Bayesian Analysis Of Demand Elasticity in the Italian Wholesale Eletricity Market.

Maria Chiara D'Errico

Università degli Studi di Perugia
Corso di Dottorato in Economia e Poltica Agroindustriale
XXVII ciclo
Prof. Relatore:
Carlo Andrea Bollino

Contents

Preface

Consumer behaviour analysis has received relevant attention in the theoretical and empirical economic studies. Demand Analysis is characterized by this peculiar position: a detailed and comprehensive theoretical framework is combined with a huge amount of empirical works. The reason of that lies on the power of the Utility Theory (the theoretical background of demand analysis) as a tool of applied economic reasoning.

Elasticity, in the demand analysis framework, is a feature that has received particular attention in the studies of consumer preference and willingness to pay, as in the institutional studies guiding policy decisions as taxation and welfare. Moreover, the consumer reactivity to changes in price can express market efficiency. Then, in strategic economic sectors, this measure can be seen as a tool leading the National Regulators in the market structure definition processes.

This thesis tries to provide an analysis of Italian Electricity Market using elasticity estimation. The Italian electricity sector undertook a deregulation process starting in the 2004 that has led to overcome the system of vertically integrated monopoly. This process led to the institution of Power Exchange (IPEX). The transition had not been simple since the definition of a proper market structure preserving competition is not an immediate task. In this context, the information provided by demand elasticity have to be exploited since the elasticity is strictly linked with the market power measured on the supply side.

Given the purpose of Italian Regulator of preserving competition and efficiency in the electricity market, investingation of demand elasticity becomes essential to identify the design factors to be used for the definition of the market structure.

The empirical questions that my thesis intends to answer are:

- Are the buyers in the Italian Wholesale Electricity Market responsive to changes in price?
- What is the extent of buyer's elasticity?
- Can buyers change their consumption profiles within the day given the rational expectation of change in price?

In this research, the approach I used differs from the tradition of the previous literature in two respects: the type of data processed and the econometric approach adopted.

Previous empirical studies used data referring the supply side of electricity market, given the assumption of oligopolistic market structure, they estimate demand elasticity using residual demand function. In this work instead, I estimated italian electricity demand elasticity using data referring the demand side.

With regard to the econometric method, my research used a Bayesian procedure, whose application in electricity demand analysis represents a novel approach.

Until recently, the Bayesian approach has been in a distinct minority in the field of econometrics, which has been dominated by the frequentist approach: computation has been the substantive reason for the minority status of Bayesian Econometrics. The computing revolution of the last twenty years has overcome this hurdle allowing to exploit the theoretical and conceptual elegance of Bayesian Statistics in the empirical studies.

The thesis is structured as follows:

- The first part is devoted to the analysis of the statistical framework of the thesis.

  The first Chapter is a discussion enlightening the motivation about the adoption of Bayesian Approach.

  The second Chapter offers a brief review of the main bayesian statistical tools, focusing on the post-estimation and simulation procedures.

- The second part refers to the empirical part of the research.

  The third Chapter analyzes the structure of the Italian Electricity Market after the liberalization process and provides descriptive statistics referring the Day-Ahead Market.

  The forth Chapter presents the multivariate linear regression model I used for the elasticity estimation and the derived results.

  The fifth Chapter presents a generalization of the linear regression model relaxing the homoskedasticity assumption and shows the derived results.

Chapter 1

The choice of Bayesian Method

There are many definition about what statistics is, and different definition underlying different purposes, different interpretations of probability and different relevant information used. Statistics can be defined as the study of how information should be employed to reflect on and give guidance for action in a practical situation involving uncertainty.

However, there are more than one definitions of Statistics, reflecting a particular philosophical view point and expressing a particular attitude towards the meaning of probability, the relevant information used and the models adopted. I decided to start with a preliminary description of the main features underlying Statistics as the concept of probability and the information used. Then, I summarised the characteristics of the three main statistical approaches: the Classical Approach, the Decision Theory and the Bayesian Approach.

Finally, I will explain the reason why I have preferred to use in my research the Bayesian method.

As I said before, Statistics concerns practical situation involving uncertainty, that means there is more than one possible outcome and the actual outcome is unknown in advance: it is undetermine. It implies the need to construct a theory, a logical model guiding the behaviour in such situation involving uncertainty.

1.1 Probability

At this stage we must dig somewhat deeper into the sub-soil of inference by examining the basic concept of probability. We have seen that different statistical approaches imply a particular probability view-point. The spectrum of views of probability is vast and complex. However, bearing in mind that our prime interest is in the way in which different view of probability interralate with different modes of statistical reasonoing, this reviews will over-simplifies the true situation. Replicating the framework proposed by Poirier [?], the concept of probability can be classified in three main categories: frequentist, logical and subjective.

1.1.1 Frequency view of probability

This approach is the earliest in terms of detailed development, nonetheless, it has been more widely discussed and used than any other. It provides the interpretative framework for classical statistical methodology.

Central in this view-point is considering probability as something as objective, avoiding any consideration of personal factors, and amenable to practical demonstration through experimentation. There is only a restricted class of phenomena that can be analysed using this narrow concept of probability. The frequency view can only unambiguosly applied in situations free from ill-defined or immeasurable factors, which are able to be repeated over and over again under the same conditions and where unique probability exists and be demonstrated empirically. The only information relevant to probability assessment comes in fact from observing outcomes during repeated realizations, that is the sample data. Probability has unique value determined by the nature of the situation under study. This uniqueness leans heavily on the assumption that basic conditions do not change. Probability becomes an unconditional concept, having no concerns for any circumstantial evidence relating to the situation, for the different environment where the experiment is performed.

The frequency approach rests on two assumptions about the behaviour of the sample.

Firstly, if a random variation occurs during repeated realizations, it means that outcomes vary from one repetition to another in an unpredictable manner.

Secondly, the relative frequency of observations shows a long-term stability, then probability can be seen as the limiting value of the relative frequency of an infinite sequence of repeatable situations and it becomes measurable and objective. Only empirical investigation of frequency allows probabilistic assessments and no assertion may make sense if it is not derived by the experience. Probability is a physical features of nature, it is seen as the means of quantifying the uncertainty of natural phenomena.

The mathematical foundation of frequentist probability has been offered by Von Mises [1] which introduced the concept of *Collective* as the reference class of phenomena that can be analysed through statistical tools. Collective are situations involving infinite sequence of uniform observations, a mass of phenomena satisfying two conditions:

1) Relative frequency of particular attribute within the collective tends to fixed limit

2) Principle of randomness: the fixed limit does not depend on the way in which the selectoin of the sequence is done.

The collective is the reference class in which the sequence of attributes has to be a convergent series.

Individual, personal and behaviouristic assessments are excluded from statistitical analysis.

The basic assumption of this approach is an empirical one: knowledge can not be augmented if they do not rest on experience.

The identification of probability with the limit of relative frequency of repeated observations does not allow to investigate single events and their probability. Reichembarch [2] awares of the practical reasonableness of probabilistic evaluation of unique events, tried to broaden the narrow empiricism which frequentist probability rest on, giving a frequentist formulation of unique events: the probability of a single event is the relative frequency of similar occurrance. This involves an inductive process, frequentist probability comes from processing particular information we have experienced in order to derive general propositions.

[1] Von Mises, R. (1964).*Mathematical Theory of Probability and Statistics*, Academic Press, New York.

[2] Reichenbach, H., The Theory of probability. *An Inquiry into the logical and mathematical Foundations of the calculus of probability*, University of California Press, Berkley-Los Angeles 1949.

1.1.2 Logical view of probability

In the logical views, probability measures the degree of the implication between two statements, which is usually intermediate between the logical impication and the complete denial. When the implication is necessary, probability is infact equal to one; while the complete denial implies probability equal to zero. The degree of confidence is supposed to be an objective measure, a formal property of the implication. Logical probability is the rational intensity of convinction, arising from the information derived from empirical evidence and subjective impressions.

This attitude is in contrast with the frequency view, since the logical probability does not expresses an empirical relationship, but a logical ones.

"Even if there is no empirical evidence, assessment on the degree of confidence through probability is not inconsistent, since probability is not a natural feature but expresses a logical relation between statements." [3]

The difference between the frequentist and the logical view concerns also the scope. The reference class of frenquency is the collective, that is a practical problem based on the idea of an infinite sequence of similar elements, while logical probability applies to propositional statement expressing its rational degree of confidence.

According to Carnap [4], the probability of a statement, with respect to a given body of evidence, is a logical relation between the statement and the evidence. Logical probability quantifies the degree to which the outcome of experiment supports or undermines an hypothesis. Probability is the logical, formal and intrinsic link between the hypotesis and the body of evidence, as it offers the link between them. Thus probability becomes the mathematical tool (the confirmation function) through which inductive logic evaluate the reliability of a hypohesis. Probability in Carnap's theory is a metalinguistic operator codifying (which is applied to) the language used to describe a given situatiton and confirm the hypothesis explaining phenomenon.

1.1.3 Subjectivist view of probability

The last formulation, the subjective probability, arises from the dissatisfaction with the way how logical probability quantify the degree of belief. Sub-

[3] Jeffreys., H. (1993) *Theory of Probability*, Clarendon Press, 3rd edition, Oxford, 1961

[4] Carnap, R.,(1967). *The Logical Foundation of Probability*, University of Chicago Press.

jectivists refuse to think probability assessments as objective and rational determined necessity between two statements; they criticize the assumption of uniqueness, since there is not only one opinion justifying the body of evidence.

De Finetti [30] have extensively deepened that intepretation arguing that probability is a personal assessment since it is related to individual judgement.

Logical probability represent rational degree of belief on an hypotesis, it is a measure of the support given by some outcome of the experiment, indipendent on the observer. For the Subjectivists, probability is a measure of individual degree of belief relying on his relevant experience. The subjective approach tries therefore to build up a formal probability theory on behaviouristic basis. It was informally advanced by Bernoulli [5] when he talked about probability as the 'degree of confidence' that we have in the occurance of an event, dependent on our personal judgement.

De Finetti formally opened the way to subjectivists quantifying personal beliefs in terms of betting behaviour. "Probability is the betting quotient at which an individual would be ready to bet a certain sum on its occurrence. The probability of an event E is the price somebody is willing to pay in order to receive an unitary amount of money if the event E occurs". Quantitative assessments of personal opinions can be operationally determined introducing a fair bet situation where it will be asked what would be the maximum price p we are willing to pay to gain a unit amount of money in case the event E occurs. The price p represents therefore the subjective probability relative to the occurrance of event E.

The fundamental criterion one must obey is to avoid sure losses. The condition of fair bet dictates that no gain or loss are guaranteed in advance. This condition is also called *consistency principle* and it admits the subjective probability to satisfy the Kolmogorov's axioms, empowering it to assume an objective meaning. The consistency principle avoids infact subjective assessment being not reasonables.

The probability is the degree of confidence that a particular individual develops in a situation of uncertainty, given a particular set of information and a specific context.

[5] Bernoulli, J., *Ars Conjectandi*, Basel, 1713.

Quoting De Finetti: "We base our judgments upon subjective circumstances, probability is not an automatic consequence of them but is a subjective assessment since the choice about which factual circumstances should affect our judgment is subjective."

"In a world without men there would be no probability, since it would not exist nor knowledge nor ignorance, but only facts. Talking about probabilities makes sense only if it is connected to human being, to his state of uncertainty and to his desire to control the world through prediction and theory.

1.2 Relevant Information

The relevant information employed in the analysis are the second component characterizing the different approaches.

1.2.1 The Sample Information

When we use a Bernoullian sample of indipendent and identically distributed observations drawn from a population we adopt the classical approach. The observations are supposed to be indipendent repetitions of a situation occurring always under identical circumstance. The evaluation of sample data is through frequentist formulation of probability and constitue the corner stone of classical approach. The sample data are supposed to come from an experiment which can be infinetely replicated under the same conditions.

This attittude is well express by Von Mises [6]:

"Statistics is the mathematical theory of repetitive events". However, sample data is not the only information relevant to a statistical study, early experience and the potential consequences are also relevant. Some current statistical approaches are designed to incorporate such alternative kind of information, than that of just sample data.

1.2.2 The Risk Function

When statistical methods are used in order to give a guidance for action in situation involving uncertainty the assessments concerning the consequences of the alternative action are vital. The quantification of the consequences

[6] Von Mises, R. (1964).*Mathematical Theory of Probability and Statistics*, Academic Press, New York.

is via what is known as the risk function, that augments the information provided by the sample data and it is crucial to the choice of the action to be undertaken. Assessment of the consequences and their formal quantification is the corner-stone of a particular statisical approach known as Decision Theory. In the decision theory much of the emphasis is on the construction of a rational model for human behaviour, in the sense of representing how individuals make choice from alternative possible actions in the face of uncertainty. However, using quantitative translation of personal assessment of consequences as relevant information, may involve incorporating subjective judgements inside the statistical model. We can present the last kind of relevant information: the prior information.

1.2.3 Prior Information

Relevant information that can be potentially included in the statistical model is represented by the general knowledge accumulated from other areas of experience as the previous observations of similar situations. Information of this type is termed Prior Information. The Bayesian Statistics is the particular branch of statistics used to combine sample data and prior information. Prior information say something about the value of the model's parameters under investigation.

Also in this particular field, evaluate previous information to be incorporated in the statistical model involves some degree of uncerainty, for this reason we assign them a probability distribution: since there is no certainty of the value of model's parameters representing the prior information, Bayesian inference assigns to them a prior probability distribution.

The three different kinds of relevant information can be seen on a temporal basis: prior information accumulated from previous external experience refers to the past, sample data arising from the current situation (the experiment) are related to the present, and the assessment of consequences refers to the (potential) future action.

Broadly speaking, Classical inference uses just sample data, Bayesian Statistics allows prior information to be parts of the statistical model, while Decision Making Theory augments the inferential knowledge incorporating also assessment of consequences.

1.3 The different approaches

Different types of information and different procedures show how the different approaches underlying different aims. Frequentist and Bayesian approaches have exentially a descriptive function, while and decision theory has the purpose of prescribing an action.

Any statistical procedure which utilizes information to obtain a description of the practical situation (through probability model) is an inferential procedure and its study will be termed as Statistical Inference. On the other hand, a procedure with the wider aim of suggesting action to be taken in a practical situation (with an action guidance function), by processing information relevant to that situation, is a decision making procedure and the study of such procedures is termed statistical decision-making.

Following the framework proposed by Bernardo [10], let rewiev the three main statistical approaches, differencing for their concept of probability, their relevant information and the purpose.

1.3.1 Classical Inference

Classical Statistics lean on frequentist concept of probability and the tecniques adopted are the point and the interval estimation and the hypotesis testing. It utilizes the sample data as its only source of relevant information. For this reason it seems to belong to an inferential scope. Sample data are managed through the likelihood function and the performance of the techniques used is assessed through criteria based on sampling distributions. Classical approach relies on the assumptionn that the sampling distribution will converge to the distribution of the population through an infinite repetition of the same experiment. Moreover, the properties of the estimate as consistency and accurancy refer only to their asymptotic behaviuor.

1.3.2 Bayesian Inference

It is again an inferential procedure admitting the processing of sample data as well as prior information. The prior information is modified by the sample data through the application of Bayes Theorem, yielding to a combined assessment of the state of knowledge of the situation of interest. There are two basic elements we have to mention.

First, subjective probabilities are assigned to the different competing hypoteses explaining the theory. The prior information are processed into the model through probability distributions which quantify the degree of confidence the researcher has in the values of model's parameters. These subjective probabilities are consistent with the Kolmogorov's axiomatic definition of probability.

Secondly, using the conditionalization rule deriving from the Bayes Theorem, it is possible to learn from the evidence and update, in a quantitative way, our credence in the ligth of the experimental results.

Inferential statements are expressed through posterior probability distributions whose variance embodies the measure of their accurancy. This approach can not lean on frequentist concept of probability, the subjective interpretation is undeniable given the use of prior informatin.

The use of Bayesian methods is not restricted to situations where prior information exist, also the ignorance has a probability interpretation and it can be processed within the inferential procedure.

Bayesian theory has become a predominant approach to confirmation in the late twentieth century. The popularity of the Bayesian approach is due to its flexibility, its apparently effortless handling of various technical problems and the injection of subjective element into the theoretic model.

1.3.3 Decision Theory

Stemming from the work of Wald [7], this approach is designed to provide action guidance under uncertainty situations, i.e. decision rules.

It embodies assessment on the consequences of alternative actions, expressed through the analytical representation of risk function. The values of any decision in favor of a specific action, on the basis of sample data and prior information, is expressed by the expected loss. The aim is to choose the decision rule with minimum risk.

No particular philosophical view of probability is implied in the decision theory, although when prior information are incorporated inside the model, a subjective view of probability is adopted.

This broad classification oversimplyfies the true structure of statistical approaches. There are not well-defined distinction between the approaches

[7] Wald, A. (1950). *Statistical Decision Functions.* John Wiley and Sons, New York.

in terms of their function. Many arbitrary and personal matters are involved in the formulation of rules for processing information in the inferential procedure, that as many approaches will exist as there are individuals with their own interpretations of what constitute reasonable rules for behaviour. In practice the boundaries are blurred and it should be recognized that the adoption of a particular procedure also depends on arbitrary criteria related to personal judgments and attitudes.

The classical approach seems more structured, but it seems a simplification.

The choice to exclude from the inferential procedure the a priori information and the consequence assessments arises from the convinction of their poor objectivity. Classical statisticians aim indeed to construct a theory which would be universal in its application, free from subjective assessment, based on the only information quantifiable. Despite the purpose, this approach also conceals subjective assessments.

Let consider the prevalent hypothesis testing procedure proposed by Neyman and Pearson whoose meaning has distanced from the significancy test of Fisher. Fisher test assumes just one hypothesis (the null hypothesis) representing the randomness of an experiment. This hypothesis can only be falsified, thus he never permits hypotesis to be accepted. Significance test does not allow to formulate an alternative hypothesis. The set of alternative hypotheses is in fact numerically infinite and qualitatively indeterminate. In the Fisher's framework, the refutation of the null hypothesis do not logically imply the acceptance of any alternative hypothesis. Fisher stated " If there is not enough evidence to reject the null hypothesis it does not mean that we have enough evidence in favour of it. The lack of evidence against a hypothesis is not an evidence for it."

In contrast, Neyman-Pearson use the terms acceptance and rejection. According to them , the outcome of a test is "an act or decision to take a particular action. This process is certainly not any sort of reasoning, it is an act".

Within the Neyman-Pearson framework, the process works differently: there are a null hypothesis and an alternative ones and after comparing p-value with the significance level, it will be decided to reject or not the null hypothesis. The rational mind did not discard a hypothesis unless it could conceive at least one plausible alternative hypothesis. For this reason Neyman proposed to replace the Fisher's inductive method with an "inductive behaviour" since the latter implies a choice.

Moreover, the standard 5% significance level doesn't really have mathematical basis for all cases. It is a result of a long-standing tradition.

Neyman's ideas on testing channelled statistical theory into new direction which culminated with Wald's general Decision Theory.

Instead of producing a procedure free from subjective assessments, some of the main procedures of classical statistical inference involve discretionary interpretations of results and they seem to be included in decision-making theory.

1.4 Why Bayesian?

1.4.1 The Problem of Induction

The previous consideratons recognize statistics as a necessary tool for kwoledge.

Knowledge relies on generalization, that is the process of deriving general statements from particular experiences and it produces propositions asserted to be true for all member of a certain class. Thus, knowledge is strictly linked with the inductive reasoning. Induction is defined as the logical link from particular statements to general assessment, it derives from particular experiences general proposition pertaining an entire class of phenomena. It is the inductive reasoning that can broaden and deepen our empiraical knowledge.

Although all generalizations are suggested by the observation of specific phenomenon, they disregard reference to particular occurrence. Mature sciences seem to be effective relying on observed evidence to establish extremely general, powerful and sophisticated theories. The shift from the observation of particular phenomena to the derivation of a general statement is not logically valid, nor necessarily implied. Considering exclusevely a logical perspective, the inductive inference, unlike the deductive inference, is not valid.

Aristotle had already exposed the fallacy of inductive inference. In the field of empiricism, we are always in an inductive situation: from a finite number of particular occurrences we draw explanatory and generalizing inference. However, induction involves a risk of error along with its potential: good induction may lead from true premises to false conclusions, inductive inferences are therefore contingent.

David Hume [8] is usually credited for having disclosed the theoretical root of these considerations in a trasparent way called "The problem of induction".

Hume divided all reasoning into deductive reasoning and probabilistic, that is the generalization of causal reasoning. Deductive logic is completely demonstrable, since the premises of an argument, constructed according to the rules of the logic, imply its argument's conclusion. Deduction is explicative: it orders and rearranges our knowledge without extending its content. On the other hand induction is ampliative, it establishes universal propositions from particular instances. Scientific theories try to derive regular and persistent pattern to the behave of phenomena.

Hume pointed out that the empirical sciences, in their confirmation through experiments, rely on two main assumptions:

- Uniformity of Nature
- Causality principle

Scientific theory need of these two principles in order to derive some kind of previsions.

According to Hume we cannot rationally justify the claim that nature will continue to be uniform. The Uniformity Principle cannot infact be proved through deduction, since it does not express a necessary connection, and it cannot be demonstrated by causal reasoning, since the principle drives itself the premises of the causal reasoning and such proof would be a *petito principi*.

The notion of causality is closely linked to the problem of induction. According to Hume, we reason inductively by associating constantly conjoined events, and this mental act founds our concept of causation. However, the relations of contiguity and succession do not necessary imply causal connexion. According to Hume, human being tends to give to contiguity relations the attribute of necessity. Causal assiociation, instead, is an "habits of the mind". Rather than reason, natural instinct explains the human ability to make inductive inferences.

In the real world there are no necessary connections, only constant conjuctions. Necessity and causality are not features of the word, they are perceptual categories. For this reason induction is a contingent inference,

[8] Hume, D. *A Treatise of Human Nature.* (1738)

probabilistic connection, no less than simple causal connection, depend upon habits of the mind and are not be found in our experience of the world. Inductive inference can yield a conclusions only with a certain probability.

Following Hume work, two centuries later Karl Popper proposed again the problem of induction formulating its controversial Falsification Theory. According to Popper, science adopts the hypotetic-deductive method formally based on four main step:

1. Formulation of hypoteses or theory
2. Derivation of consequences through deductive reasoning
3. Design of Experiment simulating the occurance of one hypothesis and observation of empirical results.
4. Acceptation or rejection of the hypotesis or theory according to the outcomes.

According to Popper it is not possible to establish the truth of a theory by empirical evidence, since scientific theories have universal scope but no finite evidence can ever adjudicate among them. Theories could be only falsified and only in that falsiabillity by counterexample lies the confidence of the knowledge. Induction has no place in the logic of science.

Theories can not be confirmed or verified. They may be falsified or tentatively accepted if corroborated by the proper kind of test.

"The best we can say of a hypothesis is that up to now it has been able to show its worth, and that it has been more successful than other hypotheses although in principle, it can never be justified or verified."[9]

1.4.2 Bayesian Confirmation Theory

Bayesian model is much richer then the tipical model assumed in classical statistics. Reasearchers, in the classical approach, can have one of three attitudes towards model's explanation: they accept the theory, they reject the theory, or they neither accept nor reject it.

A theory is accepted or rejected once the evidence in favor of one of the two decisions is sufficiently strong, if the evidence is strong in neither way, it is neither accepted nor rejected. In the Bayesian model, by contrast, the researchers' attitude toward an hypotesis is incapsulated in a degree of belief, that can take value between 0 and 1. Rather than laying down, as the

[9] Popper, K. *The Logic of Scientific Discovery.*(1934)

classical approach, a set of rules dictating when the evidence support or reject a theory, the bayesian approach prescribes a set of rules on how individual's opinion (prior information) should change in response to empirical evidence.

Let start defining Bayes' Theorem

Theorem 1 *Bayes' Theorem:*

Let be $\{H_1, H_2, ..., H_k\}$ a set of events mutually exclusive and exhaustive and A some other events of particular interest. The probabilities, $P(H_i)$, (i=1,...,k), of each of the H_i are known, as are known the conditional probabilities, P(A|H_i) (i=1,...,k) of A given that A_i has occurred. Then, the conditional ('inverse') probability of any H_i (i=1,...,k), given that A has occurred, is given by:

$$P(H_i|A) = \frac{P(A|H_i)P(H_i)}{\sum_{j=1}^{k} P(A|H_j)P(H_j)} \tag{1.1}$$

As expressed, Bayes' Theorem finds wide range of applications and there is no difficulty to extend its meaning considering the set of events H_i as the set of hypoteses representing what constitute the appropriate model explanation of a practical situation. The event A, in this way, becomes reinterpreted as an observed outcome that is the sample data.

Listing the two features of the theorem we have:

1) Prior to the observation, $P(H_i)$ is the probability that H_i is the appropriate model specification. These are named prior probabilities of the different hypoteses, and constitute the first sources of relavant information.

2) The probabilities of observing A, when H_i is the correct model specification, are $P(A|H_i)$ $(i = 1, ..., k)$ and they are simply the likelihoods of the sample data.

Bayes Theorem can be seen as a mathematical tool of updating, through the information provided by the sample data, our earlier state of knowledge expressed in terms of the prior probabilities, $P(H_i)$ (for $i = 1, ..., k$). The updated assesment is given by the posterior probabilities (or inverse probabilities) of the different hypoteses being true after using the futher information provided by observing A: $P(H_i|A)$ $(i = 1, ..., k)$.

