



Electricity derivatives: an application to the futures Italian market

Laura Casula¹ · Giovanni Masala¹

Received: 10 March 2017 / Accepted: 17 June 2020 / Published online: 29 July 2020
© The Author(s) 2020

Abstract

Since the liberalization of electricity markets, electricity prices are more volatile and expansion in electricity derivatives trading occurs. Indeed, a well-known feature of electricity prices concerns its high volatility. For this reason, operators use power futures to hedge against unexpected risk deriving from adverse fluctuations of spot prices within the planned delivering period. Indeed, futures contracts permit to fix the price of electricity in advance for the use in the scheduled period. Our paper is devoted specifically to the Italian electricity market. In this respect, we examine empirical data from IDEX, the Energy Derivatives part of the Italian derivatives market IDEM, administered by “Borsa Italiana.” We finally survey the possible connections concerning futures and spot prices and, as a consequence, we deduce information about important indicators whereof the ex-post risk premium and the net convenience yield. For this purpose, we use several regression techniques to determine suitable explanatory variables inherent the Italian market for the ex-post risk premium and the net convenience yield.

Keywords Electricity markets · Futures · Risk premium · Convenience yield · Linear regression · Partial least squares regression

JEL Classification C02 · C52 · G13 · G14

✉ Giovanni Masala
gb.masala@unica.it

Laura Casula
laura.casula@gmail.com

¹ Dipartimento di Scienze Economiche e Aziendali, Università degli Studi di Cagliari, Cagliari, Italy

1 Introduction

As a consequence of the liberalization of electricity markets, electricity prices are extremely volatile, and expansion in electricity derivatives trading occurs. Besides, the electricity markets are currently coping with important transformations with the introduction of the category of renewable energies.

Pricing electricity derivatives is a challenging task because of the peculiar characteristics of the underlying stocks and commodities (namely non-storability, generation and transmission binds, seasonality feature and weather influence for example).

A striking peculiarity of electricity derivatives pricing is that electricity production and consumption must be in equilibrium at any time, to prevent network collapses. Besides, the no-arbitrage paradigm does not hold for commodity futures' prices (unlike to typical investment assets). Indeed, storage costs, for example, are often needed.

Futures are used in the framework of energy risk management to hedge against adverse variations of spot prices occurring within the delivery period. Futures contracts aim to fix the price of electricity beforehand to manage efficiently the costs of future scheduled consumptions. The investors can also use futures to get a position in electricity markets but avoiding the problems of holding the underlying commodity.

Our paper is devoted to the Italian market. In this regard, we examine the IDEX market which is the Energy Derivatives part of IDEM (a company of London Stock Exchange Group), the Italian derivatives market administered by "Borsa Italiana." This regulated market started in 2008, and it currently negotiates both baseload and peakload power futures (on a monthly, quarterly and yearly basis).

The underlying power spot market (which is the day-ahead market) is administered by GME ("Gestore Mercati Energetici") which is a company managed by the state. Within this market, the so-called single national purchase price PUN ("Prezzo Unico Nazionale") is estimated hourly through a weighted average of the zonal prices issued by the day-ahead market. We examine then in detail the risk premium and the convenience yield associated with futures contracts. The risk premium permits to analyze the links between spot and futures prices and the convenience yield represents the advantage of holding the physical commodity.

We list hereafter some significant and recent contributions concerning electricity Markets.

We present at first some general references about electricity derivatives.

A general and exhaustive survey about electricity derivatives is presented in Aid book (2015).

Benth and Krühner (2015) examine the pricing of typical energy derivatives. Besides, they set up a theoretical framework for modeling forwards contracts in an infinite-dimensional context.

Branger et al. (2010) aim to price electricity derivatives with hourly basis. They focus then on the German market, and they emphasize that risk premiums are time-dependent. They linked this variation with the dynamic hedging fluctuations incurred by consumers and producers.

Redl and Bunn (2010) present a classification of the factors which influence the ex-post forward premium (such as fuels and higher moments). The authors apply econometric models, and they examine specifically the Western European Power market (EEX).

Diko et al. (2006) survey three important European energy markets (namely German, Dutch and French). The authors deduce then the dynamic features of the risk premiums for these markets. They highlight significant risk premiums from empirical data. Finally, they find evidence that the risk premium structure depends on the skewness and the variance of the spot prices. For this purpose, the effect of the skewness is decreasing on time to maturity. Then, the variability prevails and forces the risk premium to decrease.

Then, we present some contributions focused mainly on electricity futures.

Ballester et al. (2016) analyze futures products within the Spanish market MIBEL (Iberian Electricity Market). The authors investigate some price relationships, for example, between futures market and the forward and spot market.

Bauwens et al. (2013) examine electricity futures within the European Energy Exchange index. They take into account a particular multiplicative dynamic conditional correlation model (namely mDCC) to investigate about the dynamic feature of the volatility and the correlation.

Flasza et al. (2011) apply an econometric model for the pricing of long-term electricity futures. For this purpose, the authors use data coming from EEX AG, and they deduce some typical characteristics of the German market. Finally, the authors take into account some fuels prices (oil, coal, natural gas), emission allowances and some financial assets (for example Germany bonds with 1 and 10 years maturity and EUR/USD exchange rate).

Fleten et al. (2015) investigate, in their survey, the dynamic links involving electricity futures prices and some fossil fuels contracts. They examine data from Germany and the UK markets, and they find evidence that time-varying relations between these contracts are more suitable.

The links between spot and futures prices have also been investigated by Handika and Trück (2013). The authors analyzed some regional Australian electricity markets. They highlighted that, for several regions, the (ex-post) risk premiums are significant and positive. Another finding concerns the seasonality of the premiums.

Islyayev and Date (2015) set up a one-factor model to price electricity futures contracts. This model allows random volatility, and they perform numerical applications by using a database from the Nord Pool market.

Lucia and Torró (2011) investigate empirically in their survey how spot and futures prices in electricity markets are linked. For this purpose, they use some futures contracts data inherent the Nordic Power Exchange market. They prove the presence of average positive risk premiums when considering short-term contracts. The amplitude of the premiums exhibits seasonality as already proved in other surveys. Besides, the authors observe that the risk premiums depend on higher moments of the future spot prices (variance and skewness) in particular periods.

Shawky et al. (2003) analyze the features of the spot and futures prices coming from the New York Mercantile Exchange (for the California–Oregon Border electricity market). They also investigate the links involving spot and futures prices/returns using a

dynamic approach. For this purpose, they use an Exponential GARCH model for prices and a vector autoregression model for returns.

We conclude with the following papers which deal with risk premium and convenience yield in the Nord Pool market.

Botterud et al. (2010) examined spot and futures prices data in the Nord Pool market. They discover that futures prices are significantly higher than the corresponding spot prices. They deduce that the convenience yield is on average negative. Besides, it shares a seasonal trend and it is linked to hydro reservoirs levels as the hydroelectric component of this market is sizeable. In this survey, the authors highlight that the convenience yield and risk premium are both negative and then opposite respect to other commodities.

Weron and Zator (2014) applied correctly the linear regression to study the links between the spot and futures prices in electricity markets. In particular, they investigate the errors generated in simultaneity problems as well as correlated measurements and the consequences of seasonality on the overall results. Finally, they discover (application to the Nord Pool market) that the levels of water reservoir have a positive effect on the risk premium. Nevertheless, this expected result disproves the findings of Botterud et al. (2010).

In this contribution, we investigate about Italian futures traded on the IDEX market. This market is rather young (set up in 2008) and, at our knowledge, no academic surveys are available yet. We examine the possible interrelations concerning futures and spot prices, and we deduce some consequences concerning the ex-post risk premium and the net convenience yield. We investigate thoroughly the explanatory variables which can affect ex-post risk premium and net convenience yield by using several regression techniques (linear regression, partial least squares regression and principal components regression) which generalize the surveys of Botterud et al. (2010) and Weron and Zator (2014). The associated adequacy tests (such as the adjusted *R*-squared) showed a good accordance with the regression models proposed to explain ex-post risk premium and net convenience yield.

According to the references listed before, some typical explanatory variables are certainly fuel prices (oil, coal, natural gas) and higher spot price moments (variance, skewness, and kurtosis). We then tried to enlarge the set of variables in order to include some calendar variables or other variables specific to the Italian market.

The paper is partitioned as follows. In the current Sect. 1, we have introduced the objectives of our survey, and we have listed state of the art about electricity futures in several markets. Section 2 is dedicated to theoretical aspects involving spot and futures electricity prices and the implementation to the Italian IDEX market. The description of the database used is preliminarily presented. The factors which affect the net convenience yield and ex-post risk premium are analyzed in Sect. 3 through suitable general regression techniques. Finally, Sect. 4 concludes.

