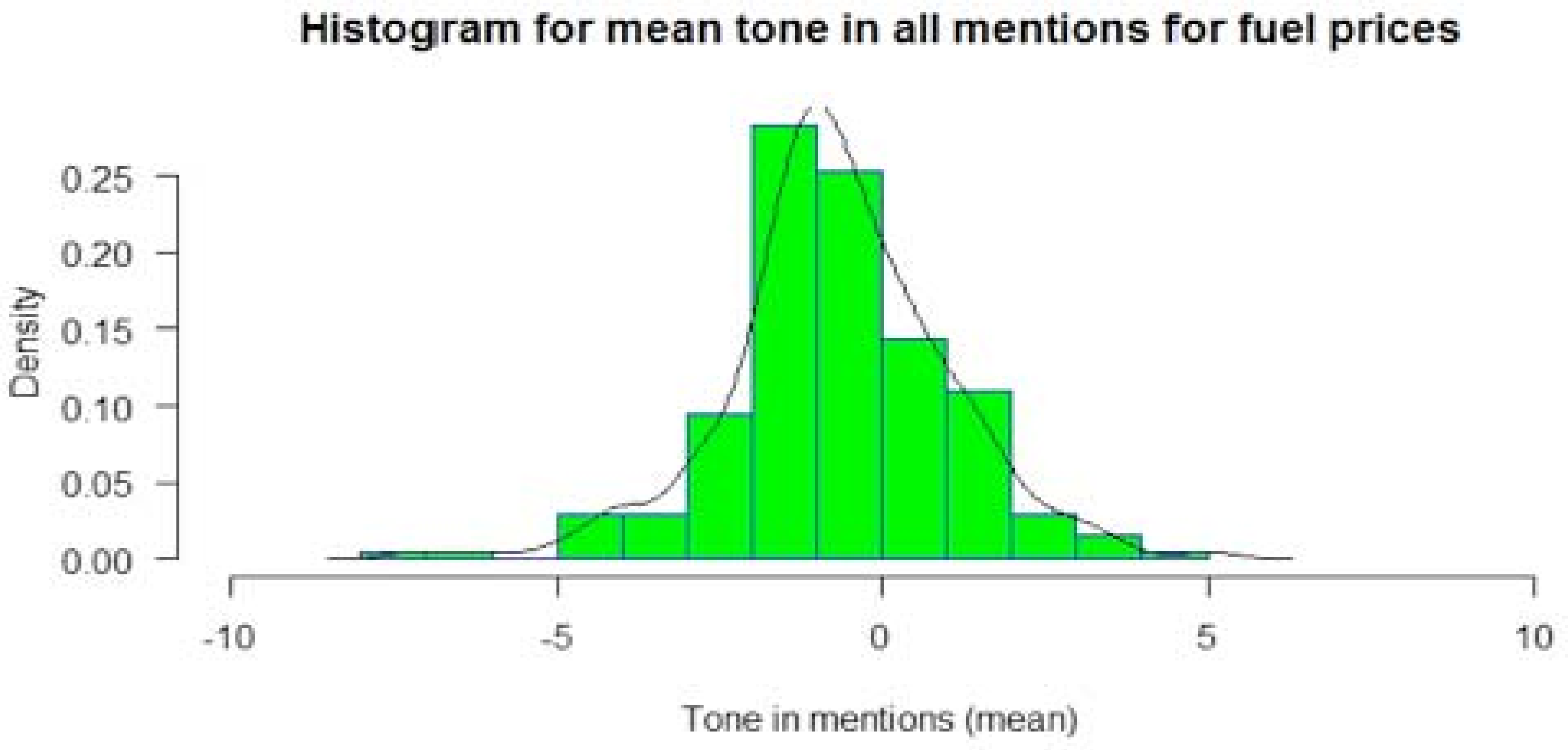

Figure 3. Histogram corresponding to all mentions to fuel prices in Spain (February, the 18th 2015 to October, the 25th 2015)

Example 2 (VII)