Bayes Theorem represents the foundation of inductive reasoning.

This is the essence of Bayesian inference: the posterior probability of H_i given A is proportional to the product of the prior probability of H_i and the likelihood of A when H_i is true. Prior information are in this way augmented by the sample data to yield a probabilistic description of the current situation The updated knowledge is fully described by the posterior distribution.

Inferences are to be made by combining the information provided by prior probabilities with information given by sample data; this combination is achieved by 'the repeated use of Bayes' Theorem' and the final infereces are expressed solely by the posterior probabilities.

It is controversial why frequentist are reluctant to assign initial probabilities to the different hypoteses about the parameters' values they investigate. They assume infact the parameters of the model are unknown constants, not a random variables which assume different values in a series of trials. In Bayesian approach prior probability does not admit any frequency interpretation; it is a subjective probability.

The idea that prior probabilities are not empirical concepts is considered by some frequentists inadmissible since these probabilities can not be tested by any experiment. However, this cannot be strongly enough refuted. In Bayes's approach the parameter values are uncertain and researchers feed some degree of confidence about their values; the quantitative traslation of this degree of belief is the assignment of a probability distributions to the parameter that become itself a random variable.

Inductive reasoning is the mental process according to which our preliminary hypoteses are modified on the basis of the data supplied by experiments. The inductive reasoning has therefore a Bayesian foundation. A set of observations may be logically consistent with several hypotheses, we all draw conclusions from the empirical results, these conclusions are uncertain and this uncertainty can be expressed in a quantitative consistent way through the posterior probability distribution. Bayes' Theorem is important because it provides a mathematical representation for this consistent choice between alternative hypotheses on the basis of quantitative evaluation of their respective uncertainties and the quantitative information provided by the data. This is the worth of Bayes' work: it provides the mathematical foundation

of inductive reasoning

Chapter 2

Bayesian Statistics

In this chapter I rewiev the main steps undertaken in Bayesian Approach, how prior knowledges are combined with sample data to derive posterior inference and the main estimation procedures.

Let suppose that sample data x arises as an observation of a random variable, X. The distribution of X, specified by the probability model, is assumed to belong to some family, $\wp$, indexed by a parameter θ. It is assumed that the probability (density) function of the random variable has a known form, $p_\theta(x)$, depending on θ; but θ is unknown, we expect it lies in a parameter space Θ. For fixed sample x, $p_\theta(x)$ is the likelihood function, it associates for any pssible value of θ a probability distribution to the sample. Both θ and x may be either univariate or multivriate, discrete or continuous variables. Knowledge of the true value of θ would be all that is needed to completely describe the current practical situation.

The aim of the inferential procedure is to investigate the unknown value of θ. This cannot be done with any certainty, all we can expect is some probabilistic statement involving θ, based on the available information. In the classical approach this is achieved by processing the sample data as the only source of information in order to produce point or interval estimates of θ. Bayesian inference indeed incorporate into the model further a-priori information. To do that, a wider view of the nature of the parameter θ is taken. It is assumed that any knowledge we have of the true value of θ, at any stage, is uncertain and it can be expressed by a probability distribution, or by some 'weighting function', over the parameter space Θ. The parameter θ is now essentially treated as a random variable, in the sense that θ can assume different values with different probabilities or weights. Prior knowledge of θ is expressed through the probability distribution function $p(\theta)$. The sampling increases this knowledge and the updating process is expressed by the posterior distriburion. Posterior distribution is derived using Bayes' Theorem combining the Likelihood function (representing information from the sample) with the prior distribution $p(\theta|x)$:

$$p(\theta|x) = \frac{p_\theta(x)p(\theta)}{\int_\Theta p_\theta(x)dF(\theta)} \tag{2.1}$$

The posterior distribution describes our assessment of where the true value of θ is likely to lie in Θ after observing the sample.

This is a more direct form of inference than that we have in the classical approach since we have the intere (posterior) distribution of θ.

There may be situations where such a full description is not nedeed. Certain summary measures about $p(\theta|x)$ may suffice. For example it may be enough to know what value of θ is most likely, or in what region θ is highly likely to fall. These concepts in Bayesian inference are parallel with the idea of point estimate and confidence region in the classical approach, but it must be remembered that their interpretation is totally different and this difference can corroborate our prefernce for the Bayesian approach.

2.1 Point Estimation

If we are interested in a crude summary of the derived posterior distribution for the parameter θ, a good choice of the point estimate can be the mode or the mean of the entire distribution. This estimate in the bayesian approach has the direct interpretation as the most likely value of Posterior distribution that constitute the complete inferential statement about θ. Bayesian estimates infact contains their own internal measure of accurancy (through the posterior dstribution). This is the most preminent reason of adopting a Bayesian approach.

This facility, indeed, is not available in the classical statistics. Properties as unbiasness, consistency precision etc., can not be ascribed to the point estimate without reference to some larger framework than the current situation. The probability model must infact be built on a sample space, which is assumed to provide the constant basis for repeated and identical experiments. This is the concept of collective: the long-run behaviour of estimates are aggragated properties referring to the population, although they are attribuited to specific realization. However, it is sometimes disputed if such collective exists, the specification of the sample space may be claimed to be arbitrary or largely subjective. In the frequentist field, if we are interesting in the final precision of our estimator, we must rest on any properties related to the long-run behaviour of the procedure itself. It is infact necessary to consider the sampling distribution defined in terms of a sequence of indipendent repetitions of the current situation.

In classical inference the situation is quite different. It may exist different values for the estimator function $\theta(X)$ and they may differ from the true fixed value of θ, and their different interpretation is in term of different sets of realization (sample data), rather than in terms of differing degree of belief about θ.

2.2 Credible Region

Given posterior distribution we can derive the credible regions, that is the region where the parameter may reasonably be expected to lie with 95% of confidence.

The expression is given by the formula:

$$\left[\underline{\theta}(x); \overline{\theta}(x)\right] : P\left\{\theta \in \left[\underline{\theta}(x); \overline{\theta}(x)\right] |x\right\} \geq 1-\alpha \tag{2.2}$$

In a frequentist approach, the confidence interval has not a direct probability interpretation. Probability interpretation only refers to the long run behaviour of a procedure and namely that a portion $1-\alpha$ of intervals derived from repeated experiments will contain the true value θ. Whether a particular interval in some current situation contains θ is totally uncertain. The assessment of its probability of actually enclosing θ is in terms of repetitions of the experimental situation. As we saw earlier there is no way of judging whether a particular confidence region does or does not include the realvalue θ.

A naive statistician will almost inevitably and implicitly attach a 95% of confidence to the specific interval; in the practical use it is hard to avoid.

This is the main fallacy of the frequentist approach. Situations admitting repetition under essentially identical conditions are the only possible object in the real of statistical enquiries.

The fallacy of the frequency approach is to not declare that the concept of a collective is untenable: that we can never really assert that an unlimited sequence of essentially identical situations exists. Bayesian confidence (credible) region for θ has instead this direct probability interpretation and it is determined solely from he current sample data x.

2.3 Hypothesis Testing

In a view of direct probability interpretation provided by the posterior distribution, the one sided Bayesian Hypotesis Test has a form simpler than its frequentist parallel.

In a practical situation we may need to assess whether some statement about θ lying in a particular region of the parameter space is reasonable or not. Posterior distrinbution offer a direct probability evaluation of the two hypoteses.

Giving the hypotesis test:

H_0=$\theta \in \Theta_0$ against $H_1 = \theta \in \Theta_0^C$

We have:

$P(H_0|x) = \int_{\theta \in \Theta_0} p(\theta|x) d\theta = 1 - P(H_1|x)$.

If $P(H_0|x)$ is smaller than the significance level we will reject H_0.

Note that the direct expression of the result of test in the form of $P(H_0|x)$ eliminates the asymmetric nature of the test observed in the classical approach. In particular there is no need to distinguish between the null and aternative hypothesis.

The theoretical and conceptual elegance of the Bayesian approach has made it attractive for many decades. However, until recently, Bayesian have been in a distinct minority of many field of econometrics, which has been dominated by the frequentist approach. The two main reasons for this reluctance rely on the use of prior information and the computational problems.

With regards to the former, many reaserchers object to the use of 'subjective' prior information in the supposedly 'objective' science as economics.

2.4 The Role of Prior Distribution

There is a long debate about the role of prior information in statistical science. Most of Bayesians would argue that an enourmous ammount of non data information need to be processed into the model (e.g. econometricians must decide which model to work with, which variable to include, which empirical results to report etc.).

The Bayesian approach is honest and rigorous about how such non-data are used, on the grounds that more information is preferred to less. Moreover, Bayesians have developed *noninformative priors* for many classes of model and, as it will be shown in the next section, the derived posterior is proportional to the likelihood, representing in this way only the information arising from data; that is, the Bayesian approach allows for the use of prior information if we wish to use it. However, if we do not wish to use it, we do not have to do so. Regardless of how researcher feels about prior information, it should in no way be an obstacles to the adoption of Bayesian Method.

Computation is the second reason for the minority status of Bayesian econometrics. That is, Bayesian econometrics has historically been computationally difficult or impossible for a large classes of model. The computing revolution of the last 20 years has overcome this hurdle and led to a blossoming of Bayesian methods in many fields. However, this has made Bayesian econometrics a field which makes heavy use of the computer. In essence, the ideas of Bayesian econometrics are simple, since they only involve the use of probability. However, using Bayesian econometrics in practice often requires advanced computational skills and software.

The next two sections rewiev how Bayesian inference has tried to answer to the main frequentist objections. Prior ignorance has been infact processed in the statistical model through the non informative priors, while computational difficults, before the computer software developments, have been overtaken using Conjugate Prior.

2.4.1 Prior Ignorance

One of the corner stone of the Bayesian approach is the principle of insufficient reason. This principle (also called principle of indifference) is a rule for assigning epistemic probabilities. It states that if there are no outcomes mutually exclusive and collectively exhaustive and the possibilities are indistinguishable, then it should be assigned equal probability to each possibility.

In Bayesian terms, this means using non-informative prior. This principle provides a statistical description of the state of prior ignorance. When our prior information are insufficient to assign precise prior probabilities to the hypoteses, we quantify this ignorance assigning equal probability to all possible outcomes.

Jeffrey [1] was the first who expressed this principle

"If there is no reason to believe one hypothesis rather than another, the probabilities are equal...to say that the probabilities are equal is a precise way of saying that we have no ground for choosing between alternative....The rule of assigning equal probabilities is not a statement of any belief about the actual composition of the world, nor it is an inference from previous experience; it is a merely the formal way of expressing the ignorance."

The prior density function assigned to parameter θ over the space Θ is therefore a costant density.

For instance, for the parameter location μ, where the parameter space is the whole real line $(-\infty;\infty)$ we would chose the prior $p(\theta)$ to be the uniform density.

The invariant density function is also named as improper prior, since its integrand cannot ensure to equal one. This presents no basic interpretative difficulty if we assume probability as a degree of belief. The prior $p(\theta)$ will infact acts as a weighting function operating on the likelihood and the posterior will be proportional to the sample density function. However, Jeffreys warned about an extensive use of improper prior and the main objection is the impossibility to carry out significance test since no odds, no probability could be assigned to a point hypotesis.

2.4.2 Conjugate prior

Although Bayes' Theorem is mathematically simple, its implementation can be troublesome. The first difficult lies in the normalizing costant, the denominator of [2.1]. This had posed the central practical problem in Bayesian Inference: finding a numerical and analytical solution for the integral. Historically, two solution to this integration problem have been sought. Before the widespread availability of computer, Bayesian researchs centered on

[1] Jeffreys, H. *Theory of Probability*, 1939, Clarendon Press, Oxford, 1961.

defining prior distributions with convenient mathematical properties including tractable analytical solution to the integral. The family of distribution found was called Conjugate Prior. Conjugate prior is a family of distribution which, combined to the likelihhod, derive a posterior which belong to the prior's same family distribution. Conjugate priors are an algebric convenience which derive a closed form expression of the posterior density. Moreover, conjugate priors give the intuition of how bayesian inference updates prior information through the likelihood function. The prior and posterior distribution are infact indexed by parameters lying in the same parameter space. That is parameter indexing prior distribution can be interpreted in terms of pseudo-observations, that is it can be seen as the sample size of prior information.

Prior and posterior parametrs are able to measure the relative contribution of prior opinion and sample data in deriving posterior inference.

In my research the properties of natural conjugate priors will not be exploited. I have preferred to use the more advanced techniques of numerical approximations through Simulations. In particular, I focused on Markov Chain Monte Carlo.

2.5 Simulation

Bayesian Computation has beem the most challenging part of the development of my research, as I said before, the integration of the normalizing costant of the posterior is not the only analytical problem a bayesian researcher has to face. If we do not use conjugate prior, the posterior it-self may not be ascribable to a standard distribution and summary statistics used for point estimation, as the expected value for instance, may not be directly derived.

However, in the last decades a myriad of posterior simulators have been developped and I exploited these statistical tools in order to approximate the posterior density through the sampling distribution of a sequence of random variables simulated to be drawn from the posterior (the target density).

The algorithms I used refer to the class of Markov Chain Monte Carlo procedures, in particular, for the linear regression model the posterior simulator I implemented is the Gibbs Sampler, while for the linear regression model with heteroskedasticity, I combined the Gibbs Sampler with the Metropolis-Hasting Algorithm.

This section want to review the main theoretical premises of posterior simulation and enlighten the main steps of the algorithm I used in my research. The section, firstly introduce Monte Carlo method, the pioneering procedure of approximation methods, then it focus on the more advanced Markov Chain Monte Carlo processes (MCMC).

2.5.1 Monte Carlo Method

It is a common practice, after deriving the posterior distribution $p(\theta|x)$, to presenting the expected value as the main point estimate and deriving its variance as a measure of the accurancy of the estimate.

All these posterior features require the evaluation of an integral of the form:

$$E\left[f(\theta|x)\right] = \int_{\Theta} f(\theta|x)p(\theta|x)d\theta$$

where $f(\theta|x)$ is the function of interest.

If we are seeking the expected value the function takes the form $f(\theta|y) = \theta$, while if we are interested in the variance $g(\theta|x) = \{\theta - E(\theta|x)\}^2$.

Thanks to the widespread availability of computer, when natural conjugate prior is not used, more recent statistical developments focus on numerical approximation. There are many methods for doing this and all of these are applications or extensions of *Law of Large Numbers* and *Central Limit Theorem*.

A straighforward implication of the Law of Large Numbers is:

Theorem 2 *Monte Carlo Integration*

Let S the numer of drawing and $\left\{\theta^{(s)}\right\}_{s=1}^{S}$ be the resulting iid random samples draw from $p(\theta|x)$, and define $\widehat{g}_S = \frac{1}{S}\sum_{s=1}^{S} f(\theta^{(s)})$

then $\widehat{g}_S$ converges to $E\left[f(\theta)|x\right]$ as S goes to infinity.

In practice, this means that if a random sample is drawn from the posterior, the theorem allows to approximate $E\left[f(\theta)|x\right]$ by simply averaging the function of interest $g(\theta)$ evaluated at the differrent random samples. Sampling from the posterior distribution is what constitutes posterior simulation and the theorem describes the simplest one.

Monte Carlo integration can be used to approximate $E\left[f(\theta)|x\right]$ only if the approximation error tends to zero as S goes to infinity. Since the procedure allows to choose any value for S (although larger values of S can increase the computational burden), there are many ways of gaunging the approximation error associated with a particular value of S. However, many of them are based on extension of central limit theorem. For the case of Monte Carlo integration, the Central Limit Theorem implies:

Theorem 3 *Numerical Standard error*
Using the setup of Monte Carlo integration Theorem
the approximation error term:
$\sqrt{S}\left\{\widehat{f}_s - E\left[f(\theta|x)\right]\right\} \to N(0, \sigma_g^2)$
as S goes to infinity, where $\sigma_f^2 = var\left[f(\theta|x)\right].$

This theorem can be used to obtain an estimate of the approximation error in a Monte Carlo integration using the properties of Normal distribution. For instance, using the exact value ± 1.96 (the values within the standard Normal distribution locates the 95% of its probability mass), we can derive the result:

$$Pr\left[-1.96\frac{\sigma_f}{\sqrt{S}} \leq \widehat{f}_s - E\left[f(\theta|x)\right] \leq 1.96\frac{\sigma_f}{\sqrt{S}}\right] = 0.95$$

By controlling S, it can be ensured that the approximation error $\widehat{g}_s - E\left[f(\theta|x)\right]$ is sufficiently small with a high degree of probability. Unless σ_g is unknown, it can be approximated through the Monte carlo procedure. In many empirical context, this may be a nice way of expressing the approximation error implicit in Monte Carlo integration. The latter theorem also implies that if $S = 10000$, then the standard erros is 1%, as big as the posterior standard devitio. Unfortunately, it is not always possible to perform Monte Carlo Integration since it is based on the the indipendence and identical distribution of the draws. Moreover for many model, as the linear regression model I proposed, the posterior (from which the sample is drawn) do not have the Normal Form. In these cases, development of the posterior simulator is more changelling and I am going to explain them in the next sections.

2.5.2 Markov Chain Monte Carlo Method

The worth of Monte Carlo Integration Theorem is to give a theoretical framework to posterior simulation. However, it is based on the assumption that the sample is the bernoullian one. When we have to approximate using a sequence of non indipendent observations, we must seek a different stochastic process to be used.

As we will see later, in the two model I proposed, posterior simulation had to handle this issue since I could simulate the target posterior distribution of the parameter using only a sequence of non indipendent realizations. Any observation infact was dependent on the previuos draw. Since this kind of dependency is ascribable to the Markov Chain Stochastic Process, I used the simulation methods exploiting the properties of Markov Chain implementing Markov Chain Monte Carlo Algorithm.

This method is based on constructing a Markov Chain process such that the target distribution (the posterior density in our case) is the stationary distribution of the chain, and such that the chain converges in probability to this invariant distribution. Thus, in Markov Chain Monte Carlo samples from the target distribution are obtained only asymptotically. When convergence occurs, realizations of the chain are realizations of the invariant target distribution.

In the standard Markov Chain theory, the purpose is to find the stationary distribution of the process, Markov Chain Monte Carlo methods instead turn the theory around: we know the invariant target distribution, $p(\theta|x)$ and we construct a chain having the target (the posterior) as its stationary distribution.

Markov Chain used for Monte Carlo purposes have to satisfy some regularity condition such as:

- The existence of a unique stationary distribution
- The convergence of the stationary distribution to the target distribution

- Ergodicity.

In the next section we are going to set out the conditions under which a Markov Chain process ensures to meet these requirements.

Let start to enlighten the basic theory about Markov Chain.

2.5.3 Markov Chain

A Marlov Chain is a sequence of random variable $\theta_1, \theta_2, ...$ such that the probability distribution of any one, given all preceding realizations, depends at most on the immediate preceding observation. Specially, if Θ is the sample space for θ and A is a subset of a collection of A on Θ then

$$P(\theta_{t+1} \in A|\theta_1, \theta_2, ..., \theta_t) = P(\theta_{t+1} \in A|\theta_t) \tag{2.3}$$

for all $t = 1, 2, ..., t$ and any such A. Here θ represents the parameter vector whose posterior ditribution need to be simulated.

Expression [2.3] contains the main features specifying a Markov Chain.

The value taken by θ_t is called the state of the chain at t which has initial distribution $P(\theta_t)$.

The Transition Probability $P(\theta_{t+1} \in A|\theta_t)$ represents the probability of the next value of the chain given the current value and it is assumed to be homogeneous, that is it does not depend on the date, t. Homogeneous Markov Chain are fully described by the initial state and byb the transition probability which describes how the chain moves from the its state at t to its state at $t+1$.

When the sample space is finite and discrete the collection of conditional probabilities, one for each θ_t, can be gathered in a stochastic matrix, called *Transition Matrix,* whose elements are:

$$k(\theta_t, \theta_{t+1}) = P(\theta_{t+1}|\theta_t) \quad \theta_t, \theta_{t+1} \in \Theta \tag{2.4}$$

Transition matrix is a square matrix of order M (where M is the dimension of the discrete state space, that is the number of possible states); whose elements are non-negative and its rows, being probability distributions, sum to one.

The probability distribution $P(\theta_{t+1})$, say p_{t+1}, can be described in terms of the transition kernel and the distribution of its previous analogous θ_t:

$$P(\theta_{t+1}) = \sum_{i=1}^{M} P(\theta_t = i)P(\theta_{t+1} = j|\theta_t = i), \quad j = 1, 2, ..., M \tag{2.5}$$

that is

$$p_{t+1} = \sum_{i=1}^{M} p_t(i)K(i, j) \tag{2.6}$$

The matrix version of [2.6] is

$$p'_{t+1} = p'_t K \tag{2.7}$$

where p'_{t+1} and p'_t are row vectors describing the unconditional probability distribution of X_{t+1} and X_t respectively.

For a Chain with continuos sample space the sum is replaced by an integral, $p(\cdot)$ is a density function and the analoguos expression is:

$$p_{t+1}(\theta_{t+1}) = \int K(\theta_t, \theta_{t+1})p_t(\theta_t)d\theta_t \tag{2.8}$$

Stationarity and Uniqueness of the invariant distribution

A Markov Chain used for Monte Carlo simulation has to convergence to the target distribution we want to simulate, then it is natural to ask if there exists a stationary distribution such that $p_{t+1} = p = p_t$. Mathematically, the invariant distribution p has to be the solution of the vector or functional equations:

$$p' = p'K \tag{2.9}$$

$$\text{or} \quad p(\theta') = \int_{\Theta} p(\theta)K(\theta, \theta')d\theta \tag{2.10}$$

The existence of this solution is subjected to some verifiable conditions.

In the discrete case, the vector equation [2.9] implies that $p(I - K) = 0$. Thus, p' is the eigenvector associated with the unit eigenvalue of the transition matrix K. Since a square matrix with non-negative entries and rows summing to one always has a real unit eigenvalues, a stationary distribution always exist. However, we have to rule out the presence of multiple unit eigenvalues, since, for our simulation purposes, the stationary distribution must be unique.

In a finite discrete state space, sufficient condition for the existence of a unique stationary distribution depends on the so called *Irreducibility condition.* A chain is irreducible if all the states can be reached from any initial state after a finite number of steps, that is all the states communicate each other.

Let have a look a little close at the algebra of the finite discrete chain. Repeated applications of [2.6] yield to:

$$p_{n+m} = p_n K^m \tag{2.11}$$

State j is said to be accessible from the state i if there must be a postive probability to move from i to j after a positive number m of steps.

Now the j-th element of p_{n+m} is the probability of the chain being in state j at time $n+m$. By the law of conditional probability this must be equal to the sum over i of the probabilities of being in state i at time n times the conditional probability of moving from state i to state j in m steps.

$$P(\theta_{n+m} = j) = \sum_i P(\theta_n = i) P(\theta_{n+m} = j | \theta_n = i) \tag{2.12}$$

Comparison of the last two expressions shows hat K^m must contain the elements of the form $P(\theta_{n+m} = j | \theta_n = i)$. K^m is called the m-step transition probability matrix and its elements are denoted by $p_{ij}(m)$.

State j is accessible from the state i if and only if a number of $m > 0$ steps exists such that $p_{ij}(m) > 0$.

If the state i is accessible from state j and viceversa, the two states are said to communicate and the sets of states communicating form a comunication class. A chain is irreducible if and only if there is only one communication class, that is all the states communicate each other and the transition matrixes $K^1, K^2, ..., K^m$ contains all positive elements.

This condition highlights the connection between Irreducible Chian and Markov Chain Monte Carlo: in MCMC, where the target distribution is the stationary distribution, the algorithm needs to visit all possible states (all possible values of θ corresponding to he parameter space), regardless of its starting point. Thus, an irreducible chain will certainly returns to any initial state because $p_{ij}(m) > 0$ for some m.

When the state space is not finite and discrete, the existence of a unique stationary distribution needs extra conditions.

In a countably infinite state space the extra condition to be mentioned is the positive recurrence. A chain is positive recurrent if any state has a finite and positive return time with probability one, where the return time is defined as the time a chain starting from i takes to be again in the state i that is :

$$T_i = \inf \{n \geq 1; X_n = i\} \tag{2.13}$$

We need to add positive recurrence when the state space is countable. This was unnecessary in the finite case because irreducibility implies positive recurrence.

Positive recurrence ensures that the chain returns to any particular state often enough to do an effective tour of the parameter space.

Convergence

Since in MCMC methods samples from the target distribution are obtained only asymptotically, the existence of unique stationary distribution is not the only requirement we need.

Convergence needs to be imposed, since the realizations of the chain must be used as sample drawn from the target distribution (the posterior we want to analyze): the chain needs to be designed in a way that, not only its stationary distribution is the target distribution, but also this invariant distribution converges to the target wherever the chain is initialized.

Convergence pertains the limiting behaviour and it can be expressed as:

$$\lim_{n \to \infty} p_{ij}(t) = p_j \quad \text{for } i, j. \tag{2.14}$$

A chain having alternating behaviour between two or more states, even if has a stationary distribution, does not converge. Convergence implies to impose the chain to be aperiodic. A chain is aperiodic if all states has probability $p_{ij}(n) > 0$ for all sufficiently large number of steps n.