2 Power derivatives

2.1 Database description

IDEX represents the Energy Derivatives part of IDEM which is the Italian derivatives market (conducted by Borsa Italiana). It is part of London Stock Exchange Group (information can be retrieved in the brochure edited by “Borsa Italiana”). The IDEX market launched in 2008 trades both baseload and peakload power futures (note that peakload products were introduced later on May 2013). A well-known characteristic of electricity price dynamics is represented by high volatility values. For this purpose, electricity futures may be used by the operators as hedging tools to face unexpected risks. This produces a more efficient business management.

We describe hereafter these products’ characteristics.

Regards the underlying power spot market:

- The day-ahead electricity market is administered by GME (“Gestore Mercati Energetici”) which is state-owned;
- The so-called single national purchase price (namely “Prezzo Unico Nazionale”—PUN), estimated hourly, is determined by the weighted mean of the zonal prices coming from the day-ahead market.

The traded futures are categorized as follows: monthly, quarterly and yearly products (both baseload and peakload). The specific features of these contracts are illustrated in the IDEX brochure edited by Borsa Italiana: http://st.formazione.ilsolo24ore.com/a/energy2014/brochure/IDEX_Brochure.pdf.

We just note that these contracts are quoted in €/MWh.

The participants of IDEX markets (suppliers/producers, energy consumers, and financial players/traders) use the futures to satisfy their specific goals.

The access to trading occurs via a financial intermediary or direct membership.

The platform used for IDEX trading is the so-called SOLA[®] platform (also used for the IDEM market). Trading also occurs on the Trayport GlobalVisionSM Portal.

Cassa di Compensazione e Garanzia (CC&G) is the main agency and ensures the execution of all the market transactions. CC&G has a separate section (Energy Derivatives Section), and it comprises a Default Fund to manage the products traded on IDEX. Additional information can be retrieved by consulting Borsa Italiana website (<http://www.borsaitaliana.it>).

Historical time series for load and prices on an hourly basis are available on the website <http://www.mercatoelettrico.org/It/Download/DatiStorici.aspx> (from April 20, 2004, to present).

Besides, futures prices for the IDEX market can be freely downloaded from the website <http://www.borsaitaliana.it/borsaitaliana/statistiche/mercati/commodities/commodities.htm> (from December 2011 to present). We will conduct our survey only for monthly contracts (with delivery period one to three months) which are the ultimate contracts.

We illustrate in Fig. 1 the baseload monthly contracts for the next three calendar months delivery. The main statistics for baseload, peakload, and spot prices are then given in Table 1.

We can highlight from Fig. 1 that the futures with delivery 1–3 months are strongly correlated with each other. The correlation coefficients for each pair of futures are 98.14% (1–2 months); 95.58% (1–3 months) and 97.71% (2–3 months). The futures are also strongly correlated with the spot prices. Indeed the correlations are, respectively: 88.13%, 85.80%, and 82.65% (for delivery 1–3 months, respectively). The baseload futures share then the same trend as the underlying spot prices and the correlation decreases with delivery time but remains strong.

The same conclusions hold for peakload futures. The correlation coefficients for each pair of futures are 92.20% (1–2 months); 82.59% (1–3 months) and 91.34% (2–3 months). The futures are also strongly correlated with the spot prices. Indeed the correlations are, respectively: 78.82%, 76.94%, and 69.85% (for delivery 1–3 months, respectively). The peakload futures share then the same trend as the underlying spot prices and the correlation decreases with delivery time but remains strong. Note that the correlations with spot prices are lower with respect to the case of baseload contracts.

We deduce, as expected, that monthly futures are rather less volatile than the underlying spot prices. Besides, baseload and peakload contracts present very similar features.

2.2 Spot and futures prices

We can investigate the links involving commodity spot and futures prices through two specific theories. The first one deals with the cost and measures the interest of keeping inventories. The second approach takes into account the standard *risk*

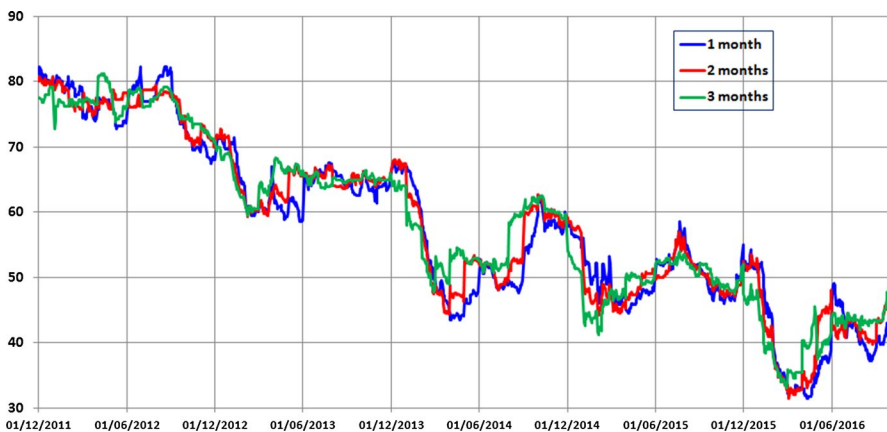


Fig. 1 Baseload monthly contracts (next three calendar months delivery) for the period Dec. 2011/Sept. 2016

Table 1 Summary statistics: monthly baseload/peakload futures contracts versus spot prices

Indicator	Baseload 1 month	Baseload 2 months	Baseload 3 months	Peakload 1 month	Peakload 2 months	Peakload 3 months	Spot price
Mean	57.58	58.08	58.32	57.56	58.32	58.45	59.72
Median	56.63	58.49	58.44	56.50	56.88	57.75	59.35
SD	13.08	12.82	12.49	11.36	11.13	10.65	15.01
Coeff. of variation	0.23	0.22	0.21	0.20	0.19	0.18	0.25
Kurtosis	-0.89	-1.00	-1.05	-0.79	-0.87	-0.99	0.99
Skewness	0.12	0.08	0.11	-0.10	0.02	0.07	0.54
Range	50.75	49.24	48.30	45.50	44.16	40.25	111.68
Minimum	31.50	31.51	33.00	32.75	34.50	37.00	24.99
Maximum	82.25	80.75	81.30	78.25	78.66	77.25	136.67
Count	1220	1220	1220	861	861	861	1220

premium to assess the link between short- and long-term prices. These theories must be adjusted carefully to fit the peculiarities of the electricity markets.

Concerning the first approach, inventories are fundamental to determine the price of commodities which can be stored. Indeed, in the light of storage theory, the divergence between spot prices and futures prices becomes clearer. The concept of convenience yield has been introduced to measure the “convenience” in storing a given commodity. Specifically, it represents a sort of liquidity premium linked to the benefit of keeping this commodity. Besides, it represents the possible supply of this commodity at a future date, and it decreases with respect to the stocks quantity.

The formula for the futures price, $F_{t,T}$, at a time t for delivery at time $t+T$ assumes the usual paradigm of no arbitrage among the spot and the futures prices:

$$F_{t,T} = S_t \cdot e^{r_T + u_T - y_T} \quad (1)$$

where S_t denotes the spot price at the time t , r_T is the usual risk-free interest rate for the holding period T , u_T denotes the storage expense and y_T is the *convenience yield* over the period T . Note also that the storage expense and the convenience yield are given as a fraction of spot price.

We say that futures market presents “*backwardation*” when the actual spot price exceeds futures price. In this case, the *net convenience yield* (defined as $y_T - u_T$) is positive and greater than the risk-free interest rate. In the opposite situation, we introduce the term “*contango*.” The percentage of backwardation for baseload is 60.25%, 56.31% and 53.36% for 1–3-month delivery. The percentage for peakload contracts is, respectively, 20.91%, 20.33%, and 24.04%.

We note that for baseload contracts backwardation status prevails, whereas contango prevails for peakload contracts.

Regards the second approach, the price of a futures contract $F_{t,T}$, at a time t for delivery at time $t+T$ is expressed as the expected future spot price, $E_t(S_{t+T})$, plus a so-called risk premium associated with the underlying commodity. We denote here i_T an opportune risk-adjusted discount rate associated with the commodity over the period T . We can then express the futures price as:

$$F_{t,T} = E_t(S_{t+T}) \cdot e^{r_T - i_T} = E_t(S_{t+T}) \cdot e^{-p_T} \quad (2)$$

where $p_T = i_T - r_T$ represents the risk premium associated with this commodity.

We deduce from this formula that futures price coincides with the expected future spot price if the risk premium vanishes. This occurs when the equality $i_T = r_T$ holds.

Finally, we observe that we can introduce the risk premium concept even when the commodity involved is not storable. Instead, regards the second approach, note that the no-arbitrage paradigm does not hold anymore. In this situation, we cannot buy a non-storable commodity in a spot market and afterward sell it in a futures market to realize a risk-free condition.

Additional details about these two approaches can be retrieved in Botterud et al. (2010) and Weron and Zator (2014).

