NNT : 2016SACLX018

THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
L'ECOLE POLYTECHNIQUE

ÉCOLE DOCTORALE N°578
Sciences de l'homme et de la société

Spécialité de doctorat : Sciences économiques

Par

Faddy Ardian

Empirical Analysis of Italian Electricity Market

Thèse présentée et soutenue à Paris, le 1 juillet 2016 :

Composition du Jury :
M. D. K. Nguyen, Professeur, IPAG Business School, Président du Jury
M. Y. Le Pen, Maître de conférences, Université de Paris Dauphine, Rapporteur
Mme. A. Creti, Professeur, Université de Paris Dauphine, Directeur de thèse
M. J. P. Ponssard, Professeur, Ecole Polytechnique, Directeur de thèse
M. F. Vallone, CEO, Cogenpower Spa, Examinateur

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

Acknowledgement

This dissertation could not have been completed without the support of many individuals.

First and foremost, I would like to express my deepest gratitude to my advisor, Anna Creti, for her full guidance and encouragement throughout my study and my research.

I would also like to thank Jean-Pierre Ponssard that has helped me in all my years at Ecole Polytechnique.

Immeasurable gratitude is also owed to Francesco Vallone and Cogenpower Spa. for all of their support and the initiation of my research.

I would also like to thank my colleagues at Ecole Polytechnique for every support, constructive discussion, and advice throughout my study.

Last but not least, I would like to acknowledge with gratitude, all support and love from my family - my wife, my grandmother, my parents, my sisters and my brothers. Thank you for all support, patience, and encouragement all these years.

Résumé

Les caractéristiques spécifiques du marché de l'électricité Italien fournissent une étude de cas intéressante pour la recherche économique. Premièrement, les gouvernements Italiens ont décidé d'appliquer un mécanisme interzonal pour formuler des prix de l'électricité. Ce mécanisme de prix ne serait pas seulement tenu un compte de courbe d'offres et de courbe de demande en Italie. Aussi, ce mécanisme prend un compte l'échange physique entre leurs zones administratives. Cette formulation des prix est déjà appliquée dans de nombreux pays comme les États-Unis, l'Australie et le Danemark. Pourtant, par rapport à d'autres pays, Italie a le plus grand nombre de zones dans leur marché de l'électricité, six zones. Deuxièmement, la politique ambitieuse par le gouvernement Italien sur énergie renouvelable a augmenté les productions intermittentes dans le marché. La dépendance de la météo a créé une incertitude dans la courbe d'offre. En conséquence, le prix est devenu plus volatil que les prix des années précédentes. Ceci est un gros problème du point de vue des fournisseurs car ils sont exposés aux risques financiers. Troisièmement, les capacités des réseaux entre les zones administratifs Italiens sont très élevées dans le nord mais ils sont très limités dans le sud (entre Italie et Sicile). Puisque les énergies renouvelables sont mises en place stratégiquement dans les régions qui ont moins des habitants. Donc, ils sont loin des régions où il y a une forte demande d'électricité. Par conséquent, la congestion devient un problème en Italie car les réseaux ont besoin de plus capacité pour transporter l'électricité. Ses caractéristiques spécifiques de marché de l'électricité Italien ont influencé notre analyse empirique du marché de l'électricité Italienne. Par ces raisons, cette thèse se concentre sur les trois points de vue du marché : la prévision du prix, l'impact des énergies renouvelables sur la congestion, et l'interdépendance des prix entre diffèrent zone géographique en Italie.

La première étude de cette thèse permet de répondre la question de la prévision de prix causée par la volatilité du marché de l'électricité. Un bon modèle de prévision est important pour les participants du marché et les gouvernements. Malheureusement, il n'y a que quelques modèles qui sont proposés dans les ouvrages économique pour prévoir le prix d'électricité en Italie. À notre connaissance, il n'y a que quatre recherches qui ont fait des études dans ce domaine. Bosco et al (2007) a étudié différents modèles Autoregressive afin de prévoir les moyens des prix quotidiens de l'électricité en Italie. Ils concluent que le modèle Autoregressive avec Garch (ARMA-GARCH) est le meilleur modèle en termes de précision. Petrella et Sapio (2011), utilisent ARMA avec les variables exogènes (ARMAX) pour modéliser la formation des prix et examiner sa précision pour prévoir des prix de l'électricité. Leurs calculs montrent que prix du gaz naturel, la demande et la température sont les déterminants du prix de l'électricité. Un an plus tard, Serinaldi (2011) applique GAMLSS (Generalized Additif Model for Location, Scale, and Shape) pour prévoir prix de l'électricité à California Power Échange (CALPX) et Italian Power Exchange (IPEX). Dans sa

recherche, il a aussi expliqué les variables exogènes qui détermine la formation de prix. Son travail conclut que GAMLSS peut être utilisé comme une technique alternative pour prévoir le prix de l'électricité. Gianfreda et grossi (2012) utilisent ARFIMAX-Garch (Autoregressive Fractionally integrated moving average and General Autoregressive Conditional Heteroskedasticity) pour prévoir les prix zonaux Italiens et explorer les variables exogènes (la demande, la technique, la congestion et la concentration du marché) qui détermine le prix.

L'objectif principal de cette étude est visé à proposer des modèles alternatifs qui est limité dans les ouvrages du marché Italien de l'électricité. Nous avons comparé sept modèles univariate et deux modèles multivariate afin de trouver le meilleur modèle pour prévoir le prix d'électricité. Nous contribuons aux ouvrages économiques en initiant des discussions sur la différence de performance entre le modèle multivariate et le modèle univariate qui est suggéré par Weron (2014). Nos résultats empiriques montrent que la saisonnalité dans les modèles améliore la précision des prévisions. Puis, conformément à Gianfreda et grossi (2012), nos estimations montrent que le prix du gaz et la demande augmentent le prix d'électricité. Notre discussion sur la comparaison entre les modelés univariate et multivariate conclut que chaque modelé a sa valeur. Les modèles multivariate sont les modèles le plus précis selon nos estimations. D'autre part, les modelés univariate sont mieux que les modèles multivariate quand il est utilisé pour gérer le risque financier.

La deuxième recherche de cette thèse examine l'impact des énergies renouvelables sur la congestion en Italie. La croissance rapide des énergies renouvelables en Italie a créé un nouveau stress dans le réseau. Pourtant, les ouvrages économiques sur l'impact des énergies renouvelables sont essentiellement axés sur l'effet de l'ordre du mérite et l'effet de la variance. En plus, un grand nombre de pays est déjà étudié afin d'établir la preuve de ces effets dans les prix d'électricité (Autriche (Würzburg et al., 2013), le Danemark (Jó nsson et al., 2010), en Allemagne (Ketterer, 2014), et l'Italie (Clo et al., 2015)). Il n'y a que quelques études qui analyse l'impact des énergies renouvelables sur la congestion. En Norvège, Førsund et al. (2008) a commencé la discussion sur cet impact. Sa recherche analyse l'impact de l'intégration de l'énergie éolienne dans le réseau norvégien. Aux États-Unis, Woo et al (2011) continuent les discussions et concluent que l'augmentation de la production d'énergie éolienne, la production nucléaire, la demande et le prix du gaz augmentent la probabilité d'avoir une congestion. En Espagne, Figuiredo et al. (2015) évaluent plusieurs déterminants de la congestion en utilisant la logique et la fonction keynésienne non paramétrique. Ses études établissent la preuve de l'impact de l'énergie renouvelable sur l'augmentation de la congestion.

Dans cette étude, nous avons estimé l'impact des énergies renouvelables sur la congestion en termes de probabilité et de coût afin de contribuer aux ouvrages limités dans ce domaine. Nous appliquons Multinomial Logit afin d'analyser les sources de congestion de réseaux et calculer son impact. Puis, pour examiner les déterminants du coût de la congestion, nous utilisons 2 SLS (two stages least squares methods). Notre analyse suggère que la production locale d'énergie renouvelable diminue la probabilité de congestion dans sa zone géographique car il réduit l'importation d'électricité depuis ses voisins. Cependant, Il augmente la probabilité d'une congestion aux zones reliées car il pourrait augmenter l'exportation d'électricité. Nos estimations sur le coût de la congestion montrent que, un choc positif de la production d'énergie renouvelable depuis une région importatrice peut changer le coût de la congestion vers la valeur négative car elle diminue le coût marginal pour mettre le system en équilibre. Mais, le coût sera plus élevé quand la production d'énergie renouvelable depuis une région exportatrice s'amplifie. Par conséquent, l'augmentation des énergies renouvelables devrait être encouragée dans les régions importatrices, mais la croissance devrait être contrôlée afin d'éviter la congestion dans le réseau Italien.

La troisième recherche de cette thèse examine l'interdépendance des prix dans six zones administratives du marché Italien de l'électricité. Cette étude permet d'identifier les régions qui est isolée dans le marché Italien. Les pays avec plusieurs marchés régionaux ont commencé à examiner l'intégration nationale de leur marché de l'électricité. En Australie, Worthingthon et al (2005), Higgs (2009) et Ignatieva et Trück (2011) ont analysé les interdépendances des cinq marchés régionaux en utilisant des méthodes différentes. Les recherches concluent que les marchés régionaux qui sont équipés avec des grandes capacités de réseaux ont forte interdépendance des prix d'électricité. Aux États-Unis, Park et al (2006) analyse le marché spot en utilisant Vecteur auto-regression. Leurs études concluent que les capacités de réseaux et le mécanisme du marché effectuent les interdépendances. Malheureusement, ce sujet a été occulté en l'Italie qui dispose de 6 marchés régionaux. En plus, Italie vient de mettre en place un nouveau réseau en 2012, entre Sardaigne et Italie, qui pourrait améliorer l'intégration du marché.

Dans cette recherche, nous analysons la corrélation croisée entre les prix zonaux Italiens en utilisant DCC-MGARCH (Dynamics Conditional corrélation - Multivariate Generalized Autoregressive conditional Heteroskedasticity) et CCC-MGARCH (Constant Conditional Corrélation - Multivariate Generalized Autoregressive conditional Heteroskedasticity) avec correction saisonnière pour calculer l'interdépendance des prix. Nous avons contribué aux ouvrages académiques en améliorant le modèle proposé par Higgs (2015) qui n'a pas des paramètres saisonniers. En plus, avec le nouveau réseau qui vient d'être mis en place, cette étude permet d'expliquer l'impact d'amélioration de capacités de réseau sur l'intégration du marché. Nos estimations avec CCC-GARCH et DCC-GARCH montrent que toutes les zones dans la péninsule Italienne ont fortes dépendances entre eux. Ce résultat est conforme à Higgs (2009) qui conclut que l'interdépendance de prix zonaux est très forte quand ils ont une grande capacité de réseau. De plus, nos calculs au nouveau réseau entre Sardaigne et Italie indiquent que les interdépendances entre ses prix sont plus fortes après avoir eu une capacité supplémentaire dans son réseau. Finalement, notre analyse suggère que Sicile est une région isolée car son interdépendance est faible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and motivation

1.1.1 Deregulation of electricity market in Europe

Prior to 1996, electricity markets in European Union state member are controlled and organized by the state-owned companies. As a consequence, natural market monopoly markets were formed with different levels of public services among EU state members. The electricity price, therefore, was generally fixed according to the government's policy and did not reflect the actual supply and demand in the market since the price did not move according to the market. This situation creates non-transparencies in gas and electricity price formation. On the other hand, in order to ensure a fluent flow of free movement of goods, services, people, and capital, EU commissions has started to consider the security supply of their electricity. All member states have also started to realize the importance on improving efficiencies of production, transmission, and distribution in their electricity market. They envision an internal electricity market with an objective to improve competitiveness and transparency of electricity market without prejudice to compliance with public service obligations. As a result, EU commission decided to push energy market towards liberalization and re-structurization with one common rule.

In 1996, the fruit of the idea was finally established with a Directive 96/92/EC. For the first time in Europe, a legal framework was born to regulate the generation, transmission and distribution of electricity through this directive. The directive was aimed to increase economic efficiency, to improve the level of pubic service and to create the freedom for customers in choosing their energy suppliers in all EU state members. The directive addresses several main issues for realizing their objectives:

- State-owned energy companies in all EU state members must separate the electricity sectors into generation, transmission, distribution and retail. They are obliged to separate each activity into different accounts in order to avoid discriminations in the market.

- Transmission and distribution, however, have to remain a monopoly and independent with non-discriminatory rules for users that require access to the transmission network. Hence, all EU state member must designate one or more independent transmission system operator who is responsible for managing and operating all the grid lines.

- In order to improve efficient productions, EU commission, dictates priority dispatching for the producers who run renewable energy production units, waste power generations, and CHP (Combined Heat and Power) plant.

- In the supply side, all EU state members must open generations for new entrants by regulations or a tender process in order to improve economic competitiveness. If a tendering procedure is chosen, specific administrations are required in order to ensure transparency.

- In the demand side, EU commission also obliges all member states to liberalize the market thus provide a possibility for customers to choose their own energy suppliers (both from inside and outside their border).

- Initial opening on the demand side starts with industrial users whose consumption exceed 40 GW annually. This is, then, improved every three years in order to complete liberalization on the demand side.

- All member states are also obliged to build a necessary mechanism to protect competition, to ensure free-collusion and to avoid predatory behavior in the market.

The directive had contributed greatly towards a creation of an internal electricity market and had given a great promise towards market liberalization. However, the levels of the re-structurization were varied among EU state members. Notable countries such as Spain, Germany, Belgium and UK have initiated competitive power market. On the other hand, some countries have to modify and to restructure their power sector in order to start competition in the power sector. Then, other EU state members were still behind as governments were still the main shareholders of the major market participants. The main problems lie in the lack of management and more detailed regulation to pursue a common level in terms of market opening. This also due to the fact that there was no ideal example for the steps towards electricity market liberalization. Hence, EU commission allows their state members to adjust the gradual opening accordingly. In addition, EU member states are still exposed to the risk of predatory behavior, market dominance, and discrimination in the transmission and distribution.

These conditions motivated EU commission to pursue another goal towards liberalization. The current objective has changed from initiation of market liberalization into creating a uniform level of the competitive market. This is, then, realized and approved in 2003 with EU directive number 54. The directive has added several changes in order to address previous issues. There were two main changes addressed in this directive:

- The obligations to ensure delivery of electricity to all household customers and small-medium companies in a clear, comparable and transparent prices. These changes were aimed to reach a demand liberalization in all EU member states. The previous directives were aimed at the industrial users with big electricity consumptions. Although EU commission have attempted to improve it every three years, many EU member states are still behind in terms of the demand liberalization.

- The obligations to appoint or to undertake the transmission and distribution system with the aim improve market efficiency and economic balance. There are also several rules and criteria that need to be applied in order to ensure the independency of the Transmission System operator. This change was made because many EU member state still allows the former monopolist to have activities in this part of the sector. Hence, they still had a lot of power in price imposition and a high market share in their country.

Besides these changes, the directive also addresses several issues deemed to be important. Firstly, the directive requires a member state to provide protection to the consumers in the markets and protection to the risk of a blackout. Hence, it obligates member states to implement a third party access system for eligible customers. Secondly,The directive also requires member states to ensure the possibility to add capacity or energy efficiency through the tendering process which is transparent. Therefore, the tendering procedure must follow the EU commission rules since it has to be published in Official Journal of the European Union. Thirdly, the directive makes another emphasize to the unbundling of accounts in order to evade competition disturbance, cross-subsidisation , and discrimination. In fact, they would enforce unbundling if there is a questionable activity in the electricity market. Finally, all member states must report EU commission on the implementation of this directive.

On the other hand, climate change arises as a global issue which requires full attention from all EU member states. This issue has motivated EU commissions to reduce GHG emission by implementing EU emissions trading system (EU ETS) which attempts to combat emission with cost-effective fashion. EU ETS facilitate emission transaction through cap and trade policy that limits the amount of greenhouse gas emission emitted from power plants and factories. In other words, companies can sell their rights to emit if they are able to reduce it and they can also buy additional rights to emit if they are not able to reduce it. The system applies to around 11000 power stations and large industries all over EU. However, even with this new system, a clear and exact target of emission reduction was never published by EU commission.

After several years of research and discussions, EU commission agreed to put a legal framework and goals as a commitment to address this new global issue. It resulted in a new energy policy,well known as the 20-20-20 policy, published in EU directive number 28 in 2009. The policy states a clear goal to reduce energy consumption by 20%, greenhouse gas emission by 20%, and 20% renewable energy mix by 2020. The directive also established several key frameworks in order to reach its target. Firstly, it strengthens the EU emission trading system by creating a single EU wide cap in order to replace national cap system that they have previously and broaden the coverage for more gasses. Secondly, it sets national targets for emitter outside EU ETS, such as transportation, housing, and agriculture for all EU members. Thirdly, it sets national renewable energy targets in order to push renewable mix in the national productions. Fourthly, it establishes a legal framework for utilizing carbon capture and storage technologies. Fifthly, it sets EU energy efficiency target which is, later, implemented in 2011 through EU energy efficiency plan and EU energy efficiency directive.

The directive has changed the tone in electricity market since it promotes intensively renewable energy generation for increasing production mix and achieving their goals in emission reductions. It ignites new policies in EU member state for facilitating new investments and installations of renewable. As a result, more intermittent generations enter the electricity markets. However, there are several issues that come with the changes in production mix. Renewable supply highly depends on the weather which varies according to the locations. Therefore, renewable production units are generally far from the demand site. This is also due to public protections from the pollution from these power generations (e.g. sound pollution). Countries without an adequate infrastructure in their electricity system would face congestion problems and isolated electricity market since it has limited access for physical exchange. Furthermore, the actual power supply from these generations can not be planned in advance since it depends on the weather. Hence, it introduces more volatility in the electricity markets.

1.1.2 Deregulation of electricity market in Italy

From 1960's, Italian electricity market was controlled, managed, and organized under a solitary vertically integrated company, Enel, which is owned by the state. Therefore, generations, transmissions, and distributions are all operated by Enel. However, with new EU directive on energy market deregulation was passed in 1996, many changes were applied in Italy. In 1999, legislative decree 79/99 was agreed by the parliament and has sparked the start of the free electricity market in Italy. [1]. In line with the European directives, the decree, now widely known as *Bersani decree*, address many issues regarding liberalization, renewable dispatch, and green certificate.

[1] The decree was passed on 16 March 1999 in compliance with 96/92/EC

The liberalization process starts with ending the national monopolies and opening the market for new entrants in order to promote competition. The decree provided specific requirements to Enel for re-structurization of the electricity sectors. The requirements are:

- Enel had to start privatization and to unbundle all their activities into different accounts. The activities are :
  - Power generations
  - Energy distributions and sales
  - The dismantling of nuclear power plants
- Transmission activity in the electricity sectors must be given up to an independent body with non-discriminatory rules.
- Enel had to sell at least 15 GW of their energy capacity in order to reduce their market dominance and to open electricity market for new entrants.

The transmissions activities, as dictated by EU directive, was designated to a new state-owned company, Terna, which was a part of ENEL group. Later in 2003, they had to be separated from Enel as a consequence of new EU directive number 54. Now, they are an independent company owned by the Italian government and private investors. As for the sales of capacities, the government facilitated Enel for the best method of these sales in order to guarantee fair market conditions and transparency. The generation units were, then, sold to major EU companies. The new owners of these capacities are:

- Endesa, Spanish utility company, and the bank of Santander, major Spanish bank, who jointly acquired 5.34 GW
- Edison, Italian utility company, and Atel, swiss utility company, who jointly bought 7 GW of the capacity.
- Energia Italia, Italian utility company, and Electrabel, Belgium utility company, who obtain 2.6 GW from the sales.

As the decree suggested, the next phase of the deregulated market was demand liberalization and creation of internal energy market, Italian government started to open the market and considered a further measure to ensure the implementation. The market was initially opened for large industrial customers that consume 30 GWh of electricity every year. Therefore, there was still large portions of customers that did not have the option to choose their energy suppliers. This is, later, improved gradually between 1999 until 2007 as follows :

- in 1999, users with a minimum 30 GWh/year of consumption is considered as eligible customers
- in 2000, users with a minimum 20 GWh/year of consumption is considered as eligible customers.
- in 2002, users with a minimum 9 GWh/year of consumption is considered as eligible customers.
- in 2003, users with a minimum 100 MWh/year of consumption is considered as eligible customers.

- in 2004, all non-household users is considered as eligible customers
- in 2007, all users are considered as customers.

As far as the creation of internal energy market concerned, operators were initially able to trade electricity through bilateral contracts. However, the decree also recommended the establishment of the independent market operator that operates the market with the objective to fully implement EU directive. The establishment of an independent body to manage the market, operators were expected to be able to buy or sell electricity by two means, bid/offer of standardizing contracts in the electricity exchange and over-the-counter contracts. Therefore, this decree initiates the legal framework for the creation of Italian power exchange and the formation of the Gestore Mercato de Energi (independent market operator) in 2004

It is also interesting to be noted that the Bersani decree also initiates green certificate mechanism which was designed as a change for feed-in price policy under CIP6/92. In this mechanism, instead of fixed guaranteed income, the new support is traded base in the market under the quota system. The quota system obliges power producers and importers to produce a certain quota of their outputs from renewable sources. The quota starts from 2% and gradually increasing annually. Green certificates were used to fulfill this obligation as they are able to purchase them from third parties. The certificates were traded on a parallel market, independent of the electricity market. Then, as encouragement for renewable energy penetration, the decree provided dispatching priority to its production and increase government support for their researchers.[2]In addition,it also introduced a renewable-energy quota system. Several years later, improvements had been made to the green certificate scheme in order to promote renewable-energy further. In 2007, under the Law 244/2007, small generators (0.2 MW -1 MW capacity) have the option to obtain a support scheme by selling their green certificates on the market or receive a feed-in tariff. In addition, the law also extended the period for the release of green certificates up to 15 years for new and refurbished installations. In order to enhance renewable growth, finance act of 2008 and 2009 ministerial decree increased the quota by 0.75 % per annum over the years between 2007 and 2012. Consequently, power producers and importers were required to produce 6.8% of their supply from a renewable resource in 2011 and 7.55 in the later year. The law also introduced another mechanism that considers the technology maturity. Hence, the quantity of the certificate granted to producers and importers, with more than 1 MW capacity, was multiplied by a coefficient depending on the technology ranging from 1.0 for the onshore wind turbine to 1.8 for wave energy conversion. Moreover, according to the law, one certificate had a contract size of one MWh instead of 50 MWh as in the previous regulation. Therefore, it can support newer installations for small renewable-energy technology. In 2011, a new legislative decree was passed on 3 March. The decree set a national target of 17% of renewable energy in the gross final energy consumption by 2020 in compliance with European Directive (2009/28/EC). With a view to realizing the goal, the decree changes the quota system back into a feed-in tariff scheme from 1 January 2013 under a given threshold and a tendering scheme for new plants (except biomass) with a capacity above the threshold. The threshold differs according to the type of technology. Therefore, a guaranteed fixed price is offered to support the renewable energy penetration. Consequently, GSE was obliged to buy all certificates that exceed the annual demand in 2012.

[2]The decree enhanced government support to the Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) for research, innovation and technology transfer for renewable-energy

1.1.3 Motivation

The deregulation of electricity markets in Italy has changed the shape of the economy in the electricity sector. Furthermore, specific characteristics of Italian electricity market makes this particular electricity market an interesting case study. Firstly, Italian governments and independent body in the electricity market have decided to apply an inter-zonal pricing mechanism for the price formation. The proposed price mechanism would not only reflect actual supply and demand in Italy but also reflect the actual net physical exchange between their administrative zones. This pricing mechanism was a normal practice as it has been applied in many countries such as United States, Australia, and Denmark. However, in comparison to other countries, Italy has the largest number of zones in their national energy market reaching six zonal prices. Secondly, recent ambitious policy by Italian government has increased the intermittent generations in the exchanges. The dependency on the weather has created uncertainty in the supply. As a result, the price has become more volatile compared to previous years. This is a big problem from the point of vies of the operators as they are exposed to financial risk. Thirdly, capacity transmission lines between Italian zones are not equally distributed. As the renewable generation units are located strategically to harvest energy, they are generally far from the population where there is a high demand for electricity. As a consequence, congestion occurrence becomes a problem in Italy since the infrastructure is not equal in all the zones.

These interesting facts in Italy have brought our attention towards analysis on the Italian electricity market. They influence our motivation to explore the literature with the aim to seek the gaps in academic and to contribute to the economic and empirical research. Our studies have brought us to three main line of researches that motivates us to explore intensively in this particular market. They are :

- Alternative models to forecast Italian electricity market
- The impact of renewable supply on congestion and congestion cost
- Examination of national integration in the Italian electricity markets.

The following subsections are dedicated to briefly describe the literature gap in these line of research and its motivational background.

Alternative models to forecast Italian electricity market

The electricity price has changed its shape after the electricity market liberalization and has sparked many researches in this field. This is mainly due to the fact that previous electricity price was not volatile since it was fixed through policy. In fact, all the activities in the Italian electricity sector were controlled, managed and organized by one state-owned company. This old framework provides no competition in the market and creates an inefficient economy. The deregulation, on the other hand, has opened new entrants in the electricity sector and has provided a fair price to reflect the actual supply and demand. The new mechanism has given the opportunity to new producers and energy retailers to compete for the right to deliver/come electricity in the exchange. Furthermore, the price would reflect the actual scarcity and surplus of electricity in the market based on the actual supply and demand.

Despite the benefits offered by the market, there were issues that come because of the uncertainties provided in the market. As the price is formed by supply and demand functions, deviations on these functions would create shocks in the market equilibrium. These shocks can be negative because oversupply in the market, which would translate into a low price, and can be positive because a scarcity in the market, which would translate into a higher price. As a consequence, the price can changes rapidly from zero (or negative) to a price jump reaching more than 100 €/Mwh. This is also due to the fact that electricity cannot be stored thus becoming vulnerable to price spikes. In Italy, the rise of intermittent energy in the power exchange has worsened the situation as it increases the volatility in the electricity price (see for instance Clo et al., 2015). On the other hand, Italy has the highest average wholesale price in comparison to the other mature EU market since gas is still their marginal technology (GME). Hence, it indicates a high deviation in the electricity price caused by demand changes, fuel price changes and supply deviation. These price changes exposed the market participant to financial risk in the electricity market. The changes in the spot price can reduce their economic gain because of the lost from price fall or jump. On the other hand, policy makers are obligated to follow the fair price in the market. Therefore, it is necessary to provide a reliable model to forecast electricity market that could also reflect the power exchanges.

Our research focused on research on the best forecasting model under statistical class [3] with the main objective to contribute to the limited literature of Italy electricity market. Our research offers an alternative model for forecasting hourly price in Italy since there is only a few researches in this particular field and market. The statistic class was chosen because it has the advantage of having an accurate quantitative prediction, having less required observations and having results that can be interpreted in an economic sense. Based on our review, we can only find four alternative models, under the statistical class, to forecast electricity price in the Italian electricity market's literature. [4] The four models existed in the literature are Periodic Autoregressive models (Bosco et al , 2007) ARMAX-GARCH (Petrella and Sapio, 2011), Generalized Additive model for Location, Scale, and Shape (Serinaldi ,2011) and ARFIMAX-GARCH, (Gianfreda and Grossi, 2012).

The impact of renewable supply on congestion and congestion cost

As it was mentioned in the previous subsection, EU directive number 28 in 2009 (20-20-20 policy) has set ambitious goals to show commitment to combating climate change. As a consequence, it has ignited the interest in renewable energy investment in Europe because many EU member state starts supporting schemes in order to meet their national renewable energy target by 2020. In 2013, the renewable energy generations in Europe has supplied 14.95% of gross final energy consumption. These changes have motivated researchers to study the impact of the renewable energy on the electricity markets. Economic literature has put their main focus on the changes in wholesale price. These researches have underlined two main changes in the wholesale price, the reduction in terms of price level and the increase in volatility. The empirical evidence of these two phenomena has been well-documented in the academic literature with various case study and methods. [5]

However, the rise of renewable supply has raised another issue in regards to the electricity network, which has not intensively explored in the academic literature. As renewable production units are generally located far from the demand site because of optimization in harvesting energy, high capacity transmissions are required to ensure

[3] This classification is based on Weron (2014)

[4] From our researches, we also found literature that utilizes other techniques with Italian electricity market as their case study. However, they are classified under different classes (see Bompard et al. (2008)); Guerci et al.(2008).

[5] see (Australia (Cutler et al., 2011), Austria (Wurzburg et al., 2013), Denmark (Jónsson et al., 2010), Germany (Wurzburg et al., 2013; Ketterer, 2014), Israel (Milstein and Tishler, 2011), Ireland (O'Mahoney and Denny, 2011), Italy (Clo et al., 2015), Spain (Gelabert et al., 2011)

electricity delivery. However, the existing network was not built for high renewable energy penetration, thus, the changes in the renewable energy mix, consequently, put additional stress on the infrastructure, amplifying transportation needs and multiplying congestion occurrence. Unfortunately, there are not many works of literature focused on this particular issue.[6]

Our studies are aimed to contribute to the limited literature in this particular field. Based on our knowledge, Sapio (2015) is the only author who studies the impact of renewable to the congestions in Italy. Utilizing regime-switching model, he estimated Italian power exchange data between 2012 and 2013 with a focus on Sicily and Southern Italy connection. The paper concludes that rising renewable production in Sicily reduce the congestion as the zone require less import from its neighbor. The increase of renewable from outside Italy (Southern Italy), on the other hand, would increase the urgency to export the electricity to its neighbor. Consequently, it results in the increase of congestion towards Sicily. The paper also concludes that wind energy supply provide a bigger impact on congestion in comparison to solar productions unit.

Isolated regional market

Deregulation of the Italian electricity market has changed the previous mechanism into a cost effective mechanism for electricity generation and transmissions. This was one of the top priority from the policy maker since an efficient market and network was the main objective of the EU directive and Bersani decree. In order to reach cost effective transmission cost, Italy adopts inter-zonal pricing mechanism which separates national electricity market into several zones based on the conditions of the electricity system. If a congestion was found in the system, the mechanism allows zonal market prices to be higher in order to balance the system and pay the cost of physical delivery between the zones. However, this could be a problem for countries with low capacities on inter-zonal transits since it will increase the congestion frequency. Consequently, the national markets would constantly splits, which results in isolation of zonal markets. Therefore, regions that do not posses adequate network capacity for physical exchange with its neighbor would fail to integrate into the national market.

This issue motivates us to examine the integration of Italian electricity market that is never explored by previous literature. Having six zonal markets on their national market, Italy offers interesting study for our research. Furthermore, new installations of transmission lines between Sardinia and Italian peninsula provide us with a unique case study for examining its impact on the regional market integration. We may examine the integration by studying the interdependencies of zonal prices since it reflects the linear relationship among the regional markets. Strong interdependencies suggest a stable transmission line and a full integration in the national market. Previous literature on this line of research is mainly used in a case study from outside Europe, Australia, and United states. In Australia, Worthingthon et al (2005), Higgs (2009) and Ignatieva and Trück (2011) have explored interdependencies of five regional markets using different methods. The researches provide a general conclusion that regional markets with better networking infrastructure display strong interdependency whereas the weaker level of interdependencies is recorded in the markets with low capacity of the transmission line. In the United States, Park et al (2006) analyze United states spot markets using Vector autoregression. Their papers conclude that interzonal transits and national market organization affect the interdependencies. Another national market examination in the United States was

[6] see Førsund et al. (2008), Woo et al. (2011), Schröeder et al. (2013) and Kuntz (2013), and Figuiredo et al. (2015) .

done by Dempster et al (2008) for California electricity market by means of granger causality test.

1.2 The Italian electricity market

Five years after the Bersani decree passed, the national electricity exchanges could be finally operated. The first stage of IPEX commenced on 1 January 2004 as technical trials with operators from supply and demand sides. The technical trials were aimed to ensure a proper infrastructure and mechanism have been installed in order to ensure a transparent and competitive electricity market in Italy as well as to avoid the risk of an unbalanced system. The seconde stage of IPEX was partial functioning of the exchange (only for a supply side) as the initial trials. Then, the beginning of 2005 was a milestone in Italian electricity sector as Italian power exchange starts complete operation of the exchanges. The market is now widely known as Italian Power Exchange (IPEX). Currently, it has 254 market participants who actively buy and sell electricity contracts with a total of 337.34 TWh of electricity was traded in 2014.

Beyond 2009, Italian power exchange did not go through any major structural changes. The market had matured enough to operate regularly in the transparent fashion, without the needs to majorly modify the structure for improving liquidity and risk mitigations. Between 2005 and 2009, Italian power exchange has seen three major structural changes that affect the exchange. The major structural changes are :

- Demand liberalization

  As it was mentioned in the previous literature, Italy adopted gradual demand liberalization between 1999 and 2007 that affects the structural changes in the market. When IPEX was opened and fully operated in 2005, only industries were allowed to enter a bid in the exchanges. Hence, all households consumers were not considered eligible by the legal framework. However, this is changed after July 2007 as all consumers become eligible in the power exchanges. This change has put 85 TWh of yearly demand in the markets which constitute 22 million families. Consequently, this change would introduce high market fluctuations in the short-term electricity price in Italy thus increasing volatility. Furthermore, these new entrants may engage in a trial-and-error learning process that produces instability in the market.

- Contract for Differences

  Starting from 2008, the ensuing price risk from the day-ahead market in IPEX can be hedged by contracts for differences (CfDs), which is held in Italian derivatives exchange (IDEX) operated by Borsa Italiana. As a result, hedging strategy can be implemented by market participants. In fact, Acquirente Unico is obliged by law to hedge against price and volumetric risks. This major structural changes reduced market participants' exposure towards financial risk. However, before December 2009, holders of the contracts are not able to physically exercise the financial contracts that they obtain from IDEX. Hence, GME, finally, introduces *Piattaforma Consegna Derivati energia* (CDE) with the purpose of converting the financial derivative contract concluded in IDEX to physical delivery of electricity.

- PCE

  The last major structural changes in the Italian structure was done by the end of 2009. Under art 17 of Annex A to AEEG's Decision 111/06, a special platform called *Piattaforma dei Conti Energia a Termine* (PCE) is started to be operated and managed by GME in order to facilitate electricity delivery from bilateral contracts concluded outside the IPEX. The platform is used by participants to register their injection and withdrawal schedules thus easing the works for managing transmission. This change has provided the market participant with options to choose between standardized contracts in physical exchange or OTC contracts. Furthermore, the platform offers the posibility to adjust easilly both long term contracts and short term contracts. As a result, the markets becomes more active and competitive which affects the Italian electricity prices.

1.2.1 The exchanges

The Italian electricity market consists of two classifications, spot, and forward market. The market structure can be seen in the Figure 1.1 below. All markets are organized by the Italian independent market organizer (GME) [7] with an exception for financial derivative contracts, which is traded by Italian derivatives exchange (IDEX) of Borsa Italiana. The IPEX organizes all the markets under GME SpA with an exception for financial contracts, which is traded in IDEX by Borsa Italiana. GME also manages the OTC registration platform, PCE, where all market participants can register their bilateral contracts concluded outside the IPEX or IDEX. Our analysis and studies, however, focus on solely on the day-ahead market where 24 periods of electricity delivery are traded. In Italy, the day-ahead market opens on the ninth day before the day of delivery and close on the day before the delivery. Between this period, all buyers and sellers are allowed to make bids for their willingness to consume/produce electricity by specifying the quantity and the maximum/minimum price they desire. [8]After the market closes, an authomatic algorithm is used optimize the transmission system and determine the electricity price.

In order to maintain and optimize Italian electricity system, IPEX divides Italy into six main geographical zones and five poles of limited production. The map of Italy in Figure 1.3 below displays zonal division of all the twenty administrative regions in Italy. NORD has the biggest territory covering eight regions followed by SUD with four. Sicily and Sardinia are the only two regions that are considered as one zonal market. The inter-zonal connections of the Italian electricity system can be found in the Figure 1.3 below. The electricity markets are coupled with four foreign markets, France, Switzerland, Austria, and Slovenia. Hence, it allows to import/export to/from Italy. The six main zonal markets are connected to the pole of limited productions or constrained zones. These zones are only a sets of generation units connected to Terna without withdrawal points (no demand in this limited zone). The maximum production that can be exported to the grid is lower than the maximum net physical exchange in the case of congestion. In other words, they are used to to balance the Italian electricity system in order to avoid congestion.

[7] Gestore Mercati Energitici

[8] For each day and each offer/bid point, a maximum of 24 bids/offers can be submitted. Three types of offer/bid exist: simple, consisting of a pair of values indicating the volume of electricity offered/bid in the market by a market participant and the price for a given hour; multiple, consisting of the division of an overall volume offered/bid in the market by the identical market participant for the same hour; pre-defined, consisting of simple or multiple offers/bids, which are submitted daily to the GME.

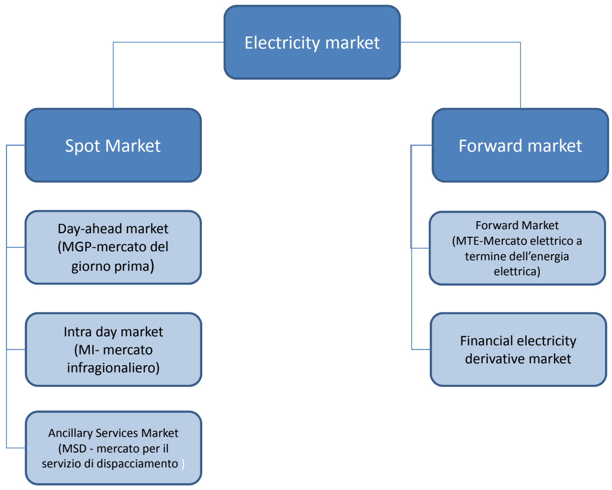

Figure 1.1: Italian geographical zones
(Source: author elaboration from GME)

As shown in the Figure 1.3 below, the poles of limited production are coupled with the close geographical markets to form six large Macro-zones: Monfalcone (MFTV) is associated with NORD, Brindisi (BRNN) and Foggia (FOGN) to the SUD, Priolo (PRGP) to SICI, and Rossano (ROSN) is a bridge that connects SUD and SICI. They are national virtual zones with a constrained set of production units whose quantity is lower than the admissible exports. Hence, they are used to balance the Italian electricity system when it is necessary.

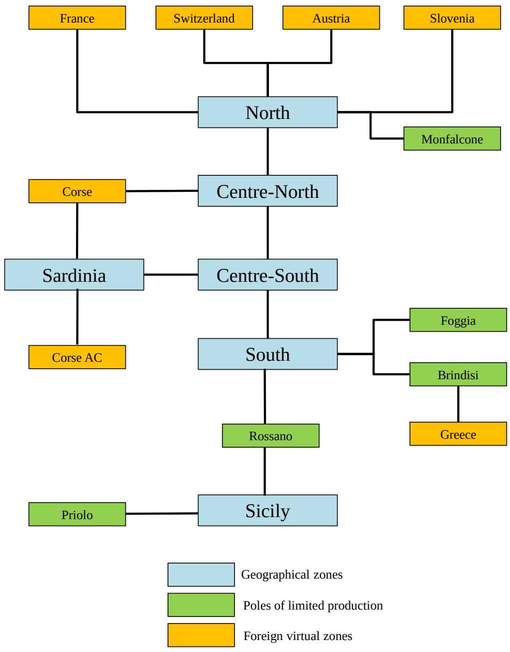

Figure 1.2: A stylized representation
(Source: author elaboration from GME)

Figure 1.3: Italian geographical zones
(Source: author elaboration from GME)

1.2.2 Liquidity

The figure below displays the liquidity and traded volume in the IPEX. Between The volumes traded in IPEX is averaging 190 TWh. This number is not far from average volume traded in EXX for Germany and Austria combined that reach 256.5 TWh. The traded volume in Italy has shown a declining trend between 2010 until 2012, from 199 TWh and 179 TWh. Then, it increases in the following year reaching 207 TWh of traded volume.This is due to the increase of liquidity in those same years where it reaches 72% of liquidity. Non-institutional participants highly contribute to the increase as they constitute 130 TWh quantity in that year. This is 42 TWh increase from the total volume traded in the previous years. It is also interesting to note that GSE offers have shown a significant increase from 2011 to 2013 as it increases from merely 39 TWh to 50 TWh in 2013. The increase of GSE offers exhibits the increase in renewable energy production since they all constitute the bids from renewable energy supply. AU , on the other hand, reduced their volume significantly from 40 TWh to only 27 TWh. In this same year, traded volume in Germany and Austria (combined) in EEX recorded 247 TWh only 40 TWh different between them. In 2014, however, IPEX saw a declining quantity and liquidity where it register, 185 TWh and 66 % respectively. Both institutional and non-institutional participants show significant declining volume. In particular, non-institutional participants reduced the quantity. On the other hand, EEX shows a positive trend with 8% changes from the previous year.