The last condition to impose to the chain in order to be used for simulattion purposes is related to the behaviour of the sample average and it is called *Ergodicity*.

Ergodicity

Let θ be the vector of parameters and $p(x|\theta), p(\theta)$ and $p(\theta|x)$ be the likelihood, the prior and the posterior respectively.

If we have a chain $\left\{\theta^{(i)}\right\}_{i=1}^{n}$ and we would want to evaluate some features of the posterior stationary distribution $p(\theta|x)$ as the expected value: $f(\theta) = E(\theta|x)$, common practice suggest to average the realizations:

$$\widehat{E_p(f)} = \frac{1}{n}\sum f(\theta^{(i)}).$$

It must be assured that the estimate above converges to the corresponding expectation with respect to the stationary distribution. This means that an ergodic theorem to be applicable to the chain need to be mentioned. The main theorem, at least for a distribution on a countable space, is the following:

Theorem 4 *Ergodicity*

An irreducible, positive recurrent chain, starting from any initial distribution, and with stationary distribution p, is such that:

$$P\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n-1} f(\theta^{(k)}) \to E(f)\right\} = 1 \tag{2.15}$$

where $E(f) = \sum_{i=1}^{n} p_i f_i$ *where* $\{f_i\}$ *are the values of the function at each possible state of the chain and* $f(\cdot)$ *is bounded.*

In words, the average of function of realized states converges (almost surely) to the expectation of the function with respect to the stationary distribution.

The effects of this results are the following. Suppose we have a target distribution p and a random variable θ and we construct a Markov Chain whose unique stationary distribution is p. If we run the chain through many steps, from any starting point deterministically or randomly chosen, we will find that the average of the function of the successive realizations $\left\{f(\theta^{(k)})\right\}$ will approach the expectation of the function $E(f)$ with respect to p. Thus, ergodicity allows to use samples of the chain in order to learn about the posterior distribution $p(\theta|y)$ and its finite expectation value. Then, MCMC turns the theory around, the relevant question is not wheter a given kernel has a unique stationary distribution but whether we can find the kernel cooresponding to the given stationary distribution.

In MCMC the posterior distribution is the target distribution and the method wants to find the transition matrix whose stationary distribution is the posterior.

2.5.4 Gibbs Sampler

In the last decades the use of Markov Chain Monte Carlo methods to simulate complex non standard multivariate distribution have been increasing since Gelfand and Smith (1990) introduced the use of Gibbs Sampler into statistical modeling. The Gibbs Sampler algorithm is one of the best known, and its impact on Bayesian statistics, following the work of Tanner and Wong [64] and Gelfand and Smith [33], was remarkable.

In Bayesian inference we are interested in finding the joint posterior distribution of parameters given data. Difficulties arise when estimation involves multi-dimensional integrations which are feasible only for small-scale models. Gibbs Sampling circumvent the problem by partitioning the parameter space and sampling from the produced full conditional posterior distrubutions.

Let be θ the p-vector of parameters and $p(\theta), p(y|\theta), p(\theta|y)$ the prior, the likelihood and the posterior respectively. The θ vector can be partitioned in different ways: $\theta = (\theta_{(1)}, \theta_{(2)}, ...\theta_{(B)})$ where $\theta_{(j)}$ for $j = 1, 2, ...B$ is a scalar or vector and B the number of partinions. For the simplicity of notation the following analysis will be limited to the case where $B = 2$. Gibbs Sampler constructs the Kernel of Markov Chain using the full conditional distributions: $p(\theta_{(1)}|\theta_{(2)}, y)$ and $p(\theta_{(2)}|\theta_{(1)}, y)$:

$$K(\theta^t, \theta^{t+1}|y) = p(\theta^{t+1}_{(1)}|\theta^t_{(2)}, y) \;\; p(\theta^{t+1}_{(2)}|\theta^{t+1}_{(1)}, y) \tag{2.16}$$

The algorithm samples from the posterior by running the chain until its realizations comes from the target distribution (the posterior), since this kernel satisfies the property guaranteing the convergence to the posterior distribution (See Appendix A.1).

$$\int K(\theta^t, \theta^{t+1}|y)p(\theta|y)d\theta = \int p(\theta_{(1)}^{t+1}|\theta_{(2)}^t, y) \;\; p(\theta_{(2)}^{t+1}|\theta_{(1)}^{t+1}, y) \;\; p(\theta_{(1)}^t, \theta_{(2)}^t|y) \;\; d\theta_{(1)}^t d\theta_{(2)}^t$$
$$= p(\theta_{(2)}^{t+1}, \theta_{(1)}^{t+1}, y) = p(\theta^{t+1}|y) \qquad (2.17)$$

The idea is to make move in the chain by sampling in turn from each of the available conditional components.

The steps of the algorithm is the following:

1. Impose an initial value $\theta_{(2)}^0$ and sample $\theta_{(1)}^1$ from $p(\theta_{(1)}|\theta_{(2)}, y)$

2. Sample $\theta_{(2)}$ from $p(\theta_{(2)}|\theta_{(1)}^1, y)$

3. Repeat the step 1 and 2 S times.

After S replications, the algorithm will lead to a sequence of drawing $\left\{\theta^i\right\}_{i=1}^S = \left\{\left(\theta_{(1)}^{i\prime}, \theta_{(2)}^{i\prime}\right)'\right\}_{i=1}^S$ which can be considerred a sample from the posterio distribution $p(\theta|y)$.

Two main issues need to be discussed: the choice of the initial condition and the number of iteration (the lenght of the sequence) to be undertaken.

Initial condition may affect the simulation results. Therefore, the procedure discards an initial random sample (the burnin sample) with size S_0 : $\left\{\left(\theta_{(1)}^{i\prime}, \theta_{(2)}^{i\prime}\right)'\right\}_{i=1}^{S_0}$ and just the remaining S_1 draw: $\left\{\left(\theta_{(1)}^{i\prime}, \theta_{(2)}^{i\prime}\right)'\right\}_{i=S0+1}^{S}$ (with $S_1 = S - S_0$) is averaged to produce an estimates of posterior features of interest.

The generic function of random parameters $g(\cdot)$ can be estimated as:

$$\widehat{g}_{S_1} = \frac{1}{S_1} \sum_{i=S_0+1}^{S} g(\theta^s) \rightarrow E(g(\theta)|y). \qquad (2.18)$$

Choosing S sufficiently large, the weak law of large number garantees that $\widehat{g}_{S_1}$ converges to the true value $E(g(\theta)|y)$.

We compute the approximation error $\left(g(\theta) - \overline{g(\theta)}\right)^2$ using Markow Chain sequence and setting the lenght S_1 allows to rearrange the probability statement to find an approximate 95% confidence interval for $E(g(\theta|y)$ of the form:

$$\left[\widehat{g}_{S_1} - 1.96\frac{\widehat{\sigma}_g^2}{\sqrt{S_1}}, \widehat{g}_{S_1} + 1.96\frac{\widehat{\sigma}_g^2}{\sqrt{S_1}}\right]. \tag{2.19}$$

Another diagnostic check, proposed by Geweke [40], is based on the intuition that, if a sufficiently large number of draws has been taking, the estimates $g(\theta)$, based on the first half of the draws should be essentially the same as the estimate based on the last half. If theese two estiamtes are different, then initial condition has not ran out of its effects on the first half of replication and more draws have to be taken.

For this reason the sample of S_1 draws is divided into three subset: $S_A = \{s : S_0 + 1, ..., S_0 + S_A\}$, $S_B = \{s : S_0 + S_A + 1, ..., S_0 + S_A + S_B\}$, $S_C = \{s : S_0 + S_A + S_B + 1, ..., S_0 + S_A + S_B + S_C\}$, usually for the convergence diagnostic the sample size assigned to the three subset is :$S_A = 0.1S_1$; $S_B = 0.5S_1$; $S_C = 0.4S_1$.

After discarding the middle set, we performe t-test for the statistic:

$$CD = \frac{\widehat{g}_{S_A} - \widehat{g}_{S_C}}{\frac{\widehat{\sigma}_A}{\sqrt{S_A}} + \frac{\widehat{\sigma}_C}{\sqrt{S_C}}} \rightarrow N(0,1) \tag{2.20}$$

where $\widehat{g}_{S_A}$ and $\widehat{g}_{S_C}$ are the two estimates of $E\left[g(\theta|y)\right]$ computing from the draws S_A and S_C respectively, and $\frac{\widehat{\sigma}_A}{\sqrt{S_A}}$ and $\frac{\widehat{\sigma}_C}{\sqrt{S_C}}$ are the numerical standard error

If we can not refuse the null hypothesis of $CD = 0$, then the convergence diagnostic approximates zero, the two extreme sub-samples may be asserted to be indipendent and the number of iterations S_1 are sufficent to calculate the final posterior results. Rejecting the Null Hypotesis of $CD = 0$ indicates that $\widehat{g}_{S_A}$and $\widehat{g}_{S_C}$ are quite different and ore replications need to be taken in order to satisfy convergence.

2.5.5 Metropolis-Hasting Algorithm

Before descrcibing Metrpolis-Hasting Algorithm, let introduce the preliminary version: the Metropolis Algorithm, proposed by Metropolis in 1953[55].

Recall $p(\theta|y)$ the target distribution which the algorithm wants to simulate. Metropolis Algorithm postulates the existence of an alternative, but similar, probability distribution over Θ: $q(\cdot)$, called *proposal distribution.*

This distributiotn depends on the current position of Markov Chain θ and represents the probability of the possible next value of the chain θ' given the current position θ; it must to be interpreted as saying that when the chain is at point θ, the density q generates a candidate θ'.This function must be symmetric, that is:

$$q(\theta|\theta') = q(\theta'|\theta) \tag{2.21}$$

Metropolis Algorithm constructs a chain having $p(\theta|y)$ as a stationary distribution through the following steps:

1. Choose an initial value for $\theta = \theta^0$ and set $t = 0$.

2. Draw a candidate θ^* from $q(\theta^*|\theta^0)$

3. Calculate the ratio $r = \frac{p(\theta^*|y)}{p(\theta|y)}$.

4. If $r > 1$ then set $\theta^{t+1} = \theta^*$, otherwise set:

$\theta^{t+1} = \theta^*$ with probability r,

$\theta^{t+1} = \theta^t$ with probability $1 - r$;

5. Repeat the step S times.

The probability that the candidate θ^* is accepted is called *acceptance probability* and can be expressed:

$$\alpha(\theta^t, \theta^*) = \min\left\{\frac{p(\theta^*|y)}{p(\theta^t|y)}, 1\right\} \tag{2.22}$$

The rational underlying the expression is that the chain is likely to accept the candidate draws that, with respect to the stationary distribution, are more probable than the current value of the sequence. The chain tends to move to the region of the parameter space where the probability is higher, that is the chain tries to simulate a sample from the target posterior distribution whose draws are the more probable realizations. However it is allowed to go "downhill", to pass from the current position to a value less probable, through the probability $1-r$. A useful feature of the algorithm is that the integral referring the normalizing constant $p(y)$ does not have to be calculated since it cancels from the ratio: $p(\theta^*|y)/p(\theta^t|y)$.

Metropolis Algorithm constructs a chain whose values can be repeated over time and it depends from the fact that candidate values θ^* can be repeatedly rejected.

Appendix 2 demonstrate how Metropolis Algorithm satisfies conditions guaranteing the existence of stationary distribution.

Let now introduce the generalization of Metropolis Algorithm, proposed by Hasting. The difference arise from the generalization of the proposal density that is allowed to be not symmetric; then the acceptance probability changes in:

$$\alpha(\theta^t, \theta^*) = \min\left\{\frac{p(\theta^*|y)q(\theta^t|\theta^*)}{p(\theta^t|y)q(\theta^*|\theta^t)}, 1\right\} \tag{2.23}$$

Choice of proposal distribution

Implementation of the M-H Algorithm required the specification of the functional form of the candidate generating density and its parameters (the location and the variance parameters). Very large number of proposal densities has been discussed in the literaure. Considerable work has been devoted to the question about how this choice should be made. Let start this section highlighting the main family proposed in the previous literature.

A common choice is to use as candidate generating densities distributions depending only on the candidate draw: $q(\theta, \theta') = q(\theta')$. These densities specify the so called Indipendence Chain [66] since proposal distribution is indipendent on the current location of θ.

Another choice is the so called *Random Walk Metropolis-Hastings Chain:* the candidate depends on the current position of the chain through the distance between the candidate draw and the current realization: $q(\theta, \theta') = q(\theta' - \theta)$, then, the candidate can be expressed as: $\theta^* = \theta + z$ where z is a noise. The acceptance probability can be be written as:

$$\alpha(\theta, \theta') = \left\{\frac{p(\theta|y)}{p(\theta\prime|y)}, 1\right\} \tag{2.24}$$

Tuning proposal's parameters is another important issue since the location and the spread of the proposal distribution affects the behaviour of the chain and the effeciency of the algorithm.

Specially the spread affects in two dimensions the efficiency of the algorithm. Firstly, the variance influences the acceptance rate (the percentage of times we accept the candidate draw); secondly the spread affects the moves of the chain and which region of the support (the parameter space) will be explored. If the spread is too small, the chain moves too slowly and it will take longer to cover all the support of the density, moreover the low probability regions will be under sampled. If the spread is too large, the candidate draw will be far from the current location, then the acceptance probability will be small and the chain will remain in the same region of the parameter space.

In the Random-Walk Metropolis Hasting Algorithm, Gelman and Gilk [36] proposed to set the acceptance rate equal to 0.25 and the spread equal to 0.23. This value yields infact to maximize the efficiency of the algorithm (measured in terms of the asymptotic variance of the sample mean of the chain).

2.6 Exchangebility

This brief rewiev can not be ended without spending few words about the concept of exchangebility. Framed originally by De Finetti [30], exchangebility is the statistical representation of the symmetry of the behaviour of the events and it is crucial to the development of subjetive probability.

Consider a sequence of random variables $\{x_1, ..., x_n\}$ whose kind of dependence among them need to be specified. The individual random quantities x_i can be supposed to be 'noninformative', in the sense that information provided by the x_i's are independent of the order in which they are collected. The simple form of dependency which accurately describes this judgement of 'similarity' or 'symmetry can be written as:

$$p(x_1, ..., x_n) = p(x_{\pi(1)}, ..., x_{\pi(n)}).$$

The joint density $p(x_1, ..., x_n)$ does not change under any kind of permutations π defined on the set $\{1, ..., n\}$. A sequence of random quantities is said to be exchangeable if this property holds for every finite subset of them. The 'similarity' assumption of exchangeability has strong mathematical implications expressed by De Finetti in its *Representation Theorem.* Formally, the theorem provides an integral representation of the joint density $p(x_1, ..., x_n)$ of any subset of exchangeable random quantities:

Theorem 5 *Representation Theorem*

If $(x_1, ..., x_n)$ is an exchangeable sequence of real-valued random quantities, then there exists a parametric model $p(x|\theta)$, labeled by some parameter $\theta \in \Theta$ which is the limit of some function of the x_i's, and there exists a probability distribution for θ, $p(\theta)$, such that

$$p(x_1, ..., x_n) = \lim_{n \to \infty} \int p(x_1, ..., x_n|\theta)p(\theta)d\theta \tag{2.25}$$

If a sequence of observations is assumed to be exchangeable, then the Representation Theorem proves that there must exist:

- a parametric model $p(x|\theta)$
- a parameter θ
- a probability distribution for θ with density $p(\theta)$.

where $p(\theta)$ describes the prior available information about the parameter, then a Bayesian approach is required. In short, exchangeability is a justification for Bayesian inference.

Any sequence that is Independent and Identically Distributed (IID) is also exchangeable, though the converse is not true: all exchangeable sequences are not IID. Exchangeable sequences are identically distributed, just not necessarily independent. Exchangeability is therefore a broader concept than IID. While frequentist inference makes heavy use of IID, Bayesian inference more commonly uses exchangeability.

A popular benefit of exchangeable sequences in Bayesian modeling is that exchangeability allows more complicated models as the hierarchial models I will proposed in the last chapter of the thesis.

Exchangebility provides a link between the frequency and the subjective view of probablity, a way of combining empiricism and pragmatism.

The starting point is a refusal of the notion of truth, and the related notions of determinism or immutable and necessary laws. De Finetti reaffirms instead the concept of science seen as a human activity, a product of thoughtwhere probability is its main tool.

"no science... - says De Finetti - will permit us to say: this fact will come about because it follows from a certain law, and that law is a asolute truth. Still less will it lead us to conclude skeptically: the absolute truth does not exist...what we can say is this: I foresee that such a fact will come about because past experience and its scientific elaboration by human thought make this forecast seems reasonable to me".

Probability is precisely what make a forecast possible and because a forecast is always refered to a subject with its experience and it convinction, the logical instrument that we need is the subjective theory of probability".

Probabilism represents an escape from the antithesis between absolutism and skepticism and its core is the subjective notion of probability.

Inference can be entirely performed by exchangebility in combination to the Bayes Therorem. If the notion of probability as degree of belief is grounded in an operational definition (betting quotient), probabilistic inference, taken the subjective sense, is grounded on the Bayes rule. Moreover, Bayesianism represents the crossroad where pragmatism and empiricism meet subjectivism. One need to be Bayesian in order to be a subjectivist, but on the other hand subjectivism is a choice to be made if one embrace pragmatism and empiricism as philosophy.

Chapter 3

Electricity Market

This chapter offers a brief analysis of the main characteristic of Italian Electrisity sector, focusing on the Italian Power Exchange (IPEX). The establishment of the Electricity Market derived from the deregulation process which has involved Electricity sector. Since 2004, the exchange of electricity has been opened to competition and the definition of a proper market structure had been the most challenging part of the deregulation process: electricity market must be infact organized in a way that preservation of competition and the cover of all demand profiles mus be constantly ensured.

3.1 Electricity Industry

Electricity industry is a leading industrial sector since it is a fundamental input for the production processes in any industrialised country. Its strategic importance for economic development and its social and environmental impact imposes an effective regulation. For this reason it is not surprising that the electric sector was regulated by public commissions and the tariffs were kept fixed over long periods of time.

The electricity is not an energy source but an energy carrier: it can be produced from any source. Some sources infact must necessarily pass through the electric vector in order to be exploited (eg, nuclear, hydro, other renewables). Hence, we can easily understand, in the energy policy, the importance of diversifying energy sources.

The provision of electrical service is a capital intensive activity both in the production phase and in the distribution phase.

The grid system, which characterizes the national electricity system, involves that the transmission and dispatching activities are subject to very stringent technical constraints, such as:

- Instantaneous and continuous balance between the amount of energy injected into and that withdrawn from the grid. It is said infact that electrical system does not admit neither storehouse nor tails. It is a

  just-in-time prototype: production and consumption must be perfectly synchronised, in order to not compromise the flow in the grid. If supply and demand are not equal at every moment, the only possible equilibrium is the black out. The instantaneous balance between supply and demand and the difficulties connected to trasmission congestions are physical problem strictly linked to market design issues and to competition policy.

- The maintainance of the frequency and the voltage of the grid within a narrow range, in order to ensure the safety of facilities.

- The need for the energy flow on each electrode to not exceed the maximum permissible transit on the electrode itself.

Any deviations from the above mentioned limits, for more than a few second, can quickly lead the system to the black-out. The characteristics of the technologies and the ways in which electricity is produced, transported and consumed make it difficult to comply with these constraints.

In particular, difficulties arise from three factors:

1. **Variability of demand**: power demand exhibits a remarkable variability both in the short term (hourly) and in the medium term (weekly and seasonal).

2. **Lack of storage and dynamic constraints in real-time adjustments of supply**. The crucial features of electricity is the impossibility of storing it in a economically feasible way. It can not be stored in significant quantities, except indirectly, just in the case of the accumulation batteries or the hydroelectric plants through the quantity of water contained in their basins. Moreover, the electrical systems have minimum and maximum limits binding the power flow as well as unit commitments need minimum time to start and change their power

output. It follows that the optimal set of unit commitments, give the available technology, is that it works with a different composition of fixed and variable costs.

3. **Network externalities**: supply and demand must be physically connected via a structure of meshed grids. The electricity is injected into or withdrawn from the grid nodes. All the nodes are connected to each other, and once energy is fed into the grid, it engages all nodes available as in a system of communicating vessels: the elettrical flow chooses a path according to the laws of Ohm (the path of least resistance), determining the balance between injections and withdrawals; this makes the path of energy no traceable, so any local imbalance not promptly corrected spread across the grid through changes in voltage and frequency.

Electrical System is a serivce, since there is simultaneity between supply and consumption and it accounts for five main activities.

- **Generation:** it pertains the identification of the source to be used, the construction of the power stations, their operating process and their maintenance. Since generation can take place through a variety of tecnologies, from steam power stations to solar etc.., each technologies has different marginal and fixed costs. For this reason, in every national markets, a wide range of plants covering electricity demand at the same time.

- **Transmission**: it is a regulated activity related to transfering electricity over long distance through the construction, management and maintenance of high-voltage grid.

- **Distribution**: it pertains construction, management and maintenance of low-voltage grid which transfers electricty to the end-consumption centers.

- **Retail**: it refers to the sale contracts and the management of relationships (connection, metering and billing) with small-customers.

- **Dispatching**: it is the main task of Indipendent Syspem Operator and involves the monitoring and the management of the transmission grid.

Moreover, System Operator coordinates generators' activity quickly responding to the fluctuation of demand and guarantees that all market partecipants have equal access to the network.

Until 15 years ago all thoose functions were regulated and subjected to the control of central gouvernement. Electricity industry has been infact characterized by economies of scale in the generation and by the necessity of an extensive trasmission grid in order to deliver the generated electricity to the end-consumers. All theese characteristics made the sector a natural monopoly and imposed electricty firms to be vertically integrated across the different functions. Only one firm was charged with all stages of the supply chain. The monopolist had all the information needed to schedule all the power plants and to minimize the generation costs.

In the last decades liberalization process started in most of the developed countries, the ownership in the electricity sector became private and industry has been split up into the different functions.

3.1.1 The Deregulation Process

The liberalization of the electricity sector has led to overcome the system of vertically integrated monopoly. Generation and retail functions have become open to competition.

Even under deregulation, instead, trasmission and distribution have remained (have been kept) monopolistic: because of their structure, no one could provide competing services in those two sectors. Trasmission and distribution networks are considered infact to be natural monopolies and the access to them must be granted in order to ensure that generators can reach their consumers. Electrical grid is an essential facility and all competitors in the other functions need to have not-discriminatory access to it. For this reason the Indipendent System Operator (ISO) has been in charge with the reliability of the system and the continuous balance between demand and supply. Finally it is important to provide incentives to the investements for new construction and development of the lines.

There are two different solutions for scheduling the unit commitment and ensure to cover demand profiles: the passing dispatching and the merit order dispatching.

In the passing dispatching all the operators who sell electricity through bilateral contracts are asked to produce. The energy price is determined by the parties.

In the merit order dispatching operators must submit offers for sell in a centralized market (power exchange). All offers for sale are selected in ascending order of price, until the demand is met. The price is determined within the stock exchange. In most electricity markets both approaches are used.

The Italian Power System is divided into portions of transmission network, this kind of configuration has been defined for the safety of the trasmission grid and for a quick removal of any congestion caused by trasmission constraints. This portion of trasmission grid are defined zones, and they can be physical geographical areas, virtual areas, or limited production poles.

These areas can be summarized in:

6 geographical zones: North, Center North, Center South, South, Sicily, Sardegna.

8 virtual zones: Austria, BSP, Corsica, Croatia, France, Greece,Slovenia, Switzerland.

4 national virtual zones: the limited production poles consisting in a single unit of production with limited interconnection capacity. They are: Brindisi, Foggia, Pirola, Rossano.

Each geographical or virtual zone is a set of offer points. The offer points is the minimum unit in respect of which injection and withdrawal programs are scheduled as a result of the acceptance of offers to sell or purchase in the Electricity Market or in execution of bilateral contracts.

When electricity is injected, usually the offer point coincides with a single plants of generation (the unit commitment) which converts the energy provided by any primary source into electrical energy.

In the case of withdrawal schedules, the offer points to buy may correspond either to single unit of consumption or to aggregate of sampling points.

3.1.2 The Design Factors

Transition from state-owned monopolies to competitive markets was not always smooth, skepticism and concernes had been raising in many countries; market structure affects infact competition and for this reason the design of deregulated electricity markets offer economists a changelling opportunity. They have been attempting to design well functioning markets that gives players the correct incentives to improve production efficiency and limit market power. In the recent years many economists have focused on the effects that market design may have on equilibrium prices. In this new context, two major problems need to be faced:

1) How to correctly schedule the different power plants given the operators' availability in order to minimize the costs of providing the service in the short term

2) How to guide investment decisions in order to minimize the generation cost in the long run.

Together with the production efficiency purpose, power market has to be designed in a way to safeguard competition, limiting the market power of supplier and assuring end-consumers of competitive prices. The market structure affects infact the consumer reactivity to change in price, that is the elasticity.

The main design factors impacting on demand elasticity are listed below.

- **Optional versus compulsory pool**: if partecipation in the pool market is compulsory, consumer are fully exposed to the pool price. On the other hand, if economic agents can make bilateral agreements with producers they are shielded from price volatility. Even in a compulsory market consumers may reduce volatility risk by financial contracts (contracts for difference).