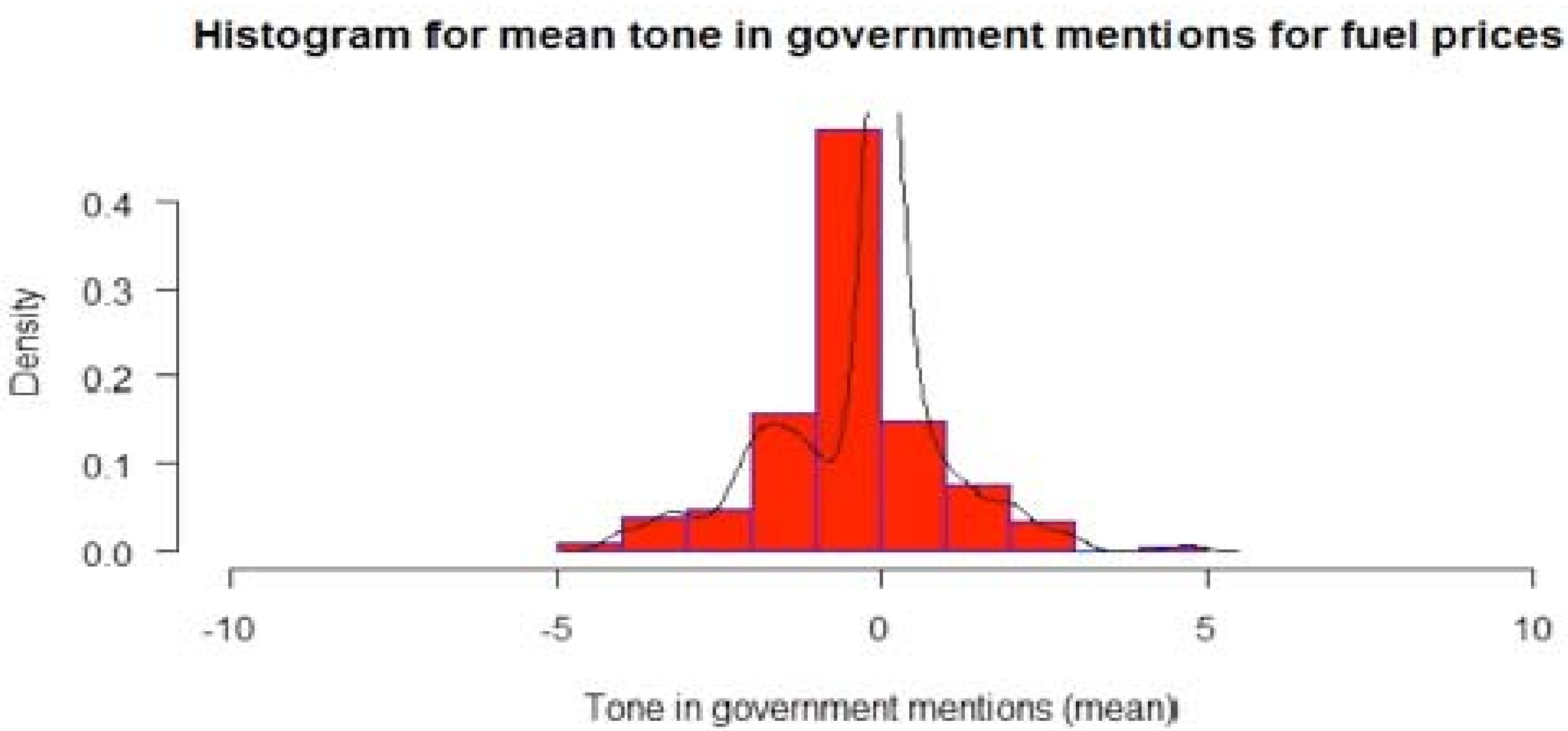

Figure 4. Histogram corresponding to all mentions to fuel prices in Spain (February, the 18th 2015 to October, the 25th 2015) filtering words (government mentions)

Example 2 (VIII)

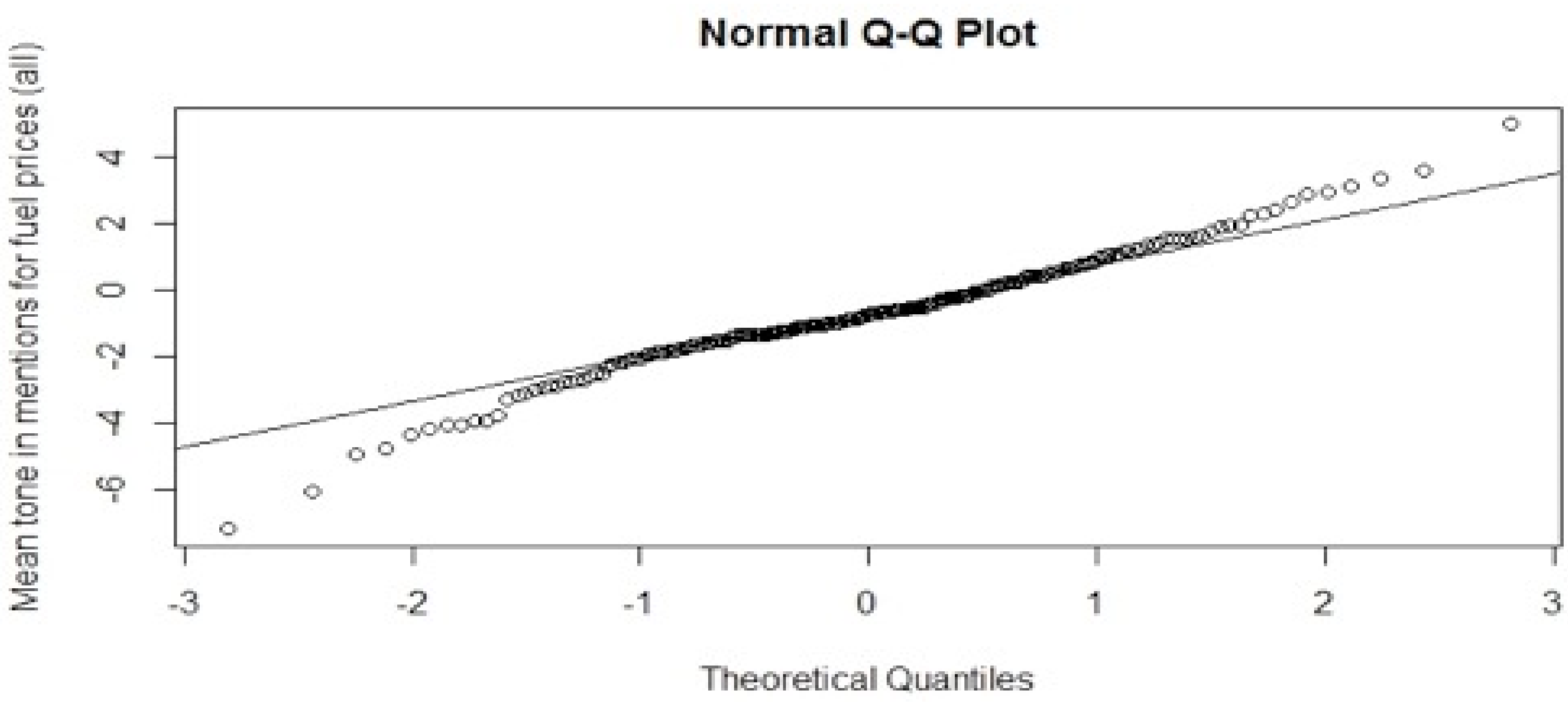

Figure 5. Q-Q plot for testing normality in mentions to fuel prices

Example 2 (IX)