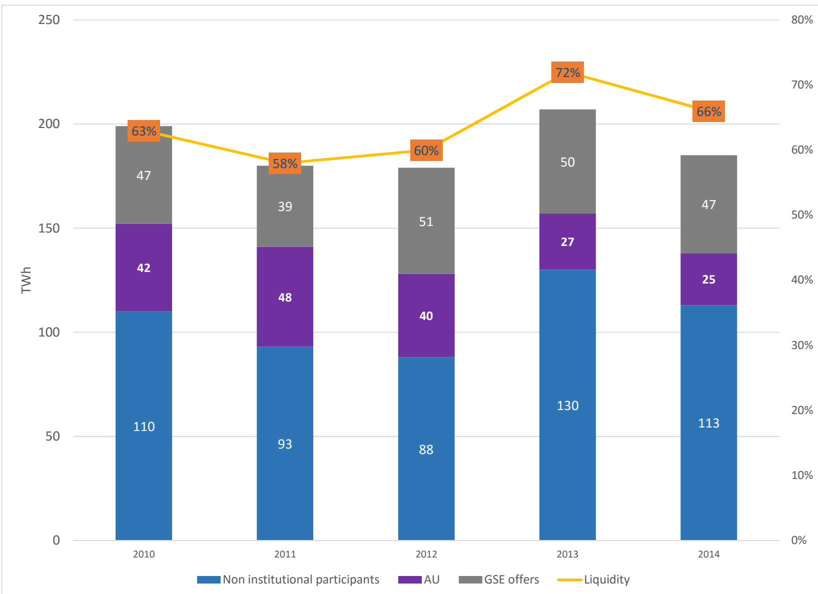

Figure 1.4: Traded volume and liquidity in IPEX
(Source: author elaboration from GME)

1.2.3 Price formation mechanism

The algorithmic procedure for determining the electricity price follows iterative steps. First, the aggregated supply curve is constructed following the merit order. Hence, all bids from the suppliers are ranked in ascending order according to the price. Then, the demand curve is built in a similar way to the descending order. Second, the algorithm will verify the adequacy of the electricity system. If there are no lines congested between the zones, then all zonal price is equal to the National Unique Price (PUN). Therefore, the price equilibrium is an intersection between national aggregate demand and supply curves. On the other hand, if the opposite case occurs, the zones are split into two big zones in the congested transmission line. The algorithm is then restarted for the two split zones and the market binds with different price equilibriums. Hence, it leads to two zonal prices, which is used as a reference for producers remuneration. The PUN price is the consumer price, and it is calculated as a weighted average of all zonal prices. Then, in the case where a transmission congestion is still found in the system, the algorithm is restarted again with more splittings until an optimization is reached. In this case, there can be more than two zonal prices in Italy but there is still one PUN price.

In this market, transactions take place between the ninth day before the day of physical delivery and the day before the day of delivery. The sellers submit hourly offers for each generating unit specifying the quantity and the minimum price at which they are willing to trade their power. The aggregated supply curve is built according to the merit order in an ascending order of price. In a symmetrical way, the market demand curve is generated through the aggregation of single bids in a descending order of price.[9] The hourly market price is determined by the intersection of the demand, and the supply curves, following an iterative procedure. Firstly, the geographical market is considered as unique: if the day-ahead production/consumption plan respects all network constraints across zones (no congestion), a single price for the whole country emerges.[10] On the contrary, if a network constraint is saturated, then the geographical market is divided into two sub-markets, each one aggregating all the zones above and below the saturated constraint. The market demand and supply curves are rebuilt for the two sub-markets (taking into account the quantity that can flow between zones up to the transmission limit), and two zonal prices result. The hourly auction is a uniform price auction which means that all accepted units are entitled to receive the system marginal price (or prices when de-zoning arises because of transmission congestion). Figures 1.5 and 1.6 illustrate through an example how the inter-zonal price mechanisms work under uniform auction rule without and with congestion respectively.

[9] For each day and each offer/bid point, a maximum of 24 bids/offers may be submitted. Three types of offer/bid exist: simple, consisting of a pair of values indicating the volume of electricity offered/bid in the market by a market participant and the price for a given hour; multiple, consisting of the division of an overall volume offered/bid in the market by the identical market participant for the same hour; pre-defined, consisting of simple or multiple offers/bids, which are daily submitted to GME (GME).

[10] The price will be in correspondence of the intersection of national demand and supply curves.

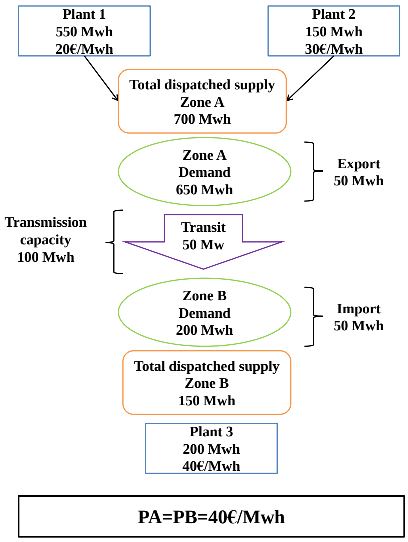

Figure 1.5: Pricing without dezoning

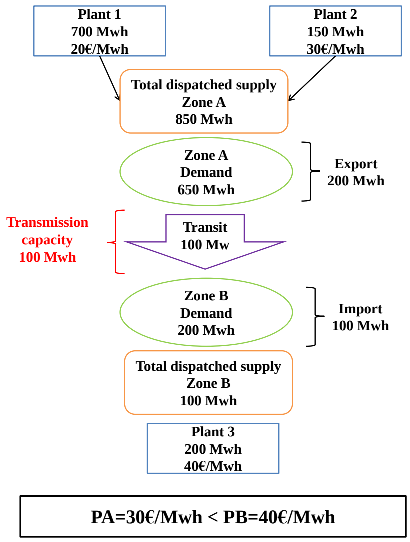

Figure 1.6: Pricing with dezoning

In the permanence of network saturation, the process of sub-setting the market continues until all constraints are satisfied (Fig. 1.7).

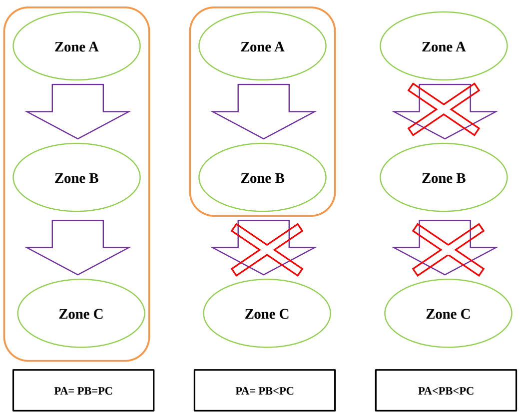

Figure 1.7: Multiple congestion

While producers receive the zonal prices in the occurrence of congestion, the buyers pay the National Single Price (PUN) for the electricity bought in the pool: the PUN is an average of zonal prices weighted for the zonal purchases.[11]

[11]The purchased quantity should be netted of purchases from pumped-storage units and from foreign zones. In the example reported in Fig. 1.6 the PUN would be equal to 32.35€/Mwh:

$$\text{PUN} = \frac{\sum P_i Q_i}{\sum Q_i} = \frac{(30 \times 650) + (40 \times 200)}{650 + 200} = 32.35$$

The electricity price formation follows an iterative procedure based on the interzonal market mechanism. Firstly, the market collects all the supply and demand bids from all the zones. The supply and demand curve are then constructed under merit order. Secondly, the algorithm verifies the transmission limitation between the zones. If there is congestion in the transmission line, then, all zonal prices are equal to National Unique Price (PUN), which is the intersection between the supply and demand curve. If the opposite case occurs, the congested transmission line split the connected zones into two zonal markets with their own supply and demand curve. If the national electricity system is still not stable, the algorithm is repeated subsequently with more market splitting until an optimum solution is obtained. As a result, it is possible to have six zonal markets if all the high transmission lines are congested.

1.2.4 Production Mix

Let us now take a look at actual production mix in the figure 1.8, the domination of gas as energy sources have gradually decreased at a rate of -49.79% as a result of intensive renewable policy support from the government. In fact, RES have surpassed total CCGT production for the first time in 2014. As for coal source, their contribution to production is increased from 24.4 to 32.3 between 2010 and 2012 but decrease afterward amounting to -22.6% changes at the end of 2014. Other sources, which include biomass, CHP, oil and other thermal sources, are observed to be stable over time with quantity only range between 29 and 30.6 TWh in this period. Even with high renewable penetration, Italy is still a constant net importer. The statistic suggests that, in average, 15.9% of the national demand is supplied by neighboring countries.

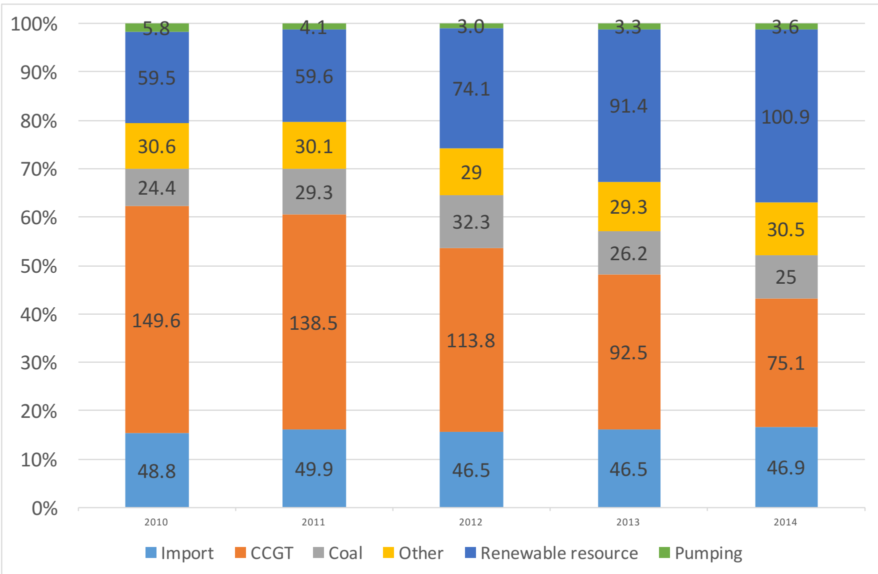

Figure 1.8: Production mix between 2010 and 2014 (Twh)
(Source: author elaboration from GME)

If we look at the breakdown of renewable supply shown in figure 1.9, wind, and solar technology had been the key to Italy's success as they had been growing at a high pace. Solar registers the highest increase in the production multiplied by 4.5 times at the end of 2014. As a result, they account 29.9 % of all renewable supply or 10.7% of the total production mix. As for wind technology, the penetration reached more than 160% changes in the end of the period, which translates to 14.4% of the renewable supply. Then, hydro production is decreasing in the green certificate scheme period, but it increases 43.4% with the introduction of the new feed-in-tariff. Although the quantity increases, hydro's share has been reduced by only a half of total renewable production in 2014. Finally, the quantity of geothermal is constant over time due to the limited resource in Italy.

This section is devoted to the description of Italian production mix, interzonal transits and zonal price differences as they result from the day-ahead ex-post market data. From 2010 to 2014, the contribution of Italy's main source of electricity, gas, has gradually decreased with the 2014 quantity (75.1 TWh) representing almost the half of 2010 figure. RES supply has surpassed total CCGT production for the first time in 2014 (100.9 TWh versus 75.1 TWh). In this year, renewable production has exceeded the target established in the National Renewable Action Plan (NREAP) to produce 100 TWh of renewable energy by 2020. Even with high renewable penetration, Italy is still a net importer. The statistics suggest that on average 15.9% of the quantity accepted in the day-ahead market is supplied by neighboring countries.[12] A detailed figure of the quantity sold in the day-ahead market by production source for the period 2010-2014 is reported in the Appendix (Figure 1.8). The breakdown of renewable supply by technology (Figure 1.9) reveals that wind and solar have experienced the strongest growth. Solar production registers the highest increase in the supply of 2014 is 4.5 times the one of 2010. As a result, in 2014 solar technology accounts for 29.9% of total renewable supply and 10.7% of the total production mix. Wind supply in 2014 is 2.6 time the production in 2010 reaching 14.6% of total renewable supply. Hydro production has decreased instead from 2010 to 2012 to rise again afterward. In 2014, it represents half of total renewable production.

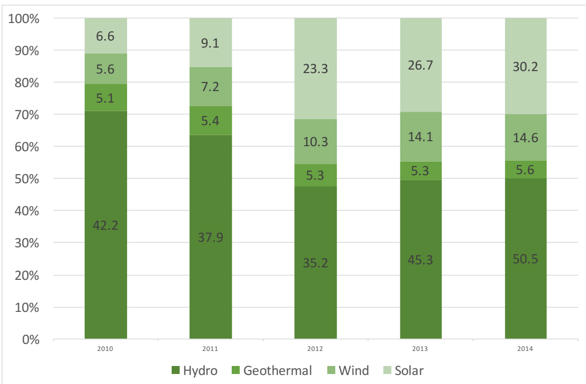

Figure 1.9: Renewable mix between 2010 and 2014 (Twh)
(Source: author elaboration from GME)

[12]This is probably due to the fact that these countries have cheaper generation mix (nuclear).

1.2.5 Physical exchange

Physical exchanges resulting from the day-ahead auction have experienced some changes over the years. Figure 1.10 shows the average net electricity flows on Italian main lines.[13] CNOR, SICI, and SARD are net importers, while CSUD and SUD act as a hub in the center and southern part of Italy, as play the role of both importer and exporter. NORD and ROSN are the main exporting regions that deliver electricity to CNOR, SUD, and SICI. However, ROSN is a virtual generation zone used for balancing the system, thus the regions do not have a withdrawal point (buyer). In terms of quantity, CSUD-SUD connection registers the highest net average physical exchange, thus cementing CSUD position as the biggest importing zone in Italy. The imports are, however, gradually decreasing. Similar patterns can be observed in NORD-CNOR and ROSN-SUD, with larger decreases in import, -76.2% for CNOR imports and -50.8% for SUD imports. CNOR imports from CSUD have, instead, increased from 2010 to 2014. Transits from CSUD to SARD display the highest increase as the quantity more than doubled in 2014 compared to 2010, as a result of the new grid connection system. SICI import continues to increase at a rate of 76% over the whole period.

Net physical exchanges between the markets have been experiencing changes over the years. Figure 1.10 below shows us the average of transferred powers between the zones for each transmission line. The columns are constructed from ex-post net flow data published by GME. First of all, it can be concluded that CNOR, SICI , and SARD are net importer regions as suggested by the direction of the quantity. CSUD and SUD, on the other hand, act as a hub in the center and southern part of Italy as they act as both a net importer of electricity and a net exporter. NORD and ROSN are the main exporting regions that deliver electricity to CNOR, CSUD, SICI and SUD. However, ROSN is a virtual generation zone used for balancing the system, thus the regions do not have a withdrawal point (buyer).

In terms of quantity level, CSUD-SUD connection registers the highest net average physical exchange, with more than two GWh, thus cementing CSUD position as the biggest importing zone in Italy. The level is gradually decreasing with -23% of changes of net flow physical exchanges from 2010 to 2014. An identical pattern is also shown in the second and the third biggest line, NORD-CNOR, and ROSN-SUD, with a much bigger decrease, 76.2% and -50.8% respectively, in the same period. In NORD-CNOR, even though it peaks in 2011 with more than one GWh, the average physical exchange plunges to mere 0.22 GWh, which implies more exports to NORD in 2014. As for physical transfer from ROSN to SUD, the net flows decrease annually until it reaches 0.44 GWh. CSUD to CNOR present fluctuation of quantities in this period with mean quantities ranging from 0.46 GWh to 0.79 GWh. Nevertheless, CNOR is still one of the biggest net importers along with CSUD. On the other hand, quantities from CSUD to SARD display the highest increase as the quantity more than doubled in 2014, 0.25 GWh, compared to 2010, 0.1 GWh. This increase is a result of the new grid connection system. Then, SICI continues to increase its electricity import from ROSN at a rate of 76%. Regardless of the increase, SARD and SICI are still the smallest net importer in terms of level because of their transmission limitation and low population compared to the other regions.

[13] Although the detailed statistics is not reported in this paper, the capacities of transmission lines are relatively constant over the years. SARD-CSUD is the only transmission line whose capacity has been reinforced in 2011, thanks to the installation of new submarine cables that started to operate in March 2011. CSUD-SUD connection has the biggest transmission line capacity, while SARD-CSUD and SICI-SUD are the most limited lines.

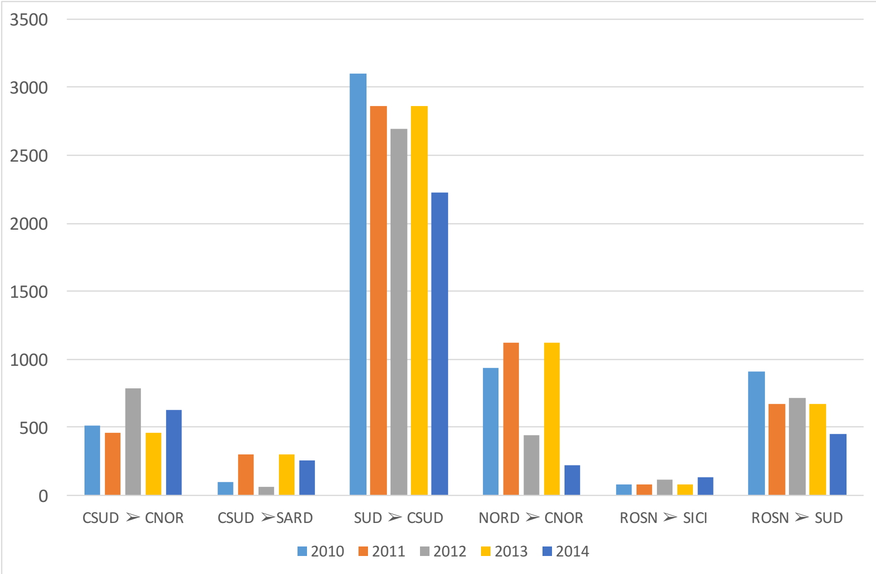

Figure 1.10: Average physical exchanges between the zones
(Source: author elaboration from GME)

1.2.6 Transmission limit

In terms of transmission limits, figure 1.11 below present the average capacity of electricity exchange between each geographical zonal market from 2010 to 2014. Every day, the information is published by GME for preliminary information to market participants. The data is also used as a reference for optimizing the Italian electricity system in the day-ahead market. Although the detailed statistic is not documented in this paper, the capacities of transmission lines are relatively constant over the years. SARD-CSUD is the only transmission line that has augmented capacity in 2011. This is due to the installation of new submarine cables that starts to operate in March 2011. In terms of quantity, we can observe in figure 1.11 that CSUD-SUD connection has the biggest transmission line capacity with SUD may import as large as 10 GWh and export up to 3.86 GWh of electricity on average. CNOR-NORD and CSUD-SUD are ranked behind as they allow lower physical exchanges between the zones. Then, SARD-CSUD and SICI-SUD are the most limited lines in Italy with an average of 0.76 GWh can be exchanged from SARD and CSUD, and no more than 0.18 GWh is transferable from SICI to SUD.[14] This limitation creates congestion problems since only limited efficient supply can be transferred towards or from SICI and SARD.

[14]Since ROSN is located as a bridge between SICI and SUD, we used transmission data of SICI-ROSN. Our assumption is based on several reasons. Firstly, ROSN is a virtual zone used for balancing the system, and it is associated with both SICI and SUD. Secondly, SICI and ROSN are generally considered as one zone based on their zonal prices. Finally, SICI-ROSN is the only transmission line that connects SICI to Italy, particularly SUD.

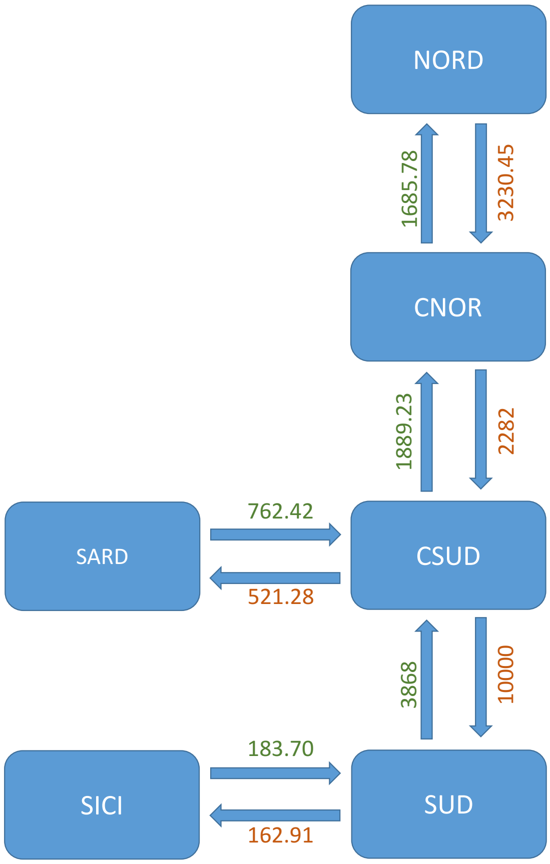

Figure 1.11: Average transmission capacity between the zones (Mwh) (Source: author elaboration from GME)

1.2.7 Zonal price difference

By studying the series of zonal prices, we expect to detect a lasting price difference between importing and exporting neighboring regions. We report the series of paired-price differences for the period 2010-2014 in Figure 1.12 for the following pairs: CNOR-NORD; CNOR-CSUD; SARD-CSUD; CSUD-SUD; SICI-SUD. It is worthy to note that during the considered period the zonal prices of SUD and ROSN have differed for less than the 2% of the time while the zonal price differences between SICI-SUD and SICI-ROSN have followed very similar patterns. This result allows us to consider SICI-SUD pair, which is formally non-bordering zones, instead of the two pairs SICI-ROSN and SUD-ROSN. For the pair CNOR and NORD we observe a substantial increase in the number of hours with negative price difference starting from 2012 (see table 1.13). This result seems to confirm that after a period characterized by a strong reliance on import from NORD, CNOR has reduced its importing needs. In CNOR-CSUD pair, CNOR has been an importer for most of the time, with rising frequency of positive price differences over time. The graph also suggests that SARD generally imports from CSUD while the frequency of congestion between these two regions has decreased at the end of 2012 as shown by many hours of the identical price. In CSUD-SUD, where the first zone is always importing, we may detect a slight decrease in the value of positive price differences. The series of price differences between SICI and SUD reveals that the negative price differences have decreased over time while the positive have substantially remained constant.

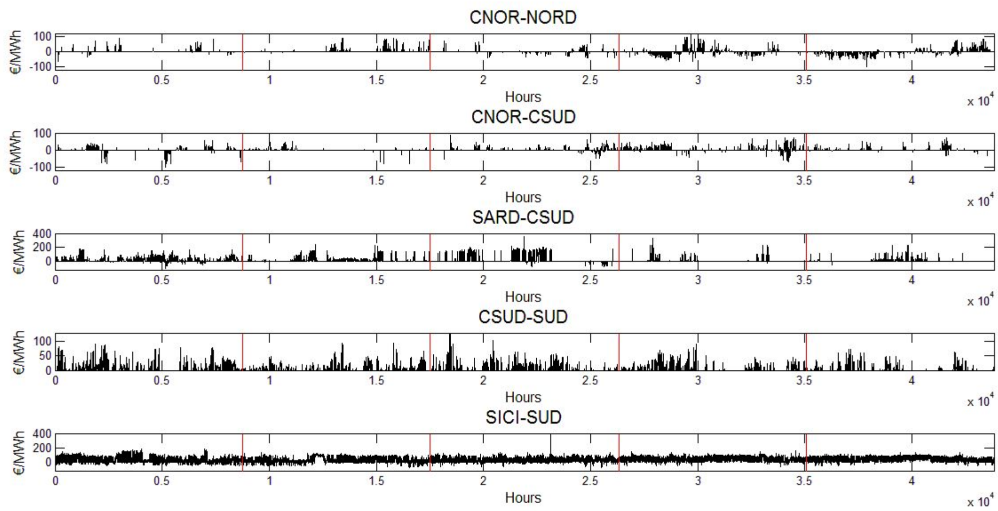

Figure 1.12: Series of zonal price differences, 2010-2014
(Source: author elaboration from GME)

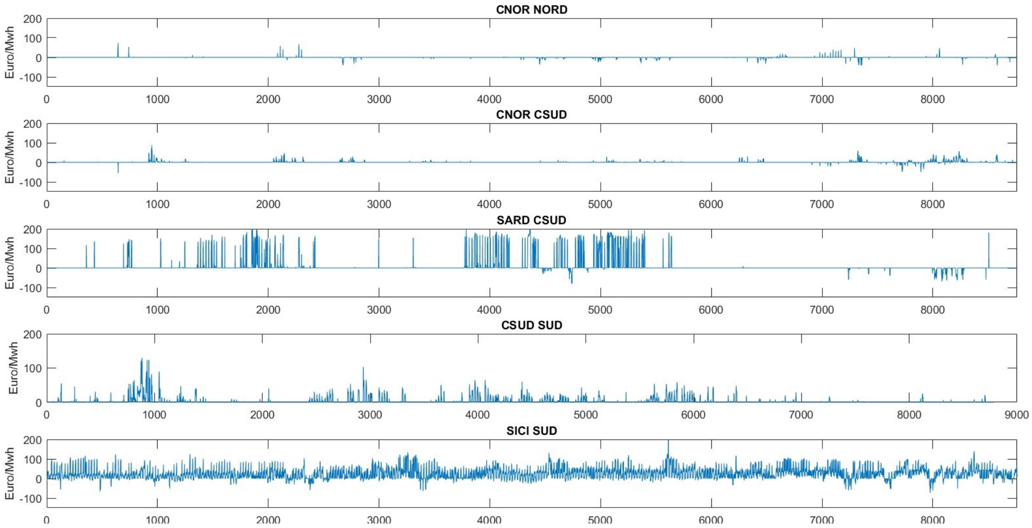

Figure 1.13: Series of zonal price differences, 2012
(Source: author elaboration from GME)

1.2.8 Congestion

To better gauge the relevance of congestion phenomenon in Italy, we have reported in Table 1.1 its frequency and the average number of zonal divisions for the five year period 2010-2014. We observe that congestion frequency has never gone below 82%, reaching a peak in 2013 with the network congested 93.6% of the time. The congestion frequency is a measure of the congestion occurrence in all transmission lines in Italy ,which is very high, since it is counted every time there is, at least, one congestion in the five transmission lines. However, if we look at the congestion occurrence in detail, only a few transmission line is highly congested. In fact, if we only look CNOR-NORD transmission, we may observe a very low frequency in our sample.The average number of sub-markets has slowly decreased from 2.416 to 2.28 between 2010 to 2012 to rise again in 2013 and 2014.

1.2.9 The database

In this study, we have built a unique database collecting bidding data for a five-year period and every hour between 1 January 2010, 00:00 and 31 December 2014, 24:00 from GME, the market operator, which publishes all the auction for the hourly offers in the day-ahead market. Table 1.2 displays a general summary of GME bids' database. The number of bids per year has reached the threshold of 8 million in 2014 while the hourly average number of participating units has slightly decreased

	Frequency	Percentage	Zonal divisions (Mean)	Std	N
2010	7210	82.3%	2.416	0.809	8760
2011	7403	84.5%	2.307	0.686	8760
2012	7921	90.1%	2.240	0.531	8784
2013	8205	93.6%	2.278	0.518	8760
2014	8044	91.8%	2.284	0.543	8760

Table 1.1: Congestion frequency
(Source: author elaboration from GME)

after 2012. The increase in the number of bids indicates that the market becomes more active. On the other hand,the changes in the number of units have suggested continuous new investment on the markets between 2010 and 2012. Beyond 2012, many units probably are not profitable enough to stay in the markets.

	Number of bids	Number of units
2010	6 975 701	976
2011	7 149 431	1 257
2012	7 090 579	1 281
2013	7 737 633	1 152
2014	8 086 282	1 189

Table 1.2: Database summary
(Source: author elaboration from GME)

Our research only uses hourly frequency for two main reasons. Firstly, our forecasting studies are aimed to make a short time prediction in order to be able to be used in high-frequency price modeling and power plant optimization. Secondly, our congestion studies require capturing the event of saturation in transmission line since we want to capture the impact of renewable and demand shock on congestion. This is cannot be captured in a lower frequency. Thirdly, our interdependence study is also aimed to examine the market splitting event which can only be observed in hourly frequency.

Since our researches are observed in hourly frequency, we extract hourly data of PUN price, zonal price, demand and renewable energy from the main database. Hourly electricity prices and demand can be downloaded easily from our database since bidding data already contains this information. However, extracting hourly renewable energy quantity requires third party data in order to match the production units identity with the type of technology. Fortunately, we are able to obtain a list of production units from Ref-E an Italian consulting group. The list provides detailed identity of the production units such as capacity, the technology, and location.

We would like to organize the following subsections as follow. We are going to describe the stationarity test that we use in order to ensure a stationary series in our research. Then, it would be followed by a general statistic summary of electricity price and demand used in our researches. Finally, it would be closed by descriptions on renewable energy data that we use.

1.2.10 Stationarity

In all our researches, it is imperative that the observation is statistically stationary since we would like to avoid spurious regression as well to ensure the assumptions in our model. Therefore, it is important that they do not exhibit seasonal trend and unit root. Most widely-known test for unit root test is Augmented Dicky and Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests. We test both ADF test and KPSS test in our observations data in order to have robust evaluation since they complement each other. ADF test examines the existence of unit root in the observation whereas KPSS test use null-hypothesis that the process is stationary. In other words, it is important to reject the null hypothesis in ADF test and accept the null hypothesis of KPSS test. The result summary can be seen in all our observation from table **??** to table 1.7 in the following subsections. In ADF test, we test down from a maximum lag, $pmax$, that are calculated as follow

$$pmax = int[\frac{12(N+1)}{100}]^{0.25}$$

Where N is number of time series observations. This rule is proposed by Schwert (1989) for $N > 100$. As for KPSS, the optimum lag order, k, is estimated with the equation below.

$$k = int(4 * (\frac{N}{100})^{\frac{2}{9}})$$

From the result of our test shown in table 1.3 to 1.7. The test has proven that there are several variables that accept the null hypothesis of the first ADF test (ADF1). Furthermore, the first KPSS test (KPSS1) do not validate the stationarity of the process in all of our variable. Hence, it is clear that either differencing or de-trending is needed to gain stationarity. Since electricity price is driven by cyclical demand, de-trending the data from seasonal pattern within a week, month and year are preferred. In addition, previous literature has also displayed seasonal characteristics in the electricity price (Contreras et al., 2003; Petrella and Sapio, 2006). Therefore, de-trending is one of the best alternatives in handling this issue. By introducing seasonal dummies for yearly, monthly and daily trend, we have succeeded to accept the null hypothesis of KPSS test and rejecting the null hypothesis of ADF test (see table 1.3 to 1.7). Therefore, seasonality needs to be addressed in our empirical modeling setup.

1.2.11 Electricity Prices

The general statistic summary of the zonal prices from our sample can be seen on table 1.3 below. In terms of mean, the highest average is recorded in SICI (90.20 €/ Mwh) and SARD (69.64 €/ Mwh). NORD, CNOR, CSUD, and SUD follow them from behind with respect to the rank of the mean price, 63.635 €/MWh, 63.632 €/MWh, 62.695 €/MWh, and 60.6 €/MWh. The standard deviation shows SICI and SARD as the regions with the highest volatility, 42.94 and 35.20. One can argue that this is caused by frequent congestion in the lines that connect these zones which subsequently makes them an importing zone. On the other hand, CSUD, CNOR, SUD, and NORD are behind these zones with lower values, 21.39, 20.65, 20.49, and 19.38. The observations are non-normally distributed, as shown by our Jarque-Bera test. Finally, from the test results, we found a stationary time series after de-trending our data.

	Mean	Minimum	Maximum	Std. dev	Jarque-Bera	ADF1	ADF2	KPSS1	KPSS2
CNOR	63.632	0	224	20.656	P<1%	P<1%	P<1%	P<1%	P>10%
CSUD	62.955	0	224	21.392	P<1%	P<1%	P<1%	P<1%	P>10%
NORD	63.635	0	224	19.382	P<1%	P<1%	P<1%	P<1%	P>10%
SARD	69.769	0	450	35.204	P<1%	P<1%	P<1%	P<1%	P>10%
SICI	90.207	0	3000	42.941	P<1%	P<1%	P<1%	P<1%	P>10%
SUD	60.6	0	212	20.494	P<1%	P<1%	P<1%	P<1%	P>10%

Table 1.3: General statististics of hourly Zonal prices
(Source: author elaboration from GME)

As for the PUN price shown in table 1.4 below, high prices can be observed in hour-20 and hour-21 with 84.74 €/Mwh and 84.7 €/MWh. In addition, these periods register the highest standard deviation compared to the other models. As a result, these hours are generally called super peak hours since the prices tend to have more fluctuations compared to the others. The maximum value is recorded in hour 21 reaching 324. 2 €/Mwh. Then, it is followed by hour 19 and hour 20 when the price reached 222.25 €/MWh and 211.87 €/MWh respectively. As for the minimum value, the floor price was reached in hour 14 and hour 15. This is due to the oversupply of renewable as they reach its peak. In general, they are non-normally distributed and requires seasonality to obtain stationarity. Nevertheless, we may conclude that each period deliver has different dynamic and volatility. Therefore, it requires specific models, such as stacked model, to capture this dynamic.

	Mean	Min	Max	Std. Dev	Jarque-Bera	ADF1	ADF2	KPSS1	KPSS2
PUN1	58.95	21.16	102.63	12.31	P<1%	P<1%	P<1%	P<1%	P>10%
PUN2	52.45	14.79	96.56	12.13	P<1%	P<1%	P<1%	P<1%	P>10%
PUN3	48.28	9.89	92.04	12.38	P<1%	P<1%	P<1%	P<1%	P>10%
PUN4	45.31	5.00	87.00	12.65	P<1%	P<1%	P<1%	P<1%	P>10%
PUN5	44.93	3.28	85.65	12.56	P<1%	P<1%	P<1%	P<1%	P>10%
PUN6	48.47	5.00	86.84	12.02	P<1%	P<1%	P<1%	P<1%	P>10%
PUN7	56.96	8.34	85.53	12.83	P<1%	P<1%	P<1%	P<1%	P>10%
PUN8	65.14	9.40	154.70	15.11	P<1%	P<1%	P<1%	P<1%	P>10%
PUN9	73.36	13.03	188.77	17.22	P<1%	P<1%	P<1%	P<1%	P>10%
PUN10	75.63	13.19	207.04	18.19	P<1%	P<1%	P<1%	P<1%	P>10%
PUN11	72.95	10.77	207.08	19.27	P<1%	P<1%	P<1%	P<1%	P>10%
PUN12	70.48	7.35	206.49	19.82	P<1%	P<1%	P<1%	P<1%	P>10%
PUN13	62.40	1.19	143.94	15.98	P<1%	P<1%	P<1%	P<1%	P>10%
PUN14	59.43	0.00	129.09	16.52	P<1%	P<1%	P<1%	P<1%	P>10%
PUN15	62.13	0.00	160.45	18.00	P<1%	P<1%	P<1%	P<1%	P>10%
PUN16	65.12	2.01	163.71	17.77	P<1%	P<1%	P<1%	P<1%	P>10%
PUN17	69.62	6.71	186.58	18.08	P<1%	P<1%	P<1%	P<1%	P>10%
PUN18	76.80	11.45	196.55	22.68	P<1%	P<1%	P<1%	P<1%	P>10%
PUN19	80.67	26.11	222.25	21.12	P<1%	P<1%	P<1%	P<1%	P>10%
PUN20	84.74	43.57	211.87	20.21	P<1%	P<1%	P<1%	P<1%	P>10%
PUN21	84.70	48.04	324.20	18.45	P<1%	P<1%	P<1%	P<1%	P>10%
PUN22	77.93	47.40	156.31	15.33	P<1%	P<1%	P<1%	P<1%	P>10%
PUN23	69.65	44.19	144.41	11.78	P<1%	P<1%	P<1%	P<1%	P>10%
PUN24	63.19	35.78	101.68	10.59	P<1%	P<1%	P<1%	P<1%	P>10%
Overall	65.39	0.00	324.20	20.09	P<1%	P<1%	P<1%	P<1%	P>10%

Table 1.4: Detailed Summary of hourly PUN Price for each period of delivery (Source: author elaboration from GME)

1.2.12 Demand

Let us look into general statistics summary of the demand for each region in the table 1.5. In terms of level, Demand in the NORD is relatively much higher in comparison to the other regions reaching 18467 MWh in an hourly average. CSUD follows from behind with less than 13 GWH of consumptions. SICI and SARD have the lowest demand reaching only 2203 MWh and 1376 MWh, respectively. In addition, the standard deviations in the Nord display a very high deviation from its mean. Therefore, it can provide a big shock in the price equilibrium. SARD and SICI exhibit the lowest standard deviation but it is still in a significant level relative to their means. Overall, the ADF test has shown that demand in SARD, SICI and SUD are all accepting the null hypothesis of the first ADF test (ADF1). This also supported by the rejection of the null hypothesis in our initial KPSS test (KPSS1). Therefore, it is necessary to add seasonality trend in our empirical setup in this research.

Variable	Mean	Std. Dev	Jarque-Bera	ADF1	ADF2	KPSS1	KPSS2
D_CNOR	3524.3	879.75	0	$< 5\%$	$< 1\%$	P<1%	P>10%
D_CSUD	5311	1189.1	0	$< 5\%$	$< 1\%$	P<1%	P>10%
D_NORD	18467	4238.2	0	$< 5\%$	$< 1\%$	P<1%	P>10%
D_SARD	1376	251.24	0	$> 10\%$	$< 1\%$	P<1%	P>10%
D_SICI	2203.2	410.19	0	$> 10\%$	$< 1\%$	P<1%	P>10%
D_SUD	2917.7	554.73	0	$> 10\%$	$< 1\%$	P<1%	P>10%

Table 1.5: Statisitic summary of hourly demand in different zonal markets (Source: author elaboration from GME)

It is also interesting to look at more detailed hours in the aggregate demand since it can show the different consumptions between peak and off-peak hours. In terms of mean, period 11 and 12 display the highest mean in terms of price with 39326 and 39293. Moreover, these periods also show a high standard deviation. As a result, in the afternoon, the price equilibrium generally reach its peak at these two periods. If we look at the means in the night period, we can observe that period 19, 20 and 21 possess high level of mean and equally high level of standard deviation. Hence, this fact can explain the high volatility and high means at these hours. Then, the minimum value of demand can be observed in the off-peak periods. In particular, period 2 to period 7 where the demand is below 20 GWh. The data are non-normally distributed and non-stationary based on our statistical test. Therefore, seasonal treatment is necessary to obtain stationarity in our time series.