- **One or two sided market**: demand can be taken costant on the basis of load forecasting programs or purchasers' bids can be ranked in a descending order forming in this way market demand. Usually, when eligible consumers take part in the pool market and directly demand their energy loads they can indicate the price they are willing to pay.

- **Simple or complex bid**: offers and bids submitted can show just a single quantity and price or they can be designed to reflect the overall structure of running costs faced by plants. In the first case generators assume the risk related to the start-up of its plants, while in the latter case this risk is borne by pool market. Theoretically, when unit commitments can structure complex offer taking into account the overall generation costs, the resulting prices are usually lower since producers, reducing the risk they bear, reduce their generation costs. On the other hand, when bids are simple, volatility is reduced and price behaviour is more predictable.

- **Market Timing**: prices can be determined ex-ante (before the delivery takes place) or ex-post (when the delivery is executed). Ex-ante prices admit consumers to react and change their consumption profiles. To an economist prespective, fixing prices ex-post is deeply unsatisfactory, since it denies any real interaction between supply and demand.

- **Capacity payments**: generators can be paid either on the basis of the generator capacity they will make available or on the basis of the electricity they actually produced. When the payment refers to capacity, the goal is to encourage producers to keep their profitable units avaliable ensuring to cover all the demand profiles. However the capacity profits and the resulting prices increase non-linearelly as the difference between the capacity and the actual load decrease during the peak periods.

- **Geographically differentiated pricing**: price paid by end-consumers can be uniform over the all market or can be geographically differentiated because of congestions due to trasmission constraints. The purpose of the single price is not to penalize the geographical areas characterized by less efficient power plants or less capacity load available. On the other hand, when price are differentiated, there are more incentives for end-users to pay more attention to their consumption profile.

- **Price capping**: limiting value of the electricity price can allow rational customers to stop consuming when the prices exceed this values, making demand more predictable.

3.1.3 The Italian Power Exchange

As in other international experiences, the creation of a market responds to two specific requirements:

• promoting competition in electricity generation, sale and purchase, under criteria of neutrality, transparency and objectivity, through the creation of a market place;

• ensuring the economic management of an adequate availability of ancillary services.

The organization and the management of the italian electricity market has been entrusted GME. Unlike other European markets, Italian Power Exchange (IPEX) is not a purely financial market aimed only to the definition of prices and quantities, but it is a physical market where injection and withdrawal profiles are scheduled and really delivered.

The Electricity Market is articulated in the Spot Electricity Market (MPE), Forward Electricity Market and the Financial Derivatives Market (IDEX).

The Spot Electricity Market is divided into three submarkets:

The **Day-Ahead Market (MGP)**, which is the venue for the trading of electricity supply offers and demand bids for each hour of the next day. All electricity operators may participate in the MGP. GME accepts Offers and Bids by the merit order, taking into account the curent trasmission constraints. Accepted supply offers are remunerated at the Zonal Clearing Price, while accepted demand bids are remunerated at the National Single Price (PUN). The accepted Offers/Bids determine the preliminary Injection and Withdrawal Schedules of each Offer Point for the next day.

The **Intra-Day Market (MI)**, which has replaced the existing Adjustment Market, it is venue for the trading of electricity supply offers and demand bids which modify the Injection and Withdrawal Schedules resulting from the Day-Ahead Market. GME accepts the Offers and Bids submitted into the MI by merit order, taking into account the Transmission Limits remaining after the Day-Ahead Market. Accepted Offers and Bids are remunerated at the Zonal Clearing Price and they Bids modify the preliminary schedules determining the revised injection and withdrawal schedules for the next day.

The **Ancillary Services Market (MSD)**, it is the venue for the trading of supply offers and demand bids in respect of ancillary services. This market is exentially used to acquire resources for relieving intrazonal congestions, procuring Reserve Capacity and balancing the injections and withdrawals in the real time. Participation in the MSD is restricted to units that are authorised to supply ancillary services and to their dispatching users. Participation in the MSD is mandatory. The MSD produces two separate results:

1) the first result (Ex-Ante MSD) concerns Offers and Bids accepted on a scheduled basis for relieving congestions and creating an adequate reserve margin;

2) the second result (ex-post MSD) concerns Offers and Bids accepted in real time for balancing injections and withdrawals (by sending balancing commands). The Offers and Bids accepted in the MSD determine the final

injection and withdrawal schedules of each Offer Point. In the MSD, Offers and Bids are accepted by economic merit order, taking into account the need for ensuring the use of the system. Offer and Bids accepted in the MSD are valued at the offered price (the Pay as Bid mehod).

3.2 The Day-Ahead Market

The Day-Ahead Market (MGP) aims at the wholesale trading of electricity where hourly blocks of power for the next day are negotiated. In this market both the injection and withdrawal programs for the next day are defined in oreder to reach the equilibrium prices and quantities.

The MGP is organized according to an implicit double auction model and the most of the transactions takes place in this market. The session opens at 8 a.m. on the ninth day before the delivery-day and closes at 9.15 a.m. on the day before the delivery is executed.

During the session, market partecipants submit offers to buy or sell that indicate the amount of energy and the maximum price (or the lowest price) at which they are willing to buy (or sell). In particular:

- The offers to buy (BID) represent the willingness to purchase an amount of energy that does not exceed that specified in the offer at a price no higher than that reported in the same offer. In the demand side operators can refer their bids only to the withdrawal unit points (sampling unit points).

- The offers to sell (OFF) express instead the willingness to sell an amount of energy not greater than that specified in the offer and at a price not lower than that indicated in the same offer. In the supply side operators can relate offers only to the injection points. If the offer is accepted, the producer undertakes to enter in the network, in a given period, the amount of electricity specified in the offer. Moreover, each offer, to sale and purchase, must be consistent with the physical constraints of the corresponding unit point.

The Day-Ahead Market is a zonal market, reflecting the structure which the national trasmission grid is divided in. Each zone is characterized by an insufficient interconnection capacity and when a congestion occurs the selling price is zonal differentiated: selling price is lower in the upstrem area of congestion and higher in the downstream ones. In depth, when the market session closes, the GME start the process for the resolution of the market. For each hour of the next day, the algorithm accepts all the bids and offers in order to maximize the value of trading, within the limits of maximum transit between zones.

The process of acceptance can be summarized as follows:

All offers to sell are sorted according an ascending price order forming aggregate supply curve,while bids are ordered by descending price order drawing the aggregate demand curve.

The intersection between the two curves derives the total quantity traded, the equilibrium price, the accepted BID and OFF.

If electricity flows resulting from the programs do not violate any transition constraints, the equilibrium price is unique for all the zones. The accepted offers to sale are those whose sale prices are not higher than the equilibrium price, while the accepted bids are those whose purchase prices is not lower than the equilibrium price.

If at least one trasmission constraint is violated, sale price are zonal differentiated and the algorithm starts the so called "Market Splitting Mechanism". It splits infact the market into two zones, one for the export, which includes all zones upstream of the bond, and one for the import, which includes all areas downstream of the bond, repeating in each of the two areas the process described above: i.e. it derives in each zone the corresponding aggregate supply and demand curve. The outcome are two equilibrium zonal price zone (p_{z_1} and p_{z_2}). In particular, p_{z_i} is greater in the area of import and is smaller in the area of export. If, within each zone, the resulting equilibrium quantities violate further transition constraints, the splitting market process goes on within the zones in order to obtain an outcome consistent with the grid constraints.

With regard to the purchase price of electricity, GME has implemented an algorithm that, given congestion and differentiated zonal sale prices, apply just a single national purchase price (PUN), that is the average of the zonal sale prices weighted with the zonal consumptions. The PUN applies only to withdrawal points belonging to national geographical areas.

The mechanism of market splitting is an 'implicit auction" for the non-discriminatory allocation of the transit rights.

Programs resulting from bilateral negotiations contribute to the derivation of the outcome of the MGP. The energy traded through bilateral contract participates to the process described above since it engages itself portion of the transmission capacity available and contributes to determine the zonal consumption weigths used for the computation of PUN.

3.3 Dataset and Descriptive Statistics.

The aim of this session is to provide an exhaustive analysis of Electricity Market, using data referring the demand side. The preliminary investigation of the Dataset I used wants to highlight the main features of the Day-Ahead Market as:

- the portion of elastic and inelastic bids
- the average relative frequency and the portion of market demand covered by the Single Buyer
- the average relative frequency and the portion of market demand covered by Bilateral Contracts
- the frequency of market segmentation
- the national single purchase price (PUN).

3.3.1 The GME Dataset

The Data Gathering Process had been disclosed an challenging task, since it had involved the download of GME daily data and their collection in monthly dataset starting from January 2011 to June 2012. At the end of the procedure, 1.5 millions of raw observations was available in each monthly Dataset.

Each raw observation is identified by the following variable:

Purpose: Purpose of BID or OFF, where BID pertaining the participant's purchases and OFF the participant's sales.

Status: the state characterizing bid and offer (Accepted or Rejected) after market execution.

Unit Reference: The identification number of the unit point respect of which the injection or withdrawal programs are defined.

Interval: The relevant period to which the bid or the offer refers.

Bid Offer Date: The flow date of the bid or the offer.

Transaction Reference: GME's identifier of the bid/offer.

Merit Order: The ranking of the offers derived by the market solution alghoritm.

Operator: The registred name of the partecipant.

Zone: The zone to which the unit belongs.

Transaction Reference: The identification number of the offer.

Grid Supply Point: The relevant nodes of grid which the unit is associated with.

Bilateral: it is categorical variable that indicates whether the bid or the offer comes from wholesale market or bilateral contract.

Quantity: The volume submitted by partecipants

Energy Price: The price submitted by partecipants.

Awarded Quantity: The volume awarded by the market

Awarded Energy Price: The price awarded by the market.

First we depurated datasets from observations referring to the supply side (OFF).

In each monthly dataset Bid are about 400-450 thousand observations, accounting for the 20-25% of the total amount of offers.

3.3.2 Dataset Preliminary Analysis

Deep investigation on bid datasets shows that there is a huge amount of observations where price is not specified, expressing the maximum willingness to pay (that is lower responsiveness to a change in the energy price) since, in principle, they are willing to pay any price resulting from market clearing mechanism. These BIDS refer to consumers which are not aware of the market price signals and have a perfectly inelastic behaviour. GME assigns to these bids a fictitious price equal to the supply price cap that is equal to 3000/MWh.

This kind of information need to be processed into the model, since higher willingness to pay resolves in lower level of elasticity and denotes the presence of different kinds of consumers. In the next chapter it will be discussed in depth how to manage the presence of heterogeneous buyers with different price reactivity. At the moment, let be shown tables comparing the hourly average absolute frequencies and the hourly average quantities referring to the both kinds of bids.

	January		Febbraio		March		April		May		June	
Hour	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic
1	437.19	6.48	437.19	6.48	432.65	5.65	422.47	8.63	419.10	6.48	423.43	10.73
2	440.94	9.26	440.94	9.26	435.68	8.65	425.23	9.40	421.77	7.81	424.27	11.83
3	440.94	10.39	440.94	10.39	435.52	10.32	424.00	10.10	421.71	8.35	424.70	12.10
4	440.74	11.45	440.74	11.45	436.13	11.90	423.53	10.70	421.77	8.68	424.63	12.53
5	440.84	11.84	440.84	11.84	435.71	11.29	423.73	10.87	421.58	8.68	424.63	12.40
6	441.61	9.65	441.61	9.65	434.97	7.65	424.60	9.93	421.94	8.97	425.53	12.20
7	440.39	9.32	440.39	9.32	433.06	5.61	421.00	8.30	419.23	8.42	421.03	11.80
8	437.97	8.03	437.97	8.03	428.97	7.26	419.30	7.53	416.71	9.32	419.13	9.93
9	431.65	7.10	431.65	7.10	423.48	3.81	413.53	5.30	411.10	5.03	415.37	9.13
10	432.29	6.23	432.29	6.23	425.10	2.90	411.77	4.93	409.94	3.90	414.97	8.20
11	432.65	6.19	432.65	6.19	424.68	3.06	412.87	5.37	410.35	4.29	415.10	7.73
12	432.74	6.77	432.74	6.77	425.16	3.48	412.67	5.53	410.29	4.74	415.27	7.73
13	434.26	9.42	434.26	9.42	424.58	5.68	411.33	6.63	410.10	9.06	414.03	10.10
14	434.58	10.42	434.58	10.42	424.48	6.23	411.27	6.83	409.61	8.13	414.13	9.97
15	434.26	8.39	434.26	8.39	423.94	5.48	410.77	6.57	409.71	5.42	414.70	9.40
16	431.06	7.23	431.06	7.23	423.81	4.94	409.47	6.57	409.42	4.97	414.93	9.20
17	432.61	7.35	432.61	7.35	424.26	4.71	409.73	6.53	409.13	5.10	414.77	8.97
18	432.74	3.84	432.74	3.84	422.00	4.52	409.47	7.53	411.00	5.81	412.67	9.63
19	431.90	4.03	431.90	4.03	422.39	2.26	409.37	7.27	410.13	7.97	410.73	9.20
20	432.03	5.71	432.03	5.71	423.42	1.16	412.07	7.17	411.39	7.74	412.60	9.13
21	435.00	6.65	435.00	6.65	428.81	2.81	420.87	4.57	418.48	4.65	420.13	9.37
22	433.84	6.42	433.84	6.42	429.65	3.90	419.17	6.13	418.55	4.94	419.67	8.27
23	435.10	5.26	435.10	5.26	429.23	4.39	419.03	5.67	418.06	5.35	420.00	8.63
24	437.45	6.81	437.45	6.81	415.55	4.45	418.47	6.63	415.90	5.94	419.17	10.20

Elastic-Inelastic Bid. Houly Average Abs. Frequency.Jul-Dec 2011.

	July		August		Septmber		October		November		December	
Hour	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic
1	432.26	10.13	425.06	9.84	425.83	10.60	428.97	11.45	424.93	12.60	422.45	13.97
2	432.87	10.26	425.42	9.74	425.20	10.97	429.74	11.39	427.23	13.90	422.52	15.52
3	432.35	10.00	424.84	9.84	424.20	11.00	430.42	11.61	425.43	15.57	421.48	16.81
4	432.03	10.23	424.16	10.00	424.03	11.00	430.45	12.26	425.37	16.33	420.58	18.48
5	432.13	10.26	424.10	10.03	424.27	11.00	430.97	12.23	426.73	16.10	421.03	18.97
6	433.35	11.29	425.32	10.94	426.57	11.57	431.45	12.29	428.77	15.67	421.65	17.94
7	426.32	10.42	416.68	10.32	427.33	10.50	426.23	11.48	426.50	15.47	420.23	15.55
8	424.03	10.55	416.84	9.97	427.27	9.73	424.97	11.61	425.93	16.07	419.74	16.48
9	421.77	9.52	411.74	10.10	423.57	9.47	421.65	11.35	421.20	15.20	416.77	16.39
10	421.23	7.65	411.52	9.52	422.13	9.17	421.13	11.61	421.17	16.10	415.39	17.45
11	420.77	6.61	409.74	9.32	421.33	9.43	420.71	12.00	420.47	17.13	415.35	18.55
12	420.58	6.58	410.52	9.35	420.53	9.63	420.90	12.16	420.80	17.33	414.58	19.06
13	419.39	8.68	410.19	9.48	417.57	10.10	417.68	12.52	416.20	18.17	411.52	19.26
14	420.77	10.32	410.13	9.81	418.60	10.63	417.71	12.26	417.73	16.57	412.55	18.19
15	420.94	9.81	410.29	9.52	419.50	10.43	418.16	11.48	419.03	15.30	412.84	17.35
16	421.03	9.35	410.45	9.35	421.00	9.93	418.87	11.10	419.20	15.03	413.00	17.32
17	422.45	8.90	412.61	9.10	422.33	9.47	419.06	11.06	421.40	16.27	416.90	17.52
18	421.39	8.29	412.68	8.77	420.73	9.83	418.39	11.26	426.93	10.50	420.03	12.68
19	420.77	9.03	412.42	8.55	419.67	9.57	419.74	11.29	424.83	11.77	419.77	14.23
20	420.06	9.10	412.52	8.71	419.33	9.13	423.42	9.19	424.70	15.37	421.68	15.16
21	426.87	9.97	419.10	9.16	424.43	9.67	425.74	10.48	426.93	17.03	422.74	17.13
22	427.06	8.42	418.03	9.42	423.07	10.53	424.61	11.58	424.03	15.13	418.48	15.65
23	425.48	8.87	416.84	9.03	422.90	10.20	421.32	11.06	419.63	13.17	414.35	12.81
24	425.81	9.61	417.19	9.19	422.67	10.17	421.52	11.29	420.40	11.77	416.26	11.58

Elastic-Inelastic Bid. Houly Average Abs. Frequency.Jul-Dec 2011.

	July		August		September		October		November		December	
Hour	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic
1	465.65	12.90	464.58	13.32	460.60	15.13	460.29	15.29	475.03	22.00	479.13	23.16
2	464.48	13.77	462.52	13.23	460.90	15.03	461.39	17.84	473.10	28.13	478.10	30.35
3	464.10	13.97	461.87	13.10	460.83	15.93	461.26	20.26	472.23	32.30	478.32	37.29
4	463.52	14.45	461.48	12.97	461.40	17.07	461.19	21.94	472.30	34.03	478.06	41.19
5	463.03	14.58	461.55	13.13	460.40	16.53	461.19	21.87	471.40	34.17	477.74	42.55
6	462.48	15.55	461.03	13.32	461.10	16.63	462.55	21.23	474.23	30.23	478.87	36.97
7	464.23	17.42	461.55	13.03	463.30	14.73	463.97	17.13	477.83	20.73	478.52	26.97
8	464.10	17.35	459.58	13.39	461.13	12.23	466.71	10.58	475.93	14.10	480.26	17.74
9	466.58	15.00	460.61	13.03	460.97	10.57	465.10	9.06	475.23	12.13	482.61	15.35
10	468.84	15.97	462.35	13.84	461.33	11.07	466.29	9.32	477.07	12.83	484.65	14.94
11	469.32	18.13	463.23	15.58	461.57	11.27	467.06	10.00	479.87	14.10	484.84	18.13
12	469.48	20.48	464.81	16.81	464.30	12.83	468.26	12.32	481.33	16.50	484.48	22.39
13	468.45	25.00	465.26	17.58	463.60	15.77	466.39	14.77	480.90	19.83	483.03	27.71
14	468.35	27.00	465.65	19.19	464.50	20.43	468.10	16.39	480.70	20.30	482.26	28.00
15	468.61	25.81	465.94	19.06	465.37	19.47	469.48	14.68	480.90	16.13	482.03	21.29
16	468.29	23.13	466.16	17.81	464.97	17.30	469.74	12.35	479.27	13.87	481.77	16.29
17	466.58	19.45	464.58	15.35	462.23	13.47	468.74	11.03	478.97	11.43	481.55	14.68
18	463.45	15.58	462.94	12.71	461.00	10.93	466.48	11.55	479.00	8.47	480.06	12.00
19	460.74	11.87	461.16	10.13	458.37	10.93	465.74	12.32	477.43	9.27	477.48	12.00
20	459.23	9.16	459.06	10.48	457.83	12.27	464.29	10.06	475.50	9.73	477.58	12.10
21	461.00	10.35	461.71	11.19	460.77	12.57	463.61	10.77	477.70	10.83	478.03	13.13
22	461.71	10.06	461.32	11.61	458.03	13.47	462.71	11.16	477.60	10.77	476.90	13.81
23	460.94	11.13	459.55	12.81	457.20	13.43	463.42	10.61	475.57	13.00	473.52	14.32
24	461.77	12.16	458.84	12.23	458.97	13.27	463.87	10.97	475.73	14.83	473.29	17.35

Elastic Inelastic Bid. Hourly Average Abs. Frequency. Jul-Dec 2012

	January		February		March		April		May		June	
Hour	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic
1	65.27	17.94	68.31	15.57	67.66	8.04	66.68	5.40	68.26	4.52	71.34	8.10
2	61.61	20.17	64.30	21.37	63.83	11.89	62.90	5.61	64.30	6.23	67.54	9.27
3	60.14	23.25	62.67	25.05	62.28	15.32	61.32	5.78	62.34	5.46	65.55	11.66
4	59.33	24.98	61.94	27.11	61.62	17.79	60.54	6.56	61.56	7.24	64.60	11.77
5	59.61	26.82	62.26	26.21	62.06	16.28	60.90	6.15	61.79	6.56	64.80	11.63
6	61.65	24.86	65.45	24.53	64.57	9.59	63.02	5.57	63.10	9.40	64.80	12.06
7	69.34	13.43	73.95	16.81	71.43	10.95	68.64	5.10	68.21	6.45	70.17	11.66
8	80.24	16.76	86.07	13.14	82.89	16.73	78.47	7.70	79.60	20.57	80.94	12.24
9	91.67	13.48	98.89	9.94	95.35	12.26	89.36	3.75	90.66	9.46	92.11	10.85
10	96.79	16.77	103.48	13.42	98.86	9.79	93.60	3.37	94.96	5.90	97.17	7.81
11	97.91	16.82	103.97	13.07	99.17	10.78	93.70	3.58	95.65	5.78	98.54	8.13
12	97.61	17.76	103.33	14.12	98.33	13.64	93.14	3.58	95.38	7.43	98.65	8.78
13	92.83	26.68	97.61	15.46	93.88	9.40	88.73	6.20	90.91	15.59	95.07	12.48
14	90.36	23.40	94.84	19.58	91.97	12.93	86.60	5.80	89.40	13.42	93.53	10.88
15	91.79	19.02	96.39	16.64	93.45	11.39	87.60	3.49	90.71	8.49	94.38	8.17
16	93.50	16.43	97.78	13.90	93.93	10.59	87.94	3.49	91.07	7.16	94.54	7.91
17	95.82	18.84	98.94	11.79	94.40	10.24	87.69	3.80	91.08	5.39	95.04	7.03
18	101.55	10.88	102.14	16.51	95.33	10.88	86.53	3.11	89.27	7.90	94.16	7.10
19	102.40	13.01	107.16	12.81	99.74	6.06	86.11	4.75	88.27	11.84	92.75	8.86
20	100.43	19.52	105.85	9.10	101.85	9.10	88.66	6.16	88.70	13.57	91.73	8.75
21	93.91	16.95	98.33	11.35	96.15	4.77	91.32	2.96	89.36	4.10	90.02	6.69
22	88.31	14.09	91.91	14.81	89.89	4.49	87.51	4.18	88.62	5.98	90.68	6.41
23	80.12	15.89	83.18	14.89	81.78	6.73	80.37	3.59	82.03	7.74	85.37	6.17
24	72.35	15.81	75.09	15.20	74.77	5.60	73.23	4.36	75.17	4.60	78.67	6.89

Elastic-Inelastic Bid. Hourly Average Quantity Bid. Jan-Jun 2011

	July		August		September		October		November		December	
Hour	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic
1	77.33	5.73	67.60	8.70	71.56	15.90	67.34	13.56	66.29	18.75	65.66	23.62
2	73.50	5.87	64.19	7.79	68.35	16.98	64.30	13.95	63.11	19.98	62.32	25.38
3	71.28	5.82	62.33	8.00	66.78	17.09	62.44	13.77	61.50	22.34	60.67	25.92
4	70.15	5.62	61.21	8.23	65.99	17.13	61.80	15.88	60.75	23.06	59.61	27.54
5	69.98	5.97	60.95	8.33	66.01	17.08	61.86	15.10	60.85	23.10	59.71	27.33
6	70.40	5.49	61.87	7.36	67.62	16.06	63.91	13.82	63.38	19.36	62.02	24.88
7	74.60	6.76	64.39	9.68	73.69	15.46	71.68	12.43	71.38	18.58	69.38	23.19
8	84.45	10.07	70.63	13.34	81.80	17.15	82.06	15.03	82.01	22.97	79.80	25.71
9	95.34	9.46	79.55	12.51	92.07	16.99	91.32	14.07	92.95	23.22	90.56	27.39
10	101.09	8.03	84.83	10.99	96.70	17.15	94.65	14.15	96.38	23.49	95.14	29.40
11	103.12	8.35	86.99	10.72	98.02	17.10	94.95	14.06	96.77	24.45	95.52	30.03
12	103.74	9.38	87.68	10.67	98.28	16.99	94.54	13.91	96.38	24.97	95.13	30.91
13	100.86	10.17	86.44	10.64	94.68	16.53	90.64	14.57	92.47	25.29	91.47	30.34
14	99.37	12.06	85.16	11.07	92.88	16.43	88.77	14.02	90.32	24.18	88.82	28.74
15	100.27	11.76	85.20	10.94	94.03	16.66	89.79	13.83	91.88	21.79	90.04	28.29
16	100.37	10.28	85.08	10.95	94.21	17.19	90.08	13.43	93.29	23.49	91.59	29.03
17	100.74	9.85	85.26	10.57	94.66	17.32	90.83	13.25	96.89	23.49	95.84	27.81
18	99.78	9.23	85.24	9.36	94.03	16.67	91.17	13.78	102.56	22.06	102.30	22.81
19	97.87	8.36	84.92	9.51	93.29	17.68	95.07	12.76	102.74	20.98	101.86	23.13
20	97.19	7.73	85.74	8.35	97.80	16.21	100.00	14.10	101.52	21.12	100.08	24.06
21	95.11	6.26	87.73	7.97	97.19	16.15	94.12	12.61	94.56	22.09	93.06	25.95
22	95.68	4.44	86.16	7.09	91.61	15.14	88.11	11.59	88.56	19.37	87.59	24.93
23	90.67	6.57	80.29	7.64	84.28	16.08	81.22	12.33	81.10	19.45	80.69	21.50
24	84.15	6.84	74.43	7.65	77.83	15.85	74.07	13.91	73.43	18.94	72.93	21.54