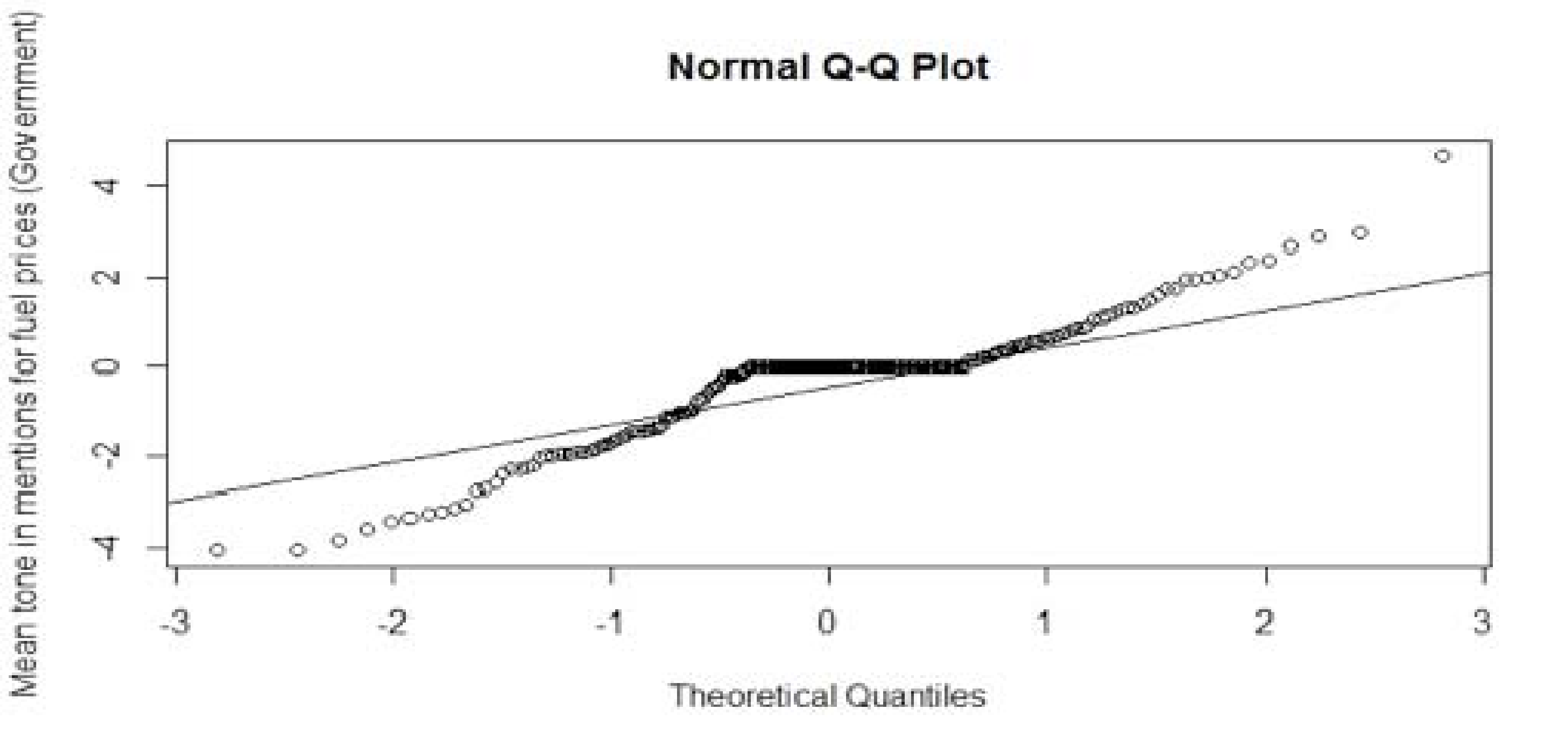

Figure 6. Q-Q plot for testing normality in mentions to fuel prices (filtering words)

Example 2 (X)