1.2.13 Renewable energy supply

By matching our database with the REF-E list, we have succeeded to collect a renewable energy observation that can be utilized for our research. Table 1.7 below display general statistics of the renewable energy supply in the day-ahead market. In this case, renewable energy supply are power productions from solar and the wind whereas hydro energy supply are power productions from pumping and run-of-river.[15]Overall, the series are non-normally distributed. Renewable energy supply from the NORD is the only exception compared to other zonal markets. The stationarity test is rejected in all of the fist ADF test (ADF1) but accepted in the first KPSS test (KPSS1). Then,

[15]More detailed renewable energy statistics for each technology can be found in the appendix.

	Mean	Min	Max	Std. Dev	Jarque-Bera	ADF1	ADF2	KPSS1	KPSS2
Demand1	28381	21142	37196	2632.7	P<1%	P<1%	P<1%	P<1%	P>10%
Demand2	27041	19940	35149	2660.7	P<1%	P<1%	P<1%	P<1%	P>10%
Demand3	26306	19028	34308	2676.4	P<1%	P<1%	P<1%	P<1%	P>10%
Demand4	25965	18643	33671	2661.7	P<1%	P<1%	P<1%	P<1%	P>10%
Demand5	26017	18761	33653	2683	P<1%	P<1%	P<1%	P<1%	P>10%
Demand6	26732	19210	34075	2769.1	P<1%	P<1%	P<1%	P<1%	P>10%
Demand7	29138	19174	38220	3714.5	P<1%	P<1%	P<1%	P<1%	P>10%
Demand8	33266	20334	44697	5315.2	P<1%	P<1%	P<1%	P<1%	P>10%
Demand9	37298	21240	49716	6563.9	P<1%	P<1%	P<1%	P<1%	P>10%
Demand10	39046	22510	51370	6580.3	P<1%	P<1%	P<1%	P<1%	P>10%
Demand11	39326	23612	51293	6373.1	P<1%	P<1%	P<1%	P<1%	P>10%
Demand12	39293	24157	51177	6247.2	P<1%	P<1%	P<1%	P<1%	P>10%
Demand13	37725	24040	49445	5480.9	P<1%	P<1%	P<1%	P<1%	P>10%
Demand14	36998	23094	49451	5758.3	P<1%	P<1%	P<1%	P<1%	P>10%
Demand15	37390	22453	50406	6328.4	P<1%	P<1%	P<1%	P<1%	P>10%
Demand16	37582	22183	50667	6537.9	P<1%	P<1%	P<1%	P<1%	P>10%
Demand17	38045	21732	51531	6638.9	P<1%	P<1%	P<1%	P<1%	P>10%
Demand18	38647	21973	54142	6453.4	P<1%	P<1%	P<1%	P<1%	P>10%
Demand19	38936	23115	53543	6006	P<1%	P<1%	P<1%	P<1%	P>10%
Demand20	39169	25490	51325	5236.7	P<1%	P<1%	P<1%	P<1%	P>10%
Demand21	38270	27700	47822	4121.7	P<1%	P<1%	P<1%	P<1%	P>10%
Demand22	36530	26844	45799	3627.9	P<1%	P<1%	P<1%	P<1%	P>10%
Demand23	33524	21510	43106	3204.2	P<1%	P<1%	P<1%	P<1%	P>10%
Demand24	30689	22607	39920	2943.5	P<1%	P<1%	P<1%	P<1%	P>10%
Overall	34221	18643	54142	7081.1	P<1%	P<1%	P<1%	P<1%	P>10%

Table 1.6: Detailed Summary of hourly demand for each period of delivery (Source: author elaboration from GME)

the second test of ADF and KPSS have shown that stationarity series can be reached by detrending the data.

In terms of level in renewable energy supply, NORD display superior productions compared to the other zones reaching 2239.8 MWh of hourly average. SUD comes in second with less than 1 GWh of average hourly production. As for hydro production, we can also observe that NORD exhibits a far superior level of hydro production averaging 3490 MWh of hourly production. Then, SICI and SARD show the lowest hydro productions averaging less than 100 MWh of hourly production. These facts show that NORD and SUD have the most efficient generations (Hydro and Renewable combined) compared to the other zones. As a consequence, these productions are required to be transferred to their neighbors in order to balance the system. However, this practice would saturate the transmission line, which will result in congestion.

Variable	**Mean**	**Std. Dev**	Jarque-Bera	ADF1	ADF2	KPSS1	KPSS2
R_CNOR	362.59	324.13	0	< 5%	< 1%	P<1%	P>10%
R_CSUD	571.98	447.92	0	< 5%	< 1%	P<1%	P>10%
R_NORD	2239.8	1168.9	0	> 10%	< 1%	P<1%	P>10%
R_SARD	184.39	172.71	0	< 1%	< 1%	P<1%	P>10%
R_SICI	403.03	308.82	0	< 1%	< 1%	P<1%	P>10%
R_SUD	984.91	768.21	0	< 1%	< 1%	P<1%	P>10%
H_CNOR	273.83	182.25	0	< 5%	< 1%	P<1%	P>10%
H_CSUD	364.44	217.17	0	< 5%	< 1%	P<1%	P>10%
H_NORD	3490.3	1721.8	0	< 5%	< 1%	P<1%	P>10%
H_SARD	47.943	55.496	0	< 1%	< 1%	P<1%	P>10%
H_SICI	12.633	15.372	0	< 1%	< 1%	P<1%	P>10%
H_SUD	192.66	168.87	0	<1%	< 1%	P<1%	P>10%

Table 1.7: General statistic of hourly renewable energy supply for different zonal markets (Source: author elaboration from GME)

1.3 Three angles of Italian electricity market

This thesis is constructed by three independent researches with an objective to provide a detailed overview of the Italian electricity market and to give an original contribution to the academic literature. As it was mentioned in the previous sections, the studies were motivated by the gap in the literature and interesting phenomenon in Italian electricity market. Hence, the main interests of our studies are, forecasting models, the impact of renewable energy supply on congestion, and interdependencies of zonal prices. The following subsections provide a brief summary of each study along with the insight of our findings.

1.3.1 Hourly electricity forecast

With the importance to provide a reliable model for forecasting electricity market, there have been many techniques proposed in the literature to predict electricity spot prices. Weron (2014) has provided a major state-of-the-art review on electricity price forecasting literature. He has classified the various technique into five different classes:

- multi-agent
- fundamental

- reduced form
- computational intelligence
- statistics.

The first technique, multi-agent models, is a special technique used to simulate interactions the agents in the markets by using historical data of the supply curve and market concentration. As a result, the methods are able to explore qualitative analysis of the markets and to provide insight on the market. Unfortunately, this method is not the best technique for accurate predictions since it is aimed to study interactions and strategic behavior in the market (see for instance Gao et al. (2008), Guerci et al. (2008), and Chatzidimitriou et al.(2012)). The second technique, fundamental models, is aimed to explore the economic relationship in the electricity market. Therefore, it is used to quantify the market mechanism and the impacts of the determinant. However, since they are not aimed for prediction, the forecasting performance is not as good as other techniques. Furthermore, the models generally utilized the high volume of data thus providing difficulty to obtain the same or similar data (Vahviläinen and Pyykkönen (2005), Coulon and Howison (2009) and Carmona et al. (2013)). The third techniques, reduced-form models, are financial models for forecasting electricity. The models are generally explored for risk analysis and derivative pricing. However, both Weron and Misiorek (2008), who use jump-diffusions, and Misiorek et al. (2006), who apply Markov regime switching model, have reported poor performances of the models for forecasting the next day's hourly prices. These results are in line with researchers by Dacco and Satchell (1999), Bessec and Bouabdallah (2005), and Heydari and Siddiqui (2010). The fourth techniques, computational intelligence, has a unique ability to capture linear and non-linearity relations in the electricity spot prices. Furthermore, it can model the complexity of the market and the system. There were three main methods three main methods in the literature, ANN[16](Gonzalez, et al., 2005), Fuzzy (Hong and Hsiao, 2002), and SVM[17](Zhao, et al., 2008). Although the method generally provides a good forecasting performance, the models do not provide insight and interpretations on the market. Finally, the last technique, statistics, use exogenous variables and lagged value to forecast electricity sports prices. The models use autoregressive, moving average and generalized autoregressive conditional heteroskedasticity and its combinations to reflect the electricity market. Cuaresma et al (2004), Contreras et al. (2003), Knittel and Roberts (2005) and Garcia et al.(2005) are a few example of the literature in electricity price forecasting. Their papers exhibit a good forecasting performance, interpretations, and insight on the electricity market. Therefore, the models attract the interest of many researchers since it can be interpreted in terms of technical as well economical, thus encouraging both economist and engineer to explore researches in this direction.

Our paper focuses only on developing the best model for hourly forecasting using a statistical class because of its advantage in accurate prediction and ability to explain the role of the determinants in electricity price. In addition, it was also chosen because it has the advantage of estimation without requiring high data input (such as fundamental model or multi agent) and having results that can be easily interpreted in an economic sense. We use (S)ARMA-GARCH (Seasonal or non-seasonal Autoregressive and moving average process coupled with generalized autoregressive conditional heteroskedasticity) with exogenous variables for forecasting hourly Italian electricity prices. Researches in this particular models have been well-documented in

[16]Artificial Neural Network
[17]Support Vector Machine

academics. Contreras et al (2003) proposes Seasonal ARMA model with large parameters for forecasting electricity price in three different countries, Spain, California, and Austalia. The models are proven to have reasonable errors relative to each market thus indicating an adequate model for price prediction. A year later, comparisons of different ARMA models has been explored by Cuaresma et al. (2004) in order to forecast hourly Leipzig electricity price. The paper proposes autoregressive models, autoregressive-moving average models, and unobserved component models with an aim to obtain the best model, in terms of forecasting performance, among them. On the other hand, Bowden and Payne (2008) compares three different ARIMA-GARCH models, for predicting the hourly price of the five hubs of the Midwest Independent System Operator (MISO) in the USA. They conclude that ARIMA-EGARCH-M provides the best forecasting performance compared to the other proposed models. Their paper is , then, extended by Liu and Shi (2013) who applies different GARCH variations on the same market, MISO hub. They propose 10 different GARCH models for modeling the conditional variance in MISO hub. Their research concludes that ARMA-SGARCH-M model is the most robust model. The inclusion of exogenous variables in ARMA-GACRH model has been discussed since 2002 when Nogales et al. (2002) attempt to predict Californian and Spanish electricity prices using demand, integration, and lagged variables (ARIMAX model). Their papers provide empirical proof of the superiority in ARIMAX compared to ARIMA. In the same country, United States, Knittel and Roberts (2005) study the forecast of electricity Californian price using local temperature, squared temperature and cubed temperature as the determinant in CALPX. However, the paper also reveals that the temperature is insignificant after and during the Californian crisis. Weron and Misiorek (2008) provide alternative models by comparing 12 different models in the California power exchange. Kristiansen (2012) extends Weron and Misiorek (2008) paper by adding Nordic demand and Danish wind power in their model for predicting Nordic spot price.

The main objective of this study is to provide additional alternative models in the limited literature of Italian electricity market. Based on our knowledge, there are only four literatures under statistical class that attempts to predict Italian electricity price. Bosco et al (2007) studied different periodic autoregressive models with an aim to forecast daily electricity price in Italy. They concluded that Autoregressive model with Garch residual (ARMA-GARCH) as the best model in terms of accuracy and forecasting ability. Petrella and Sapio (2011), on the other hand, use Autoregressive moving average with exogenous variable (ARMAX) to model the price formation and forecasting electricity price. The estimations concludes that weekly trend, natural gas price, load, and temperature are the determinant of electricity price. A year later, Serinaldi (2011) applies GAMLSS (Generalized Additive model for Location, Scale, and Shape) to forecast hourly electricity price in California power exchange (CalPX) and Italian power exchange (IPEX) and explain exogenous variables that drive the price. His research concludes that GAMLSS can be utilized as an alternative technique to forecast electricity price. Gianfreda and Grossi (2012) used ARFIMAX-Garch (Autoregressive Fractionally integrated moving average and general autoregressive conditional heteroskedasticity) to forecast Italian zonal prices and to explore exogenous variables (demand, technology, congestion, and market concentration) that drive the price. The paper describes the roles and impacts of the exogenous variables and exhibits an improvement in forecasting accuracy because of the exogenous variable.

Our empirical framework starts by clustering our observation (PUN Price, Demand and Gas price) according to its segmented period with the aim to capture linear relationship from exogenous and lagged variables at the same hours. [18]Hence, we calibrate twenty-four univariate models in order to predict all delivery periods in the day-ahead market. This is based on the structure of the day-ahead market where the price of all hours is settled at the same time. Similar framework has been done in the previous literature (for example, Bordignon et al. ,2013; Cuaresma et al. , 2004; Weron and Misiorek, 2005). Using this observation, we compare the various possible configuration of non-seasonal and seasonal ARMA-GARCH models with fundamental price driver as their exogenous variable. This paper attempts to look for the best model, in terms of forecast accuracy and goodness-of-fit. We also would like to examine the accuracy of price predictions from seasonal and non-seasonal ARMA-GARCH model in order to look into the impact of the additional stochastic process. In addition, we evaluate alternative models in order to justify our assumptions in our empirical settings, utilization of exogenous variables inside our model and our stacked framework. The results also allow us to experiment with different 24 SARMA-GARCH setups for error minimization in the prediction and create combination models. In addition, we can also analyze the role of gas price and demand for electricity in Italian electricity market. Finally, we also estimate multivariate models, VAR, and SUR, in order to challenge its performance with our univariate models.

This paper provides original contributions to the literature by several means.

- We propose alternative models for forecasting Italian electricity price, which is limited in the academic.
- We examine the impacts of adding seasonal stochastic process in the model on the forecasting performance.
- We construct a combination of 24 SARMA-GARCH setups for forecasting electricity price on the next day.
- We are able to quantify the effect of the fundamental driver on the hourly price thus enlarging our view in the Italian electricity market.
- We initiate discussion on the forecasting ability difference between multivariate and univariate models.

Our results have provided us with several main findings. First, seasonal time series process increases the forecasting performance on the electricity price. This finding reflects the adjustment from the market participants (both producers and consumers) based on the previous day and previous week market result. Second, Seasonal ARMA-GARCH model is, overall, the best model since it records the second best forecasting accuracy and displays good goodness-of-fit in comparison to the other models. Third, the result confirms Gianfreada and Grosi (2012) finding as models with exogenous variables show better forecasting performance compared to models without exogenous variable in both periods, in-sample and out-of-sample. Fourth, our analysis shows that a stacked model performs much better, in terms of forecast, compared to the global model. This result is in line with Cuaresma et al. (2005) conclusion in the Liepzig electricity market. Fifth, both gas price and demand have shown to have a positive and significant impact on PUN prices. This is mainly due to the fact that both variables change the supply and demand curve, which, subsequently, increase the PUN price. Finally, we have constructed a combination of 24 ARMA-GARCH model that is selected based on the forecasting performance. Our empirical test suggest that VAR, a multivariate model, is the best model in terms of accuracty wheras

[18]This is also known as stacked model or stacked framework.

univariate models has shown to be superior in terms of risk. Therefore, utilization of the framework should depends on the modeling or the forecasting objective.

1.3.2 Impact of renewable supply on congestion

This chapter is aimed to contribute to the limited literature on the impact of intermittent generation on the congestion occurrence and cost. In Norway, Førsund et al. (2008) initiate the discussion on this relation by researching the effect of wind power integration. Their studies have concluded an increasing network congestion between northern and southern Norway whenever there is a significant different in hydro resource couple with wind generations. In the United States, Woo et al (2011) continue the discussions by studying the relations in the Texan power as the wind generations are isolated in the west zone. They use ordered logit and log-OLS model. The authors concluded that rising wind supply, nuclear generation, load from non-West zones and gas price increases the likelihood to have congestion coming from the west. In Germany, Schröeder et al. (2013) and Kuntz (2013) have provided technical and economic analysis of future congestion problems as a consequence of the high integration of wind energy. Both studies conclude an increase in the overall cost for stabilizing saturated transmission caused by the renewable if there is no strategic network expansion in Germany. In Spain, Figuiredo et al. (2015) assess several determinants that drive the congestion and market splitting using logic and non-parametric Keynesian function. The paper also concludes that large availability of baseline technology coupled with high renewable has been proven to increase the market splitting and congestion.

In order to empirically verify and quantify the impact of renewable on congestion in the Italian electricity markets. We construct our observation from our database in an hourly frequency and estimate two econometric models in five zonal pairings : a multinomial logit model, whose dependent variable has three discrete values capturing both the occurrence of congestion and its direction, and a two stage least square model (2SLS) with segmented regression which seeks to quantify the effects of renewable production on implicit congestion costs.[19] Up to our knowledge, Sapio (2015) is the only author testing the impact of rising solar and wind generation on congestion between Sicily and Southern Italy using the regime-switching model on 2012-2013 hourly data. The results provide more insight on the locational-dependent-effect of renewable on the congestion. His estimation and analysis emphasize on the congestion-relieve impact from Sicily's renewable production. This is due to the fact that solar and wind production from Sicily substitute the transmission capacity and reduce the import from the South. The estimation also reveals that wind power shows a bigger impact on congestion compared to the solar power that has less variation in the power production.

This study provides original contributions to the academic in several ways.

- We enlarge the scope of the analysis by considering all Italian neighboring zones in order to verify the consistency of the empirical models.
- We employ a multinomial logit model in order to separately capture the effect of increasing renewable production on the probability of both directional congestions (to and from) compared to the benchmark situation of no congestion.
- We estimate the impact of renewable output not only on congestion frequency but also on congestion cost, something that has never been done before in the literature.

[19]The description of this congestion costs will be explained in the following sections.

- We apply segmented regressions instead of simple OLS in order to capture the impact of renewable on the congestion cost in two congestion directions, congestion to and congestion from, which are modeled as regimes (segments).

Our analysis suggests that the effect of a larger local wind and solar supply is to decrease the probability of suffering congestion in entry and to increase the probability of causing a congestion in exit compared to no congestion case. Increasing hydroelectric production has a similar effect. A rise in local demand, on the contrary, increases the probability of congestion in the entry (due to larger import) and decreases the probability of congestion in the exit. These results hold for both importing and exporting regions, but importing regions are much less likely to cause congestion in the exit, therefore the installation of new RES capacity in these zones may have a positive effect in terms of flow balance between regions. The estimations on congestion cost reveal that, due to the merit order effect, local larger renewable tend to push the congestion cost towards negative value as it decreases the marginal cost for balancing the system. A much bigger shock of renewable quantity consequently could reduce saturated line and merge the zone that is, zero congestion costs or could widen the gap of negative congestion costs because of excessive supply in the exit. This is true for all importing zone but it is the opposite for the exporting zones. Therefore, the increase of renewable should be promoted in the importing zones, but the overall growth should be controlled in order to avoid congestion toward opposite direction.

1.3.3 Interdependency of Italian electricity market

The interest in studying the relation between zonal prices have been displayed in the academic with many researchers attempts to investigate the mechanism and determinant of zonal price differences on two connected regions. Hauldrup and Nielsen (2006) have started the discussion by investigating the non-linear dynamic between two connected zonal prices in the Nordpool spot market. The paper aims to examine the regime switching and long memory process on the zonal prices using Markov Regime Switching Model. They conclude that the mechanism of the switch in congestion direction is a result of excess demand, which subsequently increase its zonal prices. Woo et al. (2011) study the zonal price difference with an objective to investigate the impact of rising wind supply in the west zone of Texas electricity market. Their result from log-OLS estimation suggests that rising load outside the West zone would increase the zonal price difference. Figuiredo et al. (2015) look into the variable that causes the market splits into two zonal prices in Iberian spot electricity price (Spain) by employing logit and non-parametric Keynesian function. Their calculation suggests that large availability of baseline technology coupled with high renewable increases the frequency of triggering a market splitting mechanism. These literature, unfortunately, does not provide us with an insight on the interdependencies between indirectly connected zones and mean spillover in the national zonal prices.

Research from Worthington et al (2005) initiate the discussion on the interdependencies on several zonal prices by employing Multivariate GARCH on the five Australian spot electricity market (NEM). In another continent, Park et al. (2006) investigate various US spot markets with Vector Autoregression and acyclic graph method. Their estimation indicates that the transmission lines and institutional arrangement affect the interdependencies in the zonal prices. Dempster et al. (2008) analyse the California electricity markets with Granger causality tests and cointegration analysis. Their study presents a moderate level of market integration and interdependencies between the regional markets in California. Unfortunately, their proposed method does not possses the capability to show the dynamic of conditional correlation in the markets. This research, on the other hand, requires a technique that could display the dynamic of conditional correlations against time since we attempt to observe the changes in conditional correlation after new investments in Italian electricity market. Worthingon et al. (2005) research is, then, extended, by Higgs (2009) for further investigation in NEM using data from 1 January 2006 to 31 December 2007. In her paper, she applies three different multivariate GARCH (MGARCH) model, Conditional Constant Correlation-MGARCH, Tse and Tsui's (2002) and Engle's (2002) Dynamic Conditional Correlation-MGARCH. Her research is aimed to investigate the inter-relationship between the four zonal prices of NEM. She concludes that regional markets with better networking infrastrcucture displays strong interdependency whereas weaker level of interdependencies are recorded in the markets with low capacity of transmission line. Unfortunately, her proposed model still lacks one important characteristic of electricity prices, seasonality. Seasonality is an important feature that presents in the electricity price. Therefore, it is important to be addressed in the future research. Ignatieva and Trück (2011) have focused on the structural dependencies in the Australian electricity market using GARCH model coupled with copulae method.[20] Their research shows that significant tail dependence between the zonal prices in Australia. Hence, price spikes may happen jointly across the regional market.

We attempt to extend Higgs (2009) by capturing interdependence of zonal prices in different electricity markets, Italian power exchange. To the best of our knowledge, Sapio (2015) and Ardian et al (2015) are the only two literature related to our research for the case of Italian electricity market. Their paper, however, focuses solely on capturing the impact of renewable price to zonal price difference (congestion cost) and congestion between two connecting zones. Both papers reveal the same insight in the market splitting mechanism. A positive shock on renewable energy supply on the exporting zones display increasing impact on the zonal price differences in comparison to zonal price difference under no shocks. The same effect on the zonal price differences also displayed whenever there is an increase in demand in the importing zones. Their results explain the relations between connected regions. Our paper, on the other hand, aims at studying the inter-relationship and mean spillover among Italian zonal prices. Hence, we want to shed a light into the correlation between indirectly connected zones. The result will provide us with greater insight into the pricing efficiency and the impact of the transmission line on the interdependencies in the Italian electricity market.

In order to achieve our objective, we collect six series of the Italian zonal prices from its independent market operator (GME) and estimate multivariate GARCH approach under conditional mean. The estimation starts by computing the coefficients of the univariate Seasonal ARMA and GARCH, which produces conditional mean and variance respectively. Our model tries to address Higgs (2009) suggestion by adding seasonality, both trend and stochastic. This is due to the fact that seasonality effects from the weekly pattern, hourly pattern, and trader correction are important characteristics that determine electricity price. In addition, the result is also expected to provide us with insight on the mean spillovers. The second estimation is, then, calculated with the initial parameters from the univariate Seasonal ARMA-GARCH. In this stage, we employ two multivariate GARCH model to capture the cross-correlation of the zonal prices, CCC-MGARCH from Bollerslev (1990) and DCC-MGARCH from

[20] The proposed method, however, is not applied in this research since risk management are not the main objective of this research. We aim to investigate the national market integration in the Italian electricity market by evaluating interdependencies and dynamic correlation of the zonal prices. We are able to analyse this effect by employing CCC and DCC MGARCH without coupling it with copulae method.

Engle (2002). The two models allow us to examine the cross-correlation between regional markets and to investigates the efficiency of the market integration. It is noteworthy that the method is also able to analyze the impact of the New submarine installation between Sardinia and Italian peninsula in 2011 on the conditional correlation. This method will initiate discussion on the impact of new transmission installation on the conditional correlation in the academic since this is the only case study that has an increase in transmission capacity in the sample period.

This research contributes to the academic literature in several ways .

- We initiate the discussion on interdependency and mean spillover in the Italian electricity market.
- We would like to contribute to the limited literature on the cross-correlation of electricity zonal prices
- We try to address seasonality, which is part of the concern in Higgs (2009), in this paper by adding seasonal trend and process on the model
- We try to analyse the impact of new transmission installation on the dynamic conditional correlation, which has not been done in any related literature.

Our estimations deliver an additional insight on the integration of the Italian electricity market and the interdependencies among its regional market.The results indicate that high capacity physical exchange provides strong interdependency among the connected zonal markets. Indeed, all the zones in Italian peninsula have shown strong dependencies among them as it will be shown in our CCC and DCC estimations. The result is in line with Higgs (2009) where strong interdependency is shown in the well-connected markets. Moreover, this argument is also validated by our finding in the analysis of new transmission installation where stronger dependencies are found after additional capacity is put into place despite a long transition period. In addition, the low transmission line between Sicily and Italian peninsula is the main cause of weak interdependencies between Sicily and the other regional markets. *

Chapter 2

Hourly electricity price forecasting : A case study of Italy

Faddy Ardian
École Polytechnique
Anna Creti
Université Paris Dauphine and École Polytechnique
Francesco Vallone
Cogenpower SpA, Torino, Italy

Contents

Abstract

Our paper attempts to employ (S)ARMA-GARCH with exogenous variables for forecasting hourly electricity price on the Italian day-ahead market. Italy provides an interesting test ground for our research because of its high price level and volatilty compared to the other major european markets. We have constructed dataset consist of electricity price, demand, and gas price between 2010 and 2014. Our results have provided us with several main finding: (1) seasonal time series process increases the forecasting performance on the electricity price, (2) Seasonal ARMA-GARCH mode is, overall, the best model since it records the second best forecasting accuracy and displays good goodness-of-fit in comparison to the other models, (3) models with exogenous variables show better forecasting performance compared to models without exogenous variable in both periods, in-sample and out-of-sample, (4) our analysis shows that a stacked model performs much better compared to the global model (5) both gas price and demand have shown to have a positive and significant impact on the Italian wholesale electricity prices, (6) and, from our research, it can be implied that univariate models have better forecasting performance in comparison to panel models.

2.1 Introduction

Deregulation of electricity market has displayed many changes in the economy and has influenced researchers to initiate studies in this field. Before the liberalization of the market, the price of electricity in most countries is fixed by the government according to the future fuel price and tax. However, there was always a margin whenever the global market price of fuel changes. In addition, vertical integration dictates that one company control all the supply chain of electricity, from production to distribution. As a consequence, this commodity was not very competitive and had an inefficient economy. Today, the electricity market price is determined by the actual supply and demand which reflect the scarcity and the surplus of electricity on each period of delivery. This liberalization also reflects the actual fuel price of producing each MWh of electricity. Moreover, due to the unbundling of the state-owned company that controls all the sectors, the new mechanism opens new investments, increase efficiency, and improve competitions. As a result, new producers and distributors are allowed to enter the market and to compete for the rights in delivering/consuming electricity on the spot market (day-ahead market).

The issue of the deregulated market arises as the volatility of the wholesale price increases because of the new market based mechanism in price determination. Deviation in actual demands, fuel price, and supply curve create shocks in the market equilibrium. Price can go down to zero €/Mwh or even negative in some cases (in Germany for example) because of the surplus in supply. However, it can also increase to more than 100 €/Mwh in an instant. This volatility phenomenon is mainly due to the fact that electricity cannot be stored, which makes them reluctant to price jump. With more volatility, market participants are impacted greatly as they face the financial risk in buying/selling electricity. On the other hand, policy makers are also obligated to constantly analyze the electricity spot price with the aim to constantly examine the fairness in the trading. As a consequence, both market participant and policy makers need to have reliable forecast of electricity price in order to obtain insight on the future prices.

Various models and approaches have been proposed to forecast electricity prices in the day-ahead market. Recent major state-of-the-art review by Weron (2014) has classified the forecasting techniques into five main classes: multi-agent, fundamental, reduced form, statistical and computational intelligence. Multi-agent models are engineering techniques that simulate the interaction of participants in the market using rich data of the bids in the electricity market. The ability to simulate market interaction means that it can predict electricity price only from supply curve and market concentration without any prior knowledge of the historical price. They are flexible tools to analyze strategic behavior and to explore insights on the market. Examples of this technique in electricity price modeling can be seen in Gao et al. (2008), Guerci et al. (2008), and Chatzidimitriou et al.(2012). However, the models are not suitable for a precise electricity prediction since they are focused on exploring qualitative analysis. The next class is fundamental or structural model that is generally aimed at revealing the economic relationship in electricity market rather than forecasting. The models are used to reflect and present the mechanism on the market. Nevertheless, they also can be used for predicting electricity prices. Vahviläinen and Pyykkönen (2005) build a parameter-rich parameter models using weather data that explains the price formation of Nordic market. Coulon and Howison (2009) and Carmona et al. (2013) are few examples of previous literature that construct stochastic processes of the main factors (demand, weather, etc) for predicting electricity. The main problem with this model is the data availability. The data availability of fundamental prices generally has a low frequency (daily, weekly or monthly). Therefore, there is an additional challenge for constructing higher frequency price models. In addition, stochastic process of the fundamental drivers generally follow a specific assumption thus making the simulated prices is sensitive to this limitation. Reduced form is a class model inspired from finance. The models are utilized for simulating electricity price based on its main features and characteristics such as price dynamics, price correlations, and marginal distribution. As a result, the model is commonly used for derivative pricing and risk analysis. The main techniques in this class are jump-diffusions and Markov regime-switching that simulates price spikes and different price regimes, respectively, in the electricity market. Despite its strength in risk management, both Weron and Misiorek (2008), who use jump-diffusions, and Misiorek et al. (2006), who apply Markov regime switching model, have reported poor performances of the models for forecasting the next day's hourly prices. These results are in line with researchers by Dacco and Satchell (1999), Bessec and Bouabdallah (2005), and Heydari and Siddiqui (2010). The statistical class uses models of lagged variables and exogenous factors for predicting future electricity price. The models attract the interest of many researchers since it can be interpreted in terms of technical as well economical, thus encouraging both economist and engineer to explore researches in this direction. The models circled around autoregressive, moving average and generalized autoregressive conditional heteroskedasticity with their combinations and evolution. A few example of the electricity price forecasting using this method can be found in Cuaresma et al (2004), Contreras et al. (2003), Knittel and Roberts (2005) and Garcia et al.(2005). Finally, computational intelligence-class uses artificial intelligence to simulate the complex system of an electriciy market. This feature provides this class with unique capability to capture linear and non-linearity relations in the electricity prices. Academic works in electricity price forecasting have been done with three main methods, ANN[1](Gonzalez, et al., 2005), Fuzzy (Hong and Hsiao, 2002), and SVM[2](Zhao, et al., 2008).

Based on Weron (2014) main classes, our paper focuses on developing model under statistical-class. In particular, we attempt to employ (S)ARMA-GARCH[3] with exogenous variables for forecasting electricity price on the day-ahead market. There have been many researches related to this method in electricity price forecast literature. In ARMA-type literature, Cuaresma et al. (2004) compare different linear forecasting techniques for predicting the Leipzig electricity price. They compare autoregressive models, autoregressive-moving average models, and unobserved component models with an aim to obtain the best model. Contreras et al (2003). proposes large parameter Seasonal ARMA model for forecasting electricity price in three different countries, Spain, California, and Austalia. The models are proven to have reasonable errors relative to each market thus indicating an adequate model for price prediction. In ARMA-GARCH-type literature, Bowden and Payne (2008) estimates three different models, ARIMA, ARIMA-EGARCH, and ARIMA-EGARCH-M for predicting the hourly price of the five hubs of the Midwest Independent System Operator (MISO) in the USA. The paper concludes that ARIMA-EGARCH-M model outperforms the other models based on their sample period. The research is then extended by Liu and Shi (2013) who work on different Garch variation on their ARMA-Garch approach for MISO hub. This study compares ten different GARCH models to model residual with the same equation of ARMA. Five different models are calibrated, then evaluated based on the accuracy and goodness of fit. The results show that ARMA-SGARCH-M model is the most robust model. As for the discussion on adding the exogenous variable to ARMA model (widely known as ARMAX), Nogales et al. (2002) initiate the literature by attempting to predict Californian and Spanish electricity prices using ARIMAX model with the load as the explanatory variable. The result shows superiority in forecasting performance in comparison to ARIMA. Knittel and Roberts (2005) use temperature, squared temperature and cubed temperature as the exogenous variables of their model for predicting Californian electricity prices. The temperature is reported to be significant before Californian electricity crisis in 2000, but it has been shown to be non-significant during the crisis. Weron and Misiorek (2008) compare the forecasting performance of 12 different models using Californian price data with its load as the exogenous variable and Nordic prices data with its air temperature as its explanatory variable. Kristiansen (2012) extends their paper by adding Nordic demand and Danish wind power in their model for predicting Nordic spot price.

Italy, as one of the mature electricity market in Europe, provides an interesting test ground for our research. In terms of price level, Italy has the highest average wholesale price compared to other EU market between 2010 and 2014 (see figure 2.1). In 2012, Italy's electricity price reaches its peak and records 44.28 €/Mwh on the price difference between

[1]Artificial Neural Network

[2]Support Vector Machine

[3]Seasonal or non-seasonal Autoregressive and moving average process coupled with generalized autoregressive conditional heteroskedasticity

Italy and Nordpool spot. The price gap is still relatively big even if it is compared to second highest price (PowerNext), reaching 22.33 €/Mwh on the price difference between the two markets. On the other hand, the introduction of the renewable energy increases the volatility of the wholesale price (see for instance Clo et al., 2015). Consequently, the price can easily increase/decrease during the day thus making Italy an interesting common ground to test forecast models. Moreover, there is uncertainties in the Italian electricity price that comes from the changes in demand and gas prices. Finally, models to forecast Italian electricity market are quite limited in the literature. Hence, our paper is aimed to provide additional alternatives to statistical models and to expand our discussion in electricity price forecasting for this particular market.

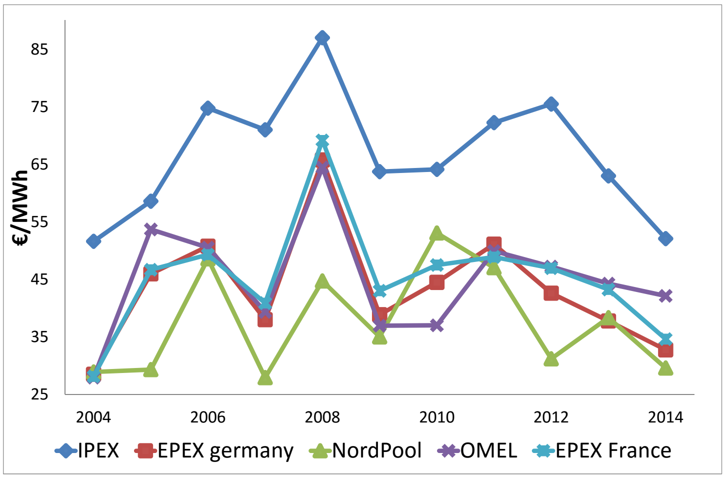

Figure 2.1: Historical price data of major electricity markets in europe

Our paper focuses only on developing the best model for hourly forecasting using a statistical class because of its advantage in accurate prediction and ability to explain the role of the determinants in electricity price. In addition, it was also chosen because it has the advantage of estimation without requiring high data input (such as fundamental model or multi agent) and having results that can be easily interpreted in an economic sense. To the best of our knowledge, four alternative techniques under statistical class have been used in the Italian electricity market.[4] Bosco et al (2007) propose different propositions of Periodic Autoregressive (AR) model to capture the mean reversion and volatility process in daily electricity price. The study compares five periodic AR model with different order and residual models (e.g using Garch or assumed to be homoscedastic). The results show that the periodic AR-Garch has the best forecasting performance compared to alternative techniques. Petrella and Sapio (2010) estimate two different baseline models for daily electricity prices, ARMAX and ARMAX-EGARCH. The paper concludes that temperature, weekly trend, natural gas price, and conditional volatility are significant regressors for price formation. The paper is, then, extended by analyzing structural changes in Italian electricity market between 2004 and 2008. Serinaldi (2011) utilized GAMLSS (Generalized Additive Models for Location,

[4]It is noteworthy that there are other techniques that have been tested in the Italian electricity market. However, they are classified under different classes (for instance Bompard et al. (2008)); Guerci et al.(2008).

Scale, and Shape) in order to represent the price as a realization of the explanatory variables (e.g historical electricity price, load, and temperatures). The model is successfully used to forecast one day of day-ahead prices in hourly frequency in two power exchanges, CalPX (California) and IPEX(Italy), and explain the exogenous variables that drive the price. After comparison of several different model propositions (such as AR and AR-GARCH), the study concludes that GAMLSS can be used as an alternative model to forecast electricity price. Gianfreda and Grossi (2012) used ARFIMAX-Garch (Autoregressive Fractionally integrated moving average and general autoregressive conditional heteroskedasticity) to capture the long memory process and volatility in the Italian zonal prices. The model forecast the electricity price accounting for technologies, market concentration, congestions and traded quantities. The paper is aimed to examine the accuracy improvement from adding these exogenous variables. By comparing the model with and without exogenous variables, the paper confirms the increase of accuracy after adding the exogenous variables.

We have constructed dataset consist of electricity price, demand, and gas price collected from GME and ICE between 2010 and 2014. The price data is then clustered according to its segmented period in order to capture linearity of the exogenous variable and the dynamic of hourly prices. [5]In other words, we estimate regressions of twenty-four univariate model that is able predict the price of all hour in the next day. This is due to the structure of the day-ahead market in which prices of all hours of delivery are determined at the same time. Therefore, the information is updated on daily basis. This framework is supported and suggested by previous literatures (for example, Bordignon et al. ,2013; Cuaresma et al. , 2004; Weron and Misiorek, 2005). Using this dataset, we compare various possible configuration of non-seasonal and seasonal ARMA-GARCH models with fundamental price driver as their exogenous variable. We aim to seek the best performing model, in terms of forecast accuracy and goodness-of-fit between them, and to compare forecast performance of seasonal and non-seasonal ARMA-GARCH model. In addition, we attempt to evaluate alternative models in order to justify exogenous variables inside our model and our stacked framework. The results also allow us to experiment with different 24 SARMA-GARCH setups for error minimization in the prediction. The estimated coefficients from our best performing model can be used to analyze the role of gas price and demand for electricity. Finally, we also estimate multivariate models, VAR and SUR, in order to challenge its performance with our univariate models.

Our paper provides contributions to the academic and business world in many means. Firstly, we add alternative models for hourly Italian electricity price forecast, which is limited in the academic literature. Secondly, theses alternative models can be used by market participants or traders for gaining an edge in their sales and trading. Thirdly, we attempt to examine the impact of the seasonal stochastic process on the forecasting performance and goodness-of-fit, which has not yet exploited in the previous literature. Fourthly, we construct a combination of 24 SARMA-GARCH setups for forecasting electricity price on the next day. Up to our knowledge, this is the first paper that starts the discussion on the univariate SARMA-GARCH combination for electricity price forecasting. Previous literatures generally focus on combining various technique (see for instance Bordignon et al., 2013). Fifthly, integrating exogenous variable in our model enables us to quantify the effect of the fundamental driver on the hourly price thus enlarging our view in the Italian electricity market. Finally, we analyze the forecasting ability of multivariate model and evaluate the difference, in terms of prediction capability, between multivariate framework and univariate framework.