2.3 An application to the Italian IDEX market

In Fig. 2 we present the spot prices' dynamics and compare them with the prices of baseload futures contracts given in Fig. 1. We observe that the futures with one to three-month delivery share the same trend as the spot price, whereas this trend slightly decreases with respect to time delivery (the same feature also holds for peakload contracts). The high values of linear correlations involving the spot and the futures prices unveiled previously confirm this finding.

We illustrate then the interrelations concerning spot and futures prices through the two approaches just described. We determine at first the characteristics concerning the net convenience yield. Afterward, we investigate the presence of eventual risk premium concerning the futures market. (as in Botterud et al. 2010; Weron and Zator 2014 surveys).

As outlined in Sect. 2.2, the convenience yield represents the advantage of holding the physical commodity. Whenever the net convenience yield is positive, this entails that the convenience yield proves to exceed the costs of storage. Besides, we deduce that if the net convenience yield exceeds the risk-free interest rate, as a consequence the spot price is greater than the futures price.

We now estimate the (net) convenience yield for monthly futures contracts with one to three months until delivery by using all available historical data. We assume here a *zero* interest rate.

Remark We can use, as a proxy for the risk-free interest rate, the return of Treasury Bills “BOT” (“Buoni Ordinari del Tesoro”) with six-month maturity. The yearly returns, retrieved from <http://www.dt.tesoro.it>, have been the following: 0.80% (2013), 0.43% (2014), 0.052% (2015) and -0.161% (2016). Besides, when considering short periods (1–2–3 months), the returns become negligible so that we can assume, without loss of generality, a zero risk-free interest rate.

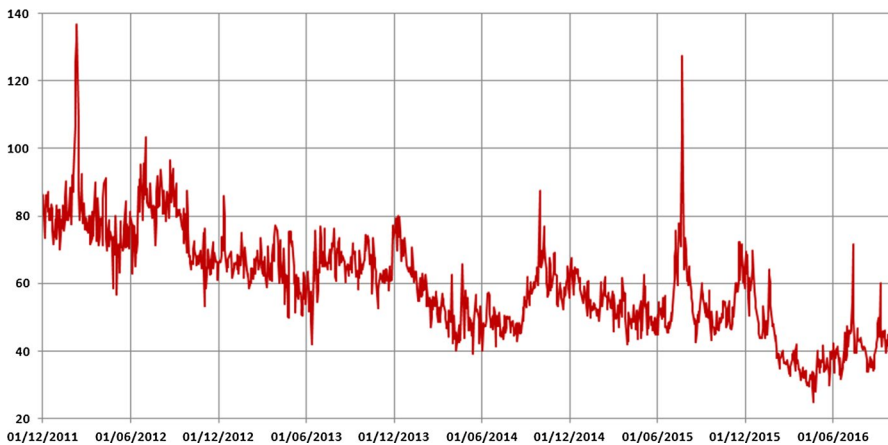


Fig. 2 Spot prices for the interval Dec. 2011/Sept. 2016

For this purpose, we deduce from Eq. (1) the formula which gives the net convenience yield:

$$cv_{t,T} = y_T - u_T = \ln \frac{S_t}{F_{t,T}} \quad (3)$$

where $cv_{t,T}$ is the net convenience yield, at day t , for holding period T months, S_t represents the spot price at day t , $F_{t,T}$ represents the futures price at day t for contract with delivery in months $t + T$.

Table 2 illustrates some statistics for the net convenience yield. We observe that for baseload contracts, the net convenience yield is on average positive for every holding period. We deduce that the spot price reveals to be generally higher than the futures price (as observed before). Besides, its standard deviation increases with respect to the time of delivery. Furthermore, the extreme values differ widely. For example, the baseload future with 3-month delivery has minimum value -0.50 and maximum value $+0.88$. On the contrary, for peakload contracts, the net convenience yield is on average negative for every holding period so that the spot price reveals to be generally lower than the futures price (indeed contango status prevails). The other characteristics are similar. Some test statistics are included in Table 2.

We present in Fig. 3 the net convenience yield only for baseload monthly futures with delivery 1, 2 and 3 months.

Figure 3 highlights some seasonal patterns (as well as for peakload version). The futures market alternates backwardation and contango states.

In spite of the previous discussion, convenience yield notion may not be suitable within electricity markets. Indeed, the profit from detaining electricity is not evident and quantifiable; this is still a challenging problem.

Let examine the second approach again. We have stated the forward price is the sum of the expected spot price with an (ex-ante) risk premium. As a consequence, the ex-ante risk premium is understood as a prevision of the spot price minus the forward price. We deduce then from Eq. (2):

$$p_T = \ln \frac{E_t(S_{t+T})}{F_{t,T}} \quad (4)$$

where $E_t(S_{t+T})$ is the expected value (performed at the time t) concerning the spot price at the time $t + T$ and $F_{t,T}$ represents the price at the time t of a futures contract with delivery period at the time $t + T$.

In practical terms, it may be suitable to use as a proxy of the ex-ante risk premium p_T the effective ex-post risk premium. The *ex-post risk premium* estimation requires substituting the expectation $E_t(S_{t+T})$ by the actual value at the time $t + T$, i.e., S_{t+T} .

Thus, the so-called ex-post risk premium can be estimated as:

$$\hat{p}_T = \ln \frac{S_{t+T}}{F_{t,T}} \quad (5)$$

Finally, as a consequence of a straightforward calculation (see Weron and Zator (2014)) we deduce that the net convenience yield is the result of the difference

Table 2 Summary statistics: net convenience yield base-load and peakload (next three calendar months)

Indicator	Base-load 1 month	Base-load 2 months	Base-load 3 months	Peakload 1 month	Peakload 2 months	Peakload 3 months
Mean	0.03** ($p=1$)	0.02** ($p=1$)	0.02** ($p=1$)	-0.08** ($p=1$)	-0.09** ($p=1$)	-0.10** ($p=1$)
Median	0.03	0.02	0.01	-0.08	-0.10	-0.11
SD	0.11** ($p=0.9892$)	0.12** ($p=0.9892$)	0.14** ($p=0.9892$)	0.13** ($p=0.9872$)	0.13** ($p=0.9872$)	0.15** ($p=0.9872$)
Coeff. of variation	3.52	5.75	8.88	1.60	1.42	1.58
Kurtosis	4.39	3.25	2.35	4.02	2.52	1.46
Skewness	0.69	0.64	0.52	0.69	0.60	0.46
Range	1.23	1.26	1.38	1.31	1.28	1.37
Minimum	-0.42	-0.43	-0.50	-0.56	-0.55	-0.64
Maximum	0.80	0.83	0.88	0.75	0.73	0.73
Count	1220	1220	1220	861	861	861

**Significance level 99.9% (p value included between parentheses)

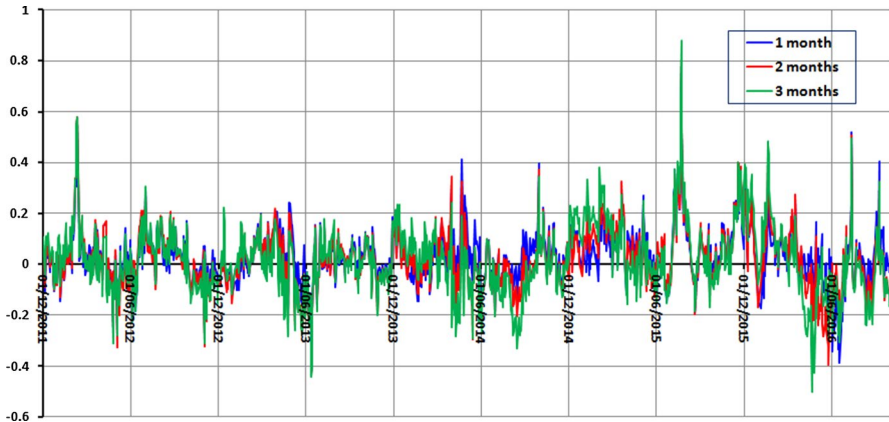


Fig. 3 Net convenience yield—baseload monthly (next three calendar months)

between two terms: the effective risk premium and the change of the spot price in the interval $[t, t + T]$. That is:

$$cv_{t,T} = \ln \frac{S_t}{F_{t,T}} = \ln \left(\frac{S_{t+T}}{F_{t,T}} \cdot \frac{S_t}{S_{t+T}} \right) = \hat{p}_T - \ln \frac{S_{t+T}}{S_t} \simeq \hat{p}_T - \Delta S_{t,T} \quad (6)$$

where $\Delta S_{t,T}$ represents the variation of the spot price in the interval $[t, t + T]$.

Next, we present in Fig. 4 the ex-post risk premium only for baseload monthly futures with delivery 1–3 months.

Table 3 reveals the main statistics for the ex-post risk premium for monthly baseload and peakload futures contracts with delivery 1–3 months.

We note that for baseload and peakload contracts, the ex-post risk premium is on average negative for every holding period and its volatility increases respect to delivery time.