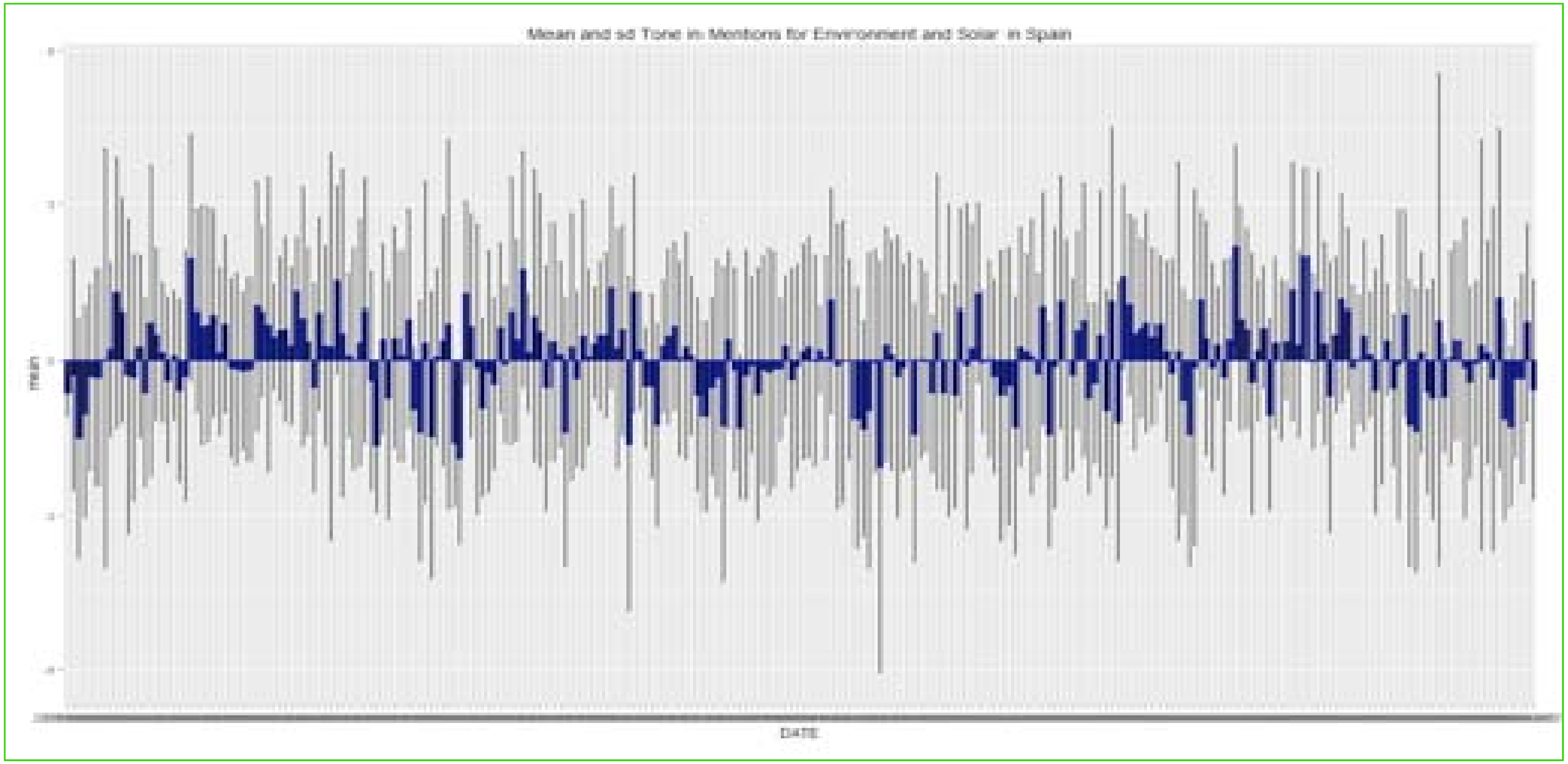

Figure 7. All mentions to environmental and solar energy in Spain (February, the 18th 2015 to October, the 25th 2015)

Example 2 (XI)

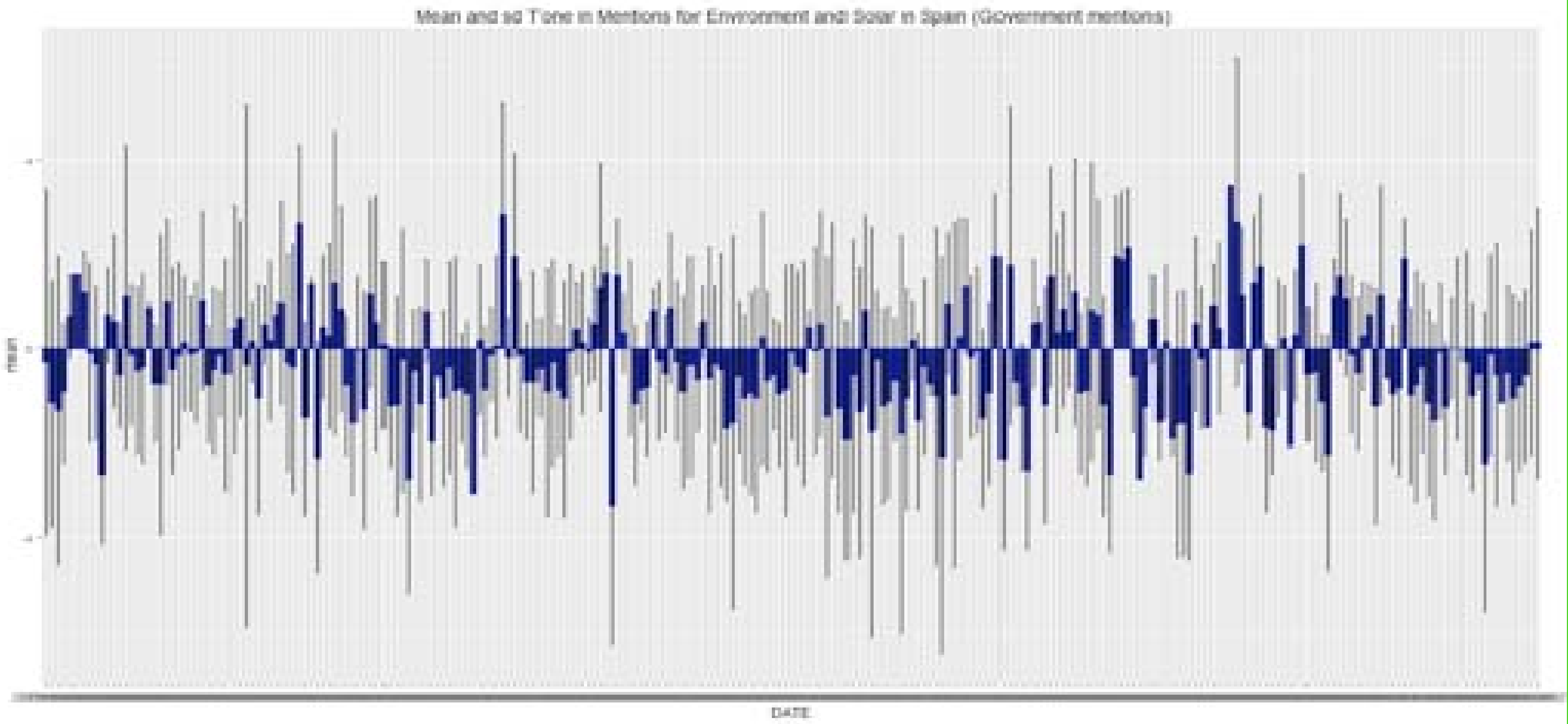

Figure 8. All mentions to environmental and solar energy in Spain (February, the 18th 2015 to October, the 25th 2015) filtering words (government mentions)

Example 2 (XII)