In the next section, we are going to present a brief description of the Italian electricity market including its structure and production mix. The section is, then, followed by a statistical report on our sample dataset. Section 4 describe our empirical method as well as the analysis on its results. Finally, the fifth section will conclude our studies with recommendations for future research extension.

[5]This is also known as stacked model or stacked framework.

2.2 Empirical framework and analysis

2.2.1 Univariate framework

In this paper, we are going to explore and examine six univariate models to forecast hourly price of the day-ahead market. Following Lucia and Schwartz (2000), a general formula of electricity price $P_h(t)$, for $h = 1, 2, ...24$,on the next day (t), consists of deterministic components $G_h(t)$ and stochastic components $F_h(t)$.

$$P_h(t) = G_h(t) + F_h(t) \tag{2.1}$$

The deterministic component $G_h(t)$ are formed by:

- Seasonal dummies ($S_h(t)$)
  As the previous subsection suggest, seasonality is part of the concern in the estimation. This is due to the fact that the assumption of our stochastic model requires a stationary series. Hence, de-trending the series from seasonal effect is obligatory to gain stationarity. In addition, it is also useful to avoid spurious regression for our exogenous variables in the model. Based on our finding from ADF test and KPSS test, we model the seasonal pattern in the series with yearly, monthly, and daily dummies.
- Price driver ($X_h(t)$)
  The price drivers are exogenous variables that form the electricity price. In this paper, two main drivers of electricity paper are analyzed, demand and fuel cost.[6] Therefore, they are formulated as follow,

  $$X_h(t) = \alpha_h^1(t) * gas_h(t) + \alpha_h^2(t) * demand_h(t)$$

  Where $gas_h(t)$, and $demand_h(t)$ are gas price and total electricity demand at our h respectively.

Then, we are going to estimate six different univariate model in the stochastic component $F_h(t)$ under Seasonal (non-seasonal) ARMA process and GARCH residuals. The general model of this process is shown in the equation below.

$$\Phi_P(B^s)\phi_p(B)F_h(t) = \Theta_Q(B^s)\theta_q(B)\omega_h(t) \tag{2.2}$$

where,

$$F_h(t) = P_h(t) - G_h(t)$$

Then, B is a backshift operator, and $\omega_h(t)$ is the residual value at day t and hour h. The term $p\ ,q\ ,P\ ,Q\ ,\ and\ S$ are integers. The term S in the model represent the seasonal length. The term p and P are orders of autoregressive process in the model for non-seasonal and seasonal, respectively, with the following form

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p \tag{2.3}$$
$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - ... - \Phi_P B^{Ps} \tag{2.4}$$

Then, the terms q and Q are the order of moving average process in the model for non-seasonal and seasonal, respectively, that is calculated with a general equaiton below.

$$\theta_q(B^p) = 1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q \tag{2.5}$$
$$\Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - ... - \Theta_Q B^{Qs} \tag{2.6}$$

From the ACF and PACF in the figure 2.4 we suggest several models as follows,

[6] It is important to be noted that we have attempted to integrate weather as the exogenous variable. However, dataset of weather under the same frequency is not available for our research. In addition, we find difficulties for aggregating the weather to become unbiased in our research as the data are not collected in the same frequency and in the same period. Then, the impact of weather data to the price would be a subjet to the location as each zone has different renewable mix. Hence, using data in current state would provide additional questions instead of answers.

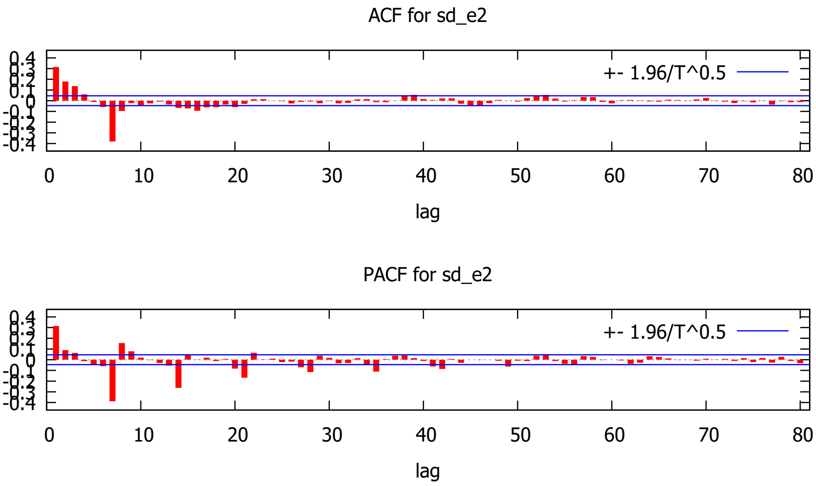

Figure 2.2: ACF and PACF of PUN at hour 24

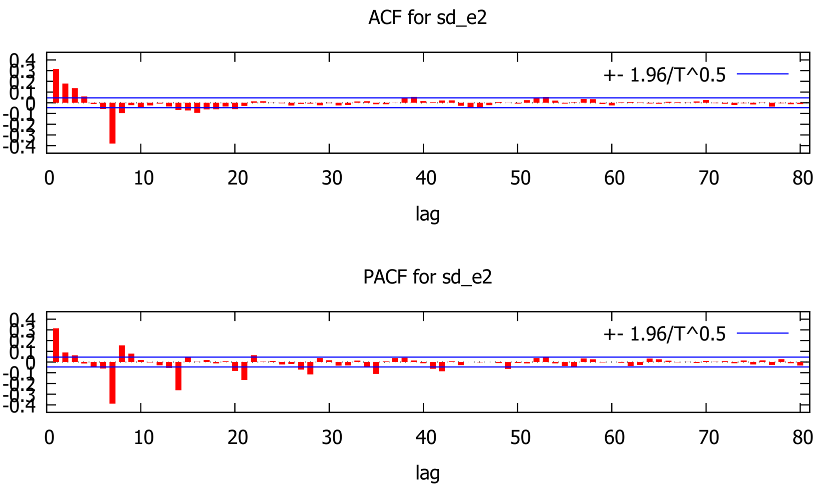

Figure 2.3: ACF and PACF of PUN at hour 24

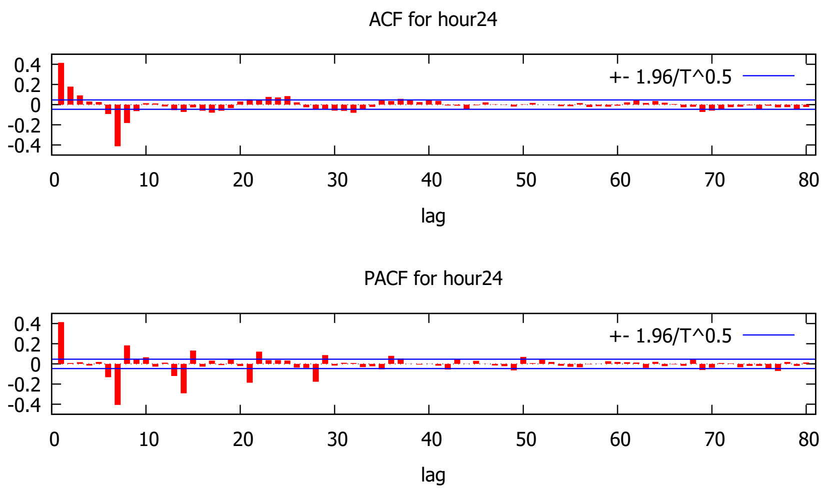

Figure 2.4: ACF and PACF of PUN at hour 24

- Model 1 (M1)
  $ARMA(1,1) - GARCH(1,1)$

$$F_h(t) = \phi F_h(t-1) + \theta\omega_h(t-1) + \omega_h(t)$$

- Model 2 (M2)
  $ARMA(1,1)x(0,1)_7 - GARCH(1,1)$
  In this model, we introduce a seasonal stochastic models in our ARMA model as shown in the notation with $(1,1)x(0,1)_7$. In this case, we have a seasonal moving average of the first order with 7 as the lagged seasonality, $(0,1)_7$, in addition to the ARMA model $(1,1)$. Therefore, the regression equation of this ARMA setup can be written as follow,

$$F_h(t) = \phi F_h(t-1) + \theta\omega_h(t-1) + \Theta\omega_h(t-7) + \theta\Theta\omega_h(t-8) + \omega_h(t)$$

- Model 3 (M3)
  $ARMA(1,1)x(1,1)_7 - GARCH(1,1)$ In m3, we introduce two seasonal stochastic models in our ARMA model, $ARMA(1,1)x(1,1)_7$. Hence, we have a seasonal autoregreession and a seasonal moving average of the first order with 7 as the lagged seasonality $(1,1)_7$ in addition to the ARMA model $(1,1)$.. We can write the regression equation of this model as follow,

$$\begin{aligned} F_h(t) = \phi F_h(t-1) + \Phi F_h(t-7) + \phi\Phi F_h(t-8) + \theta\omega_h(t-1) \\ +\Theta\omega_h(t-7) + \theta\Theta\omega_h(t-8) + \omega_h(t) \end{aligned}$$

In addition to the suggested models, we also explore other alternative models in order to validate the introduction of an exogenous variable, the seasonal process, and our stacked time series approach. Moreover, we aims at comparing our model with the similar model from Gianfreda and Grossi (2012). The alternatives models are:

- Model 4 (M4)
  In order to confirm the improvement in forecasting performance from adding exogenous variable, we are going to change the model as follow.

$$P_h(t) = S_h(t) + F_h(t)$$

$$F_h(t) = \phi F_h(t-1) + \Phi F_h(t-7) + \phi\Phi F_h(t-8) + \theta\omega_h(t-1)$$
$$+\Theta\omega_h(t-7) + \theta\Theta\omega_h(t-8) + \omega_h(t)$$

- Model 5 (M5)
  In this model, we are going to test the simplest model that is generally used in forecasting commodity, AR(1).
  $$P_h(t) = S_h(t) + F_h(t)$$
  $$F_h(t) = \phi F_h(t-1) + \omega_h(t)$$

- Model 6 (M6)
  We follow Gianfreda and Grossi (2012) calibration, $ARMA(1,7) - GARCH(1,1)$, which is proposed for modeling daily electricity price in Italy. It is noteworthy that we have adapted the model for our research by following their ARMA calibration for fitting the series. Hence, demand and fuel cost are still used as exogenous variables instead of congestion, technology, volume and market power. In addition, fractional integration is not applied in this model since we did not find any empirical evidence for long memory in our series.

  $$P_h(t) = G_h(t) + F_h(t)$$

  $$F_h(t) = \phi F_h(t-1) + \theta_1\omega_h(t-1) + +\theta_2\omega_h(t-2)$$
  $$+\theta_3\omega_h(t-3) + \theta_4\omega_h(t-4) + \theta_5\omega_h(t-5)$$
  $$+\theta_6\omega_h(t-6) + \theta_7\omega_h(t-7) + \omega_h(t)$$

- Global model
  A global model is a non-stacked time series model for forecasting hourly electricity price. In other words, the electricity price series are treated as an hourly frequency. This is the exact opposite of our model. This model has advantages in terms of simplicity and rapidity of the estimation since there is less parameter involved. However, careless empirical set up can violate the fact that day-ahead market is not a continuous market where the price is updated in every period (t). In order to avoid modeling the price based on unknown observation within the day, restricted Seasonal ARMA -GARCH(1,1) model is proposed. Hence, the price at hour t can be calculated below.

  $$P(t) = G(t) + F(t)$$

  $$F_h(t) = \phi F_h(t-24) + \Phi F_h(t-168) + \phi\Phi F_h(t-169)$$
  $$+\theta\omega_h(t-24) + \Theta\omega_h(t-168)$$
  $$+\theta\Theta\omega_h(t-169) + \omega_h(t)$$

The error term $\omega_h(t)$ is assumed to have zero mean and normal distribution in this process. However, high-frequency series generally exhibits heteroscedasticity. Therefore, the error term is estimated using generalized autoregressive conditional heteroskedasticity (GARCH), a well-known method proposed by Bollersev (1986) that allows us to model future variances. This additional method has been used to forecast electricity price forecast by many reaserchers (Knittel and Roberts, 2005; Diongue et al., 2009; Bowden and Payne, 2008).

$$\omega_h(t) = \sigma_h(t)W(t) \tag{2.7}$$
$$F_h(t) = \hat{F_h}(t) + \sigma_h(t)W(t)$$
$$\sigma_h^2(t) = \gamma_0^h + \beta_h\sigma_{(t-1)}^2 + \gamma_1^h\omega_t^2$$

2.2.2 Evaluation Method

Our data is separated into two sections, in-sample and out-of-sample. The in-sample period spans from 2010 to 2013 wheras 2014 is used as the out-ouf-sample period. In the in-sample, goodness-of-fit of each model is estimated and measured through Schwarz Criterion (Schwartz, 1978), also known as BIC (Bayesian Information Criterion), according to the equation below

$$BIC = log(\sigma^2) + \frac{klogn}{n}, \tag{2.8}$$

where n is a sample size and k is the number of parameters. Then, in both sections, in-sample and out-of-sample, several parameters are calculated in order to evaluate the forecasting performances.[7] The parameters are :

- Mean Absolute Error (MAE)
  Classical error parameter based on the absolute average of absolute values.

  $$MAE = \frac{1}{n} \sum_{i=t+1}^{i=t+n} |\hat{P}_i - P_i|$$

  with n is in-sample or out-of-sample size.
- Expected shortfall (ES)
  This measure is used to evaluate the performance based on the risk analysis. In this case, it is defined as the average error of predicted value exceeding a specified quantile of the forecasting error distribution. This research uses 97.5 % as the quantile reference.
- Maximum value of residual (Max)
  The maximum value of the residual is also used in the analysis in order to measure the model's ability to capture the spikes in the price series.

2.2.3 Univariate Model evaluation

The summary of goodness-of-fit parameters and forecast accuracy of all models can be seen in table 2.1 below. Our first finding indicates that additional seasonal process provides a better model in terms of forecasting accuracy. By comparing the performance of non-seasonal ARMA-GARCH model (M1) to seasonal ARMA-GARCH model (M2 and M3), we can observe that seasonal time series processes display better forecasting performance compared to non-seasonal. If we look at the goodness-of-fit, one can conclude that M1 is a better model since it has the lowest BIC (8411.38) compared to seasonal models (9026.0761 and 9009.630). This is due to the fact that BIC depends on the number of the parameters in the model. Hence, M2 and M3 are penalized because of the additional parameters from the seasonal processes. In terms of the forecasting performance, however, M1 has a high mean absolute error for the in-sample reaching 6.041 euro/MWh. This is higher compared to the best performer, M3, that register MAE of 5.904 €/MWh. In the same period, the different in the accuracy level can also be seen in the expected shortfall and maximum value where M1 registers lower forecasting performance, 22.67 €and 201.87€/MWh /MWh respectively. On the other hand, M2 records higher accuracy, 22.508€/MWh and 201.567 €/MWh in the same parameter. For the case of out-of-sample, non-seasonal ARMA process (M1) also displays lower accuracy as shown in its mean absolute error, 4.366 compared to 4.318 recorded by M3. However, the expected shortfall and maximum value display better performance compared to M2 and M3. Nevertheless, overall, our conclusion still stands as M2 and M3 exhibit better performance compared to M1.

It is interesting to look at the forecasting accuracy of M6, which can be considered as a non-seasonal ARMA-GARCH process. The results indicates M6 as a better model in comparison to seasonal time series process which is a contradictory on the previous finding.

[7] It is noteworthy that we use one day rolling forecast in the out-of-sample periods

Univariate	M1	M2	M3	M4	M5	M6	Global
BIC	8411.308	9026.07613	9009.6305	9114.05904	9115.19	9024.63121	9548.34
In-Sample							
MAE	6.04165541	5.9610422	5.90434926	6.13872633	6.36503663	5.86596805	6.218552
ES	22.6781573	22.5082363	22.2619261	23.8859568	24.5942959	22.1071705	22.4192292
Max	201.875935	201.56793	202.924872	203.340559	217.453152	202.618647	211.93
Out-of-sample							
MAE	4.36694571	4.39541931	4.31847927	4.83265895	4.96075787	4.26382613	6.2279
ES	14.6743948	15.280684	15.1421976	17.9743974	19.2025372	14.8914718	26.7054683
Maximum	55.1969623	57.1704275	56.5069455	58.0499622	59.1797427	54.3834864	92.92
DM	6.10(***)	3.11 (***)	-	15.17(***)	20.56(***)	4.62(***)	4.143(***)

* p-value < 10 % (DM test > 1.65)
** p-value < 5 % (DM test > 1.96)
*** p-value <1 % (DM test > 2.58)

Table 2.1: Result summary

Indeed, our estimations suggest that M6 has out-performed M2 and M3 on both in-sample and out-of-sample data with an exception in the in-sample data where M2 display the lowest maximum error value (201.56 €/MWh). The difference in the forecasting accuracy, however, is not very high in comparison to M3 (the best performer from the seasonal time series process). For example, in terms of MAE, it has 0.039 €/MWh and 0.055 €/MWh error difference in the in-sample and out-of-sample respectively. This slight difference is due to the fact that M6 uses moving average process up to seven order which considers linear relation with lagged variable at $t-7$. However, it creates a drawback in the model as it requires more parameter, which overfit the data. As a consequence, M3 has better goodness-of-fit with lower BIC (9009.6305) in comparison to M6 (9024.631). This is because more parameter will become an overfit to the data. Seasonal ARMA-GARCH, M3, captures the weekly seasonal adjustment but neglects the effect of price change within a week that is captured in M6. However, the electricity price deviation is less likely to be caused by the price shock within the week ($t-2$ to $t-6$). The price changes are highly depends on the seasonality. Indeed, market participants often adjust their bids in the market based on the previous data and previous week. They rarely make an adjustment based on price shock within the week. Therefore, we can conclude that seasonal ARMA-GARCH (M2 and M3) is the most appropriate models since they avoid overfitting the data, reflects better the characteristic of the market and has consistently good forecasting accuracy. M3, however, is the top performer in terms forecasting accuracy between the two models. Hence, we can conlude that M3 is the most robust model for forecasting hourly electricity price.

Our results also validate Gianfreda and Grossi's (2012) finding as models with exogenous variable (M1, M2,M3, and M6) clearly display superior performance compared to the models without exogenous variable (M4 and M5). In terms of BIC, models without exogenous variable register higher BIC with values bigger than 9114. This is far from our best model, M3, that reaches 9009.63. In terms of MAE, M4 and M5 have lower accuracy compared to the model with the explanatory variables in both period, in-sample period 6.138 and 6.365) and out-of-sample (4.832 and 4.960). The same conclusion is also shown on ES and Max as the forecast experiment show the highest ES and MAX compared to the other model.

Table 2.1 also reveals the forecast comparison between a stacked model (24 univariate models) and a global model (1 univariate model) for forecasting hourly electricity prices. The global model shows the worst performance in almost all parameters. Its goodness-of-fit displays the highest value compared to other models accounting to 9548.24. This finding suggests that the model is not the most appropriate method for predicting hourly electricity price. In terms of forecasting accuracy, the MAE of both in-sample and out-of-sample have passed 6 €/MWh, which is low in comparison to the other models. Although the expected shortfall in in-sample period is not the worst, the out-of-sample forecast displays a very poor performance. The same finding is also shown in the Maximum error value. The results reach the same conclusion as Cuaresma et al. (2004) where stacked model out perform global model in Leipzig Power Exchange. The main reason is because the stacked models are able to capture the price dynamics in each period. One can also argue that this is due to the structure of the day-ahead market in which prices of all hours of delivery are determined at the same time. Although the global model is clearly a much simpler model in terms of estimation and calibration, the coefficients are only able to capture limited information since we have the same coefficient for all periods. As a consequence, shocks in a specific delivery period/hour are able to cause bias in the estimation and forecast in other periods.

In order to have a robust result, we employ Diebold-Mariano test for evaluating the equality of the predictive accuracy (Diebold and Mariano (1995)). The results determine whether two models provide the same forecasting accuracy if the loss differential has zero expectation. In this research, we use Model 3 as a benchmark to compare with the other models since it is shown as the most robust model. If the comparative model displays high absolute DM test, we may reject the hypothesis that both models have an equal level of accuracy. In this case, the different mean absolute error would determine the best model since they pose a different level of accuracy. On the other hand, if we accept the null hypothesis, the result indicates that both models possess the same forecasting performance and there would be no difference in using both models. As shown in the Table 2.1, our results, and analysis still hold as the DM test shows high absolute value in all comparative model. The p-value shows that M3 has a different level of accuracy in comparison to the other models. M2 has the closest DM statistic value to the critical values (2.58) since M2 is similar to M3 as they both use seasonal stochastic process and GARCH. However, the statistic result still falls outside the range in order to reject the null hypothesis of no difference. On the other hand, if we look at M5, DM statistic display a very high value. This high statistic value is probably due to the utilization of exogenous variable and seasonal process that are not employed in M5.

2.2.4 Main fundamental drive

Our exogenous variables from our model reveal the role of gas price and demand in the formulation of electricity prices. We are going to use M3 estimations for the analysis of coefficient since it records the best forecasting performance coupled with relatively low BIC. The estimation result is shown in table 2.2 below. From the coefficient of demand, we can conclude that demand has a positive and significant role in PUN price formation. This is shown in all period of delivery with hour-21 and hour-22 as exceptions. Hence, positive shock in demand will increase PUN price. The result reflects the price mechanism built by the intersection of supply and demand curve. The increase of demand will shift the demand curve thus forming a higher PUN price. The level of impact on the PUN price, however, is different for each period. The table shows that hour 9 is the period that is highly sensitive to demand with 0.0016. Therefore, a positive shock in demand for 1 MWh would increase a PUN price for 0.0016 €/Mwh. Hour 16, 12 and 8 are behind with lower impact, 0.000611952, 0.000581841, and 0.000573967 respectively. Finally, demand in Hour 20 displays the lowest significant impact compared to the other period.

On the other hand, the natural gas price is also proven to be essential in driving the PUN price as they are shown to be positive and significant in most of the cases. This is mainly due to the dominance of natural gas in the Italian production mix thus making fossil-bassed power plant as the marginal technology in most period. As a consequence, the PUN is sensitive to the natural gas price. The value of coefficient reflects the impact level of the gas price in each period. Gas price shows the highest impact on PUN price at Hour-9 with 0.26€/Mwh of PUN price change. Hour-6 and hour-2 follow from behind with only 0.106107 €/MWh and 0.103959 €/Mwh of PUN price increase. On the contrary, hour-13 presents to be the period with the smallest impact of gas price reaching only 0.0242673 €/Mwh in price changes.

	Gas		Demand	
Hour1	0.101397	***	0.000381	***
Hour2	0.103959	***	0.000433	***
Hour3	0.103647	***	0.000446	***
Hour4	0.095696	***	0.000336	***
Hour5	0.101927	***	0.000348	***
Hour6	0.106107	***	0.000368	***
Hour7	0.078273	***	0.000483	***
Hour8	0.086923	***	0.000574	***
Hour9	0.293802	***	0.001681	***
Hour10	0.051238	***	0.000557	***
Hour11	0.022519	*	0.000473	***
Hour12	0.027587	**	0.000582	***
Hour13	0.024267	***	0.000403	***
Hour14	0.019381		0.000393	***
Hour15	0.02592	**	0.000499	***
Hour16	0.038734	***	0.000612	***
Hour17	0.033622	***	0.000368	***
Hour18	0.065944	***	0.000517	***
Hour19	0.065715	***	0.000419	***
Hour20	0.060826	***	0.0002	***
Hour21	0.008157	*	-1.4E-06	
Hour22	0.008865	*	6.18E-05	*
Hour23	0.028633	***	0.00023	***
Hour24	0.043279	***	0.000284	***

*** P< 1% ** P< 5% * P< 10%

Table 2.2: Coefficient estimation of fundamental driver

2.2.5 Multivariate Framework

In order to improve robustness of our studies, we attempt to compare the forecasting ability of univariate models with multivariate models. The complexity of electricity market has encouraged researchers to develop multivariate models for gaining extra information that cannot be tackled by univariate models. Multivariate framework offers exploration in the relation between the variables and its dependence structure. On the other hand, stacked univariate models have limitations in terms of capturing dependency between hourly price. Regardless of their advantages, the question regarding the difference in forecasting ability still remains.

Focusing on three parameters that reflect their forecasting capability, MAE, ES and MAX, we compare several models under univariate and multivariate framework with an equal test ground. The models are:

- Univariate Models
  As the base of comparison, we propose the best univariate model based on our previous discussion, M3. However, it is also interesting to experiment with the capability of a combination model and to examine the difference in its forecasting ability with multivariate framework.

Hour	Model	Description
1	M6	$ARMA(1,7) - GARCH(1,1)$
2	M6	$ARMA(1,7) - GARCH(1,1)$
3	M4	$ARMA(1,1)x(1,1)_7 - GARCH(1,1)$ without exogenous variable
4	M4	$ARMA(1,1)x(1,1)_7 - GARCH(1,1)$ without exogenous variable
5	M6	$ARMA(1,7) - GARCH(1,1)$
6	M4	$ARMA(1,1)x(1,1)_7 - GARCH(1,1)$ without exogenous variable
7	M6	$ARMA(1,7) - GARCH(1,1)$
8	M3	$ARMA(1,1)x(1,1)_7 - GARCH(1,1)$
9	M3	$ARMA(1,1)x(1,1)_7 - GARCH(1,1)$
10	M6	$ARMA(1,7) - GARCH(1,1)$
11	M6	$ARMA(1,7) - GARCH(1,1)$
12	M6	$ARMA(1,7) - GARCH(1,1)$
13	M3	$ARMA(1,1)x(1,1)_7 - GARCH(1,1)$
14	M3	$ARMA(1,1)x(1,1)_7 - GARCH(1,1)$
15	M3	$ARMA(1,1)x(1,1)_7 - GARCH(1,1)$
16	M3	$ARMA(1,1)x(1,1)_7 - GARCH(1,1)$
17	M6	$ARMA(1,7) - GARCH(1,1)$
18	M6	$ARMA(1,7) - GARCH(1,1)$
19	M6	$ARMA(1,7) - GARCH(1,1)$
20	M6	$ARMA(1,7) - GARCH(1,1)$
21	M6	$ARMA(1,7) - GARCH(1,1)$
22	M6	$ARMA(1,7) - GARCH(1,1)$
23	M6	$ARMA(1,7) - GARCH(1,1)$
24	M6	$ARMA(1,7) - GARCH(1,1)$

Table 2.3: Combination of ARMA model based on the best forecasting performance

- Vector Autoregression
  Vector autoregression is one of well-known econometric model based on multivariate framework. The method generates autoregressive process for more than one variables. Hence, the method allows us to study causality and interdependencies between the variables. With this positive feature, this method gains attention from researchers in

the electricity market. In literature, Raviv et al. (2013) has attempted to employ this method in the Nordic electricity market. The paper uses VAR with dimension reduction techniques, shrinkage and forecast combinations for forecasting Nordic electricity prices. In this paper, we use a general equation for all h, as follow,

$$P_h(t) = G_h(t) + F_h(t)$$

The stochastic component, $F_h(t)$ can be modeled as a VAR process with a general notation matrix below.

$$\begin{bmatrix} F_1(t) \\ F_2(t) \\ \vdots \\ F_24(t) \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{24} \end{bmatrix} + \begin{bmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,24} \\ A_{2,1} & A_{2,2} & \dots & A_{2,24} \\ \vdots & \vdots & \dots & \vdots \\ A_{24,1} & A_{24,2} & \dots & A_{24,24} \end{bmatrix} \begin{bmatrix} F_1(t-1) \\ F_2(t-1) \\ \vdots \\ F_{24}(t-1) \end{bmatrix} + \begin{bmatrix} \omega_1(t) \\ \omega_2(t) \\ \vdots \\ \omega_{24}(t) \end{bmatrix}$$

- Seemingly unrelated regression
  Another alternative of multivariate framework is Seemingly unrelated regression or a constrained Vector autoregression. Just like our model, Huismann et al. (2007) model also use the same general equation proposed by Lucia and Schawarz (2002). Therefore, for all h, the price is calculated from deterministic, $G_h(t)$, and stochastic variable, $F_h(t)$.

$$P_h(t) = G_h(t) + F_h(t)$$

Then, $F_h(t)$ is structured as a multivariate framework with the matrix below.

$$\begin{bmatrix} F_1(t) \\ F_2(t) \\ \vdots \\ F_{24}(t) \end{bmatrix} = \begin{bmatrix} A_{1,1}F_1(t-1) \\ A_{2,1}F_2(t-1) \\ \vdots \\ A_{24,1}F_3(t-1) \end{bmatrix} + \begin{bmatrix} \omega_1(t) \\ \omega_2(t) \\ \vdots \\ \omega_3(t) \end{bmatrix}$$

You may notice that in this model we are not considering mean spillover from the other variables like in the VAR. Consequently, we are not able to quantify the exact impact of a lagged variable to the others. However, with an estimation using seemingly unrelated regressions (SUR) method, we can obtain cross-covariance matrix from the residuals.[8] The result will provide us with information regarding the interrelations between prices in different delivery periods.

2.2.6 Empirical Result and analysis

The summary of the comparison between univariate and multivariate models can be seen on the table 2.4 below. In the in-sample period, mean absolute error presents VAR as the top performer reaching forecast accuracy of 5.786. The 24 combination of ARMA-GARCH model has slightly lower accuracy with 5.860 €/Mwh thus cementing it as the second choice of the model in terms of accuracy. Huissman et al. (2007) model (SUR) has the worst accuracy with 6.39€/Mwh of MAE. The same pattern is also shown in the expected shortfall as VAR comes as the top performer with 20.13 €/MWh and combination model ranked as the second rank with 22.05 €/Mwh. Then, SUR is ranked the last with 22.78 of ES. As for the maximum error value, however, VAR presents the highest value of maximum error (216.06) followed by SUR with 212.93. On the other hand, combination model has the lowest maximum value of error (202.61). In the out-of-sample period, we see, again, the same pattern as in-sample in terms of MAE. VAR and Combination record 4.22€/MWh and 4.26 €/MWh in mean absolute error, respectively, thus making VAR the best performer. Then, M3 and SUR follow these models from behind with 4.31 €/MWh and 4.66 €/MWh of MAE. The

[8] We use Feasible Generalized Least Squares (FGLS) as the regression technique for handling heteroskedasticity and autocorrelation in SUR estimation.We refer to Baltagi(1995) and Greene(2003) for details on SUR and FGLS.

expected shortfall and maximum error value, however, show combination models as the best performer reaching 14.93 and 54.38 respectively. As for VAR, they are estimated at 15.57 and 58.95 for ES and Maximum error value respectively. SUR shows the worst performance in expected shortfall and the maximum error value. As for the DM statistic values , we may observe that our result and analysis on the mean absolute error still holds since different level of accuracy is shown from the diebold and mariano test. Combination models shows the highest DM statisic value thus indicates its superiority in forecasting electricity price in comparison to M3.

Multivariate	VAR	SUR	M3	Combination
Mean	5.786	6.390	5.904	5.860
ES	20.135	22.788	22.261	22.052
Maximum	216.06	212.937	202.924	202.618
Mean	4.226	4.665	4.318	4.260
ES	15.578	16.036	15.142	14.936
Maximum	58.952	57.060	56.506	54.383
DM	2.97 (***)	4.08(***)	-	6.42(***)

* p-value < 10 % (DM test > 1.65)
** p-value < 5 % (DM test > 1.96)
*** p-value <1 % (DM test > 2.58)

Table 2.4: Multivariate framework forecasting summary

Let us look figure 2.5 and 2.6 that show MAE of in-sample and ou-of-sample period respectively. The graph can provide us with a better comparison of univariate and multivariate models, in terms of accuracy, as we can observe the different performance for each period. In the in-sample period, SUR clearly displays a poor performance in comparison to the other models as they show high MAE for all periods of delivery. On the contrary, VAR exhibits great performance in predicting price in most of the cases. The model displays a superior performance particularly between hour 1 and hour 12. The performance beyond hour 12, however, is lower in terms of accuracy compared to univariate models. The combination model exhibits a good model in terms of mean absolute error where it is consistently ranked as second prior to period 13. Period 13 and beyond show this model as the best performer, which is shown by its low MAE. M3, on the other hand, appears to have a comparable forecast accuracy with the combination model. As for the MAE in out-of-sample, SUR presents a very poor accuracy in the first 7 periods of delivery. The model, however, displays a superior performance between hour-18 and hour-20 as well as hour-22 and hour-23. On the other hand, VAR has, again, shown to be the best model for predicting electricity price in the morning periods (between hour 1 and hour 8). However, the model has higher MAE in comparison with univariate models and SUR beyond period 9. The 24 combination of ARMA-Garch model displays relatively good forecasting accuracy between hour-1 and hour-8 as well as hour-22 and hour-23 since it follows closely the top performer from behind in these periods. Outside these hours, the model displays the lowest error. Finally, M3 is constantly behind combination model in most of the time.

To sum up our results, we may conclude that both univariate and multivariate framework have their own merits. Therefore, the best models would depend on the objective of the model. If we look into the most accurate model, Vector autoregression seems to be the most appropriate model. This is due to the fact it exhibits the lowest MAE in comparison to the other models both in-sample and out-of-sample. However, we need to be cautious as the model provide poor accuracy in out-of-sample for the peak hours between hour-9 and hour-15. In the peak period, univariate models display far better performance as the MAE display superior results. On the other hand, if the objective is to have a forecasting model with lower risk, univariate models have exhibited better results as shown in the expected shortfall and maximum value of error. The combination model, in particular, has an out-of-sample MAE that is slightly higher from VAR thus suggesting that this model has an adequate prediction of electricity with lower risk.

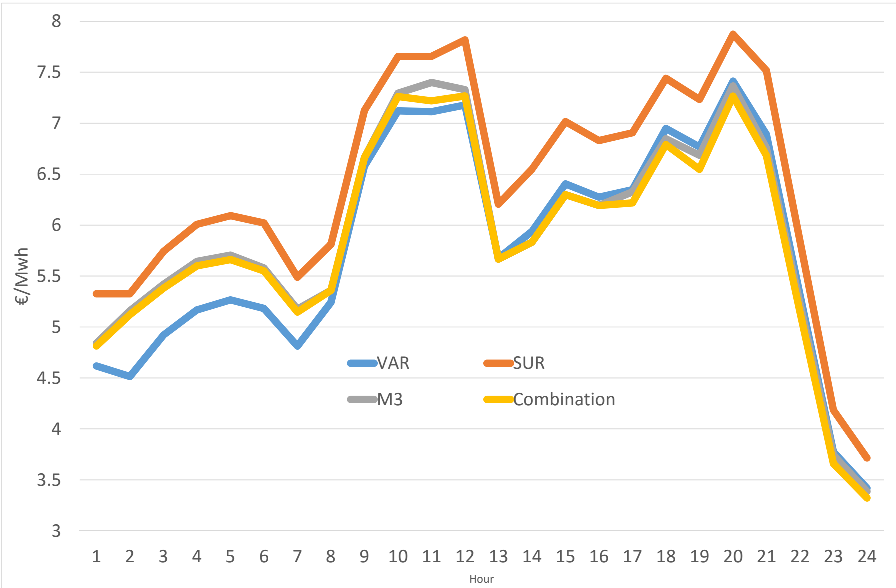

Figure 2.5: MAE of in-sample

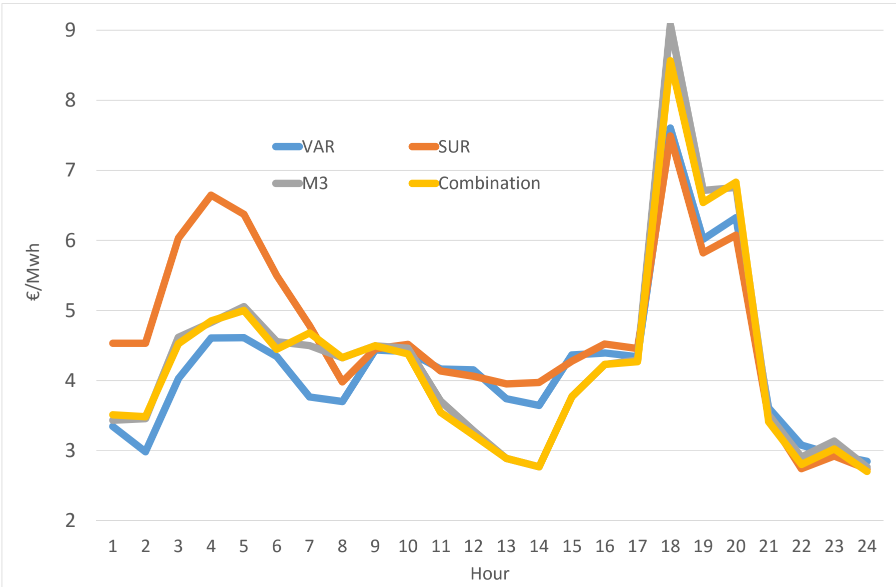

Figure 2.6: MAE of out-of-sample

2.3 Conclusion

From our forecasting results, we can conclude several important findings:

- Seasonal process adds an improvement on the forecasting accuracy.
  Our forecasting exercise has shown that seasonal ARMA process has a better MAE in both in-sample and out-of-sample periods in comparison to non-seasonal ARMA process thus indicating an increase in accuracy. This is due to the fact that market participants make a weekly adjustment on their bidding prices. Therefore, the coefficient of the seasonal process can be interpreted as the speed of reversion from seasonal deviation.
- Seasonal ARMA-GARCH model (M3) is, in overall, the best model.
  This particular model is chosen as the best model because of several reasons. Firstly, the model records the second best forecasting accuracy, 5.904 in the in-sample and 4.318 in the out-of-sample. Secondly, it displays low BIC in comparison to the other models. Thirdly, it captures the physical features of electricity prices which include traders adjustment, mean reversion and seasonality.
- Exogenous variable improve the forecasting accuracy
  Our study also validates Gianfreada and Grosi (2012) finding since models with exogenous variables display clear superiority in terms of forecasting accuracy than models without the exogenous variable. This is true in both periods, in-sample and out-of-sample.
- Stacked model is better than a global model
  From the forecasting results, we confirm the superiority of a stacked model in comparison to the global model. This is in line with Cuaresma et al. (2005) finding in the Leipzig electricity market. In addition, one can also argue that this is due to the structure of the day-ahead market in which prices of all hours of delivery are determined at the same time.
- Gas price and demand provide an increasing impact on PUN price
  The estimations of our model have shown that gas price and demand have positive and significant coefficients in almost all periods with several exceptions. This is due to the fact they both shift supply and demand curve in positive price direction which, subsequently, increase the PUN price.
- Both univariate and multivariate model has their own merit
  Our empirical test suggest that VAR, a multivariate model, is the best model in terms of accuracty wheras univariate models has shown to be superior in terms of risk. Therefore, utilization of the model should depends on the modeling or forecasting objective.

We believe that further studies on the Italian electricity market should explore price modelling with the weather as its exogenous variable. Our study is limited by the data availability of the weather since frequency is the main issue. Historical weather data is generally limited to the daily data which cannot explain the hourly dynamic on electricity price. In addition, the data is generally presented for each city thus additional work is needed to make a weighted adjustment for every aggregated zonal market. Other challenges will occur in terms of the lagged value of the weather. The shock in temperature does not produce a sudden impact in demand since generally people would adjust their heater/air conditioner after a certain period of time. The same idea is also applied to hydro supply curve, a sudden increase in rainfall is generally not equal to a sudden increase in supply because traders adjust hydro outputs according to the water level and the electricity price.