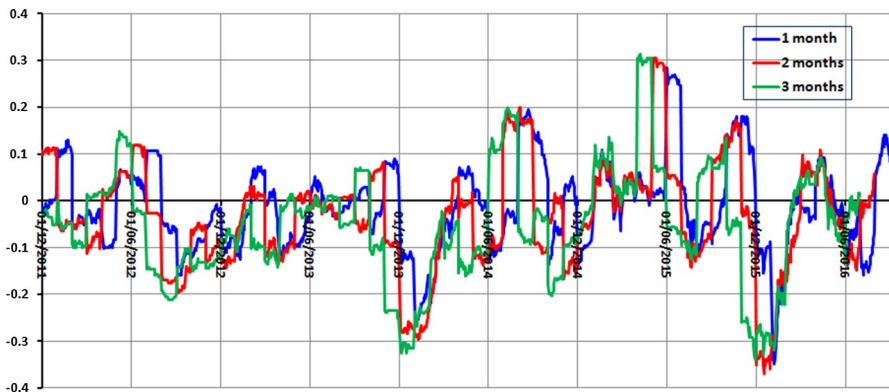


Fig. 4 Ex-post risk premium—baseload monthly (delivery next three calendar months)

Table 3 Summary statistics: ex-post risk premium—baseload and peakload monthly (delivery next three calendar months)

Indicator	Baseload 1 month	Baseload 2 months	Baseload 3 months	Peakload 1 month	Peakload 2 months	Peakload 3 months
Mean	-0.01	-0.03	-0.05	-0.02	-0.03	-0.05
Median	-0.02	-0.03	-0.05	-0.03	-0.04	-0.04
SD	0.10	0.11	0.12	0.11	0.13	0.14
Coeff. of variation	7.25	3.38	2.26	7.16	3.77	3.07
Kurtosis	0.79	0.79	0.56	1.80	1.52	1.03
Skewness	0.19	-0.04	0.15	0.60	0.33	0.38
Range	0.63	0.67	0.66	0.75	0.81	0.78
Minimum	-0.35	-0.37	-0.34	-0.39	-0.40	-0.38
Maximum	0.28	0.30	0.31	0.37	0.41	0.40
Count	1198	1176	1155	839	817	796

See Botterud et al. (2010) and Weron and Zator (2014) again for further details.

3 Linear regression for risk premium and convenience yield

3.1 Multiple linear regression

We perform a regression analysis to explain net convenience yield and ex-post risk premium for monthly contracts (both baseload and peakload) for 1–3-month delivery. The purpose is to set up the more suitable explanatory variables and their relative impact on the dependent variable.

As a preliminary investigation, we can take into account the following groups of regressors:

- *Spot price* (with variance, skewness, and kurtosis; lagged spot prices values may also be considered).

Spot price values are freely available from <http://www.mercatoelettrico.org/It/Download/DatiStorici.aspx>. The data are available from April 20, 2004, on an hourly basis and can be then converted on a daily basis. Then, the daily variance, skewness, and kurtosis can be estimated for each day by considering the 24 hourly values given for this day. We also consider lagged spot price values up to 5 lags.

- *Calendar variables* (months, seasons, day of the week, time to delivery).

The regressors representing the calendar variables (12 months, four seasons and the five opening market days of the week) are simply given by Boolean vectors with ones and zeroes (except for the time to delivery which represents the exact number of days before the start of the delivery period). A preliminary analysis highlighted that seasons and days of the week are never significant.

- *Energy and fuel prices* (coal, carbon emissions, oil, gas, gasoline, heating oil).

We consider as a proxy for coal price the Dow Jones Coal (DJUSCL). The historical values can be retrieved at the following link: <http://www.investing.com/indices/dj-coal-historical-data>.

The carbon emission historical data can be retrieved at the following link: <http://www.investing.com/commodities/carbon-emissions-historical-data>.

Finally, the other data concerning oil, gas, gasoline and heating oil prices are available on the U.S. Energy Information Administration page <http://www.eia.gov>. Specifically, the sources for these data are, respectively: <http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RB RTE&f=D> (available from May 20, 1987); <http://www.eia.gov/dnav/ng/hist/rngwhhdd.htm> (available from January 7, 1997); http://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EER_EPD2F_PF4_Y35NY_DPG&f=D (available from March 11, 2003); <http://www.>

Table 4 Correlation matrix energy products' prices

	Oil (%)	Gas (%)	Gasoline (%)	Heating oil (%)	Carbon (%)	Coal (%)
Oil	100	53	92	99	17	90
Gas	53	100	45	55	-32	33
Gasoline	92	45	100	91	22	81
Heating oil	99	55	91	100	18	90
Carbon	17	-32	22	18	100	39
Coal	90	33	81	90	39	100

Table 5 Summary statistics (oil, gas, gasoline, heating oil, carbon, coal) for the period Dec. 2011/Sept. 2016

Indicator	Oil	Gas	Gasoline	Heating oil	Carbon	Coal
Mean	85.15	3.21	2.45	2.35	6.96	112.05
Median	103.09	3.04	2.71	2.75	6.80	127.74
SD	30.09	0.90	0.68	0.76	1.89	60.70
Kurtosis	-1.48	2.42	-0.95	-1.39	-0.31	-0.46
Skewness	-0.47	0.91	-0.51	-0.51	0.51	0.18
Range	102.13	6.66	3.56	2.49	9.86	262.34
Minimum	26.01	1.49	0.62	0.82	3.15	12.33
Maximum	128.14	8.15	4.18	3.30	13.01	274.67
Count	1220	1220	1220	1220	1220	1220

eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EER_EPMRR_PF4_Y05LA_DPG&f=D (available from June 2, 1986).

These data (available on a daily basis for opening market days) are highly correlated. We present in Table 4 their correlation matrix.

Some statistics for these variables are presented in the subsequent Table 5.

The data are characterized by a high variability as we can deduce from standard deviation and range.

- *Economic indicators* (industrial production, consumer price index, load with some lagged values).

Economic indicators for Italy can be downloaded from the National Institute of Statistics <http://www.istat.it/en>. We take into account the following two indices: the industrial production IP and the consumer price index CPI. The data are accessible on a monthly basis.

These two series are weakly correlated (the correlation coefficient is 15.28%). The source for load (consumption) data is the same as for the spot price values. These data are available on an hourly basis. Besides, we consider lagged load values up to 5 lags.

- *Climatic indices* (temperature, solar radiation, wind intensity).

Climatic indices may be suitable regressors. For example, solar radiation and wind intensity are essential for the production of renewable energies. Nevertheless, these variables vary considerably throughout the country so that it is rather difficult to capture their overall effects.

Regards the temperature we could, for example, perform the mean daily values measured for the main Italian cities. Besides, we can construct an alternative temperature index τ as a mixture of heating and cooling degree days (HDD and CDD) as follows (where T is the mean daily temperature):

$$\tau = \begin{cases} T - 20 & T > 20 \\ 0 & \text{if } 16 < T < 20 \\ 16 & \text{if } T < 16 \end{cases}$$

The daily temperature values and wind intensity data are available for example through the WeatherData command of Mathematica software.

See for this purpose the following page: <https://reference.wolfram.com/language/ref/WeatherData.html>.

This temperature index is weakly correlated with futures contracts. For example, the correlations with monthly baseload futures (delivery 1–3 months) are, respectively, 22.79%, 13.82%, and 5.23%.

We have then obtained at this stage a set of 37 explanatory variables.

The choice for explanatory variables comes, in part, from literature references discussed before (for example skewness, variance, temperature index, calendar variables), and in part from some characteristics of the Italian Market (for example economic indicators and load).

The general equation of the multivariate regression is:

$$y_t = \sum_{i=1}^{12} \alpha_i \cdot M_i + \sum_{i=1}^6 \beta_i \cdot F_i + \sum_{i=1}^2 \delta_i \cdot E_i + \sum_{i=1}^3 \lambda_i \cdot I_i + \sum_{i=0}^5 \mu_i \cdot P_{t-i} + \sum_{i=0}^5 \nu_i \cdot L_{t-i} + \xi \cdot \tau + \kappa \cdot T + \varepsilon$$

where the dependent variable y_t denotes, in turn, the convenience yield and the ex-post risk premium. Besides, the explanatory variables are M_i (months), F_i (energy and fuel prices), E_i (economic indicators), I_i (spot price moments), P_t (spot price with some lags), L_t (load with some lags), τ (temperature index) and T (time to delivery).

Remark The electricity production sources for the Italian market (concerning the year 2014) are given in Fig. 5. These data can be retrieved from Terna annual report (see <http://www.terna.it>; “Statistical Data on Electricity in Italy—2014”), and they may be useful to list the factors which affect the Italian electricity markets.