Elastic-Inelastic Bid. Hourly Average Quantity Bid. Jul-Dec 2011

Hour	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic
1	62.15	20.48	65.21	38.11	64.28	16.30	58.66	14.05	59.57	14.50	63.94	11.27
2	58.61	23.98	61.58	40.72	61.30	16.48	55.47	13.45	56.66	13.69	60.66	10.32
3	56.95	24.07	59.75	40.69	59.57	18.53	53.84	12.80	55.02	13.11	58.77	10.96
4	56.20	24.78	59.20	39.17	58.96	18.63	53.27	13.54	54.39	13.17	57.99	10.90
5	56.37	22.76	59.59	38.44	59.20	18.25	53.42	13.75	54.66	13.08	57.98	10.60
6	58.73	22.97	62.18	38.84	61.42	18.03	55.47	13.10	55.74	11.78	58.15	11.36
7	66.52	20.07	70.69	31.78	68.37	18.62	60.15	15.17	59.85	13.81	62.41	11.75
8	78.29	21.27	81.91	31.57	78.83	20.18	68.81	16.83	69.83	15.94	71.50	11.92
9	86.72	26.98	91.77	34.60	89.52	18.74	78.39	18.78	79.09	17.90	81.53	14.24
10	90.83	26.59	95.91	35.03	92.64	18.98	81.41	20.02	82.18	19.09	85.71	16.12
11	91.64	27.02	96.26	35.72	92.43	20.81	81.59	20.96	82.64	18.16	87.05	19.00
12	91.31	27.50	95.81	36.05	91.57	21.63	80.76	19.93	82.44	16.74	87.43	19.57
13	87.12	27.77	90.80	36.11	86.75	23.77	77.02	21.18	78.47	17.76	84.19	19.47
14	84.96	27.94	88.76	36.98	84.77	24.71	74.93	22.64	76.79	16.80	82.72	18.77
15	86.46	26.90	90.38	30.99	86.31	22.33	75.95	20.47	77.94	15.46	83.63	17.83
16	87.64	25.30	91.37	27.02	86.78	19.13	76.26	18.12	78.04	13.64	83.83	15.15
17	90.12	24.37	92.94	24.68	87.76	18.62	76.18	16.10	77.49	11.91	83.77	13.55
18	96.38	20.17	96.21	23.22	89.08	18.36	75.27	14.32	76.35	11.46	82.34	12.76
19	97.49	22.84	100.98	31.42	93.83	18.58	75.65	16.13	75.63	15.01	81.02	13.76
20	96.05	23.54	100.47	29.36	97.49	20.58	78.62	15.44	76.53	17.06	80.78	13.09
21	90.53	23.68	94.78	31.65	92.62	20.57	81.86	15.79	77.82	13.58	79.11	10.32
22	84.93	22.39	88.73	26.82	86.98	19.37	78.29	14.19	77.58	11.73	80.06	9.80
23	77.60	23.20	80.58	33.32	78.69	20.98	71.55	14.93	71.21	13.10	75.33	9.42
24	69.81	19.76	72.81	32.51	75.62	19.85	64.47	14.67	64.20	12.95	69.49	11.51

Elastic-Inelastic Bid. Hourly Average Quantity Bid. Jan-Jun 2012

	July		August		September		October		November		December	
Hour	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic	Inelastic	Elastic
1	69.57	13.20	62.76	12.53	61.34	13.22	58.18	14.84	55.13	19.68	55.08	30.22
2	66.57	11.65	59.63	10.76	58.54	11.57	55.40	17.15	52.48	23.22	51.77	30.50
3	64.51	9.11	57.53	9.33	56.82	11.42	53.64	20.06	50.78	23.35	49.61	30.25
4	63.41	8.18	56.21	8.37	55.86	11.82	52.92	20.81	50.14	24.66	48.79	29.37
5	63.22	7.97	55.75	8.31	56.05	11.73	53.10	21.07	50.51	23.02	48.90	29.39
6	63.69	8.44	56.78	7.91	57.81	10.84	55.16	19.16	52.66	22.43	51.11	29.09
7	67.01	9.32	58.07	8.64	63.17	10.28	62.08	18.54	59.03	16.47	57.60	21.33
8	75.58	13.65	63.34	12.78	69.59	13.91	70.73	23.82	66.82	19.51	65.98	20.57
9	85.89	15.62	71.66	13.60	78.39	13.44	77.62	26.00	73.96	18.26	73.29	20.32
10	90.87	16.95	75.75	14.61	82.25	14.76	79.89	26.20	76.26	18.01	76.48	19.52
11	92.35	19.04	77.12	15.74	82.72	14.82	79.78	26.24	76.28	18.78	76.72	22.10
12	93.18	19.88	77.96	17.21	82.70	17.32	79.46	26.71	75.84	20.04	76.32	24.11
13	90.43	18.51	77.04	18.26	79.34	19.94	75.96	25.55	72.48	25.83	73.08	31.20
14	89.43	19.12	76.33	18.53	78.00	21.27	74.43	28.24	71.23	24.20	71.12	29.15
15	90.48	19.56	76.62	19.03	79.07	20.35	75.60	22.00	72.63	17.83	72.11	22.76
16	90.85	19.97	76.82	17.93	79.65	19.30	76.38	20.81	74.14	16.74	73.49	17.98
17	90.95	19.94	77.70	15.74	80.38	16.47	76.76	21.62	77.18	12.86	77.34	16.46
18	90.18	17.19	78.05	13.91	80.13	13.44	77.25	20.48	83.48	9.61	83.98	15.45
19	88.73	15.91	78.63	10.51	80.00	13.36	80.79	18.04	83.87	16.86	83.85	19.26
20	87.61	16.35	79.10	7.59	84.06	11.13	86.07	18.01	82.03	14.24	82.62	17.79
21	85.52	12.15	80.30	6.17	83.16	11.40	81.08	20.45	76.74	14.90	77.70	16.22
22	85.90	11.09	78.36	6.13	77.92	11.29	75.40	17.52	71.34	14.90	72.72	16.99
23	81.11	12.88	72.64	8.74	71.54	12.04	68.56	20.40	65.23	16.13	66.44	16.54
24	75.25	14.58	67.52	11.63	65.53	13.70	62.30	21.84	59.33	15.25	60.18	20.14

Elastic Inelastic Bid. Hourly Average Quantity Bid. Jul-Dec 2012.

Italian Single Buyer is one of the most preminent economic agent operating in the wholesale electricity market. This company was created by GSE with the task of guaranteeing the availability of electricity in order to cover the demand of captive customers. Agents submitting bids are not infact necessary final users. Single Buyer operates by purchasing the required power capacity, by reselling it to distributors on non discriminatory terms and by making possible the application of a single national tarif to customers.

Its institutional role imposes to derive and discuss the main summary statistics. In particular, it was derived the measure referring the hourly average demand share that proxies the portion of electricity demand satisfying captive and domestic consumers. Moreover, it was computed the hourly average relative frequency in order to represent the behaviour variability. Tables enlighten how Single Buyer represents a consistent portion of hourly demand, since its average bids oscillate between the 10% and the 20%. On the other hand, relative frequency is very small, around the 0.5% on average. Tables confirm the role of Single Buyer as an intermediary between power generators and distributors: with a restricted number of bids, Single Buyer covers a big portion of demand related to the captive consumers. Moreover, quantity shares vary considerably across hours, suggesting that captive consumers have a flexible behaviour and they change their consumption profile during the day. The institutional role of Single Buyer as intermediary between the market and captive consumers may suggest the possibility of moral hazard. In the next chapter it will be shown how this issue has been faced in the model.

	January	February	March	April	May	June
1	0.45%	0.45%	0.35%	0.41%	0.42%	0.46%
2	0.35%	0.28%	0.22%	0.23%	0.23%	0.28%
3	0.26%	0.22%	0.22%	0.23%	0.23%	0.23%
4	0.23%	0.22%	0.22%	0.23%	0.23%	0.23%
5	0.22%	0.22%	0.22%	0.23%	0.23%	0.23%
6	0.22%	0.22%	0.23%	0.23%	0.23%	0.23%
7	0.43%	0.42%	0.29%	0.30%	0.31%	0.32%
8	0.47%	0.48%	0.45%	0.45%	0.46%	0.46%
9	0.63%	0.68%	0.47%	0.48%	0.48%	0.47%
10	0.67%	0.70%	0.51%	0.54%	0.53%	0.51%
11	0.68%	0.70%	0.50%	0.52%	0.52%	0.52%
12	0.68%	0.69%	0.50%	0.50%	0.48%	0.54%
13	0.67%	0.68%	0.50%	0.49%	0.48%	0.54%
14	0.61%	0.60%	0.46%	0.48%	0.48%	0.48%
15	0.63%	0.61%	0.47%	0.48%	0.48%	0.47%
16	0.66%	0.64%	0.47%	0.48%	0.48%	0.47%
17	0.67%	0.68%	0.47%	0.48%	0.48%	0.48%
18	0.69%	0.70%	0.56%	0.48%	0.48%	0.51%
19	0.73%	0.73%	0.67%	0.49%	0.48%	0.51%
20	0.75%	0.79%	0.69%	0.63%	0.52%	0.58%
21	0.77%	0.86%	0.69%	0.71%	0.71%	0.69%
22	0.70%	0.69%	0.69%	0.70%	0.64%	0.70%
23	0.68%	0.68%	0.50%	0.48%	0.47%	0.52%
24	0.45%	0.45%	0.46%	0.47%	0.47%	0.47%

Single Buyer. Hourly Average Relative Frequency. Jan.-Jun. 2011.

	July	August	September	October	November	December
1	0.45%	0.46%	0.45%	0.24%	0.23%	0.34%
2	0.45%	0.46%	0.26%	0.23%	0.23%	0.24%
3	0.36%	0.44%	0.23%	0.23%	0.23%	0.23%
4	0.29%	0.36%	0.23%	0.23%	0.23%	0.23%
5	0.27%	0.33%	0.23%	0.23%	0.23%	0.23%
6	0.29%	0.33%	0.23%	0.23%	0.22%	0.23%
7	0.42%	0.46%	0.28%	0.23%	0.23%	0.23%
8	0.46%	0.47%	0.44%	0.41%	0.43%	0.43%
9	0.46%	0.47%	0.46%	0.41%	0.45%	0.46%
10	0.48%	0.52%	0.46%	0.46%	0.46%	0.47%
11	0.54%	0.54%	0.46%	0.46%	0.46%	0.48%
12	0.55%	0.57%	0.46%	0.46%	0.46%	0.47%
13	0.56%	0.58%	0.47%	0.46%	0.46%	0.48%
14	0.50%	0.54%	0.47%	0.38%	0.46%	0.46%
15	0.50%	0.53%	0.47%	0.36%	0.46%	0.46%
16	0.50%	0.51%	0.46%	0.38%	0.46%	0.46%
17	0.51%	0.55%	0.46%	0.46%	0.46%	0.50%
18	0.56%	0.57%	0.46%	0.46%	0.46%	0.61%
19	0.58%	0.57%	0.50%	0.46%	0.49%	0.62%
20	0.66%	0.70%	0.68%	0.47%	0.54%	0.66%
21	0.69%	0.83%	0.69%	0.48%	0.58%	0.67%
22	0.69%	0.70%	0.66%	0.46%	0.46%	0.63%
23	0.63%	0.70%	0.48%	0.46%	0.46%	0.49%
24	0.49%	0.56%	0.46%	0.31%	0.45%	0.47%

Single Buyer. Hourly Average Relative Frequency. Jul.-Dec. 2011.

	January	February	March	April	May	June
1	15.13%	13.27%	11.69%	11.21%	11.25%	11.25%
2	12.23%	10.18%	8.40%	8.02%	8.27%	8.86%
3	10.19%	8.25%	6.68%	6.49%	6.87%	7.35%
4	9.19%	7.53%	6.00%	5.83%	6.25%	6.73%
5	8.90%	7.52%	6.01%	5.74%	6.20%	6.57%
6	9.94%	8.69%	7.07%	6.55%	7.00%	7.59%
7	13.52%	12.23%	11.00%	9.60%	9.57%	9.01%
8	17.95%	17.76%	16.15%	14.57%	13.62%	12.27%
9	20.51%	19.59%	16.88%	14.93%	13.89%	13.16%
10	21.82%	20.72%	18.21%	16.46%	15.39%	14.91%
11	22.10%	20.89%	18.30%	16.59%	15.56%	15.37%
12	22.39%	21.00%	18.37%	16.60%	15.62%	15.68%
13	22.25%	20.76%	18.11%	16.41%	15.42%	15.73%
14	21.08%	19.47%	16.86%	15.07%	14.30%	14.66%
15	21.23%	19.71%	16.98%	15.00%	14.30%	14.40%
16	21.68%	20.04%	17.21%	15.04%	14.41%	14.46%
17	23.16%	20.89%	17.89%	15.41%	14.79%	14.95%
18	25.25%	23.15%	19.80%	16.07%	15.37%	15.54%
19	26.36%	25.04%	22.39%	17.00%	16.08%	16.13%
20	27.33%	26.37%	23.77%	20.98%	18.78%	18.41%
21	27.34%	26.08%	24.14%	22.53%	21.29%	19.78%
22	26.06%	24.63%	22.85%	21.44%	20.56%	20.12%
23	23.42%	21.86%	20.05%	19.09%	18.37%	17.97%
24	19.21%	17.46%	15.57%	15.03%	14.75%	14.79%

Single Buyer. Hourly Average Quantity Bid compared to the total amount of Bid accepted. Jan.-Jun. 2011.

	July	August	September	October	November	December
1	14.00%	15.06%	10.21%	7.76%	8.17%	11.88%
2	11.54%	12.58%	7.96%	4.79%	5.12%	8.28%
3	10.35%	10.99%	6.59%	3.24%	3.56%	6.14%
4	9.45%	10.34%	6.01%	2.63%	2.90%	5.26%
5	8.99%	9.93%	5.90%	2.57%	2.92%	5.18%
6	9.32%	9.99%	6.43%	3.45%	4.11%	6.32%
7	10.45%	10.94%	7.98%	6.58%	7.47%	9.59%
8	12.68%	13.45%	12.02%	12.16%	13.76%	14.67%
9	13.46%	13.91%	11.76%	12.04%	12.99%	15.27%
10	15.34%	15.98%	13.19%	12.95%	14.08%	17.08%
11	16.12%	16.64%	13.54%	13.09%	14.25%	17.50%
12	16.56%	17.10%	13.77%	13.18%	14.42%	17.62%
13	16.88%	17.50%	13.79%	12.94%	14.31%	17.55%
14	16.01%	16.66%	12.66%	11.95%	13.14%	16.27%
15	15.84%	16.28%	12.50%	11.87%	13.37%	16.34%
16	15.88%	16.15%	12.63%	12.10%	13.86%	16.98%
17	16.50%	16.78%	13.22%	12.54%	15.41%	18.99%
18	17.28%	17.60%	14.02%	13.58%	17.52%	20.95%
19	17.86%	18.35%	15.11%	15.84%	18.69%	21.79%
20	19.88%	20.86%	19.05%	19.20%	21.01%	23.78%
21	21.23%	23.59%	19.68%	19.81%	21.17%	24.03%
22	21.45%	22.32%	18.76%	18.40%	19.60%	23.11%
23	19.90%	20.66%	16.71%	15.86%	16.88%	20.53%
24	17.17%	18.33%	13.34%	11.96%	12.57%	16.24%

Single Buyer. Hourly Average Quantity Bid compared to the total amount of Bid Accepted. Jul.-Dec. 2011.

	January	February	March	April	May	June
1	0.44%	0.38%	0.43%	0.40%	0.43%	0.40%
2	0.29%	0.21%	0.26%	0.29%	0.26%	0.32%
3	0.24%	0.19%	0.22%	0.22%	0.22%	0.25%
4	0.23%	0.19%	0.22%	0.22%	0.22%	0.21%
5	0.22%	0.19%	0.22%	0.22%	0.22%	0.21%
6	0.22%	0.19%	0.22%	0.22%	0.22%	0.22%
7	0.41%	0.36%	0.38%	0.35%	0.38%	0.27%
8	0.44%	0.39%	0.45%	0.44%	0.45%	0.41%
9	0.45%	0.42%	0.49%	0.40%	0.49%	0.31%
10	0.49%	0.44%	0.51%	0.48%	0.51%	0.39%
11	0.52%	0.45%	0.50%	0.49%	0.50%	0.39%
12	0.52%	0.46%	0.50%	0.46%	0.50%	0.39%
13	0.53%	0.46%	0.49%	0.47%	0.49%	0.39%
14	0.49%	0.44%	0.47%	0.39%	0.47%	0.34%
15	0.49%	0.44%	0.48%	0.39%	0.48%	0.34%
16	0.51%	0.44%	0.48%	0.40%	0.48%	0.36%
17	0.52%	0.44%	0.48%	0.41%	0.48%	0.40%
18	0.63%	0.55%	0.50%	0.45%	0.50%	0.42%
19	0.67%	0.61%	0.56%	0.49%	0.56%	0.43%
20	0.68%	0.61%	0.69%	0.50%	0.69%	0.46%
21	0.76%	0.75%	0.73%	0.67%	0.73%	0.60%
22	0.71%	0.59%	0.69%	0.67%	0.69%	0.61%
23	0.66%	0.58%	0.64%	0.58%	0.64%	0.47%
24	0.45%	0.39%	0.46%	0.44%	0.46%	0.43%

Single Buyer. Hourly Average Relative Frequency. Jan.-Jun. 2012

	July	August	September	October	November	December
1	0.41%	0.42%	0.38%	0.20%	0.20%	0.21%
2	0.31%	0.42%	0.24%	0.20%	0.20%	0.20%
3	0.21%	0.42%	0.21%	0.20%	0.20%	0.19%
4	0.21%	0.37%	0.21%	0.20%	0.20%	0.19%
5	0.21%	0.34%	0.21%	0.20%	0.20%	0.19%
6	0.21%	0.35%	0.21%	0.20%	0.20%	0.19%
7	0.22%	0.41%	0.22%	0.20%	0.20%	0.20%
8	0.38%	0.42%	0.41%	0.36%	0.36%	0.35%
9	0.27%	0.41%	0.33%	0.27%	0.27%	0.33%
10	0.40%	0.42%	0.42%	0.31%	0.31%	0.37%
11	0.40%	0.43%	0.43%	0.30%	0.30%	0.37%
12	0.40%	0.42%	0.42%	0.29%	0.29%	0.37%
13	0.40%	0.44%	0.42%	0.29%	0.29%	0.37%
14	0.39%	0.42%	0.34%	0.28%	0.28%	0.33%
15	0.39%	0.42%	0.34%	0.28%	0.28%	0.34%
16	0.40%	0.42%	0.38%	0.30%	0.30%	0.37%
17	0.40%	0.43%	0.42%	0.40%	0.40%	0.41%
18	0.42%	0.46%	0.43%	0.41%	0.41%	0.43%
19	0.42%	0.46%	0.45%	0.42%	0.42%	0.43%
20	0.43%	0.53%	0.49%	0.42%	0.42%	0.43%
21	0.60%	0.78%	0.62%	0.44%	0.44%	0.50%
22	0.58%	0.63%	0.55%	0.41%	0.41%	0.44%
23	0.48%	0.64%	0.44%	0.41%	0.41%	0.41%
24	0.42%	0.54%	0.42%	0.31%	0.31%	0.40%

Single Buyer. Hourly Average Relative Frequency. Jul.-Dec. 2012.

	January	February	March	April	May	June
1	13.63%	11.03%	11.43%	11.96%	10.20%	12.07%
2	10.68%	8.28%	8.41%	8.97%	6.95%	9.48%
3	8.64%	6.70%	6.58%	7.08%	5.45%	7.82%
4	7.64%	6.08%	5.86%	6.37%	4.78%	6.84%
5	7.35%	6.04%	5.80%	6.26%	4.69%	6.59%
6	8.40%	6.98%	6.64%	6.96%	5.49%	7.45%
7	12.07%	10.14%	10.19%	9.76%	7.75%	8.51%
8	16.61%	14.86%	14.80%	13.32%	11.67%	11.10%
9	16.15%	14.61%	13.71%	12.25%	9.94%	9.70%
10	17.98%	16.04%	15.04%	13.43%	11.07%	11.04%
11	18.48%	16.34%	15.06%	13.46%	11.14%	11.38%
12	18.74%	16.53%	15.08%	13.24%	11.06%	11.59%
13	18.74%	16.58%	14.97%	13.07%	10.86%	11.73%
14	17.38%	15.50%	13.68%	11.92%	9.59%	11.05%
15	17.38%	15.45%	13.68%	11.71%	9.52%	10.89%
16	17.87%	15.78%	14.06%	11.85%	9.75%	10.99%
17	19.39%	16.84%	14.94%	12.39%	10.51%	11.82%
18	21.67%	18.98%	16.81%	13.39%	11.66%	12.99%
19	22.80%	20.59%	19.15%	15.08%	13.10%	13.94%
20	24.84%	22.67%	22.06%	19.08%	16.02%	16.67%
21	26.03%	23.55%	23.88%	22.87%	20.98%	20.05%
22	24.94%	22.23%	22.67%	21.83%	20.19%	20.46%
23	22.50%	19.82%	20.42%	19.90%	17.92%	18.80%
24	18.00%	15.29%	15.57%	15.79%	14.09%	15.67%

Single Buyer. Hourly Average Quantity Bid compared to the total amount of Bid accepted. Jan.-Jun. 2012.

	July	August	September	October	November	December
1	12.15%	15.58%	10.41%	7.61%	7.17%	9.37%
2	9.70%	13.18%	7.92%	4.78%	4.15%	5.85%
3	7.86%	11.55%	6.40%	3.33%	2.58%	3.83%
4	6.99%	10.85%	5.74%	2.74%	1.93%	2.91%
5	6.57%	10.38%	5.54%	2.67%	1.93%	2.77%
6	6.69%	10.28%	6.07%	3.59%	3.15%	3.90%
7	7.61%	10.93%	7.68%	6.69%	6.28%	7.31%
8	10.02%	12.70%	11.59%	11.71%	12.35%	12.34%
9	9.12%	11.20%	10.04%	9.91%	10.66%	11.94%
10	10.54%	12.42%	10.89%	10.58%	11.50%	13.41%
11	11.05%	12.61%	11.03%	10.36%	11.36%	13.50%
12	11.37%	12.70%	11.05%	10.19%	11.37%	13.46%
13	11.63%	13.15%	11.03%	9.88%	11.23%	13.28%
14	11.06%	12.70%	10.04%	8.60%	10.11%	12.40%
15	11.03%	12.66%	9.88%	8.59%	10.46%	12.78%
16	11.24%	12.85%	10.08%	9.02%	11.33%	13.74%
17	12.15%	13.84%	11.07%	10.19%	13.42%	15.89%
18	13.08%	15.17%	12.49%	11.73%	15.89%	18.19%
19	14.04%	16.42%	14.14%	14.42%	17.02%	19.23%
20	16.38%	19.61%	18.34%	18.43%	19.47%	21.35%
21	19.49%	23.61%	20.05%	19.70%	20.55%	22.50%
22	20.04%	22.51%	19.07%	18.46%	19.18%	21.20%
23	18.35%	21.18%	16.89%	15.82%	16.15%	18.64%
24	15.52%	18.99%	13.36%	11.89%	11.80%	14.51%

Single Buyer. Hourly Average Quantity Bid compared to the total amount of Bid accepted. Jul.-Dec. 2012.

Bilateral Contracts are another feature of the Day-Ahead Market to be deeply examined. They are contracts of supply of electricity made off the Power Exchange between Wholesalers and Eligible Customers. The price, as well as the withdrawal profiles, are freely agreed by the parties. However, transaction and related withdrawal schedule have to be recorded in order to check their consistency with the transmission constraints on the National Trasmission Grid. These contracts may be ascribed to the demand of economic agents using electricity for industrial production purposes and they account on average for 50% of the total bid accepted. Their shares had been increasing in the two years analized. During the first semester of 2011 the share was on average about the 46%, reaching the 50% between July and December. In 2012, the growth was slower; the average quantity submitted in the first semester was about 48% of the total amount of Bid accepted, reaching the 50.8% in the second Semester.