Chapter 3

Intermittent renewable generation and network congestion: an empirical analysis of Italian Power Market

Faddy Ardian
École Polytechnique
Silvia Concettini
École Polytechnique
Anna Creti
Université Paris Dauphine and École Polytechnique

Abstract

This article aims at contributing to the scant literature on the effect of increasing renewable power production on congestion frequency and cost.The economic literature has emphasized the likely reductions of wholesale prices entailed by increasing renewable supply. Beside the effects on prices, renewable supply has raised some concerns regarding network functioning and congestion management. This is due to the fact that increasing renewable output may put an additional stress on the infrastructure, amplifying transportation needs and multiplying congestion occurrence, when production and consumption sites are far from each other. The impact of renewable on network congestion may be explicitly investigated in national electricity markets organized as two or more inter-connected sub-markets (or bidding zones) where transmission rights are assigned through implicit auctions. The analysis of the links between renewables and congestion results to be extremely relevant in the path toward the implementation of the European Electricity Target Model which envisages the creation of bidding zones (defined or not by national borders) within a single EU market. We have estimated then two econometric models performed on Italian case study: a multinomial logit model, whose dependent variable has three discrete values capturing both the occurrence of congestion and its direction, and a two stage least square model (2SLS) with segmented regression which seeks to quantify the effects of renewable production on congestion costs. Our analysis suggests that the effect of a larger local wind and solar supply is to decrease the probability of suffering congestion in entry and to increase the probability of causing a congestion in exit compared to no congestion case. The estimations on congestion cost reveal that, due to the merit order effect, local larger renewable tend to push the congestion cost towards negative value as it decreases the marginal cost for balancing the system.

Contents

3.1 Introduction and literature review

The interest in alternative energy has sparked in Europe as the climate change problem emerged. The 2009 Climate and Energy package has motivated European governments to stimulate renewable energy penetration through supporting schemes in order to meet the target of a 20% share of EU energy consumption produced from renewable sources by 2020. According to the more recent figures from Eurostat, the share of renewables in gross final energy consumption has reached 14.95% in the EU-28 in 2013. The economic literature has emphasized the likely reductions of wholesale prices entailed by increasing renewable supply and originated from the displacement of higher variable cost production in the merit order ranking. This phenomenon is referred to as "merit order effect". A larger renewable production has also determined an increase in wholesale price variance as a consequence of technological dependency on exogenous variables. Evidence of these effects have been empirically analyzed, for instance, in Australia (Cutler et al., 2011), Austria (Wurzburg et al., 2013), Denmark (Jónsson et al., 2010), Germany (Wurzburg et al., 2013; Ketterer, 2014), Israel (Milstein and Tishler, 2011), Ireland (O'Mahoney and Denny, 2011), Italy (Clo et al., 2015), Spain (Gelabert et al., 2011).

Besides the effects on prices, renewable supply has raised some concerns regarding network functioning and congestion management. Some geographical locations seem particularly well suited for the installation of new capacity due to the abundance of natural resources (e.g. the North for wind and the South for solar in both Germany and Spain). Nevertheless, the existing networks may not be adequately developed to guarantee a constant and smooth flowing of more efficient RES production[1] toward consumption sites. When production and consumption sites do not coincide and are, on the contrary, very far from each other, increasing renewable output may put an additional stress on the infrastructure, amplifying transportation needs and multiplying congestion occurrence. The opposite happens if renewable supply relieves deficits of production in historical importing regions. Hence, depending on the location of supply and demand, a larger renewable production may have a positive or negative effect on congestion occurrence and, as a consequence, on congestion cost.

The impact of renewable on network congestion may be explicitly investigated in national electricity markets organized as two or more inter-connected sub-markets (or bidding zones) where transmission rights are assigned through implicit auctions.[2] A sub-market or bidding zone is defined as the largest geographical area within which market participants can offer and buy energy in the intra-day, day-ahead, and longer time frame markets; its boundaries are generally settled based on physical transmission limits in order to achieve an efficient use of the infrastructure. In the absence of transmission constraints, prices are equal across zones; when inter-zonal constraints are binding, zonal market prices diverge. With an implicit auctioning for transmission rights, transmission capacity is included in the auctions of electricity. In other words, the resulting electricity prices per area reflect both the cost of energy in each internal bidding area and the cost of congestion. Implicit auctions ensure that power flows from the surplus areas (low price areas) towards the deficit areas (high price areas) since it uses nodal pricing system (see Harvey and Hogan (2000)). The analysis of the links between renewables and congestion results to be extremely relevant in the path toward the implementation of the European Electricity Target Model which envisages the creation of bidding zones (defined or not by national borders) within a single EU market. Because of heterogeneous generation mix, geographical conditions, RES support schemes and national network configurations across EU Members, the European Target Model may face on a larger scale the same challenges of those Countries with bidding areas having experienced a significant renewable penetration.

This article aims at contributing to the scant literature on the effect of increasing renewable power production on congestion frequency and cost. Førsund et al. (2008) started the discussion on this link by studying the impact of wind power penetration in Norway. Their

[1] In terms of marginal cost.

[2] The same analysis can be applied to market coupling.

model has concluded the increase of network congestion between northern and southern Norway where there is a significant different in the hydro resource. In the another continent, Woo et al. (2011) pioneer econometric analysis on this particular relation using the case of ERCOT power market with an ordered logit model and a log-OLS model. The analysis stems from the observation that wind generation is mostly concentrated in the West zone, which is scarcely populated, whereas generation capacity in Houston zone falls short of its zonal load. The authors show that rising wind supply, nuclear generation, load from non-West zones and gas price increases the likelihood and the size of strictly positive paired-price differences between the West and the other zones.[3] In Germany, Schröeder et al. (2013) and Kuntz (2013) have provided technical and economic analysis of future congestion problems as a consequence of high integration of wind energy. Both studies conclude an increase in the overall cost for stabilizing saturated transmission caused by the renewable if there is no strategic network expansion in Germany. Figuiredo et al. (2015) assess variables that determine the occurrence of congestion using logit and non-parametric Keynesian function for the case of Iberian spot electricity price (Spain). Similar to previous studies, the paper concludes that large availability of baseline technology coupled with high renewable has been proven to increase the market splitting.

In order to assess the impact of increasing renewable output on congestion frequency and cost, we use Italian electricity market as a case study. For its particular features, Italy provides an interesting case study. Firstly, the Italian Power Exchange is composed of six regional sub-markets, which aggregate in macro-zones all administrative regions. Since each of the zones has its own specific generation mix, they provide heterogeneity in our samples. Secondly, the ambitious support policies for the development of renewable power sources have generated a significant amount of new investments in solar and wind power plants. According to the latest available data,[4] the supply from these power plants has covered 15.89% of the electricity purchased in the day-ahead market in 2014. Solar and wind production sold through the day-ahead market have registered an increase of 267.2% from 2010 to 2014. Southern regions have shown the highest growth rate due to the favorable weather conditions. This rapid growth is an essential characteristic for studying RES impact on congestion. Thirdly, the inter-zonal transmission capacities are not equally distributed in the Italian electricity system. In particular, transmission lines that connect the islands to the Italian peninsula have limited capabilities. With high renewable penetration in some regions and transmission limitations, Italy has the ideal conditions for a case study.

To empirically test the effect of renewables on congestion in Italy, we have built a unique database collecting and matching data with hourly frequency for a five-year period (2010-2014) from two sources: GME, the market operator, which publishes the hourly offers in the day-ahead market together with equilibrium prices, quantities and inter-zonal transits; REF-E, a consulting group, who has created a list of Italian power plants classified by technology and geographical location. We have estimated then two econometric models performed on five zonal pairings: a multinomial logit model, whose dependent variable has three discrete values capturing both the occurrence of congestion and its direction, and a two stage least square model (2SLS) with segmented regression which seeks to quantify the effects of renewable production on implicit congestion costs.[5] Up to our knowledge, Sapio (2015) is the only author testing the impact of rising solar and wind generation on congestion between Sicily and Southern Italy using regime-switching model on 2012-2013 hourly data. The results provide more insight on the locational-dependent-effect of renewable on the congestion. His estimation and analysis emphasize on the congestion-relieve impact from Sicily's renewable production. This is due to the fact that solar and wind production from Sicily substitute the transmission capacity and reduce the import from the South. The estimation also reveals that wind power shows a bigger impact on congestion compared to the solar power that has less variation in the power production.

This article originally contributes to the literature in several ways. First, we enlarge

[3]A strictly positive paired-price difference occurs when the West price is lower than the price in the other zones and vice-versa, meaning that the congestion is "coming from the West".

[4]GME, Annual Report 2014.

[5]The description of this congestion costs will be explained in the following sections.

the scope of the analysis by considering all Italian neighboring zones in order to verify the consistency of the empirical models beyond the specificities of each pair. Second, we employ a multinomial logit model in order to separately capture the effect of increasing renewable production on the probability of both directional congestions (to and from) compared to the benchmark situation of no congestion. Third, we consider zonal figures on production and demand instead of aggregated figures to isolate the contribution of each zone to the occurrence of congestion. Fourth, we estimate the impact of renewable output not only on congestion frequency but also on congestion cost, something that has never be done before in the literature, taking into account hydroelectric endogeneity issues. Finally, we apply segmented regressions instead of simple OLS in order to capture the impact of renewable on the congestion cost in two congestion directions, congestion to and congestion from, which are modeled as regimes (segments).

Our analysis suggests that the effect of a larger local wind and solar supply is to decrease the probability of suffering congestion in entry and to increase the probability of causing a congestion in exit compared to no congestion case. Increasing hydroelectric production has a similar effect. A rise in local demand, on the contrary, increases the probability of congestion in the entry (due to larger import) and decreases the probability of congestion in the exit. These results hold for both importing and exporting regions, but importing regions are much less likely to cause congestion in the exit, therefore the installation of new RES capacity in these zones may have a positive effect in terms of flow balance between regions. The estimations on congestion cost reveal that, due to the merit order effect, local larger renewable tend to push the congestion cost towards negative value as it decreases the marginal cost for balancing the system. A much bigger shock of renewable quantity consequently could reduce saturated line and merge the zone that is, zero congestion costs or could widen the gap of negative congestion costs because of excessive supply in the exit. This is true for all importing zone but it is the opposite for the exporting zones. Therefore, the increase of renewable should be promoted in the importing zones, but the overall growth should be controlled in order to avoid congestion toward opposite direction.

The remainder of the paper is organized as follows. Next section briefly describes Italian electricity market and the rules for congestion management. The third section provides an overview of the day-ahead market transactions in terms of the generation mix, interzonal transits and price differences between neighboring zones. The fourth section is dedicated to the econometric analysis. The last section concludes our study.

3.2 Empirical strategy

The descriptive statistics of the series used in the multinomial logit model are reported in Tables B.1 and B.2 in Appendix B. The descriptive statistics for the series used in the 3 stage least square analysis are shown in Tables **??** and **??** in Appendix C. Demand and price series are directly collected from GME database. Supply series have been constructed by aggregating bid data to build the hourly market supply curves resulting from the market splitting algorithm. GME bids have then been matched with REF's database containing a mapping of power plants from bidding units to technology. The empirical models have been estimated at five zonal pairs using observations from the 2010-2014 period. The five zonal pairs are:

1. CNOR-NORD
2. CNOR-CSUD
3. SARD-CSUD
4. CSUD-SUD
5. SICI-SUD

The first zone of the pair is generally an importing region.

3.2.1 Multinomial logit model

For each zonal pair (ZONE1-ZONE2) the dependent variable in the multinomial logit model, y, may assume three values:[6]

- $y = -1$ when the zonal price in ZONE1 is lower than the zonal price in ZONE2: in this case we say that there is "congestion from" ZONE1 (which is exporting power) or a "negative price difference";
- $y = 0$ when the zonal prices in ZONE1 and ZONE2 are equal: in this case, we say that there is "no congestion" (the flows between the two zones respect the transmission constraint) and hence no price difference;
- $y = 1$ when the zonal price in ZONE1 exceeds the zonal price in ZONE2: in this case we say that there is "congestion to" ZONE1 (which is importing power) or a "positive price difference".

On average, the zonal prices of the neighboring zones paired for the five-year period differ about 27% of the time; however, this figure hides a large heterogeneity (see Table B.2). If the link CNOR-NORD has been uncongested for 92% of the time, the zones SICI and SUD have registered the same price only 18% of the total hours. Quite surprisingly, congestion coming from CNOR have been more frequent that those coming from NORD, indicating a change in flows direction between these two zones. In CNOR-CSUD, CNOR confirms to be an importer with congestion to this region accounting for 7% of the hours, while most of the time the two zone have experienced no congestion (91% of the time). The lines SARD-CSUD and CSUD-SUD have followed similar patterns with congestion to the first zone occurring 16% and 15% of the time, respectively. The frequencies of no congestion have been also similar (83% and 85% of the time respectively). Is is worthy to note that while congestion from SARD to SUD has occurred, although quite rarely (1% of the hours), CSUD has never exported to SUD. Finally flows to SICI have congested the line SICI-SUD 75% of the time while flows from SICI have done so for 7% of the hours.

Fo each zonal pair we are going to estimate the following two equations [7]:

$$\log \frac{Pr(y=-1)}{Pr(y=0)} = \alpha_1 + \sum_{i=1}^{4} \eta_i \boldsymbol{Y_i} + \sum_{r=1}^{4} \beta_r \boldsymbol{Y_r} + \epsilon \tag{3.1a}$$

$$\log \frac{Pr(y=1)}{Pr(y=0)} = \alpha_2 + \sum_{i=1}^{4} \eta_i \boldsymbol{Y_i} + \sum_{r=1}^{4} \beta_r \boldsymbol{Y_r} + \epsilon \tag{3.1b}$$

where:

- $\alpha_{1,2}$ are the constants
- $\boldsymbol{Y_i}$ is the matrix of yearly dummies
- $\boldsymbol{Y_r}$ is the matrix of regressors and includes:
  - Hydro generation in the pairing zones (ZONE_Hydro)
  - Wind generation in the pairing zones (ZONE_Wind)
  - Photovoltaic generation in the pairing zones (ZONE_PV_tot)
  - Demand in the pairing zones (ZONE_D)

[6] CSUD-SUD pair is characterized by the occurrence of only two outcomes; in this case, we estimate a logit model with a binary dependent variable.

[7] It is important to be noted that we have experimented with several different dummies and determinant. However, the econometric estimation is not satisfactory. Yearly trend is utilized to capture the increasing renewable production in our dataset period.

Multinomial logit results

The results of the estimations in terms of log-odds and marginal effects are reported in Tables 3.1 and 3.2 respectively.[8] Standard errors are reported below the coefficients. In general we may observe (see Table 3.1) that the coefficients associated to renewable generation in ZONE1, when significant, have positive signs indicating that a rise in the production is associated with an increase in the relative log odds of a congestion coming from ZONE1 ($y = -1$) with respect to no congestion due to improved export possibilities. The coefficients of RES production in ZONE2 are on the contrary negative, indicating that a rise in production in these zones decreases the likelihood of congestion from ZONE1 since less import are needed. An opposite reasoning works for the demand: when the demand increases in ZONE1 there is less export, therefore, less probability of congestion from ZONE1. The reverse is true when the demand rises in ZONE2 since more import is needed and hence the probability of congestion from ZONE1 increases.

The results for congestion to ZONE1 ($y = 1$) are generally symmetric, with the coefficients displaying opposite signs with respect to the previous case. When ZONE1 is importing power, a larger renewable generation in ZONE1 reduces importing needs and thus the relative log odds of a congestion to ZONE1 with respect to no congestion declines. A larger RES production in ZONE2, on the contrary, increases the likelihood of congestion to ZONE1 due to the improved export possibilities. An opposite pattern is again followed by the demand: when the demand is larger in ZONE1 there is more need to import, therefore, a higher probability of congestion to ZONE1. Finally, when the demand rises in ZONE2 less production can be exported and as a consequence the probability of causing a congestion to ZONE1 decreases.

It is worthy to analyze more closely the marginal effects of each regressor on the probability of directional congestion (see Table 3.2). Marginal effects indicate how the probability of an outcome increases when the regressor increases by a megawatt hour, all the other regressors kept at their average. The results for congestion from ZONE1 ($y = -1$) may be summarized as follow.[9]

In ZONE1:

- Rising wind production increases the probability of congestion in all pairs: a Mwh increase in wind generation in ZONE1 raises the probability of congestion by 0.024% in CNOR-NORD pair, by 0.014% in CNOR-CSUD pair and by 0.013% in SICI-SUD pair; this effect, although positive, is smaller in magnitude in SARD-CSUD pair;
- Increasing photovoltaic production has a positive impact on congestion in CNOR-NORD and SICI-SUD pairs (the coefficient is not significant in CNOR-CSUD and SARD-CSUD pairs): a Mwh increase in production rises the probability of congestion by 0.0026% in CNOR and by 0.0111% in SICI;
- A rise in the hydroelectric production has a positive effect on congestion in all pairs with the exception of SARD-CSUD (probably due to the scarce hydroelectric production in SARD): producing an additional Mwh increases the probability of congestion by 0.0022% in CNOR-NORD, by 0.0011% in CNOR-CSUD and by 0.0402% in SICI-SUD;
- Overall an increase of 1 Mwh for each renewable source (wind, photovoltaic and hydroelectric) in ZONE1 increases the probability of congestion from this zone by 0.0288% in CNOR-NORD, by 0.0159% in CNOR-CSUD and by 0.064% in SICI-SUD;
- A rising demand in ZONE1 has exactly the opposite effect, i.e. it decreases the likelihood of congestion from ZONE1, in all pairs with the exception of SARD-CSUD where the coefficient is not significant: a Mwh increase in demand in ZONE1 decreases the likelihood of congestion by 0.0018% in CNOR-NORD, by 0.0008% in CNOR-CSUD, by 0.011% in SICI-SUD.

[8] The details of the estimation for each zonal pair are shown in Tables from B.3 to B.7 in the Appendix.

[9] Note that in CSUD-SUD pair the congestion never comes from CSUD.

Log-odds							
		Variable	CNOR_NORD	CNOR_CSUD	SARD_CSUD	CSUD_SUD	SICI_SUD
Zone 1	y=-1	Hydro	0.00269***	0.00221***	0.00309		0.0109***
			0.000191	0.00046	0.00268		0.0019
		Wind	0.0286***	0.0287***	0.00863***		0.00198***
			0.00213	0.00511	0.0004		0.00013
		PV	0.00313***	-0.00042	-0.00379		0.000965**
			0.000166	0.0012	0.00283		0.00048
		Demand	-0.00223***	-0.00148***	-0.00016		-0.00169***
			0.000115	0.00016	0.0004		0.00018
	y=1	Hydro	-0.00143***	0.000395*	0.00215***	0.00143***	-0.0118***
			0.000226	0.00021	0.00057	0.00015	0.00131
		Wind	-0.00615*	-0.00225	-0.00999***	0.00114***	-0.00570***
			0.00328	0.00204	0.00027	0.00019	9.3E-05
		PV	-0.00622***	-0.00016	-0.00308***	0.000270**	-0.00553***
			0.000473	0.00025	0.00109	0.00013	0.00028
		Demand	0.00295***	0.00160***	0.00411***	0.00128***	0.00487***
			0.000108	8.71E-05	0.00012	3.32E-05	0.00012
		Variable	CNOR_NORD	CNOR_CSUD	SARD_CSUD	CSUD_SUD	SICI_SUD
Zone 2	y=-1	Hydro	-0.000427***	-0.00199***	-0.00819***		-0.00132***
			2.67E-05	0.00054	0.00058		0.00019
		Wind	0.0286***	-0.00479***	-0.00100**		-0.000754***
			0.00476	0.00048	0.0004		8.6E-05
		PV	0.000148***	-0.00186*	0.00296***		-0.000643***
			5.34E-05	0.00099	0.00066		0.00017
		Demand	0.000408***	0.00144***	-0.000180**		0.000722***
			1.82E-05	0.00012	7.36E-05		0.00013
	y=1	Hydro	0.000508***	0.000799***	0.000441***	0.000601***	0.000793***
			2.89E-05	0.00024	0.00015	0.00012	0.00012
		Wind	-0.0139**	0.00355***	-0.0001	0.000706***	0.000447***
			0.00541	0.00011	0.00017	8.88E-05	5.1E-05
		PV	0.000929***	0.000900***	-0.000909***	0.00126***	0.00106***
			0.000117	0.00021	0.00024	7.93E-05	9.7E-05
		Demand	-0.000463***	-0.00150***	-0.000126***	-0.00109***	-0.000397***
			2.04E-05	6.52E-05	2.11E-05	6.55E-05	7.9E-05

*** $p<0.01$ ** $p<0.05$ * $p<0.1$

Table 3.1: Multinomial logit estimations (Log-odds), 2010-2014
$y=-1$ when congestion is from ZONE1
$y=1$ when congestion is to ZONE1

					Marginal effects		
		Variable	**CNOR_NORD**	**CNOR_CSUD**	**SARD_CSUD**	**CSUD_SUD**	**SICI_SUD**
Zone 1	**y=-1**	**Hydro**	2.27E-05***	1.13E-05***	1.50E-06		0.000402***
			2.14-E06	2.43E-06	1.40E-06		0.0000368
		Wind	0.000240***	0.000148***	4.66E-06***		0.000133***
			2.43E-05	2.73E-05	9.63E-07		0.00000453
		PV	2.68E-05***	-2.10E-06	-1.82E-06		0.000111***
			2.39E-06	6.18E-06	1.48E-06		0.00000954
		Demand	-1.89E-05***	-8.02E-06***	-2.04E-07		-0.000114***
			1.53E-06	1.01E-06	2.13E-07		0.00000468
	y=1	**Hydro**	-1.48E-05***	1.83E-05*	0.000121***	0.000103***	-0.00146***
			2.38E-06	9.82E-06	3.18E-05	1.05E-05	0.000136
		Wind	-6.50E-05*	-0.000115	-0.000563***	8.23E-05***	-0.000652***
			3.34E-05	9.77E-05	1.39E-05	1.36E-05	0.0000105
		PV	-6.36E-05***	-7.29E-06	-0.000173***	1.94E-05**	-0.000616***
			4.53E-06	1.20E-05	6.10E-05	9.20E-06	0.0000298
		Demand	3.02E-05***	7.68E-05***	0.000231***	9.20E-05***	0.000557***
			1.74E-06	4.00E-06	6.80E-06	2.43E-06	0.0000125
		Variable	**CNOR_NORD**	**CNOR_CSUD**	**SARD_CSUD**	**CSUD_SUD**	**SICI_SUD**
Zone 2	**y=-1**	**Hydro**	-3.62E-06***	-1.05E-05***	-4.16E-06***		-3.81E-05***
			3.18E-07	2.84E-06	8.90E-07		0.00000364
		Wind	0.000241***	-2.56E-05***	-5.04E-07**		-2.17E-05***
			4.30E-05	2.52E-06	2.31E-07		0.00000171
		PV	1.16E-06***	-9.80E-06*	1.52E-06***		-3.00E-05***
			4.47E-07	5.11E-06	4.48E-07		0.00000332
		Demand	3.45E-06***	7.78E-06***	-8.72E-08**		2.03E-05***
			2.70E-07	7.98E-07	4.01E-08		0.00000243
	y=1	**Hydro**	5.20E-06***	3.87E-05***	2.51E-05***	4.33E-05***	0.000108***
			3.51E-07	1.16E-05	8.48E-06	8.85E-06	0.000012
		Wind	-0.000144***	0.000171***	-5.80E-06	5.09E-05***	6.11E-05***
			5.55E-05	5.18E-06	9.31E-06	6.42E-06	0.00000534
		PV	9.44E-06***	4.36E-05***	-5.13E-05***	9.09E-05***	0.000126***
			1.18E-06	1.01E-05	1.36E-05	5.79E-06	0.0000102
		Demand	-4.75E-06***	-7.21E-05***	-7.12E-06***	-7.88E-05***	-5.52E-05***
			3.00E-07	2.93E-06	1.17E-06	4.79E-06	0.00000817

*** $p<0.01$ ** $p<0.05$ * $p<0.1$

Table 3.2: Multinomial logit estimations (Marginal effects), 2010-2014
$y = -1$ when congestion is from ZONE1
$y = 1$ when congestion is to ZONE1

In ZONE2:

- Rising wind production decreases the probability of congestion in all pairs except in CNOR-NORD (where wind production in NORD seems to increase congestion): this effect is more evident in CNOR-CSUD and SICI-SUD where an additional MWh of wind supply in ZONE2 decrease the likelihood of congestion by 0.0025% and 0.0021% respectively; the coefficient is significant and has the expected sing also in SARD-CSUD pair but its lower value seems to indicate a softer effect on congestion of a larger wind supply in CSUD;
- Rising solar production in ZONE2 decreases the probability of congestion in CNOR-CSUD and in SICI-SUD pairs by 0.0009% and 0.003% respectively for an MWh increase (a larger photovoltaic production in ZONE2 seems instead to increase the congestion in CNOR-NORD and SARD-CSUD);
- Rising hydro production in ZONE2 decreases the probability of congestion in all pairs: an additional MWh of electricity decreases the likelihood of congestion by 0.0003% in CNOR-NORD, by 0.001% in CNOR-CSUD, by 0.0004% in SARD-CSUD and by 0.0038% in SICI-SUD;
- In all pairs a larger demand in ZONE2 increases the probability of congestion; this effect is larger in magnitude in SICI-SUD pair, followed by CNOR-CSUD, CNOR-NORD, and SARD-CSUD.

We expect to observe results of opposite sign when studying the congestion to ZONE1 ($y = 1$).

In ZONE1:

- Rising wind production decreases the probability of congestion in all pairs with the exception of CNOR-CSUD (where the coefficient is not significant) and CSUD-SUD (where the coefficient is positive): an MWh rise in wind supply decreases the likelihood of congestion by 0.0065% in CNOR-NORD, by 0.0563% in SARD-CSUD and by 0.0652% in SICI-SUD;
- A larger photovoltaic production decreases the probability of congestion in CNOR-NORD, SARD-CSUD and SICI-SUD pairs (the coefficient is again not significant in CNOR-CSUD and positive in CSUD-SUD): an additional MWh from this source decreases the probability of congestion by 0.0063% in CNOR-NORD, by 0.0173% in SARD-CSUD and by 0.0613% in SICI-SUD;
- Rising hydro production decreases the probability of congestion only in CNOR-NORD and SICI-SUD pairs (in CNOR-CSUD the coefficient is not significant and it is positive in SARD-CSUD and CSUD-SUD): if the supply of hydro increases by 1 MWh in CNOR and SICI the probability of congestion would decrease by 0.0014% and by 0.0146% respectively
- An MWh rise in the demand of ZONE1 increases the likelihood of congestion by 0.003% in CNOR-NORD, by 0.0076% in CNOR-CSUD, by 0.0231% in SARD-CSUD, by 0.0092% in CSUD-SUD and by 0.0557% in SICI-SUD.

In ZONE2:

- Rising wind production increases congestion to ZONE1 in all pairs except in CNOR-NORD (where wind production in NORD seems to decrease congestion) and in SARD-CSUD (where the regressor is not significant): an MWh increase in wind production increases the likelihood of congestion by 0.0171% in CNOR-CSUD, by 0.005% in CSUD-SUD and by 0.0061% in SICI-SUD;
- Rising photovoltaic production increases congestion in all pairs with the exception of SARD-CSUD (where the coefficient is negative): when the photovoltaic supply increases by 1 MWh in the probability of congestion rises by 0.0009% in CNOR-NORD and CSUD-SUD, by 0.0043% in CNOR-CSUD and by 0.0126% in SICI-SUD;

- A larger hydroelectric production increases the probability of congestion in all pairs: an MWh increase in hydroelectric production increases the likelihood of congestion by 0.0005% in CNOR-NORD, by 0.0038% in CNOR-CSUD, by 0.0025% in SARD-CSUD, by 0.0043% in CSUD-SUD and by 0.0108% in SICI-SUD;
- A larger demand in ZONE2 decreases the probability of congestion in all pairs: by 0.00047% in CNOR-NORD, by 0.0072% in CNOR-CSUD, by 0.00071% in SARD-CSUD, by 0.0078% in CSUD-SUD and by 0.0055% in SICI-SUD for an MWh increase.

Thanks to the symmetry, the results can be easily summarized (Table 3.3). Increasing renewable production in a zone increases the likelihood of causing a congestion to the neighboring zone, due to larger export possibilities. At the same time, a larger local supply reduces import needs thus decreasing the likelihood of suffering congestion in the entry. Increasing local demand has an opposite effect: it lowers export possibilities, thus decreasing the probability of causing congestion in the exit, and it raises import needs, therefore increasing the probability of suffering congestion in the entry. It is worthy to note that these results hold for both importing and exporting regions. However, the importing regions are less likely to produce congestion in exit and more likely to suffer congestion in the entry. Therefore, a larger RES production in these regions is expected to bring more balance in the flows between regions, while a larger RES production in exporting zones may exacerbate the problem of congestion.

	Wind	**PV**	**Hydro**	**Demand**
Congestion from ZONE1	ZONE1: ↑ ZONE2: ↓	ZONE1: ↑ ZONE2: ↓	ZONE1: ↑ ZONE2: ↓	ZONE1: ↓ ZONE2: ↑
Congestion to ZONE1	ZONE1: ↓ ZONE2: ↑	ZONE1: ↓ ZONE2: ↑	ZONE1: ↓ ZONE2: ↑	ZONE1: ↑ ZONE2: ↓

Table 3.3: Multinomial logit result summary

3.2.2 2SLS with segmented regression

Having understood the impact of RES production and demand on the probability of congestion, we extend our research by examining its impact on the sensitivity of the congestion cost. This particular cost is implicitly paid by all IPEX participants (producers and consumer) through GME. Since consumer buys electricity at the national price (PUN), P_{PUN}, and producers are remunerated by zonal prices, P_i, we may find a difference between purchase value and sales value as follow.

$$\Delta = P_{PUN}Q - \sum_{i}^{n} P_i Q_i$$

with

$$\sum_{i}^{n} Q_i = Q$$

Using this offset (Δ), GME pays Terna the congestion cost for each transmission line (pair) with the calculation below

$$Cost_{1-2} = (P_{ZONE1} - P_{ZONE2}) * Q_{1-2}$$

Where $Cost_{1-2}$ is the congestion cost of a transmission line between $Zone1$ and $Zone2$, P_{ZONE1} is the zonal price of the importing zone, P_{ZONE2} is the zonal price of the exporting zone and Q_{1-2} is the quantity transferred from $Zone1$ and $Zone2$. In other words, each market participant implicitly pays Terna's service (implicit congestion cost) through GME priced as zonal price difference

$$ZONE1\_ZONE2 = P_{ZONE1} - P_{ZONE2} \tag{3.2}$$

The general summary of implicit congestion cost (ICC), $ZONE1\_ZONE2$, can be seen in table **??** under appendix C. The statistics of ICC appears to be non-normally distributed based on the Jarque-Bera test. This is mainly due to the high frequency of zero values. The means in SICI-SUD display the highest value with SARD-CSUD quite far behind. Hence, both transmission lines can be considered as the two most expensive lines in terms of congestion cost. This is due to frequent congestion and scarce efficient supply in the importing zone is available for balancing the system. Transmission lines in CSUD-SUD, CNOR-NORD and CNOR-CSUD are ranked third, fourth and fifth from the most expensive transmission line, respectively. The preliminary results of our unit root test have indicated that our observations are not stationary time series and require further treatment to avoid a spurious regression (Granger and Newbold (1974)). In our treatment, we reach stationarity process after detrending the data with its seasonality and trend, which are:

- Seasonality within a day (Hourly),
- Seasonality within a week (Daily),
- Seasonality within a year (Monthly),
- Yearly trend.

Unlike in Woo et al. (2011), in order to to capture the impact of renewable energy supply in the Italian power exchange, we use 2SLS method instead of a simple OLS.[10] In the first stage of 2SLS, we attempt to attack endogeneity problems of hydro in order to avoid bias in the estimation. Unlike renewable-energy supply, hydro production can be adjusted depending on the weather and portfolio optimization since it can be stored. In run-of-river hydro, a poundage is generally present for short term reserve whereas hydro with pumping

[10]In an attempt to capture the mechanism as a whole (one system) a 3SLS estimation was also conducted. However, correlation covariance matrix shown from the third stage of 3SLS (SUR) estimation does not indicate a correlation between the transmission lines. In addition, 2SLS estimation displays better results in terms of coefficient and goodness of fit because we may add time series process in the residual.

technology is operated fully on the price arbitrage. Hence, their output cannot be considered as fully exogenous variables. In this study, we use lagged hydro production as the instrument variables for hydro production.[11] We select $t-1$, $t-24$ and $t-168$ since the hydro production has daily seasonality, weekly seasonality and depends on the production of the hour before. Hence, our first-stage regression equation can be formulated as follows.

$$\hat{H}_t = \theta + \eta_1 H_{t-1} + \eta_2 H_{t-24} + \eta_3 H_{t-168} \ ,$$

Where $\hat{H}_t$ is the fitted value of hydro production at time t and θ is a constant.

In the second stage of the regressions, we apply segmented regressions in order to correctly capture the phenomenon on the changes in ICC. This method is significantly different compared to Sapio (2015) and Haldrup and Nielsen (2006) where regime switching model is applied to capture different congestion regime. The regime switching model may capture the impact of the exogenous variable on different regimes of prices (congested and non-congested). Hence, it uses prices as dependent variable which cannot capture the impact of the exogenous variable on congestion cost. Our segmented regression allows us to use congestion cost as a dependent variable and quantify the linear relation on a different regime. This method is also a better alternative to the reg-ARFIMA method proposed by Gianfreda and Grossi (2009, 2012) where congestion and congestion cost are considered as the exogenous variable. This is not true since congestion is an endogenous event that is caused by many variables. In addition, this method allows us to estimate the impact of exogenous variables on the different clusters of congestion direction (regime) which may explain the determinant of the cost of transmission saturation. Therefore, for each zonal pair, we have three different regression regimes as follow.

$$y_t = \begin{cases} \alpha_0 + \sum_i^3 \beta_0^i Y_i + z_t & y_t > 0 \\ 0 & y_t = 0 \\ \alpha_1 + \sum_i^3 \beta_1^i Y_i + z_t & y_t < 0 \end{cases}$$

Where α_i is a constant and X is a matrix of regressors that includes

- Renewable supply of the pairing zones, R ;
- Fitted hydro supply of the pairing zone, H ;
- Demand of the pairing zone, D.

The first segment (regime) estimate the impact of independent variables on the case of congestion to Zone1 whereas the third segment (regime) capture the regression in the opposite case (congestion from Zone1). The second segment models the zero values that occur because of no congestion in the transmission line. For the regression, the three regimes may be generalized into one equation as follow.

$$y_t = \alpha + d_i(y_t)(\sum_i^3 \beta_1^i Y_i) + z_t \tag{3.3}$$

Where $\alpha_0 = \alpha_1 = \alpha$, $d_i(.)$ is dummy variables of the regime as a function y_t and z_t is the residual.

Statistics estimations on high-frequency data such as hourly electricity price require extra treatment from researchers as heteroskedasticity and autocorrelation in the regression could provide a bias in coefficient and error in residual estimation. Our test has shown that volatility clustering occurs in the residual estimation of the second stage 2SLS. In addition, as shown in our example of ACF and PACF plot (see figure 3.1), it is clear that additional process needs to be added in our equation in order to treat autocorrelation. This result clearly violates the assumption of independent error in OLS regression. In addition, time series process of the residual reflects the error correction of the power trader in the day-ahead market. Following Shumway and Stoffer (2011), we may model the residual as

[11] Due to lower frequency in weather data available to us, we could not use weather in our instrument variables.

Seasonal ARMA process $(1,1)x(1,1)_{24}$.[12]Therefore, the regression equation of the residual may be written as follow.

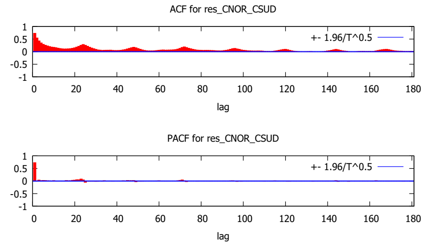

Figure 3.1: ACF and PACF of $CNOR - CSUD$

$$z_t = \phi z_{t-1} + \Phi z_{t-24} + \phi\Phi z_{t-25} + \theta\omega_{t-1} + \Theta\omega_{t-24} + \theta\Theta\omega_{t-25} + \omega_t$$

Where,

- ϕ, Φ, θ and Θ are coefficients of autoregressive (AR), moving average (MA), sesonal autoregressive and seasonal moving average respectively.
- $z_{t-k}$is the residual at lag k.
- ω_{t-k} is the error at lag k

The error term, ω_t, is estimated using generalized autoregressive conditional heteroskedasticity (GARCH), a well-known method proposed by Bollersev (1986) that allows us to model the variance as an autoregression process. The method is widely used in many electricity price modeling literature (see for instance Knittel and Roberts (2005); Garcia et al.(2005); Diongue et al. (2009)). Hence, the error term is formulated as follow,

$$\omega_t = \sigma_t Wt$$

$$\sigma_t^2 = \gamma_0 + \beta\sigma_{(t-1)}^2 + \gamma_1\omega_t^2$$

Where, σ_t is the variance at time t, Wt is a white noise, and γ_0 is a constant.

2SLS Results

We follow the intuition from Woo et al.(2011) analysis of the Texas electricity market as our basis in the expected value. In their paper, rising demand in non-west zones (less wind resource) and high wind output from the west increases price difference. Hence, in the case of congestion to Zone1 ($y > 0$), the results in importing zones are expected to display positive values on the coefficients of demand whereas negative values on the renewable and hydro production. Opposite signs are expected in the exporting zones for the respective variables. Our hypotheses are proven in our estimation displayed in Table 3.4. The results can be summarized as follow:

1) In importing zone (ZONE1):

    - Increasing demand increases the congestion cost (positive in all pairs);
    - Demand from the importing zones of SARD-CSUD and SICI-SUD have the highest sensitivity as they increase the congestion cost for 0.0687 € and 0.027 € per MWh increase of the value respectively;

[12]It is important to be noted that this model is used in order to capture trader adjustment in the previous hour and day. It is also noteworthy that we have estimated several time-series model with more parameter. However, increasing the parameters affects the goodness of fit on the model and may overfit the regression.