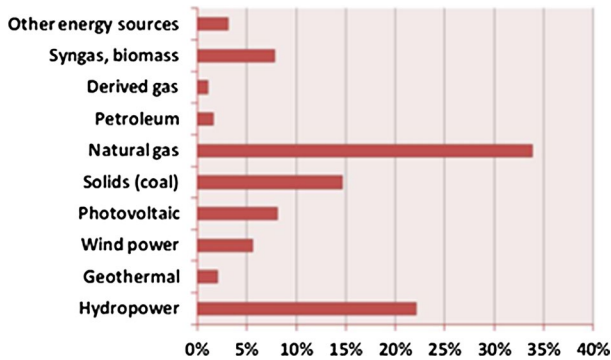


Fig. 5 Electricity production sources for year 2014—Italy (Terna annual report)

We perform now a multivariate regression analysis to explain the ex-post risk premium and the net convenience yield for monthly futures contracts with one to three-month delivery (both baseload and peakload). For this purpose, we should seek for each case, adequate explanatory variables. These regressors should be all significant at a given level (for example 5%). We then test the performance of the regression thanks to the most common indicators (namely the R -squared and adjusted R -squared).

To handle a high number of regressors, we use an algorithm developed within *xlstat* software. For this purpose, we can fix the number of regressors (from *one* to the maximum number of available regressors). The algorithm determines, for each number, the optimal selection of regressors which maximizes the adjusted R -squared (other criterions are also available, such as Akaike's AIC, MSE, Mallows C_p , Schwarz's SBC or Amemiya's PC). This selection is performed for each combination of regressors. For the final selection, we can also estimate the impact of adding or removing a particular explanatory variable (keeping the other regressors in the selection set). Besides, the determination of the normalized regression coefficients (often called *beta* coefficients) permits to establish the weight of the associated variable. Indeed a higher absolute value of the beta coefficient corresponds to a higher impact of this regressor in the explanation of the dependent variable.

We then examine in a separate analysis the two kinds of monthly futures (namely baseload and peakload) for net convenience yield and ex-post risk premium, respectively.

- Baseload monthly futures (net convenience yield).

We show in Fig. 6 the adjusted R -squared as a function of the number of regressors (the optimal regressors have been selected for each fixed number of regressors through the optimization of the adjusted R -squared). We have presented in the same plot the three delivery period types: 1-month delivery (series 1), 2-month delivery (series 2) and 3-month delivery (series 3). We note at first that the adequacy

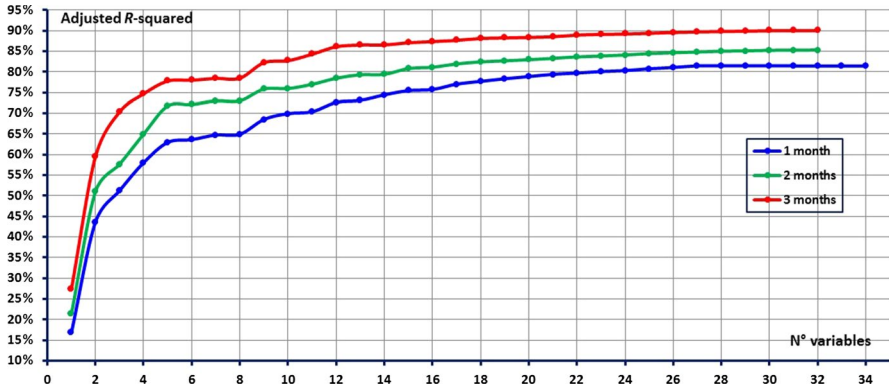


Fig. 6 Adjusted *R*-squared versus number of explanatory variables for convenience yield (baseload futures)

performance improves with longer delivery periods. Besides, the adjusted *R*-squared improves slowly if we include more than 12 explanatory variables. This is an important remark to set up the more parsimonious model. The best adjusted *R*-squared values attained are 81.40% (1 month), 85.30% (2 months) and 90.10% (3 months).

We examine now thoroughly the optimal explanatory variables. We exhibit in Fig. 7 the normalized beta coefficients (for 3-month delivery time for brevity) to visualize the impact (positive or negative) of each regressor. The coefficients are significant with a confidence level of 95%.

The meaning of each regressor is listed in Table 6.

The following explanatory variables have a *positive impact*: spot price, load, skewness, time t. d., months (except August), load 3-lags, load 4-lags, oil, gas, CPI, carbon, coal.

The following explanatory variables have a *negative impact*: variance, kurtosis, load 1-lag, spot price 1-lag, spot price 2-lags, spot price 3-lags, spot price 4-lags, gasoline, heating oil, IP.

The most important variables are spot price, months (overall), oil, heating oil, CPI.

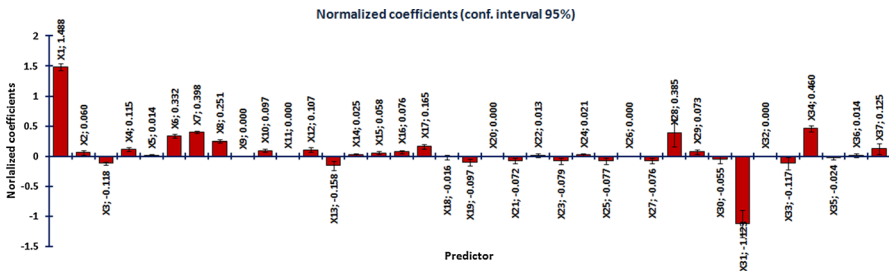


Fig. 7 Normalized regression coefficients for net convenience yield (baseload futures—3-month delivery)

Table 6 List of regressors—normalized regression for net convenience yield (baseload futures—3-month delivery)

×1	Price	×11	Month 6	×21	Price 2-lags	×31	Heating oil
×2	Load	×12	Month 7	×22	Load 3-lags	×32	Temperature
×3	Variance	×13	Month 8	×23	Price 3-lags	×33	Ip
×4	Skewness	×14	Month 9	×24	Load 4-lags	×34	Cpi
×5	Time t. d.	×15	Month 10	×25	Price 4-lags	×35	Kurtosis
×6	Month 1	×16	Month 11	×26	Load 5-lags	×36	Carbon
×7	Month 2	×17	Month 12	×27	Price 5-lags	×37	Coal
×8	Month 3	×18	Load 1-lag	×28	Oil		
×9	Month 4	×19	Price 1-lag	×29	Gas		
×10	Month 5	×20	Load 2-lags	×30	Gasoline		

Finally, we plot in Fig. 8 the predicted values (estimated from the parameters' regression, with confidence interval 95%) against the real values of the dependent variable by considering all the regressors (for 3-month delivery contract).

The results of adequacy tests for the regression estimations with all the explanatory variables are presented in Table 7. We consider several indicators (we present here futures with time to delivery 1–3 months).

In addition to (Adjusted) *R*-squared, we include also in the table DF (degrees of freedom of the model), MSE (mean squares of errors), MAPE (mean absolute percentage error), DW (Durbin-Watson statistic), Cp (Mallows' Cp), AIC (Akaike's information criterion), SBC (Schwarz's Bayesian criterion), PC (Amemiya's prediction criterion).

Remark We can set up a parsimonious model (in the case of 3-month delivery period) with these 5 explanatory variables: spot price, 2 months variables, heating oil and CPI. This model has Adjusted *R*-squared 77.8% and AIC = -6643.37.

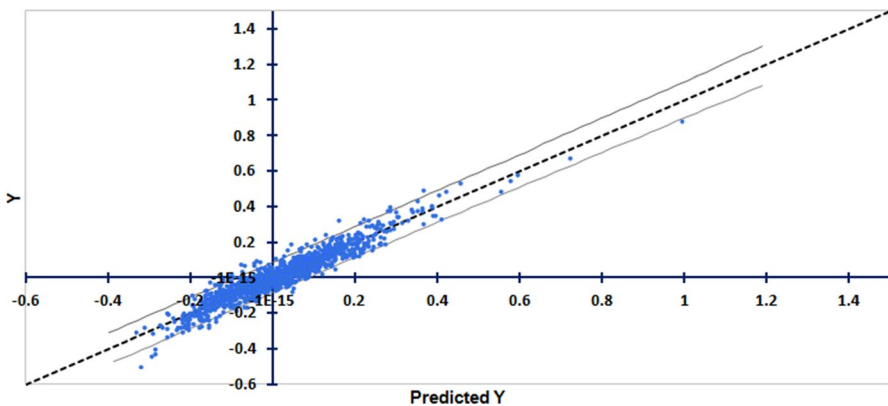


Fig. 8 Predicted values of net convenience yield (baseload—3-month delivery) versus real values