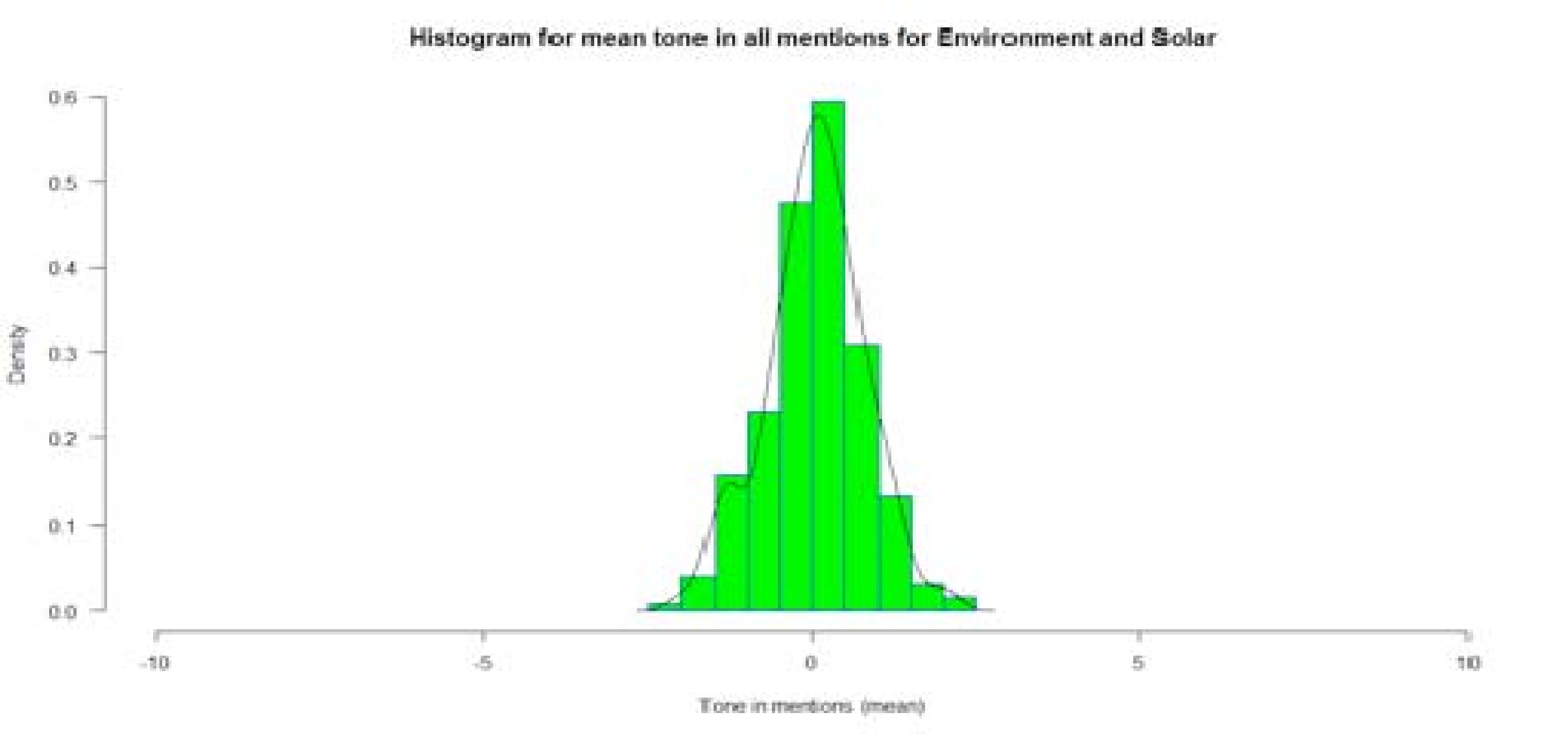

Figure 9. Histogram corresponding to all mentions to environmental and solar energy in Spain (February, the 18th 2015 to October, the 25th 2015)

Example 2 (XIII)