- Larger renewable production decreases the congestion cost towards zero (with an exception in CSUD-SUD);
- Renewable production from Importing zones of SARD-CSUD and SICI-SUD have the highest sensitivity as they decrease the congestion cost for 0.102 € and 0.032 € per MWh increase of the value respectively;
- Larger hydro production decreases the congestion cost towards zero (with exceptions in CNOR-CSUD and SARD-CSUD);
- Hydro production from importing zones of SICI-SUD and CNOR-NORD have the highest sensitivity as they decrease the congestion cost for 0.15 € andd 0.013 € per MWh increase of the value respectively;

2) In exporting zone (ZONE2):

- Increasing demand decrease the congestion cost toward zero (with exceptions in CSUD-SUD and CNOR-NORD);
- Demand from the importing zones of SICI-CSUD and CNOR-CSUD have the highest sensitivity as they decrease the congestion cost for 0.01 € and 0.008€ per MWh increase of the value respectively;
- Larger renewable production increases the congestion cost (with an exception in CNOR-NORD);
- Renewable production from Importing zones of CNOR-CSUD and SICI-SUD have the highest sensitivity as they increase the congestion cost for 0.011 € and 0.007 € per MWh increase of the value respectively;
- Larger hydro production increases the congestion cost (with exceptions in CNOR-NORD and SARD-CSUD);
- Hydro production from importing zones of CSUD-SUD and SICI-SUD have the highest sensitivity as they increase the congestion cost for 0.017 € and 0.0031 € per MWh increase of the value respectively;

In comparison to Woo et al. (2011), an identical mechanism can be found in our estimations. Lower demand or higher renewable supply in the importing zone impact the changes of congestion cost towards zero. However, the opposite impacts on congestion cost may occur if a larger production is recorded or a lower demand is displayed in the exporting zone. This phenomenon is due to the merit order impact in the zonal price. High growth of renewable energy supply in the importing zone shift the supply curve and reduce the zonal price. Consequently, the congestion cost becomes closer to zero as the cost to balance the system is reduced. This is in line with Sapio (2015) founding on the impact of renewable between Sicily and South. Our finding confirms that the conclusion holds for all the pairs in Italy's electricity system. Our econometric results also reveal that further growth of renewable energy from importing zone may result in zero congestion costs or negative congestion costs. This is due to the decrease of importing needs in the importing zone that negate the saturation on the transmission line thus creating zero congestion cost. Then, a much bigger positive shock of renewable from importing zone may create congestion to the opposite direction (towards exporting zone). In this case, the congestion cost will be negative because the reduction of zonal price caused by renewable supply. On the other hand, different behavior is shown in the exporting zones. Positive changes or shock in renewable will result in excessive efficient supply and creation of a new zonal market, importing zone. As a consequence, low-efficiency units are called in the importing zone for balancing the system and the congestion cost (ICC) increases.

In the case of congestion from Zone 1 ($y < 0$), we may expect the same sign. In this case, rising renewable or lower demand in the importing zone becomes the source of congestion as large efficient supply needs to be transferred to its neighbor. Our expectations are in line with the 2SLS estimation since we observe that:

1) In importing zone (ZONE1):

			CNOR_NORD		CNOR_CSUD		SARD_CSUD		CSUD_SUD		SICI_SUD	
			Coeff	p-value	Coeff	p-value	Coeff	p-value	Coeff	p-value	Coeff	p-value
Zone1		α	0.042661	***	0.050528	***	0.174638		0.150693	***	0.47986	***
	$y > 0$	R	-0.0148	***	-0.00463	***	-0.10336	***	0.001498	***	-0.03226	***
		H	-0.013	***	0.001533	***	0.070415	***	-0.00638	***	-0.15574	***
		D	0.010903	***	0.007397	***	0.069037	***	0.004242	***	0.028273	***
	$y < 0$	R	-0.01354	***	0.031111	***	-0.01945	***			-0.06607	***
		H	-0.02106	***	-0.03084	***	-0.0116	*			0.168236	***
		D	0.004322	***	0.008787	***	0.065962	***			0.074745	***
Zone2	$y > 0$	R	-0.00475	***	0.011325	***	0.00543	***	0.006255	***	0.007464	***
		H	-0.00257	***	0.003697	***	-0.02105	***	0.018045	***	0.002695	***
		D	0.001258	***	-0.00886	***	-0.0034	***	0.002983	***	-0.01032	***
	$y < 0$	R	0.002837	***	0.001909	***	-0.01299	***			0.003407	***
		H	0.000185	***	-0.02081	***	0.002049				0.017812	***
		D	-0.00118	***	-0.00956	***	-0.00829	***			-0.02764	***
Seasonal ARMA process		ϕ	-0.00827	***	0.640594	***	0.750439	***	0.023213	***	0.015098	***
		Φ	0.724949	***	0.014739	***	0.024473	***	0.659181	***	0.93424	***
		θ	0.225958	***	0.011568	*	-0.18209	***	0.37345	***	0.378338	***
		Θ	-0.61521	***	0.078017	***	0.183124	***	-0.55166	***	-0.78467	***
GARCH(1,1)		omega	0.010783	***	0.00904		0.192808	***	0.041412	*	1.8046	***
		alpha	0.025815	***	0.024806	**	0.052055	***	0.081793	***	0.154947	***
		beta	0.974178	***	0.976439	***	0.956378	***	0.920799	***	0.790978	***
		Log likelihood	-55094		-44346		-102634		-57532.9		-140830	
		BIC	110401.7		88905.67		205481.9		115215.5		281874.2	
		AIC	110228		88731.99		205308.2		115093.9		281700.5	
		Hannan-Quinn	110282.8		88786.74		205363		115132.2		281755.3	
		R2	0.718021		0.675919		0.526405		0.529795		0.50304	

*** P< 1% ** P< 5% * P< 10%

Table 3.4: 2sls result

- Increasing demand increases the congestion cost towards zero (positive in all pairs);
- Demand from the importing zones of SICI-CSUD and SARD-CSUD have the highest sensitivity as they increase the congestion cost for 0.075 € and 0.068 € per MWh increase of the value respectively;
- Larger renewable production decreases the congestion cost towards negative (with an exception in CNOR-SUD);
- Renewable production from Importing zones of SICI-SUD and SARD-CSUD have the highest sensitivity as they decrease the congestion cost for 0.065 € and 0.019 € per MWh increase of the value respectively;
- Larger hydro production decreases the congestion cost towards negative (with an exception in SICI-SUD);
- Hydro production from importing zones of CNOR-CSUD and CNOR-NORD have the highest sensitivity as they decrease the congestion cost for 0.03 € and 0.02 € per MWh increase of the value respectively;

2) In exporting zone (ZONE2):

- Increasing demand decrease the congestion cost toward negative (negative in all pairs);
- Demand from the importing zones of SICI-SUD and CNOR-CSUD have the highest sensitivity as they decrease the congestion cost for 0.028 € and 0.01€ per MWh increase of the value respectively;
- Larger renewable production increases the congestion cost towards zero (with an exception in SARD-CSUD);
- Renewable production from importing zones of SICI-SUD and CNOR-NORD have the highest sensitivity as they increase the congestion cost for 0.0029 € and 0.00186 € per MWh increase of the value respectively;
- Larger hydro production increases the congestion cost towards zero (with an exception in CNOR-CSUD);
- Hydro production from importing zones of SICI-SUD and SARD-CSUD have the highest sensitivity as they increase the congestion cost for 0.018 € and 0.0027 € per MWh increase of the value respectively.

Our results confirm that larger renewable supply in the importing zone reduces the congestion cost further towards negative value. Rising renewable supply in the importing zone decreases the congestion cost much further as the price equilibrium declines because of the increase in efficient supply. The same impact on the congestion cost will be shown if negative shock in demand occurs as it reduces the needs to import in this zone thus lowers the zonal price. However, it is noteworthy that lower renewable supply or high demand from Zone 1 in this regime (congestion from ZONE1) may switch congestion cost into no-congestion regime ($y = 0$) or congestion to ZONE1 regime ($y > 0$) since it reduces the congestion in the exit but increase the needs to import. In the exporting zones, higher renewable supply may also change the congestion direction (regime) as the coefficient present increasing effect on the congestion cost. The merit order effect from the renewable supply in the exporting zones decreases the cost to balance the system thus bringing the congestion cost closer to zero or switching the congestion's regime ($y = 0$ or $y > 0$). This is due to the fact that the increase of renewable in the exporting zone can satisfy the local demand which reduces the congestion from ZONE1 (congestion cost close to zero) or changes the congestion direction. As for demand, the higher quantity from the exporting zones will stimulate scarcity in supply which increases the congestion from ZONE1. Consequently, the congestion cost moves in the direction of a negative value as low efficient units have to be called in the exporting zones and increase the zonal price thus widening the zonal price difference.

We may sum up and generalize our finding as the table 3.5. Due to the merit order effect, rising renewable or hydro supply decreases the price equilibrium and subsequently push the congestion cost towards negative value. This conclusion is applied for supply from the importing zones (ZONE1), in both cases (congestion to ZONE1 and congestion from ZONE1) or regimes. We may observe the decrease in congestion from our illustration in figure 3.2 and 3.3.

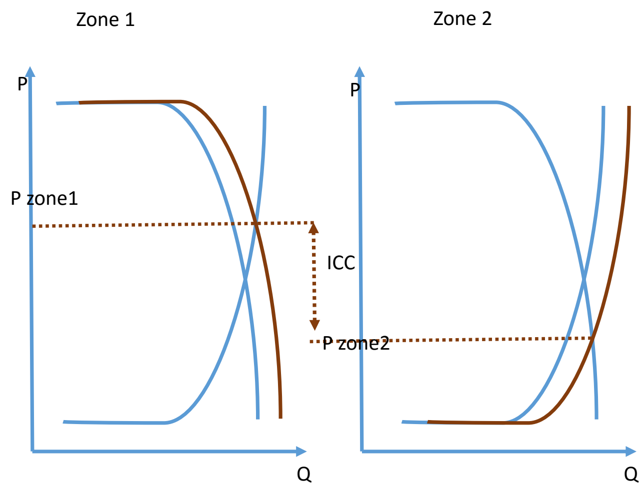

Figure 3.2: Market equilibrium without shock

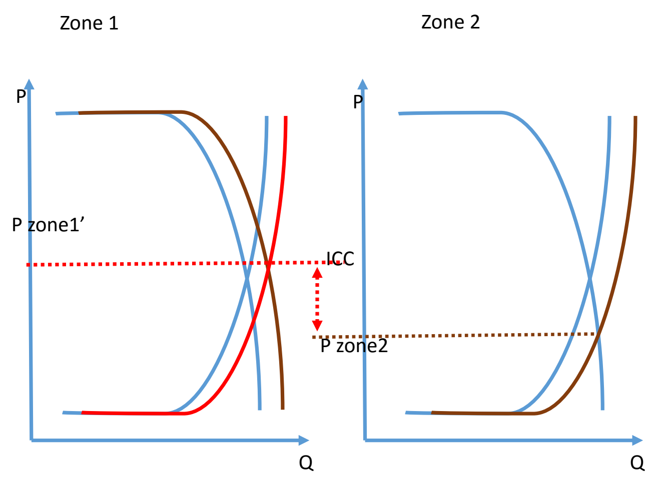

Figure 3.3: Market equilibrium with positive shock on the renewable energy supply from ZONE1

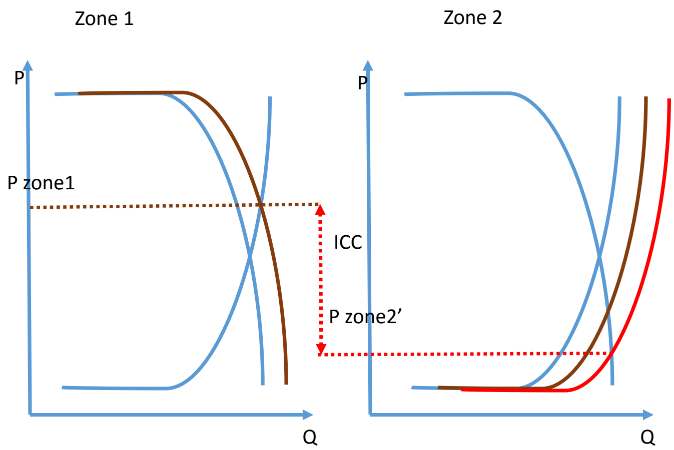

Figure 3.4: Market equilibrium with positive shock on the renewable energy supply from ZONE2

The introduction of renewable energy and hydro have shifted the supply function since there is a bigger quantity of low-cost electricity in ZONE1. However, it is important to be noted that continuous reduction of congestion cost from this regime ($y > 0$) will subsequently merge the zone (congestion cost $= 0$) since the transmission line is less saturated from the import/export. On the other hand, continuous reduction of congestion cost from the third regime ($y < 0$) will widen the price difference as congestion cost becomes more negative. As for the demand in the importing zone, their increse display the opposite impact as it push congestion cost towards positive value (increasing congestion cost). On their respective counter flow zones (ZONE2), opposite behaviors are observed for renewable and demand, since the merit order effect shift price equilibrium of ZONE2. As illustrated in figures 3.2 and 3.4, rising renewable and/or hydro will shift the congestion towards positive value as their increase will widen the gap of the price difference in the first regime and will take it closer to zero in the third regime. On the other hand, rising demand will shift the zonal price in the direction of negative value. Since less efficient supply needs to be transferred and the demand curve is shifted, the congestion cost decreases towards zero in the first regime and increases the negative value in the third regime. Overall, from the point of view of Terna, increase of renewable should be promoted in the importing zones as they tend to reduce the congestion cost or create a less saturated line ($ICC = 0$). However, excessive growth of the renewable from importing zone should be avoided since it may create a congestion to the opposite direction.

	D	**H**	**R**
Congestion to ZONE1 ($y > 0$)	ZONE1: ↑ ZONE2: ↓	ZONE1: ↓ ZONE2: ↑	ZONE1: ↓ ZONE2: ↑
Congestion from ZONE1 ($y < 0$)	ZONE1: ↑ ZONE2: ↓	ZONE1: ↓ ZONE2: ↑	ZONE1: ↓ ZONE2: ↑

Table 3.5: 2SLS result summary

3.3 Conclusion

Our empirical analysis has shown that demand and renewable supply have different impacts on the congestion occurrence and cost. The results of the multinomial logit model suggest that the effect of a larger local renewable supply is to decrease the probability of suffering congestion in entry and to increase the probability of causing a congestion in exit compared to no congestion case. Increasing hydroelectric production has a similar effect. A rise in local demand, on the contrary, increases the probability of congestion in the entry (due to larger import) and decreases the probability of congestion in the exit. This result holds for both importing and exporting regions. However, the importing regions are less likely to produce congestion in the exit. Therefore, a larger RES production in these regions is expected to bring more balance in the electricity flows between neighboring regions, while a larger RES production in exporting zones may exacerbate the problem of congestion. On the other hand, estimation on congestion cost indicates that rising renewable or hydro supply in the importing zone reduces the congestion cost toward zero in the first regime (congestion to) and widens the gap on the third regime (congestion from). The big shock of renewable or hydro quantity could switch the regime from the first to a merged zone (congestion cost $= 0$) or a third regime (congestion cost < 0) because of excessive supply in the exit of the importing zone. This is due to the merit order effect that decreases the zonal price and ,consequently, pushes congestion cost towards negative value. Finally, exact opposite impacts are displayed in the exporting zones.

Both of our estimations allow us to draw some policy recommendation.

- Additional incentive in the importing regions.
  The increase of renewable in the importing zones provides a more balanced system since it less likely to produce congestion in exit and reduces the odds for congestion in the entry. In addition, ICC could be reduced or dissipated as they shift the zonal price equilibrium towards negative value. Therefore, in the point of view of TSO and policy maker, further promotion of renewable growth in importing regions is recommended. In the current state, operators would prefer rising renewable in the exporting zones since they could profit from the high zonal price and congestion cost.
- Growth of intermittent supply should be controlled.
  Although it is true that larger renewable decreases the congestion cost and reduce the frequency, bigger shock may provide an opposite effect. Rising renewable increases the odds for congestion in exit regardless of the zones and the estimation in congestion cost validate this phenomenon as continuous increase may change the net flow direction (congestion cost < 0). Hence, excessive growth will worsen the congestion problems.
- Identical behavior will occur on the large scale.
  If a larger scale market (e.g Europe) is done under the same algorithm and bidding zone system, similar behavior should be seen. For instance, high demand in importing countries (zones) will stimulate exports of efficient supply from the neighboring countries (zones), thus increasing the odds for congestion in entry and increase its cost. However, the market would require well-organised transmission management and detailed research on bidding zones since several TSO are involved.

There are several directions that can be pursued in order to capture the full picture of the electricity market. It is important to be noticed that there are few econometric studies in this line of research, thus future extensive studies may be needed to obtain better views on RES and congestion. Our paper assumes competitive bids, which allow us to simplify the problem. Therefore, more research can be directed towards strategical bidding in the electricity market. With more renewable supply in the market, it is important to understand the renewable impact on the thermal units' bids.

Chapter 4

Interdependency of Italian zonal prices

Faddy Ardian
École Polytechnique

Abstract

Our paper captures and studies the interdependency of zonal prices in the Italian electricity market with an objective to gain greater insight on the pricing efficiency and the impact of the transmission line on the conditional correlation in the Italian electricity market. In order to achieve our objective, we collected six series of the Italian zonal prices from its independent market operator (GME) and estimate multivariate GARCH approach under conditional mean. We employ two multivariate GARCH model to capture the cross-correlation of the zonal prices, CCC-MGARCH from Bollerslev (1990) and DCC-MGARCH from Engle (2002). Our estimation gives us an additional insight on the integration of Italian electricity market and the interdependencies among its regional markets. The results suggest that high capacity physical exchange provides a strong interdependency among the connected zonal markets. Indeed, all the zones in Italian peninsula have shown strong dependencies among them. This argument is also validated by our finding in the analysis of new transmission installation where stronger dependencies are found after additional capacity is put into place despite a long transition period. In addition, the low transmission line between Sicily and Italian peninsula is the cause of weak interdependencies between Sicily region and the other regional markets.

Contents

4.1 Introduction

Electricity market has significantly changed after market deregulation. The importance of market and network efficiency in the new market was the priority of the policy maker. As a result, many countries adopt their best mechanism and structure with the aim to have electricity price that reflects their actual production and transmission cost. This motivation encourages countries to adopt inter-zonal pricing mechanism, a mechanism that divides a country into two or more regional markets connected with high transmission lines. The mechanism allows zonal markets to have higher or lower price, depending on the net import situation, whenever there is a saturation in the transmission line. The problem arises in the national market where they have limited capacities of transmission line that unites the zonal markets. Low physical limitation of the transmission line will trigger higher frequency of congestion, which is translated into higher occurance in zonal market splits. As a results, it could isolate regional markets that do not posses adequate network capacity with its neighbour, which make this region fails to be integrated nationally.

This particular angle of view can be investigated from the interdependency of the electricity market since it exhibits the correlation between regional price. Interdependencies in electricity prices shows a mutual linear relationship between zonal prices concluded in the market. The increase of one zonal price would predict an increase of price in the other zone and vice versa. Therefore, the zonal prices would vary together as they are statistically correlated. In Italian electricity market, we may interpret this interdependencies as a reflection of national integration. In other words, if two zonal prices display mutual dependency, it can be concluded that both regional prices are integrated in one zonal market. This is due to the fact that high correlation can only be obtained by having one single market without zonal splitting. Hence, strong dependency among zonal prices indicates a successful market integration whereas weak interdependency shows regional market isolation. If weak interdependencies are displayed among the zonal markets, we may question the system and the mechanism ability in fostering a complete national integration and efficient electricity market.

The interest in studying the relation between zonal prices have been displayed in the academic with many researchers attempt to investigate the mechanism and determinant of zonal price differences on two connected regions. Hauldrup and Nielsen (2006) have started the discussion by investigating the non-linear dynamic between two connected zonal prices in the Nordpool spot market. The paper aims to examine the regime switching and long memory process on the zonal prices using Markov Regime Switching Model. They conclude that the mechanism of the switch in congestion direction is a result of excess demand, which subsequently increase its zonal prices. Woo et al. (2011) study the zonal price difference with an objective to investigate the impact of rising wind supply in the west zone of Texas electricity market. Their result from log-OLS estimation suggests that rising load outside the West zone would increase the zonal price difference. Figuiredo et al. (2015) look into the variable that causes the market splits into two zonal prices in Iberian spot electricity price (Spain) by employing logit and non-parametric keynesian function. Their calculation suggests that large availability of baseline technology coupled with high renewable increases the frequency of triggering a market splitting mechanism. These literatures, unfortunately, does not provide us with an insight on the interdependencies between indirectly connected zones and mean spillover in the national zonal prices.

Research from Worthington et al (2005) initiate the discussion on the interdependencies on several zonal prices by employing Multivariate GARCH on the five Australian spot electricity market (NEM). In another continent, Park et al. (2006) investigate various US spot markets with Vector Autoregression and acyclic graph method. Their estimation indicates that the transmission lines and institutional arrangement affect the interdependencies in the zonal prices. Dempster et al. (2008) analyse the California electricity markets with Granger causality tests and cointegration analysis. Their study presents a moderate level of market integration and interdependencies between the regional markets in California. Unfortunately, their proposed method does not possses the capability to show the dynamic of conditional correlation in the markets. This research, on the other hand, requires a technique that could display the dynamic of conditional correlations against time since we attempt to observe the changes in conditional correlation after new investments in Italian electricity market. Worthingon et al. (2005) research is, then, extended, by Higgs (2009) for further investigation in NEM using data from 1 January 2006 to 31 December 2007. In her paper, she applies three different multivariate GARCH (MGARCH) model, Conditional Constant Correlation-MGARCH, Tse and Tsui's (2002) and Engle's (2002) Dynamic Conditional Correlation-MGARCH. Her research is aimed to investigate the inter-relationship between the four zonal prices of NEM. She concludes that regional markets with better networking infrastrcucture displays strong interdependency whereas weaker level of interdependencies are recorded in the markets with low capacity of transmission line. Unfortunately, her proposed model still lacks one important characteristic of electricity prices, seasonality. Seasonality is an important feature that presents in the electricity price. Therefore, it is important to be addressed in the future research. Ignatieva and Trück (2011) have focused on the structural dependencies in the Australian electricity market using GARCH model coupled with copulae method.[1] Their research shows that significant tail dependence between the zonal prices in Australia. Hence, price spikes may happen jointly across the regional market.

We attempt to extend Higgs (2009) by capturing interdependence of zonal prices in different electricity markets, Italian power exchange. To the best of our knowledge, Sapio (2015) and Ardian et al (2015) are the only two literature related to our research for the case of Italian electricity market. Their paper, however, focuses solely on capturing the impact of renewable price to zonal price difference (congestion cost) and congestion between two connecting zones. Both papers reveal the same insight in the market splitting mechanism. A positive shock on renewable energy supply on the exporting zones display increasing impact on the zonal price differences in comparison to zonal price difference under no shocks. The same effect on the zonal price differences also displayed whenever there is an increase in demand in the importing zones. Their results explain the relations between connected regions. Our paper, on the other hand, aims at studying the inter-relationship and mean spillover among Italian zonal prices. Hence, we want to shed a light into the correlation between indirectly connected zones. The result will provide us with greater insight on the pricing efficiency and the impact of the transmission line on the interdependencies in the Italian electricity market.

In order to achieve our objective, we collect six series of the Italian zonal prices from its independent market operator (GME) and estimate multivariate GARCH approach under conditional mean. The estimation starts by computing the coefficients of the univariate Seasonal ARMA and GARCH, which produces conditional mean and variance respectively. Our model try to address Higgs (2009) suggestion by adding seasonality, both deterministic trend and stochastic. This is due to the fact that seasonality effects from the weekly pattern, hourly pattern, and trader correction are important characteristics that determine electricity price. In addition, the result is also expected to provide us with insight on the mean spillovers. The second estimation is, then, calculated with the initial parameters from the univariate Seasonal ARMA-GARCH. In this stage, we employ two multivariate GARCH model to capture the cross-correlation of the zonal prices, CCC-MGARCH from Bollerslev (1990) and DCC-MGARCH from Engle (2002). The two models allow us to examine the cross-correlation between regional markets and to investigates the efficiency of the market integration. It is noteworthy that the method is also able to analyse the impact of the New submarine installation between Sardinia and Italian peninsula in 2011 on the conditional correlation. This method will initiate discussion on the impact of new transmission installation on the conditional correlation in the academic since this is the only case study that has an increase in transmission capacity in the sample period.

Our paper attempts to contribute to the academic literature in several ways. Firstly, the paper attempts to initiate the discussion on interdependency and mean spillover in the

[1] The proposed method, however, is not applied in this research since risk management are not the main objective of this research. We aim to investigate the national market integration in the Italian electricity market by evaluating interdependencies and dynamic correlation of the zonal prices. We are able to analyse this effect by employing CCC and DCC MGARCH without coupling it with copulae method.

Italian electricity market. Secondly, we would like to contribute to the limited literature on the cross-correlation of electricity zonal prices. Thirdly, seasonality is part of the concern in Higgs (2009) and we are trying to address it in this paper by adding seasonal trend and process on the model. Finally, we are trying to analyse the impact of new transmission installation on the dynamic conditional correlation, which has not been done in any related literature.

Our estimations deliver an additional insight on the integration of the Italian electricity market and the interdependencies among its regional market.The results indicate that high capacity physical exchange provides strong interdependency among the connected zonal markets. Indeed, all the zones in Italian peninsula have shown strong dependencies among them as it will be shown in our CCC and DCC estimations. The result is in line with Higgs (2009) where strong interdepency is shown in the well-connected markets. Moreover, this argument is also validated by our finding in the analysis of new transmission installation where stronger dependencies are found after additional capacity is put into place despite a long transition period. In addition, the low transmission line between Sicily and Italian peninsula is the main cause of weak interdependencies between Sicily and the other regional markets.

We organise our paper as follow. The next section is used to provide a brief description on the Italian power exchange (GME). The third section is aimed to explain our dataset and its statistical summary. The fourth section, demonstrate, our model specification and the multivariate garch. The fifth section focuses on the analysis of our estimation. The sixth section summarises our paper and suggests further researche directions.

4.2 Model Specification

We capture the interdependency of zonal electricity prices by the means of volatility modelling on Italian electricity markets using multivariate GARCH model under conditional mean. Estimation and simulation of time-varying volatility in electricity prices are challenging tasks. This is due to the fact that it is important to take into account several influential factors in specifying the model. Firstly, conditional mean needs to be considered since it reflects the characteristics of the commodity price. Therefore, stylised features of the electricity prices need to be added to the model, such as seasonality and mean reversion. This factor is also suggested by Higgs (2009) for future investigation since her model did not consider seasonality. Secondly, volatility clustering and heteroskedasticity are often displayed in electricity prices, especially in the high-frequency dataset like ours. This factor needs to be addressed appropriately using the ARCH model since it can capture the changes in the variance. Finally, modelling volatility of six zonal prices needs to be approached with a multivariate model that is able to simulate the interaction between them. Therefore, we apply multivariate GARCH model that are estimated in two stages on the six zonal prices series of Italian electricity market (CNOR, CSUD, NORD, SARD, SICI, and SUD).

4.2.1 First Regression

The first step in our estimation is a regression of univariate models that capture the characteristics of the price and the market. We initiate the model by a general formula that follows Lucia and Schwartz (2002). The general formula consist of two important features, deterministic ,$G_h(t)$, and stochastic, $F_h(t)$.

$$P_{i,t} = G_{i,t} + F_{i,t} \tag{4.1}$$

The deterministic components aims at capturing the seasonal pattern, with an objective in obtaining stationary series[2], and to capture mean spillover by the other zonal prices. Therefore, we formulate our model as follows

[2]

$$G_{i,t} = \omega_0 + \sum_{h=1}^{23} \omega_{h,1} + \sum_{d=1}^{7} \omega_{d,2} + \sum_{j=1}^{5} \omega_{j,3} P_{j,t-1}$$
$$i \neq j$$

Where, ω_0 is a constant, $\omega_{h,1}$ is a coefficient of hourly seasonality, $\omega_{d,2}$ is a coefficient of daily seasonality and $\omega_{j,3}$ is a coefficient of the zonal prices j at the $t-1$, $P_{j,t-1}$.

In the stochastic part, $F_h(t)$, we aim to model the mean reversion and the seasonal stochastic feature of the electricity prices. Mean reversion is one of the well-known features in electricity price as it was pointed by previous literature (Weron, 2014; Weron and Misiorek, 2005; Cuaresma et al., 2004;). This is due to the fact that electricity price tends to revert back to its normal value after positive or negative shock. This feature motivates us to add autoregressive features in our model. On the other hand, market participant correction in the market needs to be reflected in order to simulate the zonal price appropriately. We tackle this issue by integrating seasonal stochastic process since traders correct their trade using seasonal historical price. Consequently, seasonal autoregressive and the moving average process (SARMA) is proposed for simulating the conditional mean of zonal prices. The motivation for this approach is supported by numerous related literature in electricity price modeling (see for instance Contreras et al., 2003; Petrella and Sapio, 2007). Therefore, we can estimate $F_{i,t}$ by the regression equation below.

$$F_{i,t} = P_{i,t} - G_{i,t}$$

$$F_{i,t} = \phi F_{i,t-1} + \Phi F_{i,t-24} + \phi\Phi F_{i,t-25} + \theta\epsilon_{i,t-1} + \Theta\epsilon_{i,t-24} + \theta\Theta\epsilon_{i,t-25} + \epsilon_{i,t}$$

Where, the residual $\epsilon_{i,t}$ is assumed to have zero mean and normal distribution in this process.

In order to model the volatility clustering generally displayed in the series we simultaneously estimate generalized autoregressive conditional heteroskedasticity (GARCH), a well-known method proposed by Bollersev (1986). This approach is widely used in many electricity price modeling literature (see for instance Knittel and Roberts, 2005; Garcia et al., 2005; Diongue et al., 2009). However, it is important to be noted that the changes in the transmission capacity could create structural break in the series. Therefore, we apply Bai and Peron (2003) test in our sample dataset in 2011. The results of the test are shown in the table 4.1 below.

Zone	Date	F statistic	Critical value
CNOR	18-Aug-11	177.43	29.24
CSUD	18-Aug-11	177.81	29.24
NORD	17-Aug-11	173.564	29.24
SARD	11-May-11	92.95	29.24
SICI	16-May-11	44.456	29.24
SUD	09-Aug-11	165.65	29.24

Table 4.1: Structural break result from Bai-Peron

The result indicates several different dates of structural break in the zonal prices. Since our study focuses on the changes in interdependencies after the new tranmssion installation between SARD and CSUD, we use 11 May 2011 as the date of the structural break. As a consequence, the residual is computed with the equation below.

$$\epsilon_{i,t} = h_{i,t}^{1/2} \zeta_t$$

$$h_{i,t} = \begin{cases} \gamma_0 + \gamma_1 h_{(t-1)} + \gamma_2 \epsilon_{(t-1)}^2 & 0 < t < t_{break} \\ \gamma_0 + \gamma_3 h_{(t-1)} + \gamma_4 \epsilon_{(t-1)}^2 & t_{break} < t < T \end{cases}$$

Where, σ_t is the variance at time t, ζ_t is a white noise, γ_0 is a constant,t_{break} is the starting date of the structural break and T is the end date of our dataset. The estimation of univariate GARCH for each market is used for the second stage of our estimation aimed to build conditional covariance matrix. The first GARCH estimations, $0 < t < t_{break}$, is used for the first DCC estimation and the second DCC is estimated by using parameters from $t_{break} < t < T$.

4.2.2 Second Regression

In the last stage of our estimation, we attempt to simulate time-varying volatility by applying two MGARCH family models, dynamic conditional correlation (DCC) and constant conditional correlation (CCC), aimed at gaining insight into the interdependence of zonal prices. The discussion on the multivariate GARCH method is initiated by Bollerslev (1990) when he proposed CCC Model. Following his paper, we define conditional covariance matrix decomposed into conditional standard deviations and a correlation matrix

$$H_t = D_t P_t D_t$$

P_t denotes NxN matrix of conditional correlation and D_t denotes a diagonalNxN matrix as shown in the equation below

$$D_t = diag(h_{i,t}^{1/2}...h_{k,t}^{1/2})$$

Hence, the off-diagonal element in H_t is:

$$H_{ij}, t = h_{i,t}^{1/2} h_{j,t}^{1/2} \rho_{ij,t}$$

Where, $h_{i,t}$is the conditional variance from univariate GARCH in the market i and $\rho_{ij,t}$ is correlation between market i and j. The CCC-MGarch offers computational simplification since the conditional correlation does not change against time, $\rho_{ij,t} = \rho_{ij}$. Hence, temporal variation of the covariance matrix only depends on the conditional variance from GARCH estimation. The useful feature has attracted researcher to propose this method for empirical research (Lien and Tse, 2002 ;Higgs, 2009).

Unfortunately, CCC assumption limits researchers and explorations on economic phenomena. This limitation restricts analysis in the dynamic of the interdependence in the variable. Consequently, Engle (2002) proposes an alternative method, a Dynamic Conditional Correlation (DCC), with an assumption of time-varying conditional correlation. This improvement allows us to obtain more insight into the cross-correlations between variables but adds a computational burden in the calculation. The method follows the same general equation with different specification of correlation matrix P_t, which is specified as follow:

$$P_t = diag(Q_t)^{-1/2} diag(Q_t)^{-1/2}$$

where

$$Qt = (1 - \alpha - \beta)\bar{Q} + \alpha\zeta_{t-1}\zeta_{t-1}^T + \beta Q_{t-1}$$
$$\bar{Q} = Cov[\zeta_t, \zeta_t^T] = E[\zeta_t, \zeta_t^T]$$

$\bar{Q}$ is the unconditional covariance matrix of the standardized errors ζ_t which is defined as

$$\zeta_t = D_t^{-1}\epsilon_t \sim N(O, P_t)$$

The model has restrictions in α and β in order to obtain positive definite H_t and positive conditional variance from the univariate GARCH. The restrictions are

$$\alpha \geq 0, \ \ \beta \geq 0, \ \ \alpha + \beta < 1$$

We apply this method into two different sub-samples of our dataset, prior to the structural break and after the structural break. We attempt to look into the changes after the new installations of SARD-CSUD transmission line.

4.3 Results

The results of the first stage estimation are displayed in table 4.2 below. Let us first look into the impact of the lagged price of a zone to the other zonal prices. The regression estimation on these variable suggests the presence of mean spillover in the zonal markets. Indeed, if we look at the coefficients of each zone, we find that most of the lagged variable display significant and positive values. Therefore, the expected price increase in one market Granger cause the expected price increase in the other markets. For instance, if the price of the NORD increases for 1 €/MWh thus the price in CNOR increase for 0.32 €/MWh in the next hour. There are several main factors that cause this phenomenon.

- Technical limitation of the power plants.
  Big physical power plants generally have a technical limitation in terms of the shut down frequency and the minimum number of the operational hour. As a consequence, market participants will bid for electricity delivery in a block of hours (several period of deliveries), which affect the prices of the neighbours at time t.
- The connectivity plays a significant role in the short-run price change.
  One can argue that due to the efficient connectivity we observe positive coefficients. However, we can observe that SUD, SICI and SARD display lower lagged variable effect compared to others. This is mainly caused by the limitation on the physical transfer which reduces the effect of spillover from these zones. As a consequence, the impact of price changes in the Northen Italy (NORD, CNOR and CSUD) on the rest of Italy (SARD, SUD, and SICI) are relatively small (less than 0.01 or negative) since the transmission capacities in the direction of South (SUD and SICI) and Sardinia (SARD) is small. The same reasoning is also applied in the opposite case where the price change in the SARD, SICI and SUD have little effect on the Northern regions (NORD, CNOR and CSUD).
- Inter-zonal mechanism affects greatly the impact of lagged prices.
  High capacity transmission between NORD and CNOR, as well as CNOR and CSUD, is able to avoid congestion and transfer extra supply towards the demand. Subsequently, it makes these regions alligned with the same zonal prices in most of the time since there is no saturation in the line. This is the reason why the linear correlation are always positive with relatively high values of coefficient since there is a high mean spillover between them. The opposite case is found on the SICI, SARD and SUD as they have small coefficients on their estimation since they often have price discrepancy with their respective neighbours.

Estimations of the mean reversions, ϕ, reflect the speed of the reversion from lagged value. Hence, a high coefficient value indicates a faster reversion to its normal value. NORD displays the highest value with 0.748 thus sudden price changes will quickly revert back its normal price. The same interpretation also applies on the SARD and SUD as they display second and third highest mean reversion, 0.745 and 0.738 respectively. On the other hand, SICI's mean reversion coefficient value shows that its price deviation will come back to its normal value in a slow manner. In other words, SICI tends to sustain high price at t whenever there is a sudden increase in zonal price at $t-1$. The estimation of seasonal mean reversion, Φ , quantify the speed of the reversion from the deviation of the price of the day before and reflect the trader correction. As in mean reversion coefficient, ϕ, the value of seasonal mean reversion can be interpreted with the same idea. SICI register the biggest parameter, 0.9, making it the fastest region that corrects the deviation from the day before. CSUD follows SICI from behind with 0.71. NORD, SUD and SARD display the lowest value of seasonal mean reversion with 0.0628, 0.0619 and 0.049 respectively.

Our ARCH and GARCH estimations, shown in table 4.2, suggest that all the zonal price display volatility clustering features in the period prior o the structural break and after the structural break. Therefore, volatility, h, and own-innovation, ϵ, at $t-1$ affect its future volatility in both periods. From the GARCH estimation in the first period, CNOR and NORD show the biggest impact of previous volatility with 0.38 and 0.35. SUD follows closely from behind by registering 0.34 on its GARCH coefficient. SARD andCSUD exhibit positive and significant coefficients on GARCH. However, their values indicate a low impact on the future volatility as they are estimated at lower than 0.1. The ARCH estimation reveals that own-innovation from CSUD, SARD and SICI have big impacts on the future volatility of their respective zonal prices. This is due to the high ARCH coefficients, γ_2, on their estimation, which are higher than 0.80. NORD and CNOR, on the other hand, demonstrate the lowest impact on the future volatility with 0.29 and 0.30, respectively, on their ARCH coefficient. Second period estimation displays lower GARCH estimation on CNOR, SICI and SUD as the structural break lower the impact on the previous volatility. Then, lower ARCH caused by the structural break is shown in SUD, SARD, CSUD and CNOR in the second period of the dataset.

	CNOR		CSUD		NORD		SARD		SICI		SUD	
PCNOR			0.491541	***	0.009825		-0.05748	***	0.015099		-0.06878	***
PCSUD	0.025626	***			-0.02874	***	0.011429		-0.03265	***	0.019686	***
PNORD	0.327464	***	0.093241	***			0.090835	***	-0.0341	***	0.09064	***
PSARD	0.007013	***	0.014273	***	0.005985	***			0.00494	***	0.006543	***
PSICI	0.002211	*	0.002473	**	0.001355		0.011993	***			0.002444	**
PSUD	0.06691	***	0.197794	***	0.039075	***	0.04002	***	0.052632	***		
Conditional means												
ϕ	0.3253	***	-0.31489	***	0.748868	***	0.745917	***	0.023672	***	0.738032	***
Φ	0.114956	***	0.71317	***	0.06285	***	0.049239	***	0.904726	***	0.061991	***
θ	0.276785	***	0.467608	***	0.244835	***	0.078789	***	0.484849	***	0.218	***
Θ	0.30705	***	-0.2238	***	0.371117	***	0.280554	***	-0.65096	***	0.296311	***
First Conditional variance												
ω_0	20.6137	***	1.39257	***	16.7178	***	2.03282	***	9.27276	***	20.1134	***
GARCH 1 (ω_1)	0.383626	***	0.094377	***	0.353906	***	0.058408	***	0.215844	***	0.349138	***
ARCH 1 (ω_2)	0.305479	***	0.891263	***	0.298632	***	0.938118	***	0.809564	***	0.354465	***
Second conditional variance												
GARCH 2 (ω_3)	0.357969	***	0.124831	***	0.339046	***	0.084193	***	0.085176	***	0.3467	***
ARCH 2 (ω_4)	0.278085	***	0.866538	***	0.315056	***	0.912195	***	0.891141	***	0.250223	***

Table 4.2: Univariate GARCH estimation

Our second stage estimation is shown on the table 4.3 below. All our Constant Conditional correlation (CCC) estimation, ρ, display positive value and are proven to be statistically significant at < 0.01. From the estimation, NORD-CNOR shows the highest conditional correlation indicating high interdependency in both zonal prices. Strong interdependencies in these zones are due to the adequate capacity of transmission lines that connect them The argument is in line with the empirical result on the other directly connected zones that have a high capacity of transmission lines, CSUD-CNOR and SUD-CSUD, with conditional correlation more than 0.67 in both periods. SARD-CSUD display moderate interdependencies with 0.44 on its conditional correlation in the first period, which is increased to 0.53 in the second period. This is due to the new transmission line that increases the interdependencies on the two zonal prices. This is also supported by the increase of correlations in the other pairs after structural break. Then, SUD-SICI display the lowest conditional correlation among the connected zones (0.12 and 0.119) because the limited capacity of its transmission line.