Table 7 Adequacy tests for multivariate linear regression for all cases

Indicator	Conv. base			Conv. peak			Risk pr. base			Risk pr. peak		
	1 Month	2 Months	3 Months	1 Month	2 Months	3 Months	1 Month	2 Months	3 Months	1 Month	2 Months	3 Months
	No. observations	1220	1220	1220	861	861	861	1155	1155	1155	796	796
DF	1192	1187	1187	831	828	829	1126	1123	1124	768	765	765
R ²	81.84%	85.68%	90.36%	82.18%	86.74%	92.51%	31.76%	45.43%	56.99%	54.27%	65.27%	71.42%
Adjusted R ²	81.43%	85.29%	90.10%	81.56%	86.23%	92.23%	30.07%	43.92%	55.85%	52.66%	63.91%	70.30%
MSE	0.002	0.002	0.002	0.003	0.002	0.002	0.006	0.007	0.006	0.006	0.007	0.006
MAPE	344.80	131.77	230.82	112.32	130.20	91.86	200.37	262.51	172.32	257.64	263.49	315.29
DW	0.481	0.447	0.447	0.415	0.423	0.527	0.117	0.135	0.142	0.194	0.281	0.196
Cp	21.187	29.760	29.349	25.733	29.637	27.644	23.369	27.307	25.979	20.965	26.373	25.681
AIC	-7389	-7420	-7600	-4999	-5156	-5387	-5835	-5638	-5860	-4021	-3967	-4037
SBC	-7246	-7252	-7431	-4857	-4999	-5235	-5689	-5477	-5704	-3890	-3822	-3892
PC	0.190	0.151	0.102	0.191	0.143	0.081	0.717	0.577	0.454	0.491	0.375	0.309

- Baseload futures (ex-post risk premium)

We show in Fig. 9 the adjusted R -squared with respect to the number of regressors (the optimal regressors have been selected as before). We superimpose in the same plot the three delivery period types: 1-month delivery (series 1), 2-month delivery (series 2) and 3-month delivery (series 3). We note at first that the adequacy performance increases with the increasing of the delivery period. Besides, the adjusted R -squared improves slowly if we include more than 18 explanatory variables. The best adjusted R -squared values attained are 30.10% (1 month), 43.90% (2 months) and 55.80% (3 months).

Let examine the normalized beta coefficients for 3-month delivery time (we omit here the figure for brevity).

The following explanatory variables have a *positive* impact: kurtosis, months (except October and November), load 1-lag, load 5-lags, IP, CPI, coal.

The following explanatory variables have a *negative* impact: spot price, variance, skewness, time t.d., spot price 1-lag, spot price 2-lags, spot price 3-lags, load 4-lags, spot price 4-lags, spot price 5-lags, oil, gas, gasoline, heating oil, temperature, carbon.

The most important variables are months (overall), oil, gasoline, IP, CPI, coal.

Finally, we plot in Fig. 10 the predicted values (estimated from the parameters' regression, with confidence interval 95%) against the real values of the dependent variable by considering all the regressors (for 3-month delivery contract).

The results of adequacy tests for the regression estimations with all the explanatory variables are presented in Table 7.

Remark We can set up a parsimonious model (in the case of 3-month delivery period) with these 6 explanatory variables: 2 months variables, heating oil, temperature, CPI, and coal. This model has Adjusted R -squared 39.3% and AIC = -5515.66

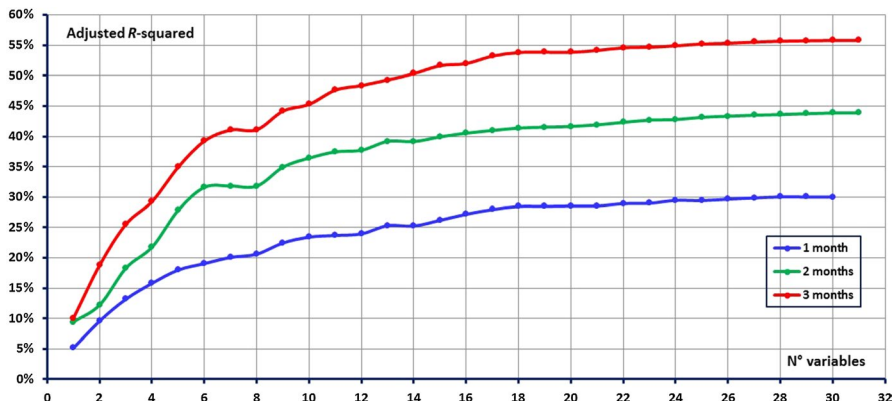


Fig. 9 Adjusted R -squared versus number of explanatory variables for ex-post risk premium (baseload futures)

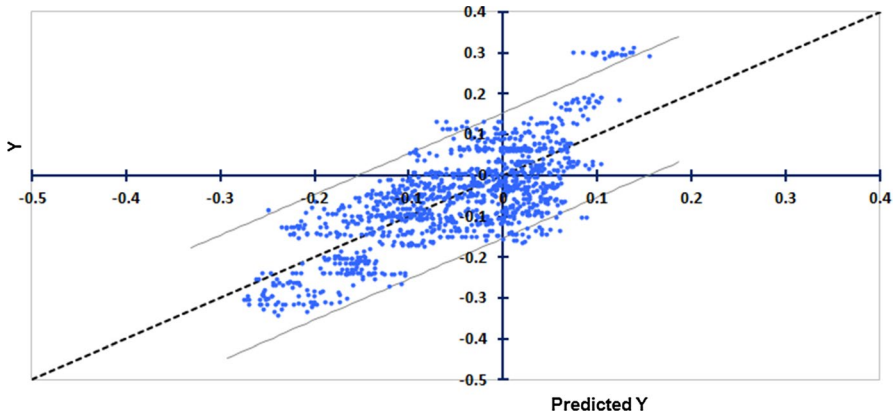


Fig. 10 Predicted values of ex-post risk premium (baseload—3-month delivery) versus real values

- Peakload futures (net convenience yield)

We show in Fig. 11 the adjusted R -squared with respect to the number of regressors (the optimal regressors have been selected as before). We have superimposed in the same plot the three delivery period types: 1-month delivery (series 1), 2-month delivery (series 2) and 3-month delivery (series 3). We note again that the adequacy performance increases with the increasing of the delivery period. Besides, the adjusted R -squared improves slowly if we include more than 18 explanatory variables. The best adjusted R -squared values attained are 81.60% (1 month), 86.20% (2 months) and 92.20% (3 months).

We examine the beta coefficients again (for 3-month delivery time for brevity).

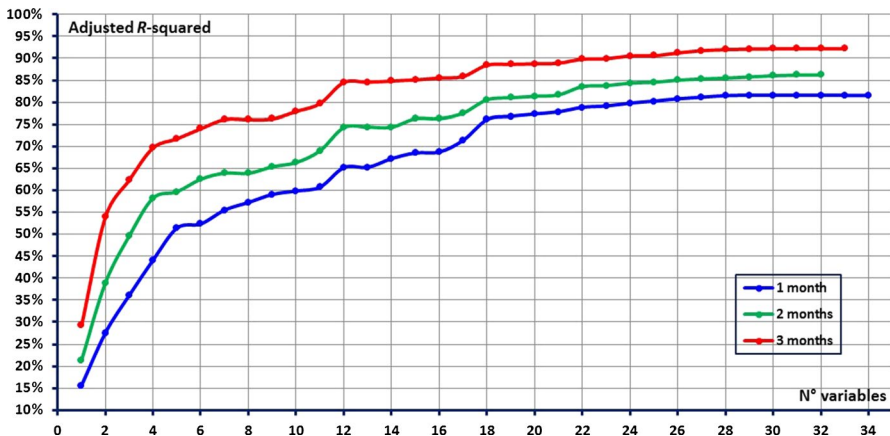


Fig. 11 Adjusted R -squared versus number of explanatory variables for net convenience yield (peakload futures)

The following explanatory variables have a *positive impact*: spot price, load, skewness, time t.d., months (except July to November), load 4-lags, oil, temperature, CPI, coal.

The following explanatory variables have a *negative impact*: variance, kurtosis, load 1-lag, spot price 1-lag, spot price 2-lags, spot price 3-lags, spot price 4-lags, spot price 5-lags, gas, gasoline, heating oil, IP, carbon.

The most important variables are spot price, months (overall), oil, heating oil, CPI, coal.

The results of adequacy tests for the regression estimations with all the explanatory variables are presented in Table 7.

Remark We can set up a parsimonious model (in the case of 3-month delivery period) with these 6 explanatory variables: spot price, 2 months variables, heating oil, gasoline, and coal. This model has Adjusted R -squared 74.1% and $AIC = -4374.29$.

- Peakload futures (ex-post risk premium)

We show in Fig. 12 the adjusted R -squared with respect to the number of regressors (the optimal regressors have been selected for each fixed number of regressors through the optimization of the adjusted R -squared). We have superimposed in the same plot the three delivery period types: 1-month delivery (series 1), 2-month delivery (series 2) and 3-month delivery (series 3). We note that the adequacy performance increases with the increasing of the delivery period. Besides, the adjusted R -squared improves slowly if we include more than 16 explanatory variables. The best adjusted R -squared values attained are 52.70% (1 month), 63.90% (2 months) and 70.30% (3 months).

We deduced some conclusions by analyzing the beta coefficients (for 3-month delivery time for brevity).