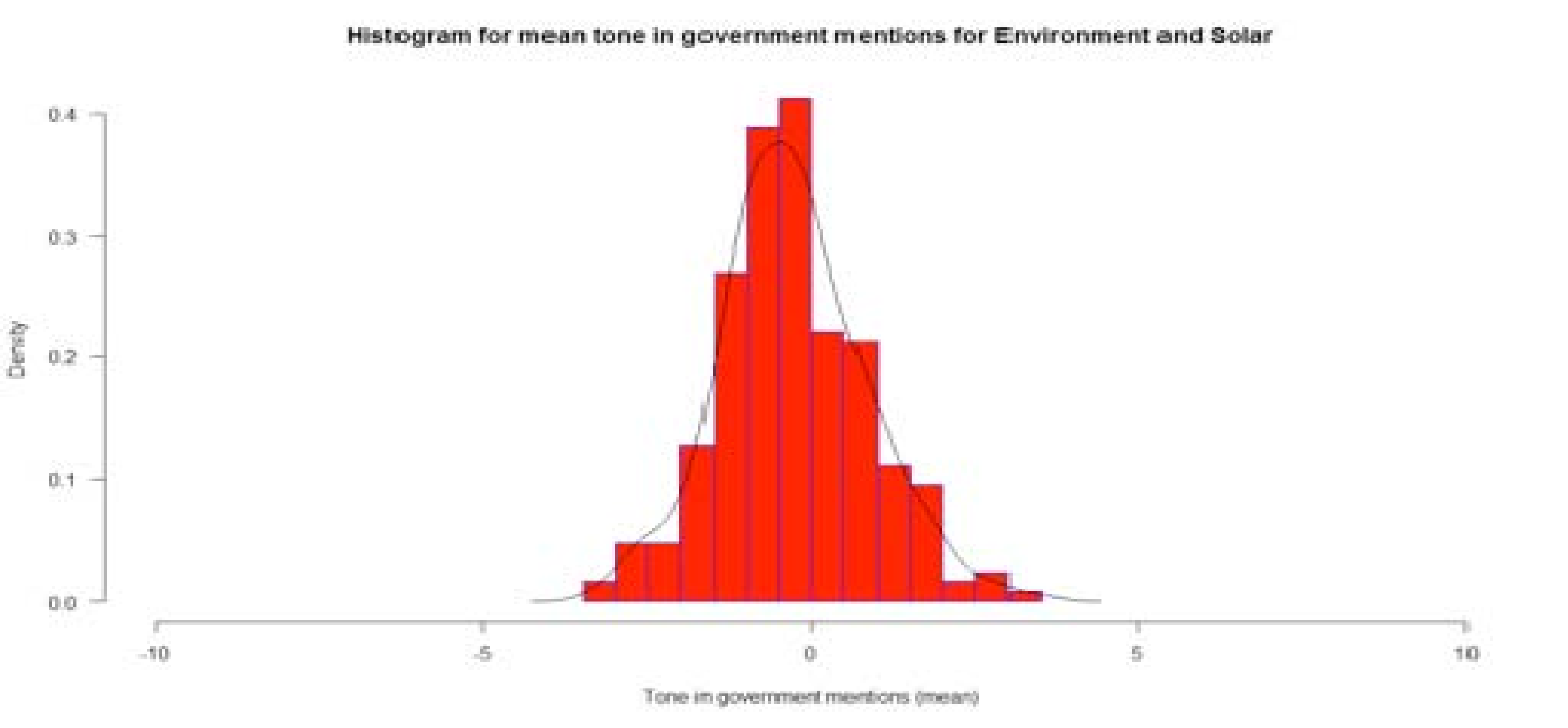

Figure 10. Histogram corresponding to all mentions to environmental and solar energy in Spain (February, the 18th 2015 to October, the 25th 2015) filtering words (government mentions)

Example 2 (XIV)

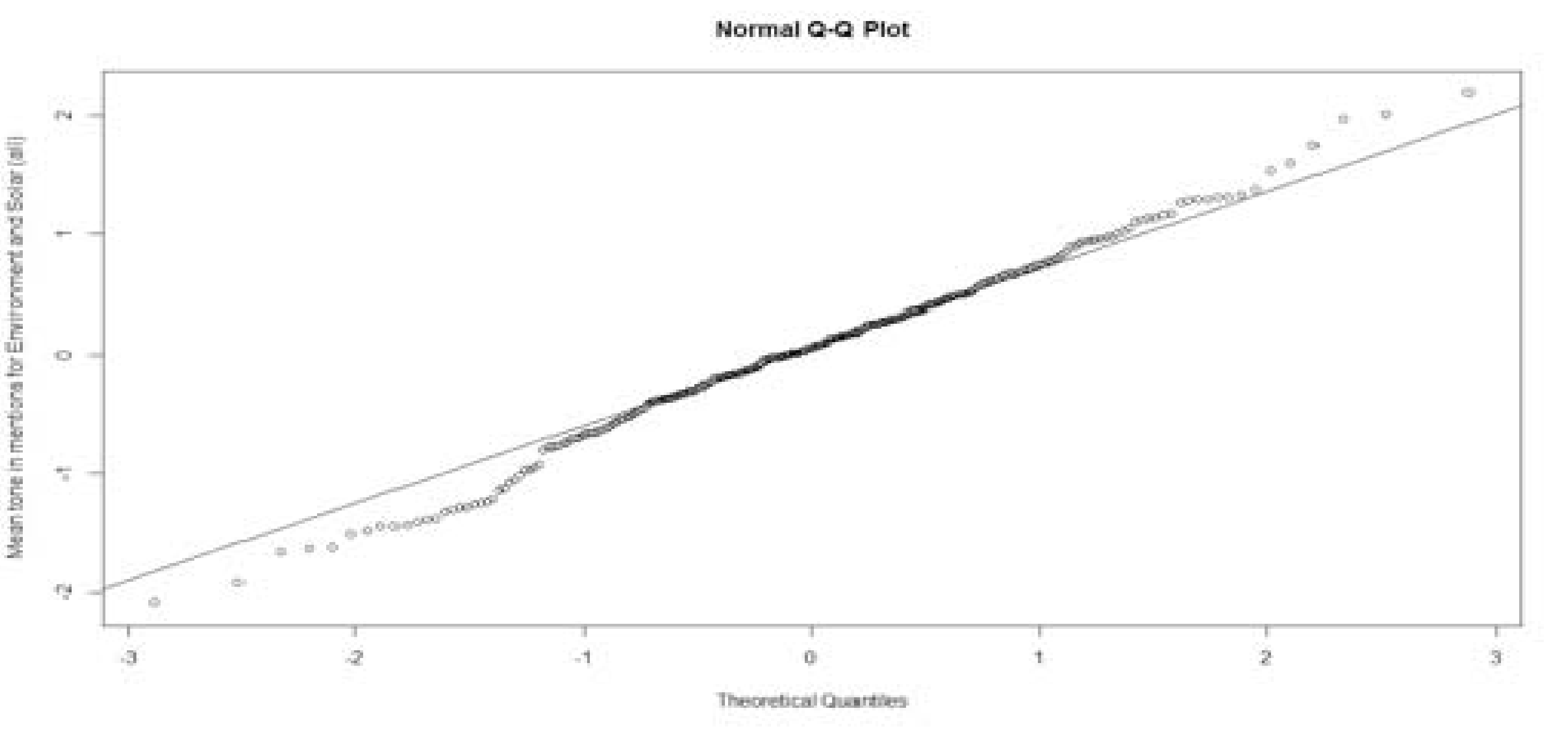

Figure 11. Q-Q plot for testing normality in mentions to environmental and solar energy

Example 2 (XV)