Interdependencies on the indirectly connected zone present an interesting discussion. Let us look into the conditional correlation of the zones in the Italian peninsula in the first period. [3] NORD-CSUD presents the highest conditional correlation among the paired indirectly connected zones, 0.71, indicating strong interdependencies between these zones. SUD-CNOR and SUD-NORD follow closely from behind, 0.69 and 0.66. The estimation suggests a strong interdependency among these zones as they are supported with high capacity of transmission lines. On the other hand, the interrelationship of SICI with its indirect neighbour display very low conditional correlation ranging from 0.1 to 0.13. The results suggest that SICI's market is isolated with the other markets. Therefore, its zonal pricess does not show high correlation with any zone. This is mainly caused by the very limited capacity of transmission line between SICI and the Italian peninsula. Then, SARD display moderate interdependencies as the coefficients are estimated between 0.42 and 0.48. The only exception is SICI-SARD whose coefficient is lower than 0.12 since SICI has a very low transmission line.

The impact of the new transmission line can be observe directly from the changes in CCC in table 4.3 below. The empirical result has proven that the increase of physical transmission capacity strengthen the interdependencies. The changes in interdependencies are shown in most of the pair in particular for SARD. SUD-SARD shows the highest changes as the correlation increases by 0.11. Then, SARD-CSUD and SARD-CNOR display the second highest improvement as the correlation increase to 0.53 and 0.55, respectively. The structural break also exhibit an improvement in the other pairs as well. SUD-CSUD and SUD-CNOR show significant changes in the interdependencies, 0.053 and 0.065 changes respectively. However, several zonal pairs displays exception as they show lower correlation in the second period.

[3] Please refer to figure **??** for the zonal division

	CCC 1		CCC 2		DCC1		DCC 2	
CSUD_CNOR	0.785397	***	0.794408	***	0.784586	***	0.810558	***
NORD_CNOR	0.908767	***	0.837273	***	0.946085	***	0.902364	***
SARD_CNOR	0.48572	***	0.557691	***	0.177683	***	0.70234	***
SICI_CNOR	0.094045	***	0.096581	***	-0.01444	***	0.015194	***
SUD_CNOR	0.690332	***	0.755512	***	0.816438	***	0.8035	***
NORD_CSUD	0.714098	***	0.662326	***	0.730249	***	0.737493	***
SARD_CSUD	0.446989	***	0.539909	***	0.152454	***	0.638898	***
SICI_CSUD	0.130005	***	0.133581	***	0.160371	***	0.0151	***
SUD_CSUD	0.670865	***	0.724118	***	0.704213	***	0.727829	***
SARD_NORD	0.4486	***	0.478045	***	0.169668	***	0.64892	***
SICI_NORD	0.101711	***	0.088098	***	-0.04355	***	0.009297	***
SUD_NORD	0.664459	***	0.67724	***	0.805864	***	0.748716	***
SICI_SARD	0.110038	***	0.069965	***	-0.04452	***	0.031816	***
SUD_SARD	0.428742	***	0.543065	***	0.133402	***	0.652813	***
SUD_SICI	0.127215	***	0.119323	***	-0.04472	***	0.127143	***
dcc alpha					0.048322	***	0.025505	***
dcc beta					0.926742	***	0.961651	***
AIC	519542.2		1417980		450968.7		1230823	
SIC	519708.8		1418435		451140.1		1231291	

Table 4.3: Estimation of CCC and DCC

Let us now look into the DCC estimation shown in Table 4.3 as well as Figure 4.1 and 4.2. In Table 4.3, we can observe the second stage estimation at time t, which includes goodness-of-fit, constant parameters, and conditional correlations. The estimation displays a better goodness-of-fit compared to CCC thus suggesting DCC as the better model for multivariate price simulation. As for the conditional correlations, their estimation at time t, are positive and significant. The coefficients exhibit the same insight with the CCC. Strong interdependencies are shown among the zones in Italian peninsula whose large capacity of transmission lines whereas weak interdependencies are shown on the correlation of any zone with SICI whose small capacity for physical exchange. However, these coefficients are the conditional correlation at time t thus they only display very few information. We need to look into the changes in conditional correlation, ρ, against time presented in Figure 4.1 and 4.3 in order to gain more information.

Let us look into the conditional correlations among zones in Italian peninsula shown in the figure 4.1below. Frequent absence of congestion should be expected in Italian peninsula as they posses high capacity of the transmission lines. Hence, DCC values close to 1 are expected through all the dataset. The changes of conditional correlation in CNOR-CSUD provide us with a good example of the dynamics in the directly connected zone. The conditional correlation shows strong interdependency at almost all t with values close to 1. This is in line with the previous finding that shows the relation between a high transmission line and strong interdependency. However, there are several periods where the regions can be considered as isolated since the values are close to or below zero. This is due to the saturation in the transmission that occurs infrequently between CNOR and CSUD which subsequently affects the market splitting between the two zones. The same insight can be observed on the other directly connected zones, CNOR-NORD and CSUD-SUD. Overall, the conditional correlation frequently displays strong interdependencies with values between 0.7 to 0.9 which indicates integrated markets with high capacity of physical transport in these zones. As for the indirectly connected zones, NORD-SUD, CNOR-SUD, and CSUD-NORD have displayed strong interdependencies in most of the sample period with values that are often close to 1. This is in line with our previous finding on CCC that shows strong interdependencies among the zones in Italian Penninsula. However, the DCC presents frequent fluctuation where they are often shown to reach below 0.5. This is due to the market split whenever extra efficient supply are flooded in the NORD or SUD thus creating two or more regional markets in the Italian peninsula. The impact of new transmission line on these zones are clearly shown in the CNOR-CSUD, CSUD-NORD CNOR-SUD as they become less volatile after the structural break. The result indicate smoother physical exchange between the zones as they are able to export extra supply to SARD. Other paired zones does not show a significant changes in the conditional correlations.

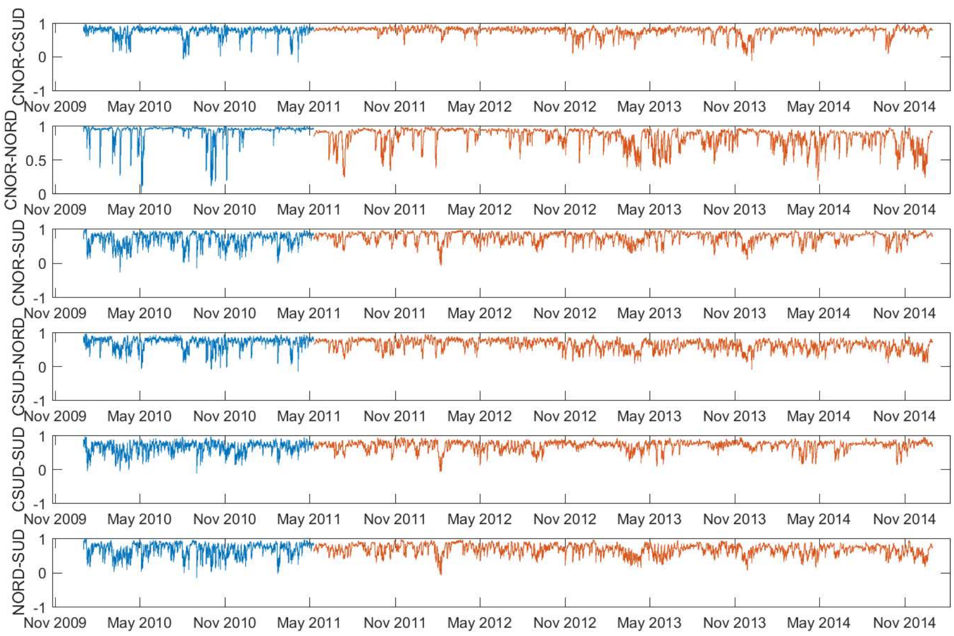

Figure 4.1: Conditional Correlations in Italian penninsula

Interesting insight can be obtained in the conditional correlation of SARD-CSUD where significant shift can be observed in the middle of our sample period. This result is important for analyzing the impact of a new transmission line which is never exploited in the previous literature. From our estimations, the conditional correlations of CSUD and SARD are fluctuating between 0 to 1 prior to August 2012 with the value frequently below 0.5. Indeed, the capacity limitation affects the interdependencies of SARD-CSUD with lower correlation thus making SARD an isolated regional market between these periods. However, it is important to note that a new installation starts to operate since March 2011 with the structural break in May 2011. Therefore, the result indicates that new capacity installation does not exhibit instant impact on the interdependencies. Hence, there is a relatively long transition period of the new installation in the zonal price behaviors. On the contrary, beyond August 2012, the interdependency start to display a stronger momentum from the new installation with correlation values that are constantly close to 1. As a result, SARD becomes more integrated into the Italian electricity market. This is mainly due to the submarine line's capability to mitigate the congestion in the physical exchange between SARD and Italian peninsula which creates frequent zonal unions. On the other hand, the impact of the new transmission line can also be seen again in the DCC of CNOR-SARD, NORD-SARD, and SARD-SUD. CNOR-SARD displays similar behavior on the interdependencies as CSUD-SARD with bigger values of conditional correlation and less frequent market isolation (Conditional Correlation < 0.5). NORD-SARD and SARD-CSUD, on the contrary, do not display a significant change in the conditional correlation although we can observe a sudden shift between August 2012 and November 2012.

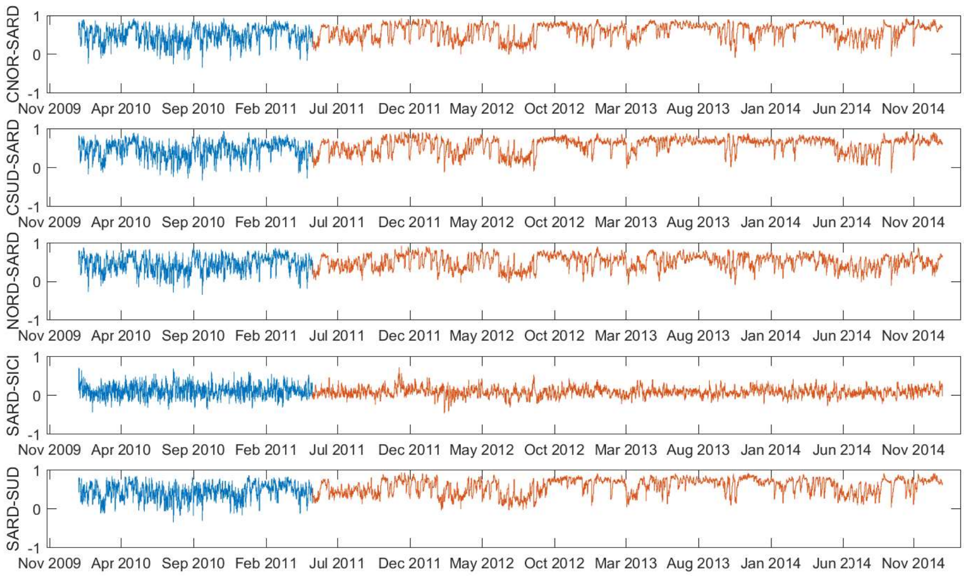

Figure 4.2: Conditional Correlations between SARD and other zonal prices

Finally, our estimations exhibit an empirical proof of market isolation in Sicily. The limitation to transport electricity between SICI and Italian peninsula are reflected in our results. Conditional correlation in SICI-SUD shows a contrast in comparison to the previous DCC plots of directly connected regions where values never reach 0.6 in all sample periods and they are frequently lower than 0.2. Frequent saturation in the transmission lines is reflected in their conditional correlations as they are often to register values lower than 0.5.This result conclude a tendency of regional market isolation thus indicating non-integration in the national market. This is mainly due to the low capacity of physical exchange between them and the rise of renewable supply in the SUD[4] which subsequently impact market splitting and interdependency. In all cases, conditional correlation between SICI and other

[4]See Sapio(2005) and Ardian et al (2015) for further reading.

regional market do not show strong interdependencies as the values generally lower than 0.2. The important changes in the physical capacity between SARD and Italian peninsula do not affect its interdependencies with SICI as its continue to display the same dynamic after the structural break.

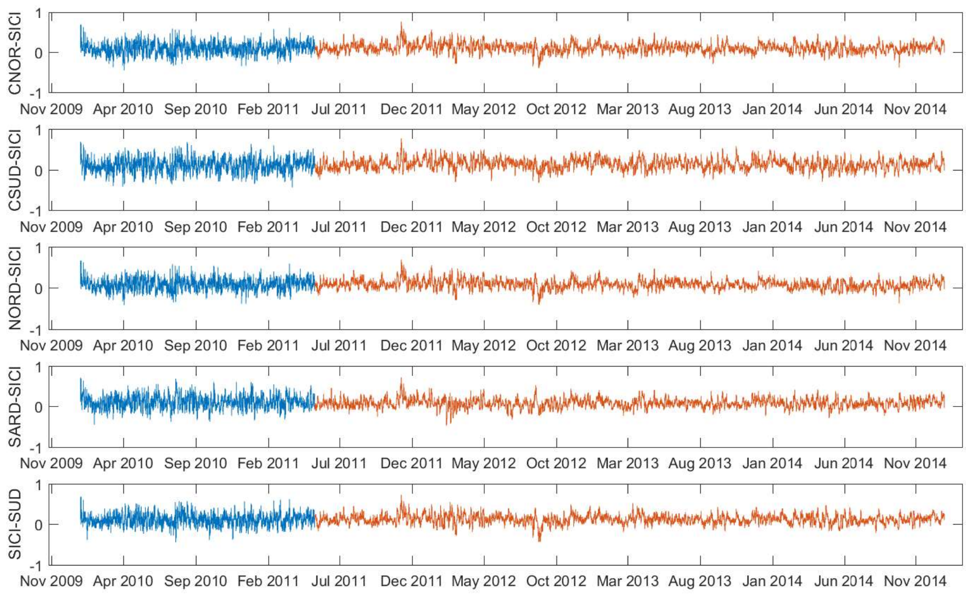

Figure 4.3: Conditional Correlations between SICI and other zonal prices

4.4 Conclusion

Our study has given us a wider view of the zonal price behaviour in Italy. Firstly, the first stage estimation has concluded a positive and significant mean spillover among the zonal prices. Hence, it can be concluded that the lagged value of one zonal price can cause positive changes to the other expected zonal price. This is mainly due to technical factor from physical operations, the connectivity between zonal prices and the interzonal mechanism that affects the Italian zonal prices at time t. Secondly, the seasonal stochastic process from our computation provides an insight that has not been examined in the previous literature. This result shows different speeds of zonal price reversion from the day before, as a reflection of the market participant correction. From our estimation, the coefficients display positive and significant value between 0.05 and 0.9 with SICI records the fastest reversion whereas SARD records the slowest reversion. Thirdly, univariate GARCH estimation shows that the previous volatility and own-innovation impacts future volatility. Furthermore, structural break is shown to lower ARCH effect in the residual. Fourthly, strong dependencies are shown among regional markets in the Italian peninsula. The result indicates a strong capability of fostering market integration in Italian peninsula from the Italian government. It is noteworthy that the result is mainly due to the high capacity of physical limitation that are shared between these zones. As a consequence, SICI ,whose transmission line is very low compared to other zones, has shown very weak interdependencies with other zonal prices. Finally, our sample period provides a special case since we are able to see the impact of new transmission line between SARD and Italian peninsula on the conditional correlation. Our finding exhibits an increase of constant conditional correlation in most of the cases after the structural break. Then, improvements in DCC are shown in CNOR-CSUD, CSUD-NORD CNOR-SUD as they become less volatile. If we look at conditional correlations between SARD and other zonal markets, we may observe long transitional period between the structural break and August 2012 when the conditional correlation starts to have constant high values. Nevertheless, higher capacity from the new installations increases the conditional correlations after the structural break which is shown in both DCC and CCC.

We believe that future research should examine the fat tail dependencies among Italian zonal prices which are important for risk analysis and management. This analysis can be explored by employing copulae method. In addition, the studies can be expanded by examining the dependencies with another commodity such as coal or gas price. Another direction of future studies would be examining interrelations between neighbouring countries with Italian zonal prices which can open a discussion regarding the price efficiency in the market coupling. Finally, studies on strategical bidding between the zonal price could give us a better understanding in this particular field.

Chapter 5

Conclusion

5.1 Main conclusions

Electricity market liberalization in Italy has raised interesting issues for economic and econometric researches because of its specifities and characterics . The changes in pricing mechanism has greatly impacted the price as it has become more volatile due to the dependenc in the demand availability in supply. Furthermore, the increase in renewable energy supply has added more uncertainty in energy supply thus affecting the volatility of electricity price. Consequently, market participants and policy makers require reliable and accurate forecasting models for benchmarking fair prices or optimising economic gain.The model should not only be accurate in predicting the price, but also reflects the market and be able to be interepreted in economic sense.

The second chapter of our this dissertation attempts to address this issue by providing discussion on the hourly price forecasting in the Italian electricity market with an aim to provide an alternative model for the Italian electricity market. We argue that the Seasonal ARMA-GARCH model with stacked framework is the best univariate model for the Italian electricity market. This is due to the fact that this model displays low mean absolute error and low bayesian information criterion (BIC). Our empirical analysis also suggests that this model provides better interpretation of the market since it is able to show the speeds of daily and weekly adjustment. In addition, stacked framework shows the daily dynamic of hourly prices that reflect the 24-simulatenous auction of day-ahead market. The results also suggest positive impacts from gas price and demand on the wholesale price as CCGT is generally their marginal technology.

The analysis serves as an important example of the significance of modeling and forecasting electricity price in the economic sense that reflect the markets. Many techniques have been documented and proposed in the literature. Advanced knowledge and technology have offered various modeling methods to forecast electricity price. However, our second chapter exhibits an important lesson as we show improvement in forecasting performance when we integrate the mechanism and the dynamic of the market in the model. Stacked framework with seasonal ARMA-GARCH is a specific method to address the characteristics of Italian electricity market. Furthermore, utilization of gas price as the exogenous variable is also aimed to reflect Italian production mix that is dominated by CCGT. This analysis is an important note for researchers in commodity market since the fundamental driver is the main determinant of price movement. Hence, integrating mathematical representation of market mechanism and its dynamic in our model would result in better forecasting performance.

Our exploration in the second chapter initiates discussion on the difference between univariate and multivariate framework. The empirical findings reveal that each framework has its own merit. Hence, framework selection would depend on the modeling's objective. This is due to the fact that VAR , a multivariate model, shows superior capability on predicting the price by recording the lowest mean absolute error. The univariate framework, on the other hand, shows better performance in the risk management as it shows low expected shortfall and the maximum error value. In addition, the analysis indicates that multivariate framework is more efficient to be used for forecasting electricity price on the off-peak periods whereas univariate framework exhibits lower error in the peak periods. These findings provide researchers, market participant and policy makers with alternative models that can be used effectively for forecasting and modeling the Italian electricity market.

EU directive number 28 in 2009 has brought significant change on EU renewable energy penetration, which subsequently create an additional problem in the electricity market. Intermittent generation in the day-ahead market has been shown to produce a merit order effect that reduces the average wholesale price. In addition, the uncertainty from the weather has increased volatility. However, rising renewable supply also give additional problem in both the market and the system. Renewable energy supply creates saturation in the interconnector between zonal market. Consequently, congestion becomes more frequent thus affecting congestion cost and zonal market splitting occurrence.

Our third chapter displays the empirical exhibits and analysis on the impact of the renewable energy supply on congestion. Our econometric estimations suggest that the flow of congestion in the inter-zonal transmission depends on the local supply and demand. Increasing local renewable energy supply decrease the probability of congestion in the entry and increase the probability of congestion in the exit. The same mechanism and logic also apply to the changes in Hydroelectric generation. This is due to the fact that they are efficient generations with priority dispatching thus their positive shock would increase the urgency for exporting electricity or decrease the imports. On the other hand, positive shock in demand increases the probability of congestion in the entry and decrease the congestion in the exit. In this case, rising local consumption increases the demand to import electricity from its neighbor. These results hold for both importing and exporting regions. However, the importing regions are less likely to produce congestion in the exit. Therefore, a larger RES production in these regions is expected to bring more balanced in the electricity flows between neighboring regions, while a larger RES production in exporting zones may exacerbate the problem of congestion. In terms of cost, positive shock from renewable and hydro supply would reduce the congestion cost in importing regions as they relieve the congestion towards the region. This is due to the merit order effect that decreases the zonal price in the importing zone and decreases the congestion cost. The opposite mechanism is applied on the exporting zone. It is also important to note that, from our analysis, big shock of renewable or hydro generation in the importing region could change the congestion flow and produce negative congestion cost as efficient supply become excessive. In this case, the importing region becomes an exporting region. Hence, the extra measure needs to be taken into account by policy maker in order to improve renewable penetration without increasing congestion cost and congestion frequency.

From our empirical result in chapter three, we may recommend several policies for reducing the impact of renewable energy supply on congestion. Firstly, an additional incentive in the importing regions would results in a system with a better balance. The increase of renewable in the importing zones provides a more balanced system since it less likely to produce congestion in exit and reduces the odds for congestion in the entry. In addition, ICC could be reduced or dissipated as they shift the zonal price equilibrium towards negative value. Therefore, in the point of view of TSO and policy maker, further promotion of renewable growth in importing regions is recommended. In the current state, operators would prefer rising renewable in the exporting zones since they could profit from the high zonal price and congestion cost. Secondly, if national transmissions are not equipped to facilitate intermittent generation, it is important to control the Growth of intermittent. Although it is true that larger renewable decreases the congestion cost and reduce the frequency, bigger shock may provide an opposite effect. Rising renewable increases the odds for congestion in exit regardless of the zones and the estimation in congestion cost validate this phenomenon as continuous increase may change the net flow direction (congestion cost < 0). Hence, excessive growth will worsen the congestion problems. Thirdly, we also believe that our empirical analysis in Italian electricity market could be replicated on a larger scale. For instance, high demand in importing countries (zones) will stimulate exports of efficient supply from the neighboring countries (zones), thus increasing the odds for congestion in entry and increase its cost. However, the larger market would require well-organised transmission management and detailed research on bidding zones since several national TSO are involved. Nevertheless, an internal energy market in Europe is the long term objective of the European commission and we believe that our proposed measures are noteworthy for the future EU energy market.

Our final angle in the dissertation has offered an interesting view on the regional integration of the Italian electricity market. The deregulation of electricity market has shown the different mechanism of price formation. In Italy, policy makers have decided to adopt interzonal pricing mechanism in order to reflect regional supply price and the cost of transmission. However, inequalities in transmission capacity could create market isolation in the national electricity market. Italy has exhibited differences in the level of transmission capacity, which could create regional market isolation in the Italian electricity market. This particular angle can be investigated from the interdependency of the Italian electricity market since it exhibits the dependency between regional price. Strong dependency indicates a successful market integration whereas weak interdependency shows regional market isolation.

The fourth chapter provides important factors that need to be taken into account in the future Italian policy by exhibiting empirical proof and analysis on the market isolation in Sicily and the success of market integration in the Italian Peninsula . Our econometric result displays high conditional correlation in the regions in Italian peninsula due to high capacity that facilitates the physical exchange between regions. On the other hand, physical exchange limitation between Sicily and the Italian peninsula have proven to produce low correlations. These results serve as an argument to improve the mechanism or the physical network in the Italian electricity market. The increase of transmission limit between Sicily and Italian peninsula could improve the integration of this region into the national electricity market. This is due to the fact that the changes in transmission capacity would mitigate the congestion which subsequently avoid market splitting that isolates Sicily most of the time. This argument is also supported by the changes in the interdependencies between Sardinia and the Italian peninsula. The new transmission installation that connects the island and Italy has proven to strengthen interdependencies in the zonal prices as the physical exchange increase between these regions.

From the empirical analysis in the fourth chapter, we may also draw recommendations for a larger scale market. Ongoing discussion on single European market has always been an interesting issue. Our empirical findings would suggest a probable market isolation if there is a limitation in the transmission capacity. Therefore, it is important to ensure adequacy of network infrastructure before implementing a single European market. This is important because, in larger scale, the system becomes more complex with many alternative and efficient supply.

5.2 Future research direction

There are several directions that can be pursued in order to obtain the full picture of the electricity market. Firstly, future study should address utilization of weather as an exogenous variable. The current study, unfortunately, is limited by the availability of weather data. Historical weather data is generally limited to daily data and it would create a bias if it is used for research in higher frequency data. Consequently, the economic phenomenon that exists in hourly data cannot be investigated from the weather perspective. Furthermore, additional research needs to be done in the time dimension, particularly in lagged value of the weather. This is due to the fact that sudden shock in weather does not produce sudden impact in demand since end user generally adjusts their consumption after a period of time. In addition, traders would have lagged time to adjust their trade based on the changes in weather. For instance, a sudden increase in rainfall is generally not equal to a sudden increase in hydro supply because traders adjust hydro outputs according to the water level and the electricity price. However, this rule does not apply to solar and wind generation units since they are not able to store their energy. Secondly, econometric studies on the relation between renewable energy supply and congestion should be further explored. Climate change issue and changes in the renewable energy policy have stimulated investments in renewable energy. Consequently, it increases the congestion frequency and market splitting. However, there is only very limited economic research in this field and they only display several angles in this research. Engineering papers are dominating this study with an attempt to provide studies on the integration of renewable energy. However, they lack economic views such as our research. Therefore, additional research may reveal more information and plan better policy for both electricity market and system. Thirdly, we realize that our research based on the assumption of competitive bids in the market. In other words, we did not take into account strategical bidding between zones or strategical bidding across the different markets. With six different zonal prices, the market participant is able to strategically allocate and bids of their assets in different locations in order to optimize their gain in all zonal markets. Moreover, our research only investigates the day-ahead market of the Italian electricity markets. In practice, market participants also use intraday and balancing markets in order to maximize their assets. This is an important feature in the market since intraday and balancing market may provide a better price for the market operator. As a result, power generators may choose the best market that can fully maximize the economic value of their production units. For example, now, CCGT operators more frequently bids in the intraday and balancing market since day-ahead market becomes unattractive to them. Finally, risk management is an important activity for all market participant and we believe that future research is needed in this direction. Our studies are limited in the research objective. Hence, risk studies are not the main output in our investigations. However, frequent extreme prices displayed in the electricity market requires more attentions. For instance, the Italian electricity market is exposed with fat tail dependencies among its zonal prices. Therefore, exploration using copulae method could give additional insight and explanations on the risk management.

Appendix A

Appendix for chapter 2

A.1 Forecast results

	Hour1	Hour2	Hour3	Hour4	Hour5	Hour6	Hour7	Hour8
M1								
BIC	8411.308	8411.308	8411.308	8411.308	8411.308	8411.308	8411.308	8411.308
MAE (in-sample)	4.9279	5.2568	5.5584	5.7508	5.8089	5.7753	5.2205	5.4895
MAE (out-of-sample)	3.416	3.4288	4.683	4.8221	5.0985	4.6141	4.7266	4.1263
M2								
BIC	8396.34	8501.146	8733.651	8851.349	8861.997	8782.323	8684.077	8978.267
MAE (in-sample)	4.8795	5.205	5.4657	5.6779	5.7416	5.6013	5.1725	5.3802
MAE (out-of-sample)	3.4059	3.4579	4.6713	4.7531	5.0839	4.5939	4.6577	4.1712
M3								
BIC	8388.553	8492.083	8727.192	8842.891	8857.297	8775.51	8698.757	8938.132
MAE (in-sample)	4.8392	5.1603	5.4187	5.6426	5.7057	5.5787	5.1753	5.3592
MAE (out-of-sample)	3.433	3.4559	4.6159	4.8286	5.054	4.5555	4.4972	4.3256
M4								
BIC	8380.351	8470.329	8696.956	8812.609	8832.425	8764.297	8834.864	9131.9
MAE (in-sample)	4.8203	5.1438	5.3815	5.5997	5.6671	5.5511	5.403	5.7708
MAE (out-of-sample)	3.4559	3.3966	4.5242	4.851	4.9405	4.4447	4.8487	5.5019
M5								
BIC	8416.262	8520.774	8739.51	8846.297	8864.175	8847.751	8906.893	9406.212
MAE (in-sample)	4.9659	5.2472	5.5226	5.7192	5.7751	5.7753	5.6708	6.1486
MAE (out-of-sample)	3.4461	3.3081	4.4504	4.7126	4.854	4.2224	4.6832	4.9995
M6								
BIC	8393.911	8490.901	8726.182	8841.102	8859.717	8779.957	8736.61	9025.253
MAE (in-sample)	4.8145	5.121	5.3945	5.6009	5.6619	5.5548	5.1459	5.4124
MAE (out-of-sample)	3.5094	3.4832	4.4901	4.8962	5.0006	4.6293	4.6789	4.0627
M7								
BIC	8598.204	8809.958	8982.353	9084.996	9112.197	9046.492	8995.161	9588.289
MAE (in-sample)	4.8259	5.1438	5.396	5.6213	5.6887	5.5621	5.3768	5.8583
MAE (out-of-sample)	3.4612	3.3993	4.5609	4.9135	4.9956	4.4729	5.1019	5.0292

Table A.1: Forecasting result on 1st to 8th period

	Hour9	Hour10	Hour11	Hour12	Hour13	Hour14	Hour15	Hour16
M1								
BIC	8411.308	8411.308	8411.308	8411.308	8411.308	8411.308	8411.308	8411.308
MAE (in-sample)	6.7585	7.3841	7.3838	7.4956	5.9495	6.2519	6.6523	6.472
MAE (out-of-sample)	4.7301	4.3443	3.5524	3.3371	3.2621	3.0755	3.6657	4.1746
M2								
BIC	9549.103	9834.505	9752.448	9739.481	8896.633	9065.51	9311.276	9269.355
MAE (in-sample)	6.662	7.3121	7.3037	7.3881	5.7838	6.0826	6.4466	6.3506
MAE (out-of-sample)	4.7822	4.4693	3.6313	3.3288	3.1683	3.0404	3.7337	4.2401
M3								
BIC	9533.988	9838.132	9793.764	9737.981	8834.841	8949.153	9232.027	9227.992
MAE (in-sample)	6.6606	7.2939	7.3984	7.3284	5.6664	5.8328	6.2993	6.1918
MAE (out-of-sample)	4.4965	4.461	3.7058	3.2803	2.887	2.7687	3.7687	4.2278
M4								
BIC	9849.266	10080.19	9940.808	9918.79	8987.145	9191.847	9497.057	9480.464
MAE (in-sample)	7.3074	7.8498	7.6798	7.7247	5.9531	6.1913	6.6911	6.6674
MAE (out-of-sample)	5.5879	5.4039	4.5449	4.1408	3.6828	3.8395	4.8052	5.1006
M5								
BIC	9975.081	10183.99	10068.7	10017.22	9155.342	9333.87	9635.067	9619.255
MAE (in-sample)	7.6088	8.1227	8.012	7.957	6.1431	6.4312	6.9708	6.9333
MAE (out-of-sample)	5.716	6.1732	5.4391	5.1624	4.5806	4.9627	5.6953	5.7502
M6								
BIC	9602.651	9860.856	9762.633	9746.116	8885.592	9045.301	9303.369	9263.84
MAE (in-sample)	6.6879	7.2601	7.2187	7.2664	5.7092	5.9472	6.3318	6.2115
MAE (out-of-sample)	4.6012	4.378	3.543	3.2194	2.9214	2.8116	3.669	4.1667
M7								
BIC	10295.53	10521.75	10499.5	10455.72	9423.042	9565.653	9971.96	10001.01
MAE (in-sample)	7.2922	7.8417	7.7255	7.702	5.9392	6.2315	6.7547	6.7037
MAE (out-of-sample)	5.664	5.5218	4.6956	4.3078	3.8873	4.0523	4.9488	5.1321

Table A.2: Forecasting result on 9th to 16th period

	Hour17	Hour18	Hour19	Hour20	Hour21	Hour22	Hour23	Hour24
M1								
BIC	8411.308	8411.308	8411.308	8411.308	8411.308	8411.308	8411.308	8411.308
MAE (in-sample)	6.4737	7.0276	6.7879	7.4485	6.7602	5.2721	3.8064	3.4689
MAE (out-of-sample)	4.0606	9.2639	6.9563	6.6839	3.5389	2.9199	3.3184	2.886
M2								
BIC	9372.812	9582.444	9611.772	9898.228	9715.193	8636.009	7467.906	7134.002
MAE (in-sample)	6.3912	6.9074	6.77	7.4552	6.7896	5.268	3.7885	3.4242
MAE (out-of-sample)	4.249	9.5809	7.0663	6.8396	3.5303	2.9261	3.2332	2.7402
M3								
BIC	9369.744	9581.097	9605.401	9881.723	9720.22	8632.822	7462.124	7109.708
MAE (in-sample)	6.3242	6.8457	6.685	7.3654	6.7569	5.2282	3.7438	3.387
MAE (out-of-sample)	4.2919	9.0793	6.7129	6.7551	3.511	2.9064	3.1352	2.7627
M4								
BIC	9586.467	9765.945	9727.143	9915.054	9691.309	8600.238	7454.997	7126.966
MAE (in-sample)	6.8152	7.3685	7.0428	7.471	6.7686	5.2272	3.7467	3.404
MAE (out-of-sample)	5.4931	10.421	7.7494	6.8573	3.3725	3.0637	3.0619	2.7226
M5								
BIC	9706.419	9892.436	9810.504	10033.36	9896.546	8772.186	7620.964	7219.131
MAE (in-sample)	7.0361	7.6582	7.3096	7.7472	7.0589	5.5218	3.903	3.5225
MAE (out-of-sample)	5.5027	8.83	7.0767	6.5986	3.7143	3.1274	3.0347	2.8443
M6								
BIC	9371.631	9594.386	9600.457	9875.33	9709.028	8598.116	7430.571	7087.639
MAE (in-sample)	6.2182	6.7897	6.5471	7.2645	6.6813	5.1458	3.6593	3.3222
MAE (out-of-sample)	4.2686	8.5656	6.5408	6.8308	3.4105	2.8017	3.023	2.702
M7								
BIC	10130.86	10465.86	10386.11	10410.76	10329.1	9197.5	7885.436	7473.098
MAE (in-sample)	6.7952	7.3518	7.0525	7.4674	6.7027	5.226	3.738	3.4114
MAE (out-of-sample)	5.43	9.7158	7.4493	6.7222	3.3893	3.0554	2.951	2.7183

Table A.3: Forecasting result on 17th to 24th period

A.2 M3 estimation

	phi	Phi	theta	Theta	omega_1	alpha	beta
Hour1	0.641568 ***	0.102711 ***	-0.15342 ***	0.063671 **	2.65274 ***	0.072348 ***	0.866814 ***
Hour2	0.642154 ***	0.088883 ***	-0.16791 ***	0.068512 **	1.27295 ***	0.078144 ***	0.896603 ***
Hour3	0.626113 ***	0.081732 ***	-0.12255 ***	0.077198 ***	1.97977 ***	0.080931 ***	0.882332 ***
Hour4	0.62376 ***	0.087539 ***	-0.09547 **	0.059306 **	1.69511 ***	0.069255 ***	0.900672 ***
Hour5	0.622211 ***	0.079631 ***	-0.09643 **	0.067549 **	1.81861 ***	0.077505 ***	0.890954 ***
Hour6	0.632129 ***	0.088922 ***	-0.18635 ***	0.112799 ***	1.10129 ***	0.066859 ***	0.913308 ***
Hour7	0.724544 ***	0.083136 ***	-0.4086 ***	0.066242 ***	32.3809 ***	0.343711 ***	0.123122 ***
Hour8	0.040774 ***	0.800193 ***	0.223117 ***	-0.55565 ***	27.7516 ***	0.500443 ***	0.19989 ***
Hour9	0.180545 ***	0.255667 ***	0.132739 ***	0.044897	202.415 ***	0.001187 ***	-1.00376 ***
Hour10	0.752227 ***	0.070141 ***	-0.45033 ***	0.125507 ***	13.1167 ***	0.176856 ***	0.719685 ***
Hour11	0.770884 ***	0.080149 ***	-0.45775 ***	0.099194 ***	9.91679 ***	0.206361 ***	0.723534 ***
Hour12	0.074748 ***	0.764559 ***	0.280324 ***	-0.5786 ***	6.78509 ***	0.186953 ***	0.768859 ***
Hour13	0.044524 ***	0.834871 ***	0.327231 ***	-0.62252 ***	3.00658 ***	0.159233 ***	0.804261 ***
Hour14	0.73215 ***	0.100008 ***	-0.36558 ***	0.122948 ***	3.77293 ***	0.124301 ***	0.826744 ***
Hour15	0.060867 ***	0.776337 ***	0.31204 ***	-0.54989 ***	6.14091 ***	0.157899 ***	0.775402 ***
Hour16	0.094481 ***	0.695049 ***	0.302427 ***	-0.45095 ***	9.93252 ***	0.162703 ***	0.719038 ***
Hour17	0.818221 ***	0.055772 ***	-0.47408 ***	0.114347 ***	11.7631 ***	0.209159 ***	0.671018 ***
Hour18	0.75156 ***	0.076739 ***	-0.38499 ***	0.134558 ***	5.35037 ***	0.163662 ***	0.797815 ***
Hour19	0.774949 ***	0.075853 ***	-0.39807 ***	0.098212 ***	5.22408 ***	0.169554 ***	0.800204 ***
Hour20	0.804452 ***	0.073194 ***	-0.49082 ***	0.067033 **	3.18499 ***	0.14812 ***	0.841285 ***
Hour21	0.965668 ***	0.009073	-0.74542 ***	-0.00975	30.0256 ***	0.607932 ***	0.274949 ***
Hour22	0.874934 ***	0.050866 ***	-0.53026 ***	0.03383	1.36023 ***	0.092496 ***	0.887343 ***
Hour23	0.095199 ***	0.785241 ***	0.338159 ***	-0.53594 ***	1.27575 ***	0.096943 ***	0.860152 ***
Hour24	0.741203 ***	0.101045 ***	-0.24521 ***	0.084386 ***	1.61832 ***	0.070736 ***	0.852713 ***

Table A.4: SARMA-GARCH Coefficient of M3

Appendix B

Appendix for chapter 3

B.1 Multinomial logit: Descriptive statistics and Tables

Variable	Mean	Std. Dev.
CNOR_Hydro	257.513	173.97
CNOR_Wind	7.934	10.38
CNOR_NRRes	353.457	320.536
CNOR_PV	1.203	2.733
CNOR_PV_Tot	354.66	322.328
CNOR_D	3524.293	879.753

Variable	Mean	Std. Dev.
SARD_Hydro	34.064	35.458
SARD_Wind	126.031	142.605
SARD_NRRes	54.567	76.789
SARD_PV	3.793	8.753
SARD_PV_Tot	58.36	84.416
SARD_D	1376.044	251.241

Variable	Mean	Std. Dev.
CSUD_Hydro	314.226	146.13
CSUD_Wind	209.619	183.134
CSUD_NRRes	344.889	347.385
CSUD_PV	17.476	34.705
CSUD_PV_Tot	362.365	378.943
CSUD_D	5310.994	1189.058

Variable	Mean	Std. Dev.
SICI_Hydro	11.008	12.143
SICI_Wind	260.465	218.838
SICI_NRRes	140.622	193.418
SICI_PV	1.94	4.398
SICI_PV_Tot	142.562	196.973
SICI_D	2203.172	410.19

Variable	Mean	Std. Dev.
NORD_Hydro	3137.343	1462.653
NORD_Wind	7.933	6.216
NORD_NRRes	2212.615	1133.996
NORD_PV	19.23	42.097
NORD_PV_Tot	2231.844	1168.411
NORD_D	18466.874	4238.214

Variable	Mean	Std. Dev.
SUD_Hydro	192.663	168.872
SUD_Wind	527.586	421.578
SUD_NRRes	443.762	564.546
SUD_PV	13.558	25.532
SUD_PV_Tot	457.32	587.877
SUD_D	2917.71	554.735

Variable	Description
Hydro	Hydroelectric production
Wind	Wind production
NRRes	RES production from Non Relevant Unit (Power<10 MVA)
PV	Photovoltaic production
PV_Tot	Large and small photovoltaic production
D	Demand

Table B.1: List of regressors and descriptive statistics, 2010-2014

Status in CNOR-NORD	Number	Per cent
Congestion from CNOR	1,992	5
No congestion	40,122	92
Congestion to CNOR	1,710	4
Total	43,824	100
Status in CNOR-CSUD	Obs	Per cent
Congestion from CNOR	620	1
No congestion	40,026	91
Congestion to CNOR	3,178	7
Total	43,824	100
Status in SARD-CSUD	Obs	Per cent
Congestion from SARD	469	1
No congestion	36,556	83
Congestion to SARD	6,799	16
Total	43,824	100
Status in CSUD-SUD	Obs	Per cent
No congestion	37,183	85
Congestion to CSUD	6,641	15
Total	43,824	100
Status in SICI-SUD	Obs	Per cent
Congestion from SICI	2,963	7
No congestion	8,060	18
Congestion to SICI	32,801	75
Total	43,824	100

Congestion from = first region has lower price
No congestion = Equal prices
Congestion to= first region has higher price

Table B.2: Network status in neigbouring regions, 2010-2014

Multinomial logit estimations are shown in Tables from B.3 to B.7. Standard errors are reported below the coefficients. The second and the third columns report the results in terms of log-odds and marginal effects respectively when the congestion is from ZONE1 ($y = -1$). The fourth and the fifth columns present the results in terms of log-odds and marginal effects when the congestion is to ZONE1 ($y = 1$).