Hour	January	February	March	April	May	June
1	50.79%	49.89%	51.56%	53.67%	54.99%	53.09%
2	50.28%	49.67%	51.29%	53.55%	54.84%	53.10%
3	50.07%	49.44%	50.94%	53.69%	54.82%	52.93%
4	49.98%	49.14%	50.75%	53.70%	54.79%	53.02%
5	49.83%	49.18%	50.86%	53.63%	54.60%	53.10%
6	50.19%	49.86%	51.34%	53.40%	54.42%	52.95%
7	50.05%	49.77%	51.79%	54.29%	54.92%	53.05%
8	50.25%	50.11%	51.59%	53.99%	54.26%	52.76%
9	50.76%	50.82%	52.75%	54.75%	55.76%	53.71%
10	50.62%	50.46%	52.35%	54.68%	55.82%	53.80%
11	50.49%	50.45%	52.39%	54.39%	55.63%	53.74%
12	50.45%	50.49%	52.39%	54.45%	55.58%	53.66%
13	50.49%	50.81%	52.45%	54.59%	54.83%	53.42%
14	50.60%	50.70%	52.45%	54.79%	55.11%	53.56%
15	50.89%	51.06%	52.57%	54.87%	55.47%	53.58%
16	50.78%	50.68%	52.42%	54.89%	55.53%	53.59%
17	50.63%	51.05%	52.62%	54.99%	55.52%	53.57%
18	51.11%	51.54%	53.09%	55.17%	55.65%	53.74%
19	51.19%	51.77%	53.12%	55.82%	55.69%	54.29%
20	51.30%	51.77%	52.97%	55.61%	55.69%	54.31%
21	50.48%	50.16%	51.44%	53.62%	54.39%	52.74%
22	50.79%	50.42%	51.52%	54.11%	54.38%	52.98%
23	50.91%	50.27%	51.78%	54.85%	54.98%	53.47%
24	50.64%	49.98%	51.72%	54.78%	55.33%	53.41%

Bilateral Contracts. Hourly Average Relative Frequency. Jan.-Jun. 2011.

Hour	July	August	September	October	November	December
1	52.32%	51.20%	50.35%	52.73%	53.39%	51.94%
2	52.32%	51.38%	50.75%	52.77%	53.03%	51.86%
3	52.63%	51.55%	50.96%	52.42%	52.74%	51.76%
4	52.87%	51.74%	51.18%	52.38%	52.71%	51.63%
5	52.79%	51.61%	51.13%	52.31%	52.64%	51.43%
6	52.44%	51.11%	50.29%	52.22%	52.50%	51.38%
7	52.59%	50.81%	49.46%	52.02%	52.61%	51.51%
8	52.38%	50.67%	49.67%	52.24%	52.20%	51.22%
9	52.76%	51.63%	50.39%	53.17%	52.98%	51.99%
10	53.10%	51.77%	50.93%	53.08%	52.74%	51.75%
11	53.13%	51.94%	50.92%	53.18%	52.95%	51.55%
12	53.14%	51.78%	50.91%	53.13%	52.84%	51.55%
13	52.69%	51.70%	50.86%	53.64%	53.54%	52.16%
14	52.40%	51.83%	50.77%	53.82%	53.67%	52.20%
15	52.50%	51.86%	50.71%	53.66%	53.23%	52.16%
16	52.52%	51.85%	50.59%	53.59%	53.20%	52.02%
17	52.41%	51.62%	50.48%	53.44%	52.89%	51.98%
18	52.62%	51.62%	50.85%	53.60%	53.14%	52.23%
19	52.77%	51.75%	51.14%	53.48%	53.15%	52.09%
20	53.51%	52.07%	51.25%	53.56%	52.72%	51.81%
21	52.28%	50.68%	50.88%	52.70%	51.82%	50.71%
22	52.58%	51.31%	51.02%	52.81%	52.43%	51.53%
23	53.12%	51.84%	51.28%	53.24%	53.51%	52.78%
24	53.13%	51.75%	50.90%	52.79%	53.47%	52.60%

Bilateral Contracts. Hourly Average relative Frequency. Jul.-Dec. 2011.

Hour	January	February	March	April	May	June
1	45.96%	49.52%	52.53%	50.73%	49.63%	48.26%
2	47.69%	51.85%	54.69%	52.93%	51.83%	50.33%
3	48.54%	52.73%	55.66%	54.15%	53.19%	51.40%
4	49.01%	52.90%	55.93%	54.75%	53.66%	51.96%
5	48.76%	52.72%	55.67%	54.43%	53.51%	51.81%
6	47.69%	51.57%	54.77%	52.93%	52.57%	51.60%
7	43.56%	46.66%	50.50%	49.65%	49.68%	49.08%
8	38.99%	41.73%	44.82%	44.62%	43.97%	44.52%
9	42.06%	45.22%	49.33%	46.76%	46.40%	45.91%
10	40.32%	43.66%	48.09%	45.42%	45.06%	44.44%
11	40.01%	43.56%	48.07%	45.33%	44.82%	44.02%
12	40.08%	43.78%	48.34%	45.53%	44.89%	43.93%
13	41.31%	45.37%	49.96%	47.08%	45.95%	44.91%
14	42.14%	46.22%	50.74%	48.03%	46.69%	45.46%
15	41.73%	45.73%	50.15%	47.57%	46.35%	45.20%
16	41.46%	45.54%	49.98%	47.62%	46.27%	45.14%
17	40.62%	45.22%	49.79%	47.75%	46.30%	44.90%
18	39.01%	44.21%	49.75%	48.47%	46.83%	45.26%
19	38.67%	42.32%	47.78%	48.69%	47.27%	45.85%
20	38.84%	42.35%	46.36%	46.60%	46.58%	45.48%
21	34.49%	37.68%	39.89%	39.43%	40.50%	40.83%
22	36.33%	39.71%	41.96%	40.84%	40.55%	40.47%
23	39.05%	42.18%	44.75%	43.42%	42.74%	42.22%
24	42.29%	45.90%	48.58%	47.15%	46.25%	45.02%

Bilateral Contracts. Hourly Average Quantity Bid compared to the total amount f Bid Accepted. Jan-Jun. 2011.

Hour	July	August	September	October	November	December
1	49.20%	52.76%	51.90%	58.40%	59.46%	57.51%
2	51.26%	55.06%	53.99%	60.27%	61.14%	59.55%
3	52.61%	56.40%	55.13%	61.42%	62.23%	60.68%
4	53.32%	57.31%	55.68%	61.72%	62.69%	61.25%
5	53.42%	57.54%	55.64%	61.64%	62.46%	61.15%
6	53.00%	56.85%	54.35%	60.31%	60.73%	59.86%
7	51.38%	55.60%	50.56%	56.04%	55.95%	55.59%
8	47.25%	51.29%	47.11%	50.28%	49.56%	49.61%
9	48.53%	52.39%	48.25%	49.89%	47.96%	47.75%
10	46.77%	50.37%	46.89%	48.83%	46.83%	46.21%
11	46.19%	49.70%	46.47%	48.82%	46.73%	46.08%
12	45.94%	49.34%	46.40%	48.93%	46.81%	46.24%
13	46.74%	49.65%	47.79%	50.67%	48.61%	47.71%
14	46.95%	50.11%	48.41%	51.41%	49.29%	48.59%
15	46.65%	50.08%	47.91%	50.88%	48.54%	48.07%
16	46.61%	50.12%	47.72%	50.67%	47.97%	47.41%
17	46.31%	49.79%	47.43%	50.35%	46.36%	45.58%
18	46.68%	49.74%	47.88%	50.30%	44.28%	43.77%
19	47.28%	49.76%	48.21%	48.58%	44.26%	43.83%
20	47.09%	49.03%	45.96%	46.06%	44.28%	43.80%
21	42.63%	43.27%	42.08%	45.48%	44.73%	43.90%
22	42.31%	44.14%	44.04%	47.78%	47.30%	46.16%
23	44.24%	46.79%	46.60%	51.22%	51.19%	49.64%
24	46.46%	49.71%	49.04%	55.06%	55.53%	53.72%

Bilateral Conrtacts. Hourly Average Quantity Bid compared to the total amount of Bid accepted. Jul.-Dec. 2011.

	January	February	March	April	May	June
1	45.34%	40.70%	46.88%	47.49%	48.54%	49.57%
2	45.00%	40.65%	46.42%	47.15%	48.24%	49.68%
3	45.22%	40.80%	46.51%	47.21%	48.16%	49.44%
4	45.35%	40.92%	46.62%	47.27%	48.13%	49.53%
5	45.31%	40.90%	46.59%	47.26%	48.09%	49.53%
6	45.17%	40.50%	46.54%	47.18%	47.96%	49.75%
7	44.69%	40.55%	46.59%	47.29%	47.99%	49.31%
8	44.77%	40.34%	46.73%	47.34%	47.88%	49.07%
9	46.33%	40.97%	47.70%	48.50%	48.66%	49.65%
10	46.00%	40.86%	47.56%	48.36%	48.37%	49.10%
11	46.16%	40.88%	47.34%	47.92%	48.18%	48.76%
12	46.24%	41.14%	47.24%	47.53%	47.86%	48.47%
13	46.01%	41.43%	47.18%	47.28%	47.68%	48.42%
14	46.23%	41.66%	47.33%	47.31%	47.62%	48.49%
15	46.51%	41.57%	47.57%	47.49%	47.87%	48.68%
16	46.53%	41.68%	47.95%	47.92%	48.26%	48.94%
17	46.39%	41.55%	48.07%	48.33%	48.70%	49.35%
18	46.72%	41.54%	48.41%	48.45%	49.16%	49.85%
19	47.21%	41.59%	49.19%	49.42%	50.31%	50.28%
20	47.61%	42.61%	49.46%	50.41%	50.88%	51.46%
21	45.21%	40.85%	46.87%	48.64%	48.52%	48.97%
22	45.24%	41.10%	46.90%	48.47%	48.74%	48.83%
23	45.56%	41.22%	47.20%	48.42%	48.93%	49.65%
24	45.84%	41.77%	47.57%	48.20%	48.89%	50.24%

Bilateral Contracts. Hourly Average Relative Frequency. Jan.-Jun. 2012

	July	August	September	October	November	December
1	50.29%	50.44%	50.81%	52.48%	50.07%	48.42%
2	50.61%	50.46%	51.02%	52.21%	49.60%	48.03%
3	50.69%	50.40%	50.93%	51.91%	49.37%	47.42%
4	50.74%	50.55%	50.66%	51.57%	49.01%	47.05%
5	50.80%	50.53%	50.80%	51.53%	49.01%	46.79%
6	50.79%	50.30%	50.49%	50.94%	49.12%	46.74%
7	49.77%	49.67%	50.09%	50.93%	48.68%	47.09%
8	49.22%	49.64%	50.15%	50.79%	49.52%	47.29%
9	50.46%	51.13%	51.87%	52.10%	52.22%	48.43%
10	49.57%	50.60%	51.61%	51.69%	51.87%	48.07%
11	49.25%	50.10%	51.51%	51.45%	51.24%	47.67%
12	48.89%	49.56%	50.82%	51.06%	50.83%	47.24%
13	48.39%	49.30%	50.55%	51.69%	51.12%	47.42%
14	48.27%	49.27%	49.99%	51.29%	51.21%	47.61%
15	48.30%	49.22%	49.92%	50.97%	51.24%	47.89%
16	48.57%	49.38%	50.11%	51.08%	51.57%	48.30%
17	49.49%	49.83%	50.81%	51.42%	51.70%	48.44%
18	50.46%	50.33%	51.28%	51.86%	51.92%	48.61%
19	51.15%	51.15%	51.93%	51.82%	52.70%	49.34%
20	52.47%	51.86%	52.41%	52.60%	53.34%	49.46%
21	50.42%	50.03%	50.90%	52.27%	50.02%	47.98%
22	50.51%	50.42%	51.62%	52.77%	50.37%	48.50%
23	50.86%	50.55%	51.90%	52.73%	50.97%	49.48%
24	50.96%	50.94%	51.39%	52.79%	51.24%	49.15%

Bilateral Contracts. Hourly Average Relative Frequency. Jul.-Dec. 2012.

	January	February	March	April	May	June
1	52.90%	46.61%	55.56%	51.18%	53.41%	50.56%
2	54.60%	48.02%	56.91%	52.92%	55.14%	52.77%
3	55.56%	49.00%	57.86%	54.03%	56.22%	54.06%
4	55.97%	49.35%	58.37%	54.35%	56.57%	54.65%
5	55.95%	49.14%	58.18%	54.20%	56.40%	54.62%
6	54.76%	47.81%	56.88%	52.87%	55.65%	54.43%
7	50.23%	44.63%	52.96%	50.08%	52.97%	51.60%
8	44.16%	40.00%	47.10%	45.02%	47.21%	46.13%
9	45.36%	41.20%	49.31%	48.23%	49.95%	48.74%
10	44.02%	40.03%	48.40%	46.80%	48.42%	46.68%
11	44.11%	40.06%	48.43%	46.59%	48.23%	46.11%
12	44.28%	40.33%	48.62%	46.67%	48.05%	45.67%
13	45.87%	42.03%	50.45%	48.25%	49.89%	46.99%
14	46.79%	42.73%	51.13%	49.19%	50.68%	47.58%
15	46.30%	42.21%	50.66%	48.96%	50.39%	47.23%
16	45.95%	42.01%	50.72%	49.04%	50.58%	47.32%
17	44.84%	41.51%	50.48%	49.35%	51.07%	47.64%
18	42.93%	40.70%	50.29%	50.24%	52.04%	48.65%
19	42.73%	39.45%	48.67%	50.54%	52.79%	49.50%
20	42.68%	39.47%	47.10%	48.93%	52.01%	49.74%
21	39.86%	36.94%	42.64%	40.61%	43.84%	43.13%
22	41.95%	38.57%	44.55%	41.81%	44.11%	42.50%
23	45.33%	41.16%	47.98%	44.67%	47.18%	44.73%
24	49.40%	44.72%	52.42%	48.36%	51.12%	47.97%

Bilateral Contracts. Hourly Average Quantity Bid compared to the total amount of Bis accepted. Jan.-Jun. 2012.

	July	August	September	October	November	December
1	49.87%	47.80%	52.53%	59.94%	59.31%	55.43%
2	51.68%	49.83%	54.57%	61.82%	61.16%	57.81%
3	53.07%	51.23%	55.77%	63.00%	62.44%	59.15%
4	53.85%	52.21%	56.27%	63.25%	62.69%	59.70%
5	54.01%	52.57%	56.14%	63.04%	62.46%	59.56%
6	53.75%	52.02%	54.61%	61.20%	60.69%	57.98%
7	51.18%	51.05%	50.80%	56.06%	56.44%	54.05%
8	46.62%	47.86%	46.82%	50.23%	51.44%	48.96%
9	48.06%	49.58%	48.77%	53.03%	54.54%	50.05%
10	45.74%	47.48%	47.16%	51.91%	53.27%	48.47%
11	45.06%	46.75%	47.04%	51.96%	52.93%	48.20%
12	44.62%	45.99%	46.58%	51.79%	52.72%	48.17%
13	45.49%	46.08%	48.01%	53.67%	54.12%	49.40%
14	45.84%	46.34%	48.43%	54.30%	54.93%	50.64%
15	45.38%	46.14%	47.90%	53.66%	54.46%	50.41%
16	45.35%	46.11%	47.76%	53.30%	53.89%	49.90%
17	45.62%	46.00%	47.81%	53.21%	52.44%	48.32%
18	46.44%	46.14%	48.15%	53.02%	49.93%	46.40%
19	47.39%	46.12%	48.40%	51.29%	49.79%	46.68%
20	47.86%	45.66%	46.46%	48.83%	50.40%	46.62%
21	43.01%	40.14%	41.74%	46.16%	46.88%	43.91%
22	42.83%	41.05%	44.18%	48.96%	49.37%	45.98%
23	44.79%	43.39%	47.23%	52.56%	52.84%	49.35%
24	47.42%	45.95%	50.23%	56.52%	56.86%	53.00%

Bilateral Contracts. Hourly Average Quantity Bid compared to the total amount of Bid Accepted. Jul.-Dec. 2012.

As already mentioned, the Day-Ahead Market is a zonal market reflecting the limited interconnection capacity of transmission grid. These constraints influencing the equilibrilibrium prices both on the supply and the demand side. On the supply side congestion yields to differencing equilibrium prices. On the demand side, congestion affects the single national price, since the price paid by buyers is the weighted average of zonal prices formed on the supply side. Taking into account the role played by trasmission constraints, the tables below show the number which the national market was split in during 2011 and 2012. During the 2011 National Single Market occurred 15% of times. Division in the two zones was the most frequent, while division in four zones was marginal and occurred on average just 2% of times. In the 2012, instead, congestions were more frequent since in just 9% of times Single Market occured. Also the division in four zones increases in frequency which reached the 6%. Division in two and three zones remained the more frequent, the relative frequency of two zone division augmented to 54% of times, while the three division zones relative frequency decreased to 27%. Moreover, it must be mentioned that market segmentation in five zones has not been taken in consideration since it has never occurred in the 2011 and just once in the 2012. The five zones market segmentation has deeply decreased, comparing these results with the relative frequency recorded in the 2010, where this kind of division occurred around in 60 hours.

Zone	1	2	3	4
January	18%	52%	28%	2%
February	18%	47%	33%	3%
March	21%	47%	31%	1%
April	16%	47%	36%	1%
May	7%	52%	36%	4%
June	9%	50%	36%	5%
July	6%	41%	46%	6%
August	12%	59%	29%	1%
September	12%	31%	55%	2%
October	18%	48%	32%	2%
November	25%	39%	36%	0%
December	24%	48%	27%	1%

Monthly Frequency Of Division Zone. 2011.

Zone	1	2	3	4
January	12%	65%	23%	0%
February	0%	5%	37%	48%
March	13%	48%	13%	13%
April	16%	59%	24%	1%
May	16%	67%	17%	0%
June	10%	57%	31%	1%
July	3%	41%	50%	7%
August	3%	54%	40%	2%
September	6%	52%	41%	0%
October	9%	76%	15%	0%
November	11%	65%	23%	1%
December	10%	64%	25%	1%

Monthly Frequency of Zone Division. 2012.

Finally, the investigation of main features of electricity market ends with a descriptive analysis of PUN, the price paid by buyers. Tables highlight that in the 2012, during peak hours, the price is more dispersed than that recorded in the 2011. Moreover, tables relate PUN to the market segmentation. When congestion occurs, demand equilibrium price tends to increase. Congestions are infact representative of high level of demand and this is particular evident in the 2012. On the other hand, in the 2011, the PUN behaviour related to segmentation reachs higher values when segmentation in three zone occurred. (both in peak and off-peak hours). Neverthless, the lowest average level of PUN has been recorded when single market occurred. confirming that the absence of congestions improves market efficiency. Furthermore, tables highlight how investments in trasmission capacity are one of the major concern of Indipendent System Operator, given the crucial role played by electricity in national economic activities.

Percentile	mean	sd	min	max
1	44.21	6.40	10.00	51.35
2	55.77	2.31	51.36	59.65
3	62.45	1.52	59.66	64.94
4	66.80	0.98	64.96	68.40
5	69.93	0.85	68.41	71.39
6	73.15	1.06	71.40	74.97
7	76.82	1.15	74.98	78.92
8	81.47	1.51	78.93	84.11
9	87.34	2.10	84.12	91.25
10	104.47	13.92	91.27	164.80

Summary Statistics of PUN by its daily Percentile. 2011.

Percentile	mean	sd	min	max
1	60.69	4.61	31.00	65.69
2	67.34	0.90	65.70	68.90
3	70.22	0.72	68.92	71.46
4	72.84	0.81	71.48	74.21
5	75.56	0.77	74.22	76.82
6	78.57	1.03	76.83	80.46
7	82.25	1.11	80.47	84.14
8	86.34	1.37	84.15	88.85
9	92.25	2.34	88.87	96.89
10	111.51	15.08	96.92	164.80

Summary Statistics of Peak Hour PUN by its Peak Hour Percentile. 2011.

Percentile	mean	sd	min	max
1	39.66	6.11	10.00	46.09
2	48.94	1.57	46.12	51.86
3	54.29	1.30	51.87	56.37
4	58.64	1.30	56.38	60.76
5	62.73	1.07	60.80	64.47
6	66.30	1.04	64.48	68.00
7	69.61	0.97	68.01	71.29
8	73.74	1.52	71.30	76.48
9	79.91	2.16	76.48	84.07
10	93.42	9.98	84.08	137.84

Figure 3.1: Summary Statistics of Off-Peak Hour PUN by its Off-Peak Hour Percentile. 2011.

Percentile	mean	sd	min	max
1	39.15	6.27	12.14	47.77
2	53.87	3.46	47.80	59.15
3	63.11	2.02	59.17	66.26
4	68.60	1.32	66.27	70.76
5	72.59	1.07	70.77	74.52
6	76.78	1.27	74.53	79.03
7	81.29	1.30	79.04	83.73
8	86.54	1.68	83.74	89.81
9	94.15	2.78	89.85	99.58
10	118.78	21.68	99.61	324.20

Summary Statistics of PUN by its daily Percentile. 2012.

Percentile	mean	sd	min	max
1	45.88	9.54	12.14	58.14
2	63.22	2.29	58.15	66.75
3	69.10	1.33	66.77	71.29
4	73.20	1.16	71.30	75.37
5	77.26	1.10	75.38	79.24
6	81.02	1.00	79.25	82.82
7	85.12	1.27	82.84	87.32
8	90.04	1.66	87.34	93.11
9	97.70	2.88	93.12	103.58
10	126.33	23.17	103.66	222.25

Figure 3.2: Summary Statistics of Peak Hour PUN by its Peak Hour Percentile. 2012.

Percentile	mean	sd	min	max
1	36.67	5.17	14.88	42.75
2	47.43	2.53	42.77	51.69
3	56.18	2.37	51.70	59.93
4	63.00	1.90	59.94	65.96
5	68.15	1.32	65.97	70.25
6	72.02	1.05	70.27	73.92
7	76.20	1.46	73.93	78.71
8	81.78	1.79	78.72	85.01
9	89.56	2.94	85.02	95.20
10	109.99	17.83	95.21	324.20

Summary Statistics of Off-Peak Hour PUN by its Off-Peak Hour Percentile. 2012.

	Month	mean	sd	min	max
2011-I Semester	January	73.09	6.69	54.00	91.72
	February	75.50	10.39	55.79	110.40
	March	77.29	11.54	53.65	142.96
	April	70.91	11.62	47.83	118.07
	May	76.50	8.50	47.45	98.01
	June	73.65	8.51	44.40	102.83
2011-II Semester	July	76.02	13.68	31.00	142.67
	August	77.52	12.20	47.00	123.58
	Septembe	88.63	12.41	57.59	123.54
	October	85.73	14.32	50.22	151.09
	Novembe	90.12	20.63	45.90	160.62
	Decembe	91.77	17.81	59.21	164.80
2012-I Semester	January	91.05	19.53	50.31	165.76
	February	104.71	34.55	44.31	222.25
	March	82.70	25.49	34.94	176.37
	April	73.42	20.59	12.14	138.56
	May	70.90	16.56	20.52	140.02
	June	80.16	17.56	31.61	130.66

Monthly Average PUN. Peak Hour. 2011-2012.

	Month	mean	sd	min	max
2011-I Semester	January	56.92	13.30	10.00	84.07
	February	57.07	11.54	27.00	90.37
	March	59.06	11.60	36.18	115.13
	April	59.45	12.63	22.94	105.33
	May	66.05	11.34	39.98	97.43
	June	63.18	11.17	37.49	100.24
2011-II Semester	July	63.45	14.59	16.14	101.78
	August	71.50	15.60	31.06	132.99
	Septembe	73.99	15.30	40.01	131.71
	October	71.49	17.10	23.53	137.84
	Novembe	66.82	15.98	28.00	115.66
	Decembe	66.96	18.53	28.00	133.16
2012-I Semester	January	68.66	17.65	32.47	123.34
	February	73.37	19.17	31.71	161.78
	March	67.90	22.63	32.41	140.43
	April	72.01	21.59	23.50	152.82
	May	69.02	17.74	23.94	119.53
	June	75.60	20.18	28.09	133.38

Monthly Average PUN. Off-Peak Hour. 2011-2012

After preliminary analysis of dataset, market demand was derived. In each hour of the two years I ranked the bids according to the merit order (price descending order); I included also the rejected bids in order to have the estimation of elasticities relative to the prices of demand curve lower than equilibrium price. These latter elasticities represent infact the real responsiveness to change in price of purchaser less incline to buy. Then, I aggregated all inelastic bids (bids with submitted price equal to 3000), computing in this way the market point of demand corresponding to the intercept. Finally I derived the remaining downward sloping market demand curve by horizontal sum of bids characterized by the same price.

Finally, each monthly Dataset accounts for a sample size ranging from 15558 observations (for February 2011) to 29496 observations for July 2012.

Chapter 4

Multivariate Regression Model

4.1 Electricity Demand Model: A Review

Since the early 1970s, when energy caught the attention of policy makers in the aftermath of the first oil crisis, research on energy demand has vastly increased in order to overcome the limited understanding of the nature of energy demand and demand response due to the presence of external shocks encountered at that time. Moreover, the increase in population and the pressure for better living standards, the emphasis on large scale industrialization in developing countrires and the need to sustain positive economic growth rates had led worldwide to a fast increase in energy consumption. In the last 20 years, the strong and constant increase in energy consumption has imposed an accurate planning in order to avoid eletricity shortage and guarantee adeguate infrastructures.

Given this fact, economic model and estimation tecniques become essential features for energy planning, formulating strategies and recommending energy policies.

The main limitations and problems to be faced in the estimation models concern:

1) Modeling tools do not always give representation of a scenario able (good) to deal with uncertainty. The same problem comes up again with the characterization of the behaviour of economic agents.