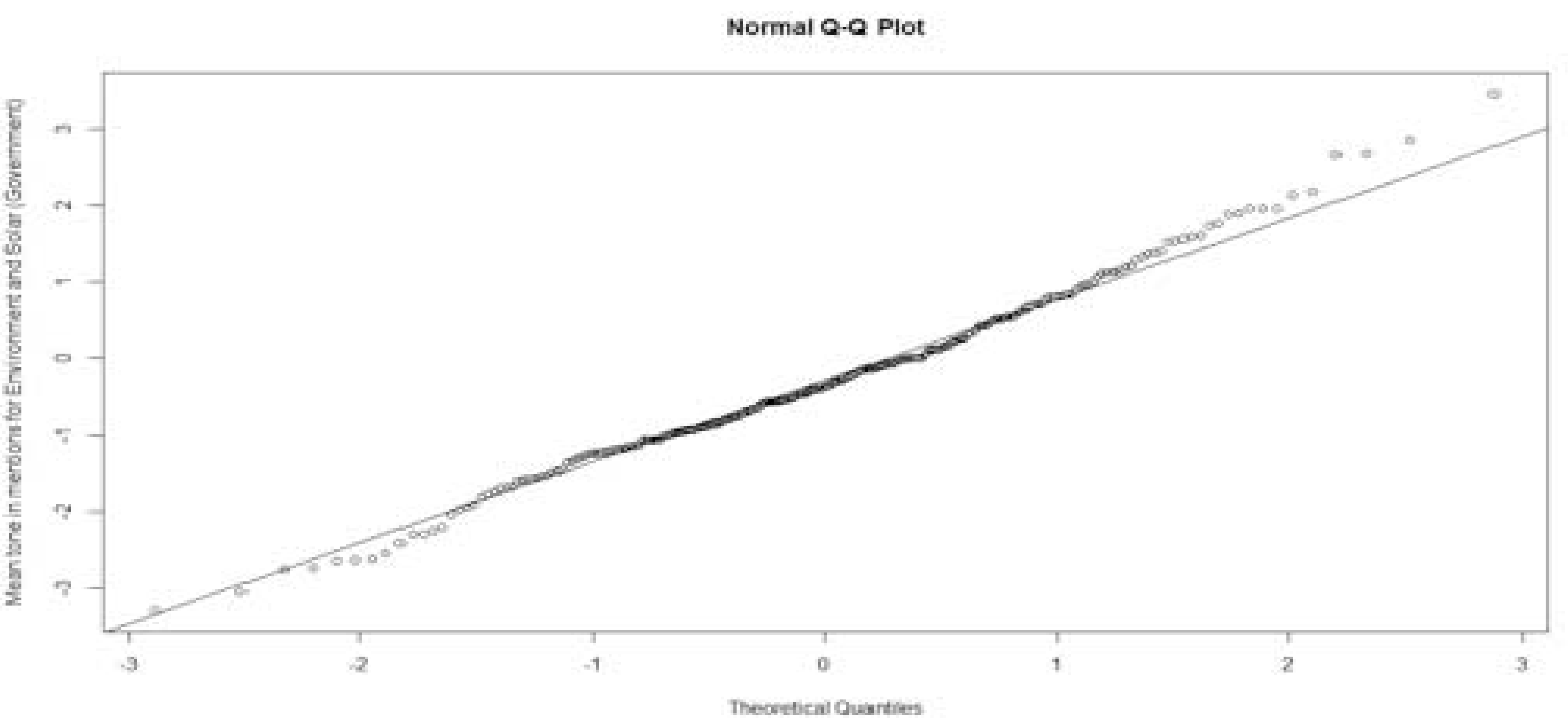

Figure 12. Q-Q plot for testing normality in mentions to environmental and solar energy (filtering words)

Example 2 (XVI)

- From all these analysis, I will only present at this moment some correlation (but I have also done some more formal correlation analysis through regression, i.e., demand models including sentiment indicators)
  - Coefficient of correlation log average prices of energy and meangovsolar index = -0.042 (significant at 0.32)
  - Coefficient of correlation average log prices of energy and meanallfuel index = -0.071 (significant at 0.09)
  - Coefficient of correlation log demand of energy and meanallfuel index = -0.045 (significant at 0.30)

Example 2 (XVII)