Variables	Congestion from CNOR		Congestion to CNOR	
	Log-odds	mfx	Log-odds	mfx
CNOR_Hydro	0.00269***	2.27e-05***	-0.00143***	-1.48e-05***
	-0.00019	-2.14E-06	-0.00023	-2.38E-06
CNOR_Wind	0.0286***	0.000240***	-0.00615*	-6.50e-05*
	-0.00213	-2.43E-05	-0.00328	-3.34E-05
CNOR_PV_Tot	0.00313***	2.68e-05***	-0.00622***	-6.36e-05***
	-0.00017	-2.39E-06	-0.00047	-4.53E-06
CNOR_D	-0.00223***	-1.89e-05***	0.00295***	3.02e-05***
	-0.00012	-1.53E-06	-0.00011	-1.74E-06
NORD_Hydro	-0.000427***	-3.62e-06***	0.000508***	5.20e-06***
	-2.67E-05	-3.18E-07	-2.89E-05	-3.51E-07
NORD_Wind	0.0286***	0.000241***	-0.0139**	-0.000144***
	-0.00476	-4.30E-05	-0.00541	-5.55E-05
NORD_PV_Tot	0.000148***	1.16e-06***	0.000929***	9.44e-06***
	-5.34E-05	-4.47E-07	-0.00012	-1.18E-06
NORD_D	0.000408***	3.45e-06***	-0.000463***	-4.75e-06***
	-1.82E-05	-2.70E-07	-2.04E-05	-3.00E-07
Year2	-1.463***	-0.00867***	0.537***	0.00658***
	-0.345	-0.00121	-0.0883	-0.0013
Year3	0.289*	0.00268	-0.499***	-0.00444***
	-0.168	-0.00166	-0.145	-0.00112
Year4	-0.567***	-0.00414***	0.928***	0.0129***
	-0.168	-0.0011	-0.151	-0.0029
Year5	-0.743***	-0.00531***	1.935***	0.0397***
	-0.188	-0.00113	-0.175	-0.00678
Constant	-5.183***		-7.917***	
	-0.204		-0.19	
Observations	43,824	43,824	43,824	43,824
Log-Lik Intercept Only:	-31133.193		Log-Lik Full Model:	-20880.1
D(43798):	41760.267		LR(24):	20506.12
			Prob > LR:	0
McFadden's R2:	0.329		McFadden's Adj R2:	0.328
ML (Cox-Snell) R2:	0.374		Cragg-Uhler(Nagelkerke)R2:	0.493
Count R2:	0.812		Adj Count R2:	0.253
AIC:	0.954		AIC*n:	41812.27
BIC:	-426349.993		BIC':	-20249.6
BIC used by Stata:	42038.154		AIC used by Stata:	41812.27

*** p<0.01
** p<0.05
* p<0.1

Table B.3: Estimations for CNOR-NORD

Variables	Congestion from CNOR		Congestion to CNOR	
	Log-odds	mfx	Log-odds	mfx
CNOR_Hydro	0.00221***	1.13e-05***	0.000395*	1.83e-05*
	-0.00046	-2.43E-06	-0.00021	-9.82E-06
CNOR_Wind	0.0287***	0.000148***	-0.00225	-0.00012
	-0.00511	-2.73E-05	-0.00204	-9.77E-05
CNOR_PV_Tot	-0.00042	-2.10E-06	-0.00016	-7.29E-06
	-0.0012	-6.18E-06	-0.00025	-1.20E-05
CNOR_D	-0.00148***	-8.02e-06***	0.00160***	7.68e-05***
	-0.00016	-1.01E-06	-8.71E-05	-4.00E-06
CSUD_Hydro	-0.00199***	-1.05e-05***	0.000799***	3.87e-05***
	-0.00054	-2.84E-06	-0.00024	-1.16E-05
CSUD_Wind	-0.00479***	-2.56e-05***	0.00355***	0.000171***
	-0.00048	-2.52E-06	-0.00011	-5.18E-06
CSUD_PV_Tot	-0.00186*	-9.80e-06*	0.000900***	4.36e-05***
	-0.00099	-5.11E-06	-0.00021	-1.01E-05
CSUD_D	0.00144***	7.78e-06***	-0.00150***	-7.21e-05***
	-0.00012	-7.98E-07	-6.52E-05	-2.93E-06
Year2	-2.444***	-0.00759***	-0.636***	-0.0256***
	-0.248	-0.00062	-0.0724	-0.0025
Year3	-0.338***	-0.00154***	-0.183**	-0.00827***
	-0.13	-0.00057	-0.0711	-0.00309
Year4	0.0498	0.000375	-0.485***	-0.0205***
	-0.125	-0.00067	-0.0766	-0.00283
Year5	-0.660***	-0.00262***	-1.531***	-0.0518***
	-0.169	-0.00065	-0.0904	-0.00224
Constant	-5.299***		-1.382***	
	-0.261		-0.125	
Observations	43,824	43,824	43,824	43,824
Log-Lik Intercept Only:	-14607.4		Log-Lik Full Model:	-12640.4
D(43798):	25280.78		LR(24):	3934.012
			Prob > LR:	0
McFadden's R2:	0.135		McFadden's Adj R2:	0.133
ML (Cox-Snell) R2:	0.086		Cragg-Uhler(Nagelkerke)	R2: 0.176
Count R2:	0.913		Adj Count R2:	-0.008
AIC:	0.578		AIC*n:	25332.78
BIC:	-442829		BIC':	-3677.5
BIC used by Stata:	25558.66		AIC used by Stata:	25332.78

*** p<0.01
** p<0.05
* p<0.1

Table B.4: Estimations for CNOR-CSUD

Variables	Congestion from SARD		Congestion to SARD	
	Log-odds	mfx	Log-odds	mfx
SARD_Hydro	0.00309	1.50E-06	0.00215***	0.000121***
	-0.00268	-1.40E-06	-0.00057	-3.18E-05
SARD_Wind	0.00863***	4.66e-06***	-0.00999***	-0.000563***
	-0.0004	-9.63E-07	-0.00027	-1.39E-05
SARD_PV_Tot	-0.00379	-1.82E-06	-0.00308***	-0.000173***
	-0.00283	-1.48E-06	-0.00109	-6.10E-05
SARD_D	-0.00016	-2.04E-07	0.00411***	0.000231***
	-0.0004	-2.13E-07	-0.00012	-6.80E-06
CSUD_Hydro	-0.00819***	-4.16e-06***	0.000441***	2.51e-05***
	-0.00058	-8.90E-07	-0.00015	-8.48E-06
CSUD_Wind	-0.00100**	-5.04e-07**	-0.0001	-5.80E-06
	-0.0004	-2.31E-07	-0.00017	-9.31E-06
CSUD_PV_Tot	0.00296***	1.52e-06***	-0.000909***	-5.13e-05***
	-0.00066	-4.48E-07	-0.00024	-1.36E-05
CSUD_D	-0.000180**	-8.72e-08**	-0.000126***	-7.12e-06***
	-7.36E-05	-4.01E-08	-2.11E-05	-1.17E-06
Year2	-4.913***	-0.00133***	-0.775***	-0.0360***
	-0.44	-0.00026	-0.0445	-0.00183
Year3	-4.018***	-0.00108***	-1.897***	-0.0714***
	-0.208	-0.00022	-0.0655	-0.00223
Year4	-8.216***	-0.00254***	-1.873***	-0.0706***
	-0.647	-0.00042	-0.0753	-0.00226
Year5	-7.746***	-0.00235***	-0.745***	-0.0348***
	-0.654	-0.0004	-0.0629	-0.00254
Constant	-0.635		-5.095***	
	-0.433		-0.124	
Observations	43,824	43,824	43,824	43,824
Log-Lik Intercept Only:	-21426.2		Log-Lik Full Model:	-16034.4
D(43798):	32068.87		LR(24):	10783.6
			Prob > LR:	0
McFadden's R2:	0.252		McFadden's Adj R2:	0.25
ML (Cox-Snell) R2:	0.218		Cragg-Uhler(Nagelkerke) R2:	0.350
Count R2:	0.845		Adj Count R2:	0.065
AIC:	0.733		AIC*n:	32120.87
BIC:	-436041		BIC':	-10527.1
BIC used by Stata:	32346.76		AIC used by Stata:	32120.87

*** $p<0.01$
** $p<0.05$
* $p<0.1$

Table B.5: Estimations for SARD-CSUD

Variables	Congestion from CSUD		Congestion to CSUD	
	Log-odds	mfx	Log-odds	mfx
CSUD_Hydro			0.00143*** -0.00015	0.000103*** -1.05E-05
CSUD_Wind			0.00114*** -0.00019	8.23e-05*** -1.36E-05
CSUD_PV_Tot			0.000270** -0.00013	1.94e-05** -9.20E-06
CSUD_D			0.00128*** -3.32E-05	9.20e-05*** -2.43E-06
SUD_Hydro			0.000601*** -0.00012	4.33e-05*** -8.85E-06
SUD_Wind			0.000706*** -8.88E-05	5.09e-05*** -6.42E-06
SUD_PV_Tot			0.00126*** -7.93E-05	9.09e-05*** -5.79E-06
SUD_D			-0.00109*** -6.55E-05	-7.88e-05*** -4.79E-06
Year2			-0.558*** -0.0451	-0.0351*** -0.00251
Year3			-1.043*** -0.055	-0.0591*** -0.00248
Year4			-1.910*** -0.0697	-0.0925*** -0.00248
Year5			-1.405*** -0.077	-0.0742*** -0.00305
Constant			-6.927*** -0.119	
Observations	0	0	43,824	43,824

Log-Lik Intercept Only:	-18641.31	Log-Lik Full Model:	-14165.7
D(43811):	28331.303	LR(12):	8951.316
		Prob > LR:	0
McFadden's R2:	0.24	McFadden's Adj R2:	0.239
ML (Cox-Snell) R2:	0.185	Cragg-Uhler(Nagelkerke)	R2: 0.322
Count R2:	0.859	Adj Count R2:	0.067
AIC:	0.647	AIC*n:	28357.3
BIC:	-439917.9	BIC':	-8823.06
BIC used by Stata:	28470.246	AIC used by Stata:	28357.3

*** $p<0.01$
** $p<0.05$
* $p<0.1$

Table B.6: Estimations for CSUD-SUD

Variables	Congestion from SICI		Congestion to SICI	
	Log-odds	mfx	Log-odds	mfx
SICI_Hydro	0.0109*** -0.0019	0.000402*** -3.68E-05	-0.0118*** -1.31E-03	-0.00146*** -0.000136
SICI_Wind	0.00198*** -0.000126	0.000133*** -4.53E-06	-0.00570*** -9.30E-05	-0.000652*** -1.05E-05
SICI_PV_Tot	0.000965** -0.000484	0.000111*** -9.54E-06	-0.00553*** -2.83E-04	-0.000616*** -2.98E-05
SICI_D	-0.00169*** -0.000181	-0.000114*** -4.68E-06	0.00487*** -1.16E-04	0.000557*** -1.25E-05
SUD_Hydro	-0.00132*** -0.000193	-3.81e-05*** -3.64E-06	0.000793*** -1.16E-04	0.000108*** -1.20E-05
SUD_Wind	-0.000754*** -8.59E-05	-2.17e-05*** -1.71E-06	0.000447*** -5.08E-05	6.11e-05*** -5.34E-06
SUD_PV_Tot	-0.000643*** -0.000173	-3.00e-05*** -3.32E-06	0.00106*** -9.74E-05	0.000126*** -1.02E-05
SUD_D	0.000722*** -0.000127	2.03e-05*** -2.43E-06	-0.000397*** -7.89E-05	-5.52e-05*** -8.17E-06
Year2	0.0642 -0.0605	-0.00661*** -0.000968	0.503*** -0.0461	0.0479*** -0.00385
Year3	0.195*** -0.0747	-0.0205*** -0.000951	1.942*** -0.0552	0.142*** -0.003
Year4	-0.772*** -0.105	-0.0359*** -0.00135	3.014*** -0.0649	0.198*** -0.0032
Year5	-0.708*** -0.105	-0.0399*** -0.0015	3.700*** -0.0709	0.226*** -0.00346
Constant	0.333** -0.152		-7.747*** -0.118	
Observations	43,824	43,824	43,824	43,824
Log-Lik Intercept Only:	-31133.193		Log-Lik Full Model:	-20880.1
D(43798):	41760.267		LR(24):	20506.12
			Prob > LR:	0
McFadden's R2:	0.329		McFadden's Adj R2:	0.328
ML (Cox-Snell) R2:	0.374		Cragg-Uhler(Nagelkerke)R2:	0.493
Count R2:	0.812		Adj Count R2:	0.253
AIC:	0.954		AIC*n:	41812.27
BIC:	-426349.993		BIC':	-20249.6
BIC used by Stata:	42038.154		AIC used by Stata:	41812.27

*** p<0.01
** p<0.05
* p<0.1

Table B.7: Estimation for SICI-SUD

Appendix C

Appendix for chapter 4

C.1 Seasonal trend estimation

	coefficient	std.error	t-ratio	
Mon	62.7784	0.453689	138.4	***
Tue	63.0223	0.453689	138.9	***
Wed	63.2403	0.453689	139.4	***
Thu	63.2819	0.453893	139.4	***
Fri	62.5943	0.453689	138	***
Sat	59.9366	0.453689	132.1	***
Sun	54.1446	0.453689	119.3	***
h1	-2.9975	0.573913	-5.223	***
h2	-9.12352	0.573913	-15.9	***
h3	-13.2338	0.573913	-23.06	***
h4	-16.2046	0.573913	-28.24	***
h5	-16.591	0.573913	-28.91	***
h6	-12.9058	0.573913	-22.49	***
h7	-5.23837	0.573913	-9.127	***
h8	2.41809	0.573913	4.213	***
h9	10.0625	0.573913	17.53	***
h10	12.4075	0.573913	21.62	***
h11	9.70671	0.573913	16.91	***
h12	7.34135	0.573913	12.79	***
h13	-1.33789	0.573913	-2.331	**
h14	-4.54836	0.573913	-7.925	***
h15	-0.8577	0.573913	-1.494	
h16	2.61482	0.573913	4.556	***
h17	6.59304	0.573913	11.49	***
h18	13.139	0.573913	22.89	***
h19	16.6033	0.573913	28.93	***
h20	20.5117	0.573913	35.74	***
h21	19.449	0.573913	33.89	***
h22	12.9142	0.573913	22.5	***
h23	5.622	0.573913	9.796	***

Table C.1: CNOR estimation

	coefficient	std.error	t-ratio	
Mon	62.5651	0.467468	133.8	***
Tue	62.5143	0.467468	133.7	***
Wed	62.8968	0.467468	134.5	***
Thu	63.2362	0.467678	135.2	***
Fri	62.3087	0.467468	133.3	***
Sat	60.0145	0.467468	128.4	***
Sun	54.2544	0.467468	116.1	***
h1	-3.71806	0.591343	-6.287	***
h2	-10.4299	0.591343	-17.64	***
h3	-14.7268	0.591343	-24.9	***
h4	-17.537	0.591343	-29.66	***
h5	-17.849	0.591343	-30.18	***
h6	-14.256	0.591343	-24.11	***
h7	-5.92388	0.591343	-10.02	***
h8	1.95141	0.591343	3.3	***
h9	9.26223	0.591343	15.66	***
h10	11.5796	0.591343	19.58	***
h11	8.79195	0.591343	14.87	***
h12	6.12688	0.591343	10.36	***
h13	-1.85994	0.591343	-3.145	***
h14	-5.24834	0.591343	-8.875	***
h15	-1.93196	0.591343	-3.267	***
h16	1.67216	0.591343	2.828	***
h17	6.15486	0.591343	10.41	***
h18	13.3799	0.591343	22.63	***
h19	17.0103	0.591343	28.77	***
h20	21.465	0.591343	36.3	***
h21	20.6832	0.591343	34.98	***
h22	13.5968	0.591343	22.99	***
h23	6.0425	0.591343	10.22	***

Table C.2: CSUD Estimation

	coefficient	std.error	t-ratio	
Mon	62.6581	0.423314	148	***
Tue	62.7927	0.423314	148.3	***
Wed	63.1055	0.423314	149.1	***
Thu	63.203	0.423504	149.2	***
Fri	62.5052	0.423314	147.7	***
Sat	59.8067	0.423314	141.3	***
Sun	53.6588	0.423314	126.8	***
h1	-2.9527	0.535488	-5.514	***
h2	-8.92766	0.535488	-16.67	***
h3	-12.9608	0.535488	-24.2	***
h4	-15.9389	0.535488	-29.77	***
h5	-16.3128	0.535488	-30.46	***
h6	-12.6565	0.535488	-23.64	***
h7	-4.01588	0.535488	-7.499	***
h8	2.902	0.535488	5.419	***
h9	10.5285	0.535488	19.66	***
h10	12.7368	0.535488	23.79	***
h11	10.573	0.535488	19.74	***
h12	8.63868	0.535488	16.13	***
h13	0.267025	0.535488	0.4987	
h14	-1.92395	0.535488	-3.593	***
h15	0.966248	0.535488	1.804	*
h16	3.56692	0.535488	6.661	***
h17	7.08616	0.535488	13.23	***
h18	12.7051	0.535488	23.73	***
h19	15.368	0.535488	28.7	***
h20	17.7342	0.535488	33.12	***
h21	16.8203	0.535488	31.41	***
h22	11.5034	0.535488	21.48	***
h23	5.0632	0.535488	9.455	***

Table C.3: NORD Estimation

	coefficient	std.error	t-ratio	
Mon	71.9195	0.832086	86.43	***
Tue	73.2223	0.832086	88	***
Wed	73.0747	0.832086	87.82	***
Thu	74.1035	0.83246	89.02	***
Fri	72.1837	0.832086	86.75	***
Sat	67.8587	0.832086	81.55	***
Sun	61.4243	0.832086	73.82	***
h1	-6.72424	1.05258	-6.388	***
h2	-15.5588	1.05258	-14.78	***
h3	-20.6255	1.05258	-19.6	***
h4	-24.1576	1.05258	-22.95	***
h5	-24.4439	1.05258	-23.22	***
h6	-20.4564	1.05258	-19.43	***
h7	-12.6271	1.05258	-12	***
h8	-1.67619	1.05258	-1.592	
h9	7.9309	1.05258	7.535	***
h10	10.4894	1.05258	9.965	***
h11	5.55078	1.05258	5.273	***
h12	2.20739	1.05258	2.097	**
h13	-5.13559	1.05258	-4.879	***
h14	-10.068	1.05258	-9.565	***
h15	-7.48395	1.05258	-7.11	***
h16	-4.28341	1.05258	-4.069	***
h17	0.989007	1.05258	0.9396	
h18	9.93265	1.05258	9.436	***
h19	16.8603	1.05258	16.02	***
h20	24.742	1.05258	23.51	***
h21	27.84	1.05258	26.45	***
h22	17.8245	1.05258	16.93	***
h23	10.3872	1.05258	9.868	***

Table C.4: SARD Estimation

	coefficient	std.error	t-ratio	
Mon	85.062	0.862226	98.65	***
Tue	85.4619	0.862226	99.12	***
Wed	83.8214	0.862226	97.22	***
Thu	86.3074	0.862613	100.1	***
Fri	86.9568	0.862226	100.9	***
Sat	82.2043	0.862226	95.34	***
Sun	75.1473	0.862226	87.15	***
h1	-16.1779	1.09071	-14.83	***
h2	-24.9442	1.09071	-22.87	***
h3	-29.3631	1.09071	-26.92	***
h4	-32.2272	1.09071	-29.55	***
h5	-32.7816	1.09071	-30.06	***
h6	-30.8156	1.09071	-28.25	***
h7	-20.9995	1.09071	-19.25	***
h8	0.143777	1.09071	0.1318	
h9	20.5194	1.09071	18.81	***
h10	25.2177	1.09071	23.12	***
h11	20.5038	1.09071	18.8	***
h12	14.5048	1.09071	13.3	***
h13	5.29	1.09071	4.85	***
h14	-2.10056	1.09071	-1.926	*
h15	-2.89799	1.09071	-2.657	***
h16	0.823962	1.09071	0.7554	
h17	13.1216	1.09071	12.03	***
h18	31.0504	1.09071	28.47	***
h19	42.7392	1.09071	39.18	***
h20	54.6698	1.09071	50.12	***
h21	59.9782	1.09071	54.99	***
h22	45.7944	1.09071	41.99	***
h23	17.3649	1.09071	15.92	***

Table C.5: SICI Estimation

	coefficient	std.error	t-ratio	
Mon	61.871	0.452365	136.8	***
Tue	61.8864	0.452365	136.8	***
Wed	61.688	0.452365	136.4	***
Thu	61.7114	0.452568	136.4	***
Fri	61.6452	0.452365	136.3	***
Sat	61.2355	0.452365	135.4	***
Sun	55.684	0.452365	123.1	***
h1	-3.67642	0.572238	-6.425	***
h2	-10.3605	0.572238	-18.11	***
h3	-14.615	0.572238	-25.54	***
h4	-17.397	0.572238	-30.4	***
h5	-17.697	0.572238	-30.93	***
h6	-14.172	0.572238	-24.77	***
h7	-6.03006	0.572238	-10.54	***
h8	1.25931	0.572238	2.201	**
h9	5.95005	0.572238	10.4	***
h10	5.69166	0.572238	9.946	***
h11	2.00292	0.572238	3.5	***
h12	-0.74333	0.572238	-1.299	
h13	-5.47537	0.572238	-9.568	***
h14	-8.4685	0.572238	-14.8	***
h15	-6.71406	0.572238	-11.73	***
h16	-3.17546	0.572238	-5.549	***
h17	2.62445	0.572238	4.586	***
h18	11.1279	0.572238	19.45	***
h19	15.5149	0.572238	27.11	***
h20	20.2984	0.572238	35.47	***
h21	19.9492	0.572238	34.86	***
h22	13.0245	0.572238	22.76	***
h23	5.88455	0.572238	10.28	***

Table C.6: SUD Estimation

Bibliography

[1] Ardian, F., Concettini, S., and Creti, A., 2015, "Intermittent renewable generation and network congestion: an empirical analysis of Italian Power Market", *SSRN working paper series.*

[2] Bai,J. and Perron,P., 2003, "Computation and Analysis of Multiple Structural Change Models" , *Journal of Applied Econometrics*, 6, 72-78.

[3] Baltagi, B.H., 1995, *Econometric Analysis of Panel Data*, Wiley Publishers, New York.

[4] Bessec, M., and Bouabdallah, O., 2005," What causes the forecasting failure of Markov-switching models? A Monte Carlo study",*Studies in Nonlinear Dynamics and Econometrics*, 9(2), Article 6.

[5] Bollerslev, T., 1986, "Generalized autoregressive conditional heteroscedasticity", *Journal of Econometrics*, 31, 307-327.

[6] Bompard, E., Carpaneto, E., Ciwei, G., Napoli, R., Benini, M., Gallanti, M., and Migliavacca, G., 2008, "A game theory simulator for assessing the performances of competitive electricity markets". *Electric Power Systems Research*, 78, 217 - 227.

[7] Bordignon, S., Bunn, D.W., Lisi, F., and Nan, F., 2013,"Combining day-ahead forecasts for british electricity prices", *Energy Economics*, 35, 88 - 103.

[8] Bosco, B., Parisio, L., and Pelagatti, M., 2007, "Deregulated wholesale electricity prices in italy: an empirical analysis ", *International Advances in Economic Research*, 13, 415 -432.

[9] Bowden, N., and Payne, J.E., 2008, "Short term forecasting of electricity prices for fMISOg hubs: Evidence from arima-egarch models" *Energy Economics*, 30, 3186-3197.

[10] Carmona, R., Coulon, M., and Schwarz, D., 2013, "Electricity price modeling and asset valuation: a multi-fuel structural approach", *Mathematics and Financial Economics*, vol. 7(2), 167-202.

[11] Chatzidimitriou, K. C., Chrysopoulos, A. C., Symeonidis, A. L., and Mitkas, P. A., 2012), "Enhancing agent intelligence through evolving reservoir networks for predictions in power stock markets", *Lecture Notes in Computer Science*, Vol. 7103, pp. 228-247.

[12] Clo, S., Cataldi, A. and Zoppoli, P., 2015, "The merit-order effect in the Italian power market: The impact of solar and wind generation on national wholesale electricity prices", *Energy Policy*, 77, 79-88.

[13] Contreras, J., EspÃnola, R., Nogales, F. J., and Conejo, A. J., 2003, "ARIMA models to predict next-day electricity prices", *IEEE Transactions on Power Systems*, 18(3), 1014-1020.

[14] Coulon, M., and Howison, S., 2009, "Stochastic behaviour of the electricity bid stack: from fundamental drivers to power prices", *Journal of Energy Markets*, 2(1), 29-69.

[15] Cuaresma, J. C., Hlouskova, J., Kossmeier, S., and Obersteiner, M., 2004, "Forecasting electricity spot prices using linear univariate time-series models", *Applied Energy*, 77, 87-106.

[16] Cutler, N.J., Boerema, N.D., MacGill, I.F., and Outhred, H.G., 2011, "High penetration wind generation impacts on spot prices in the Australian national electricity market", *Energy Policy*, 39, 5939-5949.

[17] Dacco, R., and Satchell, C., 1999,"Why do regime-switching models forecast so badly?",*Journal of Forecasting*, 18(1), 1-16.

[18] Dempster, G., Isaacs, J., and Smith, N., 2008,"Price discovery in restructured electricity markets", *Resource and Energy Economics*, 30(2).

[19] Dickey, D., and Fuller, W., 1987, "Distribution of the estimators for autoregressive time series with a unit root", *Journal of American Statistical Association*, 55, 277-301.

[20] Diebold, F.X. and R.S. Mariano. ,1995, "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, 13, 253-63.

[21] Diongue, A. K.Guegan, D., and Vignal, B., 2009, "Forecasting electricity spot market prices with a k-factor GIGARCH process", *Applied Energy*, 86(4),505-510.

[22] Engle, R., 2002, "Dynamic conditional correlation: a sample class of multivariate generalized autoregressive conditional heteroskedasticity models", *Journal of Business and Economic Statistics*, 20 (3), 339-350.

[23] Figueiredoa, N.C., da Silva, P.P., and Cerqueira,P.A., 2015, "Evaluating the market splitting determinants: evidence from the Iberian spot electricity prices", *Energy Policy*, Volume 85, 218-234.

[24] Førsund, F.R., Singh, B., Jensen, T. and Larsen, C., 2008, "Phasing in wind-power in Norway: Network congestion and crowding-out of hydropowers", *Energy Policy*, 36 , 3514- 3520.

[25] Gao, C., Bompard, E., Napoli, R., and Zhou, J., 2008," Design of the electricity market monitoring system", *Proceedings of DRPT 2008*, pp. 99-106, art. no. 4523386.

[26] Garcia, R. C., Contreras, J., van Akkeren, M., and Garcia, J. B., 2005, " A GARCH forecasting model to predict day-ahead electricity prices.", *IEEE Transactions on Power Systems*, 20(2), 867-874.

[27] Gelabert, L., Labandeira, X., and Linares, P., 2011, "An ex-post analysis of the effect of renewables and cogeneration on Spanish electricity prices", *Energy Economics*, 33 (S1), 59-65.

[28] Gianfreda, A. and Grossi, L., 2009, "Zonal price analysis of the Italian wholesale electricity market", *Energy Market, 2009. EEM 2009. 6th International Conference on the European*, IEEE, Leuven, Belgium.

[29] Gianfreda, A., and Grossi, L., 2012, "Forecasting italian electricity zonal prices with exogenous variables",*Energy Economics*,34, 2228 -2239.

[30] Gonzalez, A. M., San Roque, A. M., and Garcia-Gonzalez, J., 2005.,"Modeling and forecasting electricity prices with input/output hidden Markov models", *IEEE Transactions on Power Systems*, vol. 20(1), 13-24.

[31] Granger, C.W.J., and Newbold, P., 1974, "Spurious regressions in econometric", *Journal of Econometrics*, 2 ,111-120.

[32] Greene, W. , 2003, *Econometric Analysis, 5th ed.*, Prentice Hall, New Jersey.

[33] Guerci, E., Ivaldi, S., and Cincotti, S., 2008, "Learning agents in an artificial power exchange: tacit collusion, market power and efficiency of two double-auction mechanisms", *Computational Economics*, 32, 73-98.

[34] Haldrup, N., and Nielsen,M., 2006, "Directional congestion and regime switching in a long memory model for electricity prices", *Stud. Nonlinear Dyn. Econom*, 10(3).

[35] Harvey, S.M., and Hogan, W.W., 2000, "Nodal and zonal congestion management and the exercise of market power", *Harvard working paper series.*

[36] Hemdan, G.A.N., Kurrat, M., Schmedes, T., Voigt, A., and Busch, R., 2014, "Integration of superconducting cables in distribution networks with high penetration of renewable energy resources: Techno-economic analysis", *International Journal of Electrical Power & Energy Systems*, Volume 62.

[37] Heydari, S., and Siddiqui, A., 2010,"Valuing a gas-fired power plant: A comparison of ordinary linear models, regime-switching approaches, and models with stochastic volatility". *Energy Economics*, 32, 709-725.

[38] Hickey, E., Loomis, D.G., and Mohammadi, H., 2012, "Forecasting hourly electricity prices using armaxgarch models: An application to MISO hubs", *Energy Economics*, 34, 307-315.

[39] Higgs, H., 2009, "Modelling price and volatility inter-relationships in the australian wholesale spot electricity markets", *Energy Economics*,31(5), 748-756.

[40] Hong, Y.-Y., and Hsiao, C.-Y, 2002, "Locational marginal price forecasting in deregulated electricity markets using artificial intelligence", textitIEE Proceedings: Generation, Transmission and Distribution, vol. 149(5), 621-626.

[41] Huisman, R., Huurman, C., and Mahieu, R., 2007, "Hourly electricity prices in day-ahead markets", *Energy Economics*, 29, 240 -248.

[42] Ignatieva, K., and Trück, S., 2016, "Modeling Spot Price Dependence in Australian Electricity Markets with Applications to Risk Management", *Computers & Operations Research*, Vol. 66, 415-433.

[43] Jónsson, T., Pinson, P., and Madsen, H, 2010, "On the market impact of wind energy forecasts", *Energy Economics*, 32 (2), 313-320.

[44] Ketterer, J.C., 2014, "The impact of wind power generation on the electricity price in Germany", *Energy Economics*, 44, 270-280.

[45] Knittel, C. R., and Roberts, M. R., 2005, "An empirical examination of restructured electricity prices", *Energy Economics*, 27, 791-817.

[46] Kunz, F., 2013, "Improving Congestion Management: How to Facilitate the Integration of Renewable Generation in Germany", *Energy Journal*, International Association for Energy Economics, vol.0 (number 4).

[47] Lien, D., and Tse, Y.K., 2002, "Evaluating the hedge performance of the constant correlation GARCH model", *Applied Financial Economics*,12 (11), 791-798.

[48] Liu, H., and Shi, J., 2013, "Applying armagarch approaches to forecasting short-term electricity prices", *Energy Economics*, 37, 152 -166.

[49] Lucia, J.J., and Schwartz, E.S., 2000," Electricity prices and power derivatives. - evidence from the nordic power exchange", *Review of Derivatives Research*, 5, 5-50.

[50] Milstein, I. and Tishler, A., 2011, "Intermittently renewable energy, optimal capacity mix and prices in a deregulated electricity market", *Energy Policy*, 39 (7), 3922-3927.

[51] Misiorek, A., TrÃ$\frac{1}{4}$ck, S., and Weron, R., 2006," Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models", *Studies in Nonlinear Dynamics and Econometrics*, vol 10(3), Article 2.

[52] Nogales, F. J., Contreras, J., Conejo, A. J., and Espinola, R., 2002, "Forecasting next-day electricity prices by time series models", *IEEE Transactions on Power Systems*, 17, 342-348.

[53] O'Mahoney, A. and Denny, E., 2011, "The merit-order effect of wind generation in the Irish electricity market", Proceedings of the 30th USAEE/IAEEE North American Conference, USAEE, Washingotn D.C., USA.

[54] Park, H., Mjelde, J., and Bessler, D., 2006," Price dynamics among U.S. markets",*Energy Economics*, 28(1).

[55] Petrella, A., and Sapio, S., 2006, "No PUN Intended: A time series analysis of Italian Day-Ahead electricity price", *EUI Working papers*, RSCAS 2010/03.

[56] Raviv, E., Bouwman, K.E., and van Dijk, D. , 2013, "Forecasting Day-Ahead Electricity Prices: Utilizing Hourly Prices",*Tinbergen Institute Discussion Paper*, TI 2013-068/III.

[57] Sapio, A., 2015, "The effects of renewables in space and time : A regime switching model of the Italian power price", *Energy Policy*, vol. 85, 487-499.

[58] Schwartz, G., 1978, "Estimating the dimension of a model", *Ann. Stat*, 6, 461-464.

[59] Schröeder, A., Oei, P.Y., Sander, A., Hankel, L., and Laurisch, L.C., 2013, "The Integration of Renewable Energies into the German Transmission Grid: A Scenario Comparison", *Energy Policy*, 61, 140-150.

[60] Schwert, 1989, "Test for Unit Roots: A Monte Carlo Investigation",*Journal of Business and Economics Statistics.*

[61] Serinaldi, F., 2011, "Distributional modeling and short-term forecasting of electricity prices by generalized additive models for location, scale and shape", *Energy Economics*, 33, 1216 -1226.

[62] Shumway, R. H., and Stoffer, D.S., 2011, "Time series analysis and its applications", Springer.

[63] Tse, Y.K., and Tsui, A.K.C., 2002, "A multivariate generalised autoregressive conditional heteroscedasticity model with time-varying correlations", *Journal of Business and Economic Statistics*,20 (3), 351-362.

[64] Vahviläinen, I., and Pyykkönen, T., 2005, "Stochastic factor model for electricity spot price - the case of the Nordic market", *Energy Economics*, vol. 27(2), 351-367.

[65] Weron, R., and Misiorek, A., 2005, "Forecasting spot electricity prices with time series models", *IEEE Conference Proceedings-EEM05*, pp. 133-141.

[66] Weron, R., and Misiorek, A., 2008. "Forecasting spot electricity prices: a comparison of parametric and semiparametric time series models", *International Journal of Forecasting*, 24, 744-763.

[67] Weron, R., 2014, "Electricity price forecasting: A review of the state-of-the-art with a look into the future", *International Journal of Forecasting*, 30 (2014),1030-1081.

[68] Woo, C.K., Zarnikau, J., Moore, J., and Horowitz, I., 2011, "Wind generation and zonal-market price divergence: Evidence from Texas", *Energy Policy*, 39(7), 3928-3938.

[69] Worthington, A.C., Kay-Spratley, A., and Higgs, H., 2005, "Transmission of prices and price volatility in Australian electricity spot markets: a multivariate GARCH analysis", *Energy Economics*, 27 (2), 337-350.

[70] Wurzburg, K., Labandeira, X., and Linares, P., 2013, "Renewable generation and electricity prices: Taking stock and new evidence for Germany and Austria", *Energy Economics*, 40 (1), Supplement Issue: Fifth Atlantic Workshop in Energy and Environmental Economics, 159-171.

[71] Zhao, J. H., Dong, Z. Y., Xu, Z., and Wong, K. P. , 2008, "A statistical approach for interval forecasting of the electricity price", *IEEE Transactions on Power Systems*, 23(2), 267-276.

université PARIS-SACLAY | ÉCOLE DOCTORALE Sciences de l'homme et de la société (SHS)

Titre : Analyse empirique du marché Italien de l'électricité

Mots clés : prediction, congestion, et interdépendance

Résumé : Dérégulation du marché de l'électricité a montré de nombreux changements dans l'économie et a influencé les chercheurs à initier des études dans ce domaine. Italie fournit une étude de cas intéressante pour explorer le marché de l'électricité en raison de ses spécifications. Notre projet se compose de trois études quantitatives indépendant pour voir le marché de l'électricité Italien en trois angles différents. La première étude permet de répondre la question de la prévision causée par la volatilité du marché de l'électricité. Le résultat suggère une méthode de prévision alternative pour la modélisation de prix de l'électricité sur l'Italie.

La deuxième recherche examine l'impact des énergies renouvelables sur l'apparition de la congestion et ses coûts. Nous analysons les propriétés quantitatives de l'estimation économétrique afin de comprendre le mécanisme économique et d'en tirer la suggestion de la politique. Enfin, la recherche finale analyse l'interdépendance des prix dans six macro-zones du marché italien de l'électricité.

.

Title : Empirical analysis of Italian electricity market

Keywords : forecast, congestion, interdependency

Abstract : Deregulation of the electricity market has displayed many changes in the economy and has influenced researchers to initiate studies in this area. Italy provides an interesting case study to explore the electricity market due to its specifications. Our project consists of three independent quantitative studies to view the Italian electricity market in three different angles. The first study helps us to answer the question of the forecast caused by the volatility of the electricity market. The result suggests an alternative forecasting method for modeling electricity prices on Italy.

The second study examines the impact of renewable energy on the appearance of congestion and its costs. We analyze the quantitative properties of the econometric estimation in order to understand the economic mechanism and draw the suggestion of the policy. Finally, the final research analyzes the interdependence of prices in six macro-areas of the Italian electricity market.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France