2) Models are not concerned about the social and environmental issues conditioning energy policy. Moreover models do not include an adequate and detailed mathematical representation of the effects of technology employed.

3) Models are based on the unrealistic assumptions of perfectly functioning markets, fully employed and efficiently allocated resources, rational individuals and optimizing firm behaviour.

Energy demand analysis may be referred to different kinds of approaches: on one hand it should be mentioned the traditional approach which relies on optimizing behaviour within the neoclassical framework. On the other hand we have the engineering tradition that assumes a different behavioural assumptions based on the satisfying approach (in the sense of Hebert Simon or evolutionary approach for technological change) and beliefs. This divergence in the views has dominated the energy literature in the past and led to two distinct traditions of energy analysis: econometric approach and engineering end-use approach [9].

Econometric approach has seen significant developments over the past four decades. In the 1970s, the main aim was to understand the relationship between energy and other economic variables.

The enginnering approach (also known as bottom-up approach) is a energy demand forecasting methhod that focuses on end-uses or final needs at disaggregated level. The first systematic elaboration of the method was reported by Lapillone (1978) [1]. Since then, this approach gained prominence through works at Institute of Applied System Analysis (IIASA), International Atominc Energy Agency (IAEA) and elsewhere has emerged as an alternative method of demand forecastig. The motivation of this bottom-up investigation raised from the imputation of high level of energy demand registered in 1970s, even if prices were deeply increased, to end-consumers needs. Prices clearly have significant influence in energy use decicisions, but they are not all that matter. This method is infact based on the assumption that the standard neoclassical economic framework is insufficient for energy models aiming to explore the different dimension of potential policy impact. The engineering approach involves the following step:

- Disaggregation of total energy demand into homogeneous end use categories
- Systematic analysis of social, economic and technological factors
- Organization of determinants into a hierarchical structure

[1] Lapillonne, B. (1978). Long-term Energy Demand Forecastin. A New Approach. *Enetgy Policy*, **6**:2, pp 140-157.

- Formlization of the structure in mathemathical relationships
- Scenario design for the future
- Quantitative forecating using mathemathical relations and scenarios.

This end-use aproacch can not be used in our investigation. In the end-use tradition, the aggregated demand is obtained by summing demand at the disaggregated levels.

Data I am going to analyzed refers to Electricity Wholesale Market and they are not able to identify for each observation the homogeneous end use category which belong to. Moreover, the demand side of MGP market is essentially represented by industrial demand and eligible customers (natural or legal persons entitled to choose their own supplier of electricity producer, distributor and wholesaler) and traders, while the demand of domestic consumers (the so called captive customers) is usually coverd by the Single Buyer.

On the other hand, in the econometric approach, the majority of the studies have focused on the aggragated level of demand; for this reason our approach will belong to this last technique.

The econometric approach is a standard quantitative approach for economic analysis that establishes relationship between the dependent variable and certain chosen indipendent variables by statistical analysis of historical data. The relationship so determined can the be used for forecasting simply by considerig changes in the indipendent variables and determining their effect on the dependent variable.

Griffin [2] has identified three major developments since 1970s, namely the trans-log method, panel data methodology and the discrete choice method.

Wirl and Szirucsek [3] remarked that the trans-log function emerged as the preferred choice of researchers due to its flexible properties.

A large numbers of studies appeared in the 19070s and 1980s that applied the trans-log model at the aggregated and disaggregated level, including Brendt and Wood [12] and Pindyck [4]. On the other hand, the Panel

---

[2] Griffin, J., (1993), Methodological advances in energy modelling: 1970-90. *The Energy Journal*, **14**:11, pp 111-124.

[3] Wirl, F. and E. Szirucsek, (1990). Energy modelling – a survey of related topics, *OPEC Review*, pp. 361-378.

[4] Pindyck, R. S. (1979). The structure of world energy demand, *The MIT Press*, Cambridge, Massachusetts.

Data analysis approach allows capturing interregional variations that can be considered to reflect the long-term adjustment process as opposed to short term adjustment process in the time series data.

Finally, the discrete choice method determines demand relying on the capital stock and its utilization decisions. Despite its appeal, this method found limited econometric use due to lack of stock data.

The main partecipants in the Italian Electricity wholesale Market are industrial consumers using power as an input in the production function to produce goods and services and traders. Industrial agents choose the amount of electricity input which minimizes their cost funtion given the technological constraint, for this reason our econometric approach will lie inside the neoclassical framework and will be grounded on rational optimizing behaviour theory.

4.1.1 Theoretical Background: The Duality Approach

Although data available refers only market prices and demand, the duality approach gives us a theoretical justification, allowing to legitimately switch from agent's preference (optimization theory) to market demand (The Marshallian demand) in which quantities are functions of prices and total expenditure. We assume all the agent taking part in the MGP rationally behave minimizing a cost function.

Recalling the tradition introduced by Brendt and Wood the cost function assumed is the translog cost function, that is the the second order approximation of an agent's cost function. Its general form can be written as follow:

$$\ln C = \alpha_0 + \sum \alpha_i \ln p_i + \frac{1}{2} \sum_i \sum_j \gamma_{ij} \ln p_i \ln p_j + \alpha_Q \ln Q + \frac{1}{2} \gamma_{QQ} (\ln Q)^2 + \sum_i \gamma_{Q_i} \ln Q \ln p_i \tag{4.1}$$

where C is the total cost, i and j are the inputs for industrial consumers, p_i is the factor or good prices, Q is the objective variable (the objective variable to be maximized: it can be the output quantity).

This cost function must satisfies certain properties:

- Homogeneous of degree 1 in prices;
- Satisfying all the conditions guaranteeing a well-behaved production function

- Homothetic (separable function of the objective variable and prices).

Then, we have to impose the following parameter restrictions:
$\sum \alpha_i = 1$
$\gamma_{ij} = \gamma_{ji} \quad , i \neq j$
$\sum_i \gamma_{ij} = \sum_j \gamma_{ji} = 0$
$\sum_i \gamma_{Q_i} = 0$
$\gamma_{Q_i} = 0$
$\gamma_{QQ} = 0$

Minimization problem is usually solved using Lagrangian techniques, leading to the first order condition:

$$\frac{\partial C(Q,p)}{\partial p_i} = h_i(Q,p) = q_i \text{ for all } i \tag{4.2}$$

Under the given assumptions, solving the problem yields to a demand functions expressed in terms of prices and the objective variable : $q_i = h_i(Q,p)$.These functions are known as Hicksian demands and sometimes they are called compensated demand equations because they consider the objective variable Q as a constant parameter.

For empirical works the optimization model need to be linked to economical model in which quantities are a function of prices and total expenditure. The duality approach is the theoretical framework allowing to shift from the production possibility sets (and the system of preferences) to the market demand function.

Given the convexity of production possibility sets, the Roy Identity allows to derive Marshallian demand from the Hicksian demand substituting the objective variable Q in the Hicksian demand with its inverse function.

First we derive the Minimum Expenditure function and we put it into indirect production function $V(m,p)$, substituting m with $C(Q,p)$ evalueted at the optimum level. This lead to the trivial identity:

$$V(C(Q,p),p) = Q(m,p) \tag{4.3}$$

where $Q(m,p)$ is the production function of the maximization problem, p is the price vector and m is the budget constraint.

This says that the indirect function $V(C(Q,p),p)$, that minimizes the cost for achieving a given level of production given a set of prices, is equal to Q evaluated at those prices . Taking the derivative of both sides of this equation with respect to the price of a single input/good p_i (with the Q's level held constant) gives:

$$\frac{V(C(Q,p),p)}{\partial Q} \cdot \frac{\partial C(Q,p)}{\partial p_i} + \frac{V(C(Q,p),p)}{\partial p_i} = 0 \tag{4.4}$$

Rearranging what we obtain is:

$$\frac{\partial C(Q,p)}{\partial p_i} = -\frac{\frac{V(C(Q,p),p)}{\partial p_i}}{\frac{V(C(Q,p),p)}{\partial Q}} = h_i(Q,p) = g_i(m,p) \tag{4.5}$$

The function $g_i(m.p)$ represents the Marshallian demand which express quantity demaded for an input or good as a function of its own price, the budget constraint and the price of all the other goods.

Given the Marshallian demand function of electricity the multidimensional model need to be reduced into a two dimensional problem. For this reason, all the other goods and inputs will be bundled in a numeraire good. The numeraire is evaluated at a price proxied by the monthly consumer price index (adjusted excluding from its computation the energy consumption).

4.2 The Statistical Model

The model uses a log-linear demand function: the dependent variable is the logarithm of aggregated demand and the explanatory variables are the corresponding logarithm of prices, adjusted by the monthly consumer index price (representing the price of the numeraire) and dummy variables (relative to the day the zone etc...) which approximate the total expenditure.

Analitically, the model is:

$$\log y_i = \alpha_i + \beta_i \log(\frac{p_i}{\overline{p}}) + \sum \gamma_{ki} d_{ki} \tag{4.6}$$

where y_i represents a point of aggregated demand and i index the hour of the day.

Given this functional form β_i represents the hourly elasticity of electricity.

Regressors d_{ki} refer both to daily and zone intercept dummies and daily and zone interaction dummies which allow to derive the hourly elasticity for each day.

Recalling the issue arised from the prelyminary analysis of datasets in the previous chapter, the model has to process the information pertaining the presence in the market of heterogeneous consumers. As I said before, bids with no price cover a considerable portion of market and express the maximum willingness to pay. Following the Bigerna and Bollino approach [13], for simplicity, all bids are divided in just two groups representing two categories of consumers.

- Bids specifying the price are considered referring to an elastic consumer

  (aware of the wholesale market price signals) with demand $y_1 = f(p_1)$ having $\varepsilon_1 < 0$.

- Bids with no price are gathered into a consumer category denoted by a demand function: $y_2 = f(p_2)$ with elasticity $\varepsilon_2 = 0$.

Given different price responsiveness, the two kinds of consumers also differ for their reservation price (the price make the demand equal zero) $p_1^* = f_1^{-1}(0)$ and $p_2^* = f_2^{-1}(0)$ respectively with $p_2^* < p_1^*$. Given the equilibrium price p^*, the aggregate demand can fall in two cases:

- If $p_2 < p^* \leq p_1$, then the market demand is $y = y_1 + y_2 = f_1(p_1)$, with $y_2 = 0$ the market demand is expresed only by the type 1 with elasticity ε_1.

- If $p_2 < p^* \leq p_1$, then the market demand is $y = y_1 + y_2 = f(p^*)$, the market demand is given by the aggregation of both type of consumers.

How the role of Single Buyer should be processed into the model is another open question. Single Buyer is an intermediary agent buying elasticity from the market and reselling it to distributors. Partecipants who submit demand bids on benhalf of final customers can be seen as the agents in a principal-agent relationship where the principals are the final users. In a principal-agent framework, problems of moral hazard and conflicts of interest may arise because of asymmetric information.

Opportunistic behaviour, as arbitraging between Day-Ahead Market and Infra-Day Market, may be expected since Single Buyer holds high share of quantity demanded. However, institutional market factors suggest the impossibility for agents to have strategic behaviour. Arbitraging are not convenient because National Market Regulator imposes penalty charges if the real loads deviate from the withdrawal prifiles defined in the Day-Ahead Market. For this reason traders have to submit bids reflecting the real willingness to pay. Since agent's incentives are consistent with principal's interest, the derived equilibrium is Pareto-Optimal.

The above discussion shows that the theoretical model is able to represent the rational economics behaviour of agent, as final consumers and traders, presenting purchase bids into the electricity market.

Let divide the day into two groups of hours (peak and off-peak hours), one ranging from 9 a.m. to 8 p.m. (the time period in which the majority of consumption and economic activities take place), the second instead goes from 21 p.m. to 8 a.m.. We expect that partecipants, within these two groups of hours can affect the market price sensitivity: setting prices in advanced gives purchasers the time to react to high prices, postpone their electricity consumption, reshedule their activities and their demand profiles, flattening in this way the load curves. Moreover, hourly average demand gives evidence of the assumption of differentiated group of hours; the off-peak electricity demand profiles substantially differ from those recorded during peak hours; since some economic ativities can not be run, electricity demand is lower. As tables below show, the total quantity submitted in a off-peak hours is on average the 25% lower than the total quantity recorded in a peak hour.

Hour	January	February	March	April	May	June	July	August	Septeber	October	November	December	Average
Peak	41865.36	43244.02	40983.41	36767.31	37570.82	39353.03	42168.54	35166.48	40148.07	39132.19	41001.70	40035.42	39786.36
Off-Peak	31378.59	32682.33	31806.83	30134.98	30313.60	31614.65	34330.47	29644.57	32502.48	31245.96	31083.64	30348.19	31423.86
Diff	10486.77	10561.69	9176.58	6632.33	7257.22	7738.37	7838.07	5521.91	7645.58	7886.23	9918.06	9687.24	8362.50

Hourly Average Demand (MWh). Peak/Off Peak Hour. 2011

Hour	January	February	March	April	May	June	July	August	Septeber	October	November	December	Average
Peak	39725.59	42808.40	38228.20	34306.98	35172.65	39112.37	42415.71	35995.87	37519.21	36997.81	37052.36	37534.23	38072.45
Off-Peak	30187.59	32794.87	30370.54	28070.33	28185.90	30696.38	33420.18	29728.15	30023.83	29241.93	28653.32	28835.81	30017.40
Diff	9538.01	10013.53	7857.65	6236.65	6986.75	8415.99	8995.53	6267.72	7495.38	7755.89	8399.04	8698.42	8055.05

Hourly Average Demand (MWh). Peak/ Off Peak Hours. 2012.

Differences can be noticed even in Market Equilibrium Prices (PUN): during peak hours, as demand is higher, equilibrium price is higher. Tables below show summary statistics of hourly average PUN aggregated within peak/off-peak group of hours.

Hour	Frequency	Mean	St. Dev.	Min	Max
Off-peak	4380	64.71	15.30	10.00	137.84
Peak	4380	79.75	14.66	31.00	164.80

Average PUN by Peak/Off-Peak Hours. 2011.

	Frequency	Mean	Std. Dev.	Min	Max
Off-Peak	4392	70.08	20.99	14.88	324.20
Peak	4392	80.88	22.02	12.14	222.25

Average PUN by Peak/Off-Peak Hours. 2012

Given the differences in the main economics variables between peak and off-peak hours, we assume that the hourly demands and the hourly spot prices are correlated within each group. If the derived peak hour elasticities will be higher than off-peak elasticities, the assumption of economic agents conditioning market elasticity will be confirmed. On the other hand, if price responsiveness during peak hours do not significantly differs from nigth hour elasticities, we can conclude that purchasers have small market power and, given their stiff consumption profiles, they can not influence market equilibrium prices and quantities.

Given this market structure, we apply a Seemingly Unrelated Regression model. SUR model is a multiple equations regression model, in our case regression equations are 12, one for each hours.

The SUR can be written as:

$$y_{mi} = \beta_{m1}x_{mi1} + \beta_{m2}x_{mi2} + ... + \beta_{mik}x_{mik} + \varepsilon_{mi} \tag{4.7}$$

with $i = 1, ..., N$ observations for $m = 1, ..., M$ equations. M represents the number of hours whoose electricity prices and loads are considered correlated). y_{mi} is the ith observation of the dependent variable (the log-demand) in equation m, x_{mik_m} (with $k = 1, ..., K$) is the ith obeservation of the of explantory variable of the mth equation and β_{mk} is the k regression coefficient of the $m-$th equation.

Model can be written in a compact form. Let denote $y_m = (y_{m1}, ..., y_{mN})', \varepsilon_m = (\varepsilon_{m1}, ..., \varepsilon_{mN})'$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ \beta_M \end{bmatrix}$$

$$X_m = \begin{bmatrix} x'_{m1} & x'_{m2} & . & . & x'_{mk_m} \end{bmatrix}$$

and define $k = \sum_{m=1}^{M} k_m$.
Stack all vectors together as:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_M \end{bmatrix}$$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ \varepsilon_M \end{bmatrix}$$

$$X = \begin{bmatrix} X_1 \\ X_2 \\ . \\ . \\ X_M \end{bmatrix}$$

the model obtained takes the following form:

$$y = X\beta + \varepsilon$$

The SUR model can be written as a familiar linear regression model. If we assume ε_{mi} to be i.i.d $N(0, \sigma^2)$ for all m and all i, the model would simply turn into the normal linear regression model. However, we have assumed that market partecipants can reprogramming their activities and reschedule their demand profiles (within the group of hours) if they suppose changes in electricity prices. Therefore, ε_i must be i.i.d. $N(0, \Sigma)$ just for $i = 1, ..., N$ where Σ is an $N \times N$ full variance covariance matrix. With this assumption it can be seen that ε is $N(0, \Omega)$ where $\Omega = \Sigma \otimes I_M$ is an $NM \times NM$ block diagonal matrix gien by:

$$\Omega = \begin{bmatrix} \Sigma & 0 & . & . & 0 \\ 0 & \Sigma & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & \Sigma \end{bmatrix}$$

In the stacked model the number of all the explanatory variables (the economic variables, the intercept dummies and the interaction dummies) goes from 816 to 1032.

Bayesian technique imposes to set Prior distributions for all parameters of interest, then, the next step will be to choose adeguate prior distribution for βs parameters.

4.2.1 Prior Distributions

Usually Bayesian technique suggests to use natural conjugate priors, making the β's distribution be dependent on Ω, in this way the joint posterior distribution would become: $p(\beta, \Omega) = p(\beta|\Omega)p(\Omega)$.

This joint prior has the advantage to derive analytically tractable joint posterior distributions whoose main summary statistics are available, sparing in this way the use of posterior simulator: However, the natural conjugate prior for the SUR model has been found by many to be too restrictive. The prior covariances between coefficients in each pair of equations are infact all proportional to the same matrix. For this reason, following the mainstream literature, here I apply the extended version of the natural conjugate prior: the indipendent Normal - Wishart prior:

$$p(\beta, \Sigma) = p(\beta)p(\Sigma) \tag{4.8}$$

where:

$$p(\beta) = N(\beta_0, V_0) \tag{4.9}$$

and

$$p(\Sigma) = IW_{\nu_0}(\Lambda_0^{-1}) \tag{4.10}$$

where the prior distribution for the Variance-Covariance Matrix is an Inverse Wishart distribution (the inverse matrix of Σ has a Wishart distribution, that is the matrix generalization of the Gamma distribution).

$$f_W(\Sigma^{-1}) = \frac{1}{c_W}\Lambda_0^{\frac{v_0}{2}} \left|\Sigma\right|^{-(\frac{\nu_0 - N}{2})} \exp\left\{-\frac{1}{2}tr(\Lambda_0\Sigma^{-1})\right\} \tag{4.11}$$

with

$$c_W = 2^{\frac{\nu_0 NM}{2}} \pi^{\frac{NM(NM-1)}{4}} \prod_i \Gamma(\frac{\nu_0 + i - 1}{2}) \tag{4.12}$$

When there are no pre-experimental information, this model allows to use *non-informative prior* simply setting ν_0 (which can be interpreted as the number of pseudo-observation, i.e. the size of the fictious sample) equal to 0 and allowing to the prior variance of β (V_0) to go to infinity. However, in this study, prior hyperparameter elicitation comes from the previous empirical study of Bigerna & Bollino [13]. Then, for the beta parameters I used a Normal Prior distribution centered on the frequentist hourly estimates referring to the previous year (2010 and 2011).

4.2.2 Posterior Distributions

The posterior is proportional to the prior times the likelihood. Hence, if we multiply the two priors with the likelihood function and we discard the terms that do not depend upon β and Σ, we obtain:

$$
\begin{aligned}
p(\beta,\Sigma|y) \;\propto\; & \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(y_i - X_i\beta)'\Sigma^{-1}(y_i - X_i\beta)\right\} \times \\
& |V_0|^{-1/2}\exp\left\{-\frac{1}{2}(\beta-\beta_0)'V_0^{-1}(\beta-\beta_0)\right\} \times \\
& |\Sigma|^{-\frac{1}{2}(\nu_0+M+1)}\exp\left[-\frac{1}{2}tr(\Lambda_0\Sigma^{-1})\right] \qquad (4.13)
\end{aligned}
$$

This joint posterior density for β and Σ does not take any well-known functional form and, hence, it can not be directly used for posterior inference given that integrals as the expected values for β are not analytically derivable. Posterior simulation is required since there is not an analytical formula. However, we can recover the well-known kernel of the full conditional posterior denities: $p(\beta|y,\Sigma)$ and $p(\Sigma|y,\beta)$.

$$
\begin{aligned}
p(\beta|y,\Sigma) &= N(\beta_n, V_n) \qquad (4.14)\\
p(\Sigma|y,\beta) &= IW_{\nu_n}(\Lambda_n^{-1}) \qquad (4.15)
\end{aligned}
$$

where:

$$
\begin{aligned}
V_n &= (V_0^{-1} + n\Sigma^{-1})^{-1}, \qquad (4.16)\\
\beta_n &= V_n(V_0^{-1}\beta_0 + n\Sigma^{-1}\overline{y}), \qquad (4.17)\\
\nu_n &= \nu_0 + n, \qquad (4.18)\\
\Lambda_n &= \Lambda_0 + \sum_{i=1}^{n}(y_i-\beta)(y_i-\beta)'. \qquad (4.19)
\end{aligned}
$$

For β the full conditional posterior distribution is a Normal, while for Σ is

an Inverse Wishart. Also these two distributions combine data and prior information. However, given the two full conditional posteriors, the joint posterior distribution $p(\beta,\Sigma|y)$ keeps remaining untractable since $p(\beta,\Sigma|y) \neq p(\beta|\Sigma,y)\cdot p(\Sigma|\beta,y)$. Inference related to the joint posterior distribution of the random parameters β and Σ has not been possible yet and this computational issue needs to be resolved since it is essential to derive various numerical summaries of the parameters (such as their posterior expected values and variances).

A broad class of computational alghorithms can help to perform simulation aiming to approximate a general density (in this case the posterior distribution $p(\beta, \Sigma|y)$) when it can not directly integrated.

When posterior full conditional distributions take a well-known form, Gibbs sampler is the Markov Chain Monte Carlo algorithm usually adopted. It imposes to factor the posterior distribution in the normal $p(\beta|y, \Sigma)$ and in the Wishart $p(\Sigma|\beta, y)$ and then simulate the posterior using these two conditional distributions.

The algorithm samples from the joint posterior distribution by constructing a Markov Chain whose transition kernel adopts the two full conditional distributions. The chain will converge to the posterior distribution since the stationarity conditions are satisfied.

The Markov Chain is constructed using the following kernel:

$$\begin{aligned} K(\theta^{(t+1)}, \theta^{(t)}) &= K([\beta^{(t+1)}\ \Sigma^{(t+1)}], [\beta^{(t)}\ \Sigma^{(t)}]) \\ &= p(\beta^{(t+1)}|\Sigma^{(t+1)}, y)p(\Sigma^{(t+1)}|\beta^{(t)}, y) \end{aligned} \tag{4.20}$$

The algorithm undertakes the following steps:

1. Set an intial value for $\Sigma = \Sigma^{(0)}$ and then sample β from $p(\beta|y, \Sigma^{(0)})$ and obtain the realization $\beta^{(1)}$.

2. Given $\beta^{(1)}$, sample Σ from the $p(\Sigma|y, \beta^{(1)})$,

3. Repeat the step *1* and *2* 11000 times.

After the algorithm was performed, the first subsample of 1000 realization was discarded in order to avoid the chain to be dependent on the starting value. The remaining sequence of draws $\left\{\left(\beta^{(i)\prime}, \Sigma^{(i)\prime}\right)'\right\}_{i=1001}^{11000}$ simulates a sample from $p(\beta, \Sigma|y)$.

Averaging the simulated sequence $\{\beta_i\}$ allows to derive a point estimates for the beta coefficients. *By the law of Iterated Expectations,* these averages can be infact considered as an estimatates of the expected values of the marginal posterior distributions: $p(\beta|y)$.

Setting the number of replication equal to 11000 guarantees that the chain takes enough steps to cover all the parameter space, diagnsotic procedure shows satisfactory results: the convergence diagnostic tests performed do not refuse the null hypotesis of convergence to the posterior density.

4.3 Empirical Results

The statistical model has been applied to data referring the whole 2011 and the first semester of 2012, the whole computation has been performed using MatLab, we have derived the full posterior distributions of the coefficients and the variances for each hour within the group of equations.

For each system of equations we derive the hourly-elasticity for each day of the month by simply adjusting the coefficient related to the log-price regressors through the coefficient related to the daily (iteration) dummy variables. In this way we have derived all beta elasticity for each hour and each day.

In order to have some statistical summaries we aggregated the later estimates in the hourly average elasticity for each month.