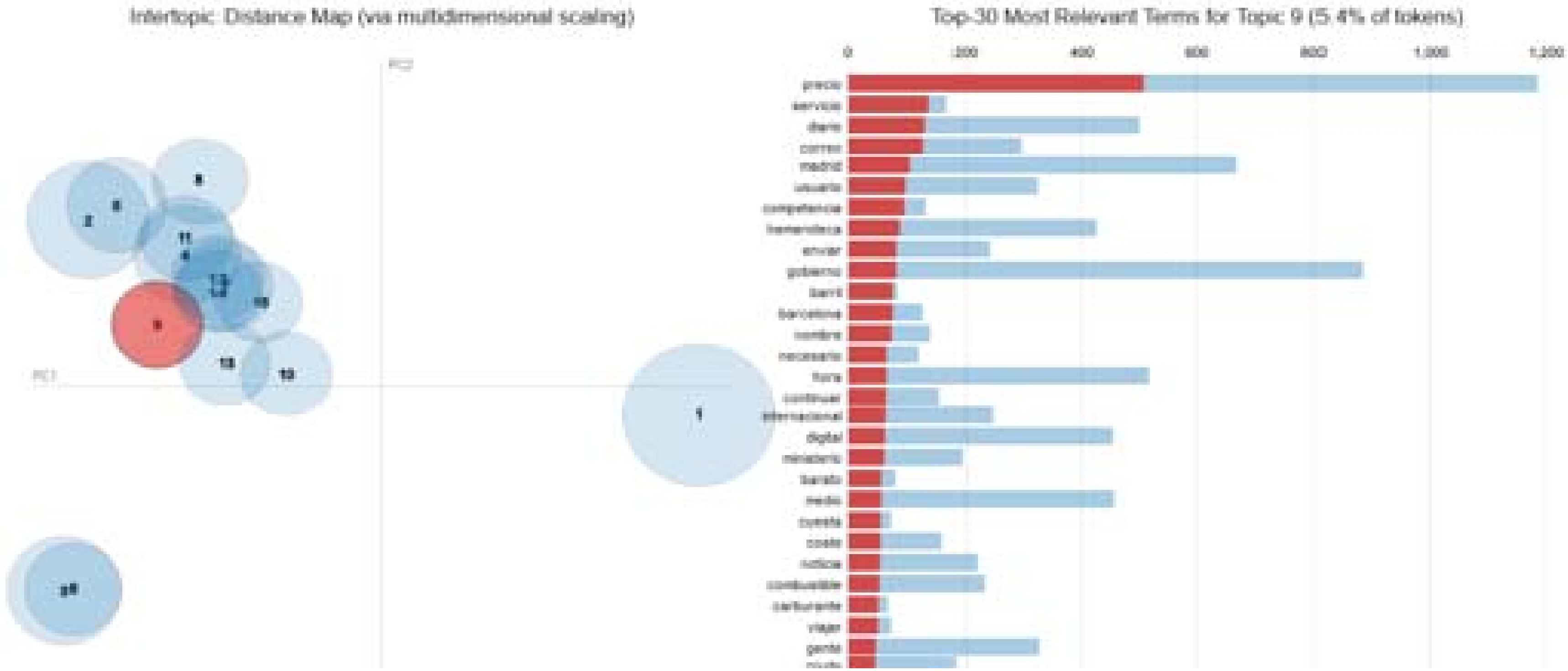

Figure 13. Analyzing “FUELPRICES” theme using topics model, text written in Spanish

Example 2 (XVIII)

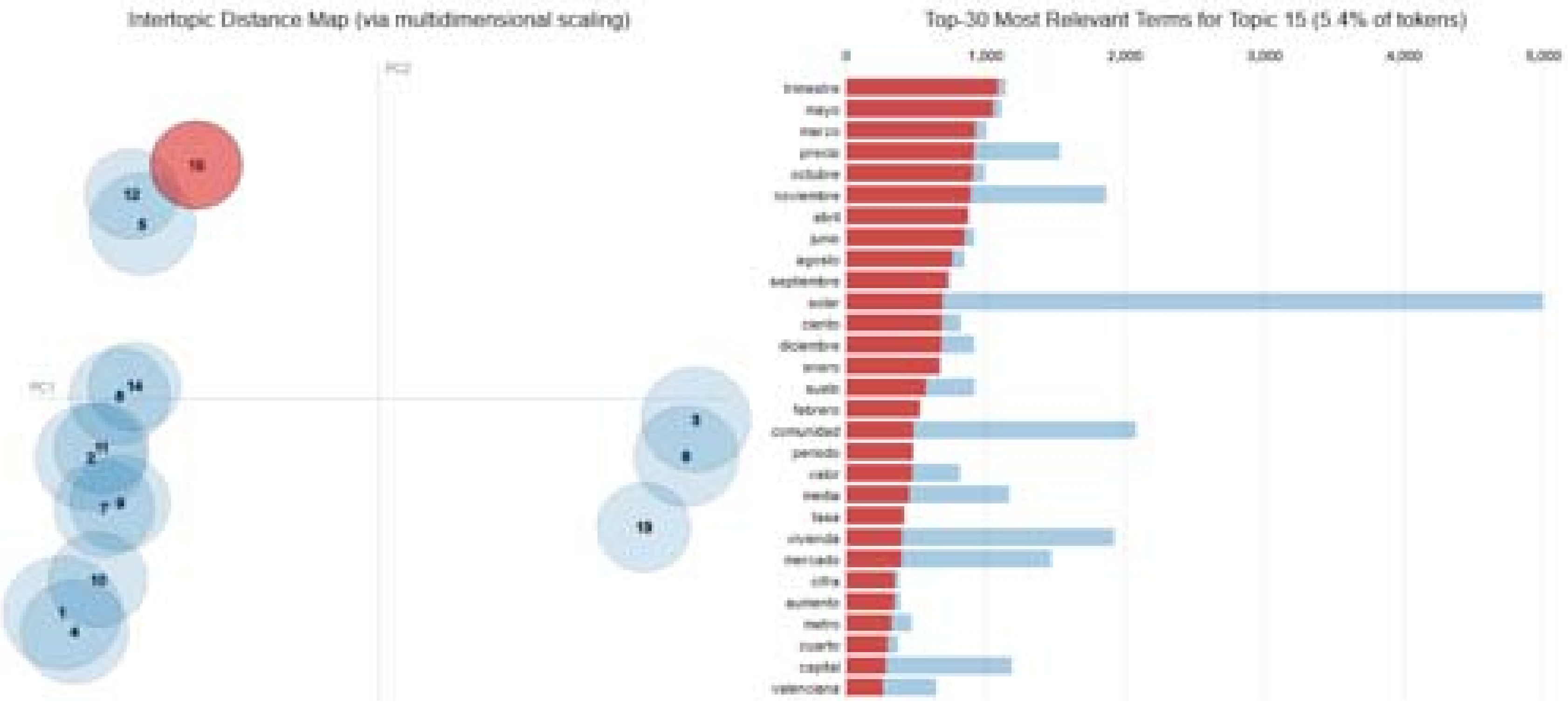

Figure 14. Analyzing “ENVSOLAR” theme using topics model, text written in Spanish

Example 2 (XIX)

- Our final aim:
  - To estimate demand models with daily data (coming from OMIE) including sentiment indicators on a daily base and see how news, media, opinion, etc. do affect energy demand (directly or indirectly through prices) to be able to:
    - Inform about the effects of policy decisions on demand through “non-market variables”
    - Analyze whether price-elasticities (or income elasticities) are affected
    - Etc.
  - Preliminary results:
    - Some negative effects on demand at short run and they influence market variables at specific dates

New methods to evaluate key variables in energy markets

Many thanks

Madrid, January the 27th 2016

FUNDACIÓN RAMÓN ARECES

economics for energy