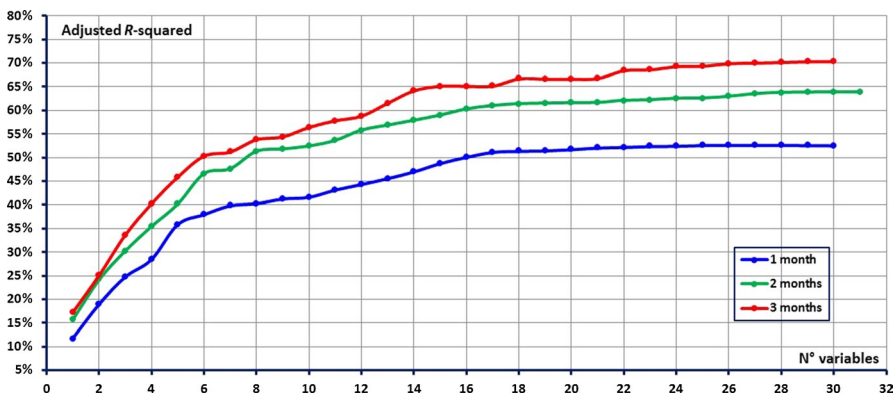


Fig. 12 Adjusted R -squared versus number of explanatory variables for ex-post risk premium (peakload futures)

The following explanatory variables have a *positive* impact: load, kurtosis, load 1-lag, load 3-lags, load 5-lags, spot price 5-lags, heating oil, temperature, CPI, carbon, coal.

The following explanatory variables have a *negative* impact: spot price, variance, skewness, months (except April, June to September), load 1-lag, load 4-lags, spot price 4-lags, oil, gas, gasoline, IP.

The most important variables are months (overall), oil, gas, gasoline, heating oil, coal.

The results of adequacy tests for the regression estimations with all the explanatory variables are presented in Table 7.

Remark We can set up a parsimonious model (in the case of 3-month delivery period) with these 8 explanatory variables: 4 months variables, gas, kurtosis, carbon, and coal. This model has Adjusted R -squared 53.9% and $AIC = -3707.94$.

We note that the regression performance increases with longer delivery periods. We find evidence that the following explanatory variables are always significant:

- Calendar variables (months) which explain the seasonality feature;
- Load values;
- Spot price with higher moments (variance, skewness, kurtosis);
- Fuel and energy prices (oil, gas, heating oil, carbon emission, coal);
- Economic variables (IP, CPI).

Instead, some regressors are generally not significant or produce a low impact on the dependent variable. For instance:

- Lagged variables (load and spot prices);
- Time to delivery;
- Temperature.

3.2 Other regression techniques

In the previous section, we have applied a multivariate linear regression model with a high number of explanatory variables, some of which were highly correlated. In this situation, other regression techniques such as partial least squares regression (denoted as PLSR) and principal components regression (denoted as PCR) may be more suitable.

These two methods consist in determining new explanatory variables, denoted as components, obtained as linear combinations of the starting variables. Nevertheless, these components are constructed differently by PLSR and PCR. Indeed, the components created by PCR aim to explain only the fluctuations of the explanatory variables while PLSR also takes into account the dependent variable. The overall goal of these techniques is to set up the more parsimonious model. Since PLSR is a mixture of multivariate regression and PCR, we present first PCR and then PLSR.

- Principal components regression

The principal components regression follows two main steps. At first, we perform a principal components analysis (PCA) on the set of the explanatory variables. At this stage, the so-called principal components are determined. Then, we compute a multivariate regression of the dependent variable on the components just created, and we determine the parameters of the model. These tasks can be achieved thanks to the *xlstat* software.

The adequacy tests of these regressions are presented in Table 8 for net convenience yield and ex-post risk premium, both for baseload and peakload contracts with one to three-month delivery (all the components are considered).

Finally, we plot in Fig. 13 the predicted values (estimated from the parameters' regression, with confidence interval 95%) against the real values of the dependent variable. We consider, for brevity, contracts with 3-month delivery.

We note that the fitting for net convenience yield is better than the fitting for ex-post risk premium. In both cases, the fitting for peakload contracts is better.

- Partial least squares regression

Partial least squares regression technique applies to handle data which comprise correlated or sometimes collinear explanatory variables. The new components built with this technique come from linear combinations of the given list of explanatory variables. The settlement of these components takes into account the observed dependent variable. We get then a more parsimonious model which shares a good predictive power.

This technique can be viewed as a mix of the classical multiple linear regression with a principal component analysis. Indeed, the multiple linear regression aims to determine a suitable combination of the explanatory variables which correctly predicts a dependent variable, whereas principal component analysis aggregates explanatory variables (the so-called components) to reduce correlations without taking into account the dependent variable.

We can perform a PLSR by using several software tools given for example by *MATLAB* or *xlstat*. These tools determine automatically the components (after fixing its number), the coefficients of the linear combinations of the original predictors needed to construct the components and finally, the percentage of variance explained in the dependent variable with respect to the number of components. These results can be used then to determine the optimal number of components useful for a reliable and parsimonious model.

We finally exhibit in Figs. 14 and 15 the proportion of variance explained as a function of the number of components (only baseload case for brevity).

We note that 10 components are sufficient to explain more than 60% of the variability of the dependent variable. The proportion remains quite unchanged whenever we add more than 22 components. The proportion of variance explained with all the components is, respectively: 83.84% (1 month), 85.71% (2 months) and 90.40% (3 months).

Table 8 Adequacy tests for principal components regression for all cases

Indicator	Conv. base			Conv. peak			Risk pr. base			Risk pr. peak		
	1 Month	2 Months	3 Months	1 Month	2 Months	3 Months	1 Month	2 Months	3 Months	1 Month	2 Months	3 Months
	No. obs.	1220	1220	1220	861	861	861	1155	1155	1155	796	796
DF	1183	1183	1183	824	824	824	1118	1118	1118	759	759	759
R ²	81.88%	85.68%	90.36%	82.24%	86.75%	92.51%	31.91%	45.44%	57.03%	54.39%	65.34%	71.44%
Adjusted R ²	81.33%	85.25%	90.07%	81.46%	86.17%	92.18%	29.72%	43.69%	55.65%	52.22%	63.69%	70.09%
MSE	0.002	0.002	0.002	0.003	0.002	0.002	0.006	0.007	0.006	0.006	0.007	0.006
MAPE	339.63	130.70	233.71	112.30	131.44	92.56	203.18	262.96	173.07	260.81	273.01	316.00
DW	0.481	0.448	0.446	0.415	0.423	0.524	0.116	0.133	0.143	0.194	0.275	0.197
Cp	37.000	37.000	37.000	37.000	37.000	37.000	37.000	37.000	37.000	37.000	37.000	37.000
AIC	-7373	-7413	-7592	-4988	-5149	-5378	-5821	-5629	-5849	-4005	-3957	-4026
SBC	-7184	-7224	-7403	-4812	-4973	-5202	-5635	-5442	-5662	-3832	-3783	-3853
PC	0.193	0.152	0.102	0.194	0.144	0.082	0.726	0.582	0.458	0.501	0.380	0.313

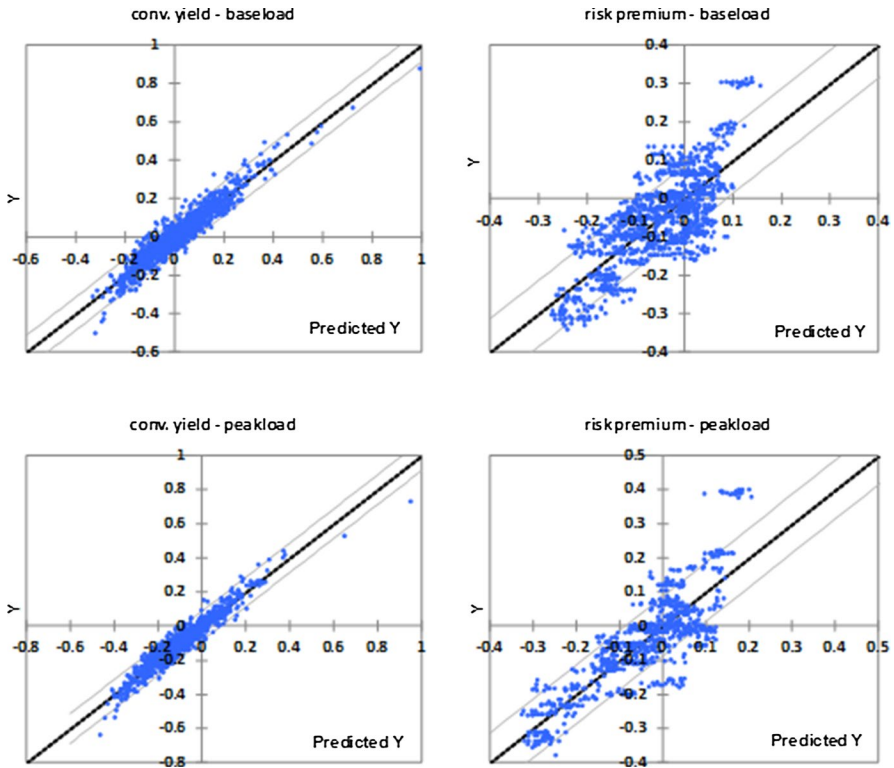


Fig. 13 Predicted values of net convenience yield and ex-post risk premium (baseload/peakload with 3-month delivery) versus real values

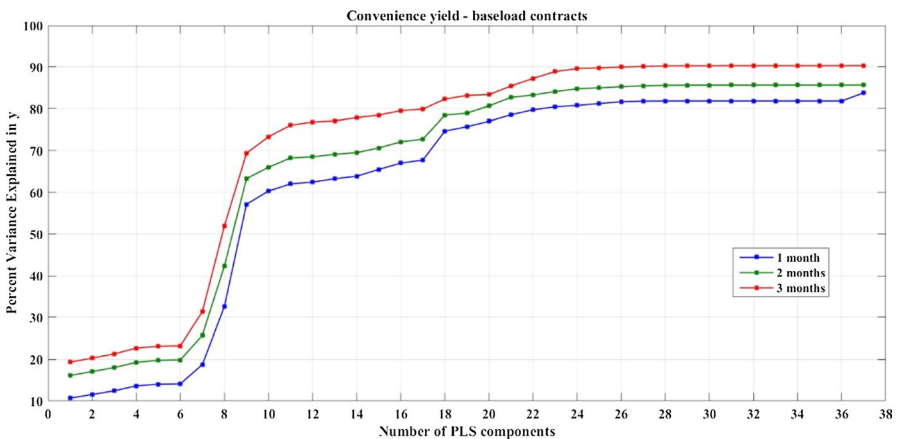


Fig. 14 Variance explained (%) versus number of PLSR components for net convenience yield (baseload contracts)

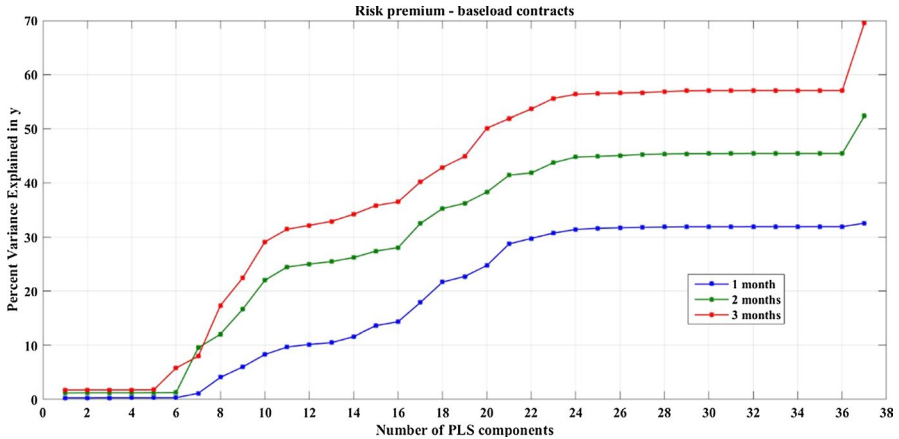


Fig. 15 Variance explained (%) versus number of PLSR components for ex-post risk premium (baseload contracts)

For peakload case, we find evidence that 16 components are sufficient to explain more than 60% of the variance of the dependent variable. The proportion remains quite unchanged whenever we add more than 20 components. The proportion of variance explained with all the components is, respectively: 85.37% (1 month), 86.76% (2 months) and 92.85% (3 months).

We note that 22 components are sufficient to explain more than 30% of the variability of the dependent variable. The proportion remains quite unchanged whenever we add more than 24 components. The proportion of variance explained with all the components is, respectively: 32.55% (1 month), 52.37% (2 months) and 69.56% (3 months).

For peakload case, we find evidence that 18 components are sufficient to explain more than 40% of the variance of the dependent variable. The proportion remains quite unchanged whenever we add more than 22 components. The proportion of variance explained with all the components is, respectively: 57.12% (1 month), 66.48% (2 months) and 80.15% (3 months).

Finally, Figs. 14 and 15 highlight that the performance of PLSR increases with longer time delivery periods.

The regression models used in this survey (multivariate linear regression, partial least squares regression and principal components regression) produced a good fitting with the aim of explaining the net convenience yield and the ex-post risk premium for several types of monthly futures contracts (baseload and peakload with delivery period 1–3 months). For this purpose, we used a long list of explanatory variables and we showed how to get a more parsimonious model.

Finally, the comparison between these techniques is not straightforward as the multivariate regression relies on the number of explanatory variables while PCR and PLSR rely principally on the number of components built with the available variables. Nevertheless, one can compare the adequacy tests for the two main techniques illustrated in Table 7 (for multivariate regression) and in Table 8 (for PCR). We

present also a graphical inspection concerning the percentage of explained variance with respect to the number of variables/components for the multivariate regression (Figs. 6, 9, 11, 12) and for the PLSR (Figs. 14, 15).

4 Conclusions

This paper investigates the possible relationships between futures and spot prices for the Italian electricity market through the net convenience yield and the ex-post risk premium concepts. For this purpose, we take into account a large database for monthly futures (both baseload and peakload) traded at the IDEM market. In the subsequent empirical application, we have employed these futures prices from 2011 to 2016.

At first, we find evidence that the ex-post risk premiums are negative on average. This issue is in agreement with results present in other electricity markets. Nevertheless, we further examine the dynamic evolution of the premiums, and we highlighted that its value varies seasonally.

Besides, we detect that the net convenience yield for monthly futures is on average positive within all the considered holding periods for baseload contracts (backwardation prevails) and negative for peakload contracts (contango prevails).

Finally, a multivariate linear regression conducted on the net convenience yield and the ex-post risk premium for monthly futures with 1–3-month delivery period permitted to highlight the explanatory variables which mainly affect these quantities.

We conclude our survey with more general regression techniques such as partial least squares regression and principal components regression. The scope was to reduce the number of explanatory variables to get a more parsimonious model.

Our paper followed the framework of Botterud et al. (2010) and Weron and Zator (2014) used for the Nord Pool market. We investigated here about the Italian futures market with more refined regression techniques and a greater set of explanatory variables to explain ex-post risk premium and net convenience yield. The adequacy tests for the regression models produced very reliable results.

Acknowledgements Open access funding provided by Università degli Studi di Cagliari within the CRUI-CARE Agreement.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aïd R. Electricity derivatives. Springer, 2015; ISBN 978-3-319-08394-0.
- Ballester JM, Climent F, Furió D. Market efficiency and price discovery relationships between spot, futures and forward prices: the case of the Iberian electricity market (MIBEL). *Span J Financ Account*. 2016;45(2):135–53. <https://doi.org/10.1080/02102412.2016.1144441>.
- Bauwens L, Hafner CM, Pierret D. Multivariate volatility modeling of electricity futures. *J Appl Econom*. 2013;28(5):743–61.
- Benth FE, Krühner P. Derivatives pricing in energy markets: an infinite-dimensional approach. *SIAM J Financ Math*. 2015;6:825–69.
- Botterud A, Kristiansen T, Ilic MD. The relationship between spot and futures prices in the Nord Pool electricity market. *Energy Econ*. 2010;32:967–78.
- Branger N, Reichmann O, Wobben M. Pricing electricity derivatives on an hourly basis. *J Energy Mark*. 2010;3(3):51–89.
- Diko P, Lawford S, Limpens V. Risk premia in electricity forward prices. *Stud Nonlinear Dyn Econ*. 2006;10(3), Article 7.
- Flasza R, Rippel M, Solc J. Modeling prices of electricity futures at EEX. *Int Sch Sci Res Innov*. 2011;5(6):1103–7.
- Fleten S-E, Huisman R, Kilic M, Pennings E, Westgaard S. Electricity futures prices: time-varying sensitivity to fundamentals. *J Energy Mark*. 2015;8(4).
- Handika R, Trück S. Risk premiums in interconnected Australian electricity futures markets. SSRN Working Paper. 2013. <https://doi.org/10.2139/ssrn.2279945>.
- Islyayev S, Date P. Electricity futures price models: calibration and forecasting. *Eur J Oper Res*. 2015;247:144–54.
- Lucia JJ, Torró H. On the risk premium in Nordic electricity futures prices. *Int Rev Econ Finance*. 2011;20:750–63.
- Redl C, Bunn DW. Determinants of forward premia in electricity markets: a taxonomic empirical analysis. In: 11th IAEE European conference, Vilnius; 2010.
- Shawky HA, Marathe A, Barrett CL. A first look at the empirical relation between spot and futures electricity prices in the United States. *J Futures Mark*. 2003;23(10):931–55.
- Weron R, Zator M. Revisiting the relationship between spot and futures prices in the Nord Pool electricity market. *Energy Econ*. 2014;44:178–90.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.