The tables below show the derived results:

	Hour	January	February	March	April	May	June	July	August	September	October	November	December
Peak	9	-0.0243	-0.0433	-0.115	-0.0958	-0.0728	-0.0327	-0.0569	-0.0608	-0.0363	-0.0744	-0.028	-0.0434
	10	-0.056	-0.109	-0.071	-0.0914	-0.1415	-0.0459	-0.0601	-0.0563	-0.0325	-0.0793	-0.0298	-0.0488
	11	-0.0482	-0.1099	-0.0468	-0.0855	-0.0617	-0.0365	-0.0608	-0.0626	-0.0193	-0.0876	-0.0337	-0.0432
	12	-0.0588	-0.1108	-0.1147	-0.0655	-0.0572	-0.0374	-0.0552	-0.079	-0.0214	-0.0469	-0.0407	-0.0458
	13	-0.0513	-0.097	-0.0757	-0.0684	-0.1513	-0.0349	-0.0568	-0.0648	-0.023	-0.0597	-0.1103	-0.0366
	14	-0.0618	-0.1131	-0.1206	-0.0719	-0.1479	-0.0293	-0.0675	-0.084	-0.0263	-0.0715	-0.0431	-0.043
	15	-0.0491	-0.0794	-0.1135	-0.0705	-0.1615	-0.0359	-0.0479	-0.0705	-0.0234	-0.0462	-0.0314	-0.0414
	16	-0.0438	-0.0225	-0.1082	-0.0741	-0.0772	-0.0189	-0.0699	-0.063	-0.0241	-0.0611	-0.0392	-0.0435
	17	-0.0345	-0.079	-0.0742	-0.0771	-0.1747	-0.0262	-0.0733	-0.0415	-0.0269	-0.0731	-0.0367	-0.0572
	18	-0.0486	-0.0222	-0.0928	-0.0707	-0.1654	-0.0373	-0.0792	-0.0516	-0.0127	-0.061	-0.0282	-0.0522
	19	-0.0548	-0.015	-0.1124	-0.0924	-0.123	-0.0359	-0.0706	-0.0473	-0.0095	-0.0472	-0.0352	-0.0524
	20	-0.0365	-0.0147	-0.0616	-0.0972	-0.127	-0.0265	-0.0689	-0.0516	-0.0089	-0.0532	-0.01	-0.0416
Off-Peax	21	-0.0428	-0.0176	-0.0506	-0.0467	-0.065	-0.0459	-0.0466	-0.0019	-0.0529	-0.0249	-0.0529	-0.0736
	22	-0.051	-0.009	-0.0511	-0.04	-0.1406	-0.0286	-0.0522	-0.0183	-0.0357	-0.0382	-0.0532	-0.079
	23	-0.0284	-0.0185	-0.087	-0.0546	-0.0644	-0.0336	-0.0537	-0.0446	-0.0507	-0.0237	-0.0459	-0.0694
	24	-0.0573	-0.0375	-0.1408	-0.0566	-0.0615	-0.0498	-0.0461	-0.0447	-0.0399	-0.0167	-0.0467	-0.0639
	1	-0.0648	-0.0791	-0.1214	-0.0615	-0.1524	-0.0477	-0.0439	-0.0542	-0.0916	-0.0314	-0.0502	-0.0496
	2	-0.0522	-0.0792	-0.1638	-0.0746	-0.1596	-0.0487	-0.0398	-0.0791	-0.0558	-0.0383	-0.069	-0.0591
	3	-0.0379	-0.0855	-0.109	-0.0589	-0.1606	-0.0518	-0.046	-0.0715	-0.0641	-0.0366	-0.071	-0.049
	4	-0.052	-0.0922	-0.1952	-0.0733	-0.0833	-0.0477	-0.0419	-0.0583	-0.0759	-0.0347	-0.0595	-0.0647
	5	-0.0446	-0.081	-0.1491	-0.0466	-0.1756	-0.0366	-0.0548	-0.0635	-0.0581	-0.0257	-0.0783	-0.0746
	6	-0.0682	-0.076	-0.1472	-0.0561	-0.1711	-0.0525	-0.0454	-0.0588	-0.0786	-0.0403	-0.0561	-0.0682
	7	-0.0638	-0.0504	-0.135	-0.0553	-0.1239	-0.0541	-0.037	-0.0867	-0.0847	-0.0492	-0.0383	-0.0591
	8	-0.0667	-0.0114	-0.1001	-0.0421	-0.1269	-0.0577	-0.0434	-0.0425	-0.0878	-0.0243	-0.0411	-0.0821

Hourly Average Elasticity. 2011

		January	February	March	April	May	June
Peak	9	-0.0409	-0.0454	-0.0738	-0.0585	-0.0757	-0.1305
	10	-0.0614	-0.0534	-0.07	-0.0423	-0.0787	-0.1122
	11	-0.0313	-0.0504	-0.0772	-0.0392	-0.0502	-0.1206
	12	-0.0375	-0.0829	-0.0771	-0.0555	-0.0574	-0.1303
	13	-0.0337	-0.0607	-0.0667	-0.0499	-0.0485	-0.1125
	14	-0.0358	-0.065	-0.059	-0.0654	-0.0522	-0.0974
	15	-0.0308	-0.0646	-0.0715	-0.0495	-0.0691	-0.1177
	16	-0.0353	-0.0762	-0.0793	-0.0691	-0.0714	-0.136
	17	-0.0338	-0.0717	-0.0903	-0.0713	-0.0791	-0.1178
	18	-0.0367	-0.0401	-0.0726	-0.0521	-0.0802	-0.1101
	19	-0.0458	-0.0438	-0.0762	-0.0605	-0.0761	-0.0862
	20	-0.0418	-0.0353	-0.0703	-0.067	-0.0632	-0.0863
Off-PeaK	21	-0.0474	-0.0503	-0.0762	-0.0431	-0.0207	-0.0134
	22	-0.06354	-0.0528	-0.0655	-0.0425	-0.0192	-0.0184
	23	-0.059	-0.0819	-0.0847	-0.0528	-0.0187	-0.0156
	24	-0.0541	-0.0456	-0.0912	-0.0516	-0.0206	-0.0123
	1	-0.0652	-0.0923	-0.0547	-0.0865	-0.0325	-0.0231
	2	-0.0667	-0.0549	-0.0824	-0.0828	-0.0378	-0.0171
	3	-0.0645	-0.0691	-0.0869	-0.0652	-0.0321	-0.0168
	4	-0.0729	-0.0641	-0.0732	-0.0739	-0.0265	-0.0164
	5	-0.0748	-0.072	-0.0663	-0.0833	-0.023	-0.0184
	6	-0.0832	-0.0859	-0.0619	-0.0861	-0.0195	-0.0186
	7	-0.0538	-0.0709	-0.098	-0.0796	-0.0176	-0.019
	8	-0.0537	-0.0606	-0.0732	-0.055	-0.0156	-0.0153

Huorly Average Elasticity. I Semester 2012

Firstly, it can be noticed that average elasticities vary within the hours of the day, months and the years. In 2011 estimates go from a minimum value of -0.0019 (registered in August) to -0.1952 recorded in March. In the 2012 the extent of the range between minimum and maximum values is tinier, the minimum value (equal to -0.0123) was registerd in June at 0 a.m., while the maximum, equal to -0.136, was recorded in June at 4 p.m. On average hourly elasticities in the 2011 are greater than in the 2012.

Secondly, in 2011 elasticities are higher during the off-peak hours. In the peak hours period, electricity quantities traded are greater than the average as it is shown by the previous tables and confirmed by high frequency of congestion. High quantities traded reflect high levels of expenditure and higher price level. From literature we know that a driver of elasticity is the price level which usually is positive correlated with it. On the other hand, income level is negative correlated with price sensitivity. It means that, during peak hours the majority of market partecipants are characterized by high level of income and greater expenditure availability. Other factors affecting elasticity are the availability of substitutes for the commodity under exam, the possibility of postponing its consumption and the force of habit. Demand for a commodity having valid substitute is relatively more elastic. The possibility of postponing consumption is positive correlated with demand elasticity. During the peak period of the day, economic activities use electricity as an essential commodity and for this reason they are not able to postpone their consumption, reschedule their demand profile and flat the load curve.

Despite the wider set of energy sources (renewable sources as solar) may suggest the existence of substitutes that increases demand elasticity, during the peak period, price sensitivity is lower and this suggests a stiff and less flexible consumer's behaviour.

In the off-peak hours partecipants are more responsive to changes in price. This may be explained by the more consistent partecipation of agents with lower income level. As we seen before, off-peak electricity traded is on average lower than peak quantity. The average of this economic variable proxies lower purchasers' income levels which could explain higher elasticity. Strictly expenditure constraints make infact consumers more reactive to change in prices, since it has greater impact on their budget. In the off-peak hours, electricity has become a luxury good whose consumption is not necessary and can be postponed since the market has covered demand electricity for the economic activities during the peak hours. Moreover, during off-peak hours lower equilibrium prices had been recorded, it means that a lower portion of income is allocated for electricity compared with the budget shares of other inputs and goods. The budget share allocated for a commodity affects its demand elasticity: as lower is the share for an input, as greater will be the impact of change in price on the share it-self and on its consumption level.

Let compare our elasticities with estimates derived by Bigerna and Bollino [13] for the 2011 (from January to September). Our aggregation has been different since we derive hourly average elasticities per each month of the

2011. A quick look sugggests that our estimates are not so different from the previous frequentist results, especially for what concerns the peak hour elasticities. The only difference is that in our research the maximum average elasticity (equal to -0.073) has been observed at 2 p.m. while in Bollino' paper at 12 p.m and it is equal to -0.09. Observing off-peak hours elasticities our summaries statistics show a behaviour completely different, within this group, parameters have more variability and the maximum values of average elasticities, equal to -0.076, was recorded at 6 a.m. Moreover, on average peak elasticities are less than off-peak estimates in contrast with the conclusion of the previous literature. Theoretical framework can legitimate our conclusions, since, during the off-peak period, the quantity traded, and the frequency of submit bid are smaller than those recorded during the day. As economic activities slow down during the off-peak hours, lower level of demand and the reduced risk of congestion curtails the sellers' market power. On average Bayesian estimates are (in absolute value) higher then frequentist estimates. It depend on using for the beta parameters a normal disitribution truncated at zero as a prior distribution. This density, exploiting the information derive from frequentist studies, restricts the parameter space which will be explored by the algorithm making inferential results more negatives.

During the 2012 a turning trend is instead recorded. Peak elasticities are on average higher than off-peak ones. It may depend on an increasing presence of unit commitments using renewable sources. In particular photovoltaic plants have sharply increased during the 2012 and market recognized them dispatching priority (priority in the economic merit order according to which the offers are ranked for Market Resolution). Renewable sources are effective substitutes [5] of traditional sources and they have enlarged traditional fuel mix used for generation. During these hours, the presence of a greater number of sellers and substitute can increase the price responsiveness of purchasers.

Next tables show aggregated estimates by quarter confirming the previous conclusions. In the 2011 three of four quarters (the first, the third and the fourth qaurter) show off-peak average elasticities higher than the peak elas-

---

[5] D.L. n 91/2014, Disposizioni urgenti per il settore agricolo, la tutela ambientale e l'efficientamento energetico dell'edilizia scolastica e universitaria, il rilancio e lo sviluppo delle imprese, il contenimento dei costi gravanti sulle tariffe elettriche, nonche' per la definizione immediata di adempimenti derivanti dalla normativa europea. GU n.144 del 24-6-2014

ticities. The presence of economic activities difficult to be rescheduled during the day can explain why electricity demand is less responsive to changes in price during the peak hours.

In the 2012, instead, from the second quarter it can be noticed that consumer's price responsiveness increased during peak hours. Moreover, the extent of the range between peak and off-peak elasticities became greater. This tendency confirms the descriptive statistics shown above. During the 2012 the submitted quantities and the frequencies of inelastic bids has infact decreased compared to the values recorded in the 2011 and it can be ascribed to the increasing presence of substitutes of traditional energy sources.

Average Elasticity	2011				2012	
	I Quarter	II Quarter	III Qaurter	IV Quarter	I Quarter	II Quarter
Peak	-0.0692	-0.0783	-0.0490	-0.0494	-0.0566	-0.0849
Off-Peak	-0.0755	-0.0752	-0.0542	-0.0511	-0.0687	-0.0453

Peak / Off-Peak Average Elasticity by Quarter. 2011-2012

The monthly aggregation enlightens that in 2011 winter average elasticities do not much differ from the corresponding values recorded in the summer months. Here we can see how the effects of dispatching priority recognized to photovoltaic plants (working for more hours during the summer) are not so appreciable. Although incentives and dispatching priority were introduced to promote market integration of renewable generation plants one of the result has been a distortion of market signals. Photovoltaic plants, whose marginal cost are usually null, profit by the competitive advantages of dispatching priority since they can be ensured to operate in the market offering higher price than the average. The recent attempt of Italian government to reshape the structure of incentives, reducing the amount of subsidies to photovoltaic plants, confirms a tendency to reduce the distortion in the competition measures promoting renewable sources; the decision enlighten infact how incentives to photovoltaic had been lower than benefits they had led. Subsidies were defined in order to overcome the market failure whose signals (the prices) did not include environmental esternalities. New regulation should be always preceded by preliminary benefit/cost analysis, however, this market failure would not justify in Italy the extent of subsidies and just in the last few years it has been disclosed the awareness that the environmental costs were lower than incentives financed by contributors.

Average Elasticity											
January	February	March	April	May	June	July	August	September	October	November	December
-0.0499	-0.0606	-0.1065	-0.0678	-0.1228	-0.0397	-0.0549	-0.0565	-0.0433	-0.0477	-0.0470	-0.0559

Monthly Average Elasticity. 2011.

Average Elasticity					
January	February	March	April	May	June
-0.0510	-0.0621	-0.0749	-0.0618	-0.0452	-0.0651

Monthly Average Elasticity. Jan.-Jun. 2012.

As regards to zone segmentation, we computed average elasticities within the group of hours having the same number of zone segmentation after congestion. The 2012 shows increasing tendency to congestions which can explain the reducing yearly elasticity.

In the first semester of the 2011 and 2012 elasticities were higher when congestions occurred and there was the maximum segmentations of the market (division in three and four zones), while in the second semester of 2011 the elasticity behaviour turned around: higher levels of elasticity were recorded when single market occurred. We have to remember that congestions are physical violations of transmission constraints given higher levels of quantity traded; then the derived results for the second semester may confirm the hypotesis that presence of congestions is negative correlated with elasticity.

Zone	2011-First Semester			2011-Second Semester			2012 First Semester		
	Average	Peak	Off-Peak	Average	Peak	Off-Peak	Average	Peak	Off-Peak
5	0	0	0	0	0	0	0	0	0
4	-0.076	-0.074	-0.078	-0.046	-0.041	-0.052	-0.070	-0.068	-0.064
3	-0.073	-0.075	-0.071	-0.053	-0.050	-0.056	-0.063	-0.070	-0.059
2	-0.070	-0.075	-0.066	-0.051	-0.050	-0.052	-0.064	-0.072	-0.058
1	-0.063	-0.061	-0.065	-0.053	-0.053	-0.052	-0.058	-0.082	-0.045

Average Elasticity by Zone Segmentation. 2011-2012.

The decomposition of average elasticities between peak and off-peak hours shows that within peak period, the lower elasticities were recorded when both single market and maximum segmentation market (four zone division) occurred. The lowest average elasticity was recorded in the presence of maximum segmentation of national market that gives evidence of higher levels of demand and income. As we said before, high levels of expenditure availability affect demand elasticity reducing the responsiveness to change in price. Moreover, congestions during the peak hours may suggest that electricity is an essential commodity whose demand is stiff and whose consumption can not be postponed.

Looking at off-peak hour elasticities, they show an inverted behaviour, since they are higher when there are congestions. During off-peak hours electricity is less allocated for industrial uses and it means that traders have a more flexible behaviour when the quantities traded are greater.

	January		February		March		April		May		June	
	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.
5	0	.	0	.	0	.	0	.	0	.	0	.
4	5	-0.050	11	-0.073	4	-0.105	1	-0.062	16	-0.118	17	-0.035
3	91	-0.049	146	-0.064	125	-0.094	110	-0.079	107	-0.134	133	-0.029
2	269	-0.047	178	-0.074	226	-0.091	215	-0.081	225	-0.121	205	-0.034
1	7	-0.033	1	-0.002	17	-0.104	34	-0.085	24	-0.102	5	-0.043

Peak Average Elasticity by Zone Segmentation. Jan.-Jun. 2011.

	July		August		September		October		November		December	
	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.
5	0	.	0	.	0	.	0	.	0	.	0	.
4	22	-0.065	0	.	10	-0.020	7	-0.057	1	-0.020	2	-0.044
3	152	-0.060	52	-0.067	221	-0.022	146	-0.066	163	-0.041	137	-0.045
2	196	-0.062	311	-0.060	124	-0.023	202	-0.063	156	-0.044	202	-0.046
1	2	-0.076	9	-0.088	5	-0.019	17	-0.054	40	-0.034	31	-0.048

Peak Average Elasticity by Zone Segmentation. Jul.-Dec. 2011

	January		February		March		April		May		June	
	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.
5	0	.	0	.	0	.	0	.	0	.	0	.
4	7	-0.052	10	-0.044	3	-0.187	8	-0.059	15	-0.071	18	-0.054
3	117	-0.056	73	-0.060	105	-0.128	152	-0.055	162	-0.077	123	-0.050
2	121	-0.057	135	-0.056	122	-0.122	120	-0.051	165	-0.072	157	-0.037
1	127	-0.056	118	-0.054	141	-0.120	80	-0.050	30	-0.065	62	-0.047

Off-Peak Average Elasticity by Zone Segmentation. Jan.-Jun. 2011

	July		August		September		October		November		December	
	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av Elast.
5	0	.	0	.	0	.	0	.	0	.	0	.
4	25	-0.044	5	-0.027	7	-0.079	7	-0.044	1	-0.039	4	-0.078
3	193	-0.045	162	-0.059	172	-0.074	93	-0.036	94	-0.054	67	-0.070
2	109	-0.044	128	-0.051	99	-0.060	155	-0.036	127	-0.055	153	-0.065
1	45	-0.044	77	-0.060	82	-0.058	117	-0.034	138	-0.053	148	-0.065

Off-Peak Average Elasticity by Zone Segmentation. Jul.-Dec. 2011

	January		February		March		April		May		June	
	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.
5	0	.	14	-0.060	0	.	0	.	0	.	0	.
4	0	.	186	-0.057	2	-0.085	3	-0.049	0	.	1	-0.118
3	77	-0.038	145	-0.057	88	-0.076	113	-0.058	64	-0.067	126	-0.110
2	285	-0.037	3	-0.067	254	-0.074	214	-0.056	266	-0.066	202	-0.114
1	10	-0.039	0	.	28	-0.075	30	-0.053	42	-0.070	31	-0.120

Peak average Elasticity by Zone Segmentation. Jan.-Jun. 2012.

	January		February		March		April		May		June	
	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.
5	0	.	30	-0.066	2	-0.072	0	.	0	.	0	.
4	2	-0.050	159	-0.069	32	-0.085	6	-0.051	0	.	9	-0.021
3	93	-0.060	122	-0.070	112	-0.076	60	-0.067	64	-0.021	98	-0.019
2	199	-0.059	37	-0.063	143	-0.079	212	-0.067	231	-0.020	211	-0.017
1	78	-0.063	0	.	82	-0.073	82	-0.067	77	-0.022	42	-0.017

Off-Peak average Elasticity by Zone Segmentation. Jan.-Jun. 2012.

Elasticities aggregated by PUN (the market equilibrium price) percentiles show higher values for lower perccentiles (both in peak and off peak hours). Peak elasticities show larger variability, ranging from (-0.065 to -0.046). Lower price levels means lower quantities traded and lower income level, thus, as we said before, Industrial consumers and traders with limited expenditure availability (referring essentially to domestic user) have more flexible behaviuor given changes in price.

Percentile	January		February		March		April		May		Jun	
	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak
10	-0.0536	.	-0.0282	.	-0.1024	-0.1390	-0.0870	-0.0617	-0.0976	-0.0682	-0.0276	-0.0214
9	-0.0533	.	-0.0358	-0.0103	-0.0997	-0.1007	-0.0837	-0.0517	-0.1386	-0.0780	-0.0310	-0.0502
8	-0.0470	-0.0600	-0.0759	-0.0722	-0.0925	-0.1295	-0.0889	-0.0549	-0.1324	-0.0752	-0.0328	-0.0428
7	-0.0464	-0.0593	-0.0819	-0.0649	-0.0920	-0.1308	-0.0829	-0.0484	-0.1397	-0.0673	-0.0344	-0.0456
6	-0.0503	-0.0586	-0.0666	-0.0608	-0.0926	-0.1391	-0.0797	-0.0583	-0.1229	-0.0723	-0.0350	-0.0411
5	-0.0480	-0.0505	-0.0773	-0.0601	-0.0883	-0.1133	-0.0788	-0.0528	-0.1024	-0.0754	-0.0302	-0.0409
4	-0.0437	-0.0580	-0.0728	-0.0483	-0.0879	-0.1211	-0.0820	-0.0494	-0.1088	-0.0685	-0.0328	-0.0457
3	-0.0454	-0.0547	-0.0800	-0.0602	-0.0882	-0.1276	-0.0760	-0.0527	-0.1030	-0.0735	-0.0299	-0.0388
2	-0.0433	-0.0568	-0.0918	-0.0585	-0.0868	-0.1315	-0.0756	-0.0514	-0.0800	-0.0787	-0.0305	-0.0460
1	.	-0.0567	.	-0.0524	.	-0.1163	-0.0801	-0.0536	-0.0676	-0.0715	-0.0153	-0.0483

Montlhy average Elasticity by PUN Percentile. Jan.-Jun. 2011.

Percentile	July		August		September		October		November		December	
	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak
10	-0.0611	-0.0469	-0.0480	-0.0663	-0.0250	-0.0700	-0.0608	-0.0337	-0.0360	-0.0514	-0.0449	-0.0629
9	-0.0620	-0.0455	-0.0526	-0.0576	-0.0199	-0.0726	-0.0620	-0.0373	-0.0397	-0.0580	-0.0477	-0.0670
8	-0.0614	-0.0461	-0.0616	-0.0489	-0.0198	-0.0651	-0.0639	-0.0350	-0.0506	-0.0500	-0.0463	-0.0675
7	-0.0593	-0.0447	-0.0621	-0.0605	-0.0243	-0.0699	-0.0682	-0.0340	-0.0437	-0.0583	-0.0436	-0.0670
6	-0.0606	-0.0445	-0.0646	-0.0568	-0.0290	-0.0646	-0.0672	-0.0363	-0.0420	-0.0544	-0.0477	-0.0645
5	-0.0615	-0.0454	-0.0639	-0.0545	-0.0238	-0.0629	-0.0560	-0.0367	-0.0523	-0.0547	-0.0471	-0.0649
4	-0.0630	-0.0449	-0.0692	-0.0489	-0.0145	-0.0549	-0.0695	-0.0342	-0.0415	-0.0528	-0.0445	-0.0668
3	-0.0635	-0.0425	-0.0680	-0.0595	-0.0171	-0.0730	-0.0699	-0.0349	-0.0440	-0.0567	-0.0485	-0.0719
2	-0.0620	-0.0470	-0.0754	-0.0511	-0.0141	-0.0576	-0.0782	-0.0370	-0.0429	-0.0543	-0.0505	-0.0630
1	-0.0604	-0.0424	-0.0884	-0.0525	.	-0.0732	-0.0447	-0.0355	-0.0449	-0.0519	.	-0.0668

Monthly Average Elasticity by PUN Percentile. Jul.-Dec. 2011

Percentile	January		February		March		April		May		Jun	
	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak
10	-0.0388	-0.0564	-0.0610	-0.0669	-0.0754	-0.0756	-0.0532	-0.0665	-0.0176	-0.0190	-0.1120	-0.0687
9	-0.0373	-0.0583	-0.0564	-0.0740	-0.0791	-0.0790	-0.0539	-0.0604	-0.0184	-0.0221	-0.1086	-0.0679
8	-0.0363	-0.0621	-0.0585	-0.0616	-0.0742	-0.0784	-0.0564	-0.0666	-0.0153	-0.0188	-0.1130	-0.0681
7	-0.0377	-0.0576	-0.0551	-0.0715	-0.0733	-0.0802	-0.0554	-0.0581	-0.0167	-0.0251	-0.1078	-0.0642
6	-0.0371	-0.0577	-0.0531	-0.0689	-0.0765	-0.0752	-0.0573	-0.0691	-0.0174	-0.0205	-0.1123	-0.0670
5	-0.0386	-0.0616	-0.0542	-0.0679	-0.0754	-0.0778	-0.0580	-0.0701	-0.0179	-0.0205	-0.1146	-0.0664
4	-0.0350	-0.0585	-0.0430	-0.0675	-0.0757	-0.0811	-0.0579	-0.0657	-0.0176	-0.0194	-0.1165	-0.0684
3	-0.0377	-0.0587	-0.0450	-0.0691	-0.0739	-0.0749	-0.0581	-0.0708	-0.0170	-0.0198	-0.1124	-0.0669
2	-0.0370	-0.0589	-0.0456	-0.0686	-0.0728	-0.0807	-0.0597	-0.0678	-0.0168	-0.0210	-0.1212	-0.0632
1	.	-0.0639	-0.0488	-0.0657	-0.0738	-0.0746	-0.0569	-0.0696	-0.0205	-0.0198	-0.1254	-0.0680

Monthly Average Elasticity by PUN Percentile. Jul.-Dec. 2012

Percentile	Mean		Stand. Dev		Min.		Max.	
	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak
1	-0.065	-0.062	0.031	0.032	-0.182	-0.207	-0.002	-0.003
2	-0.072	-0.063	0.032	0.030	-0.184	-0.211	0.000	-0.004
3	-0.063	-0.063	0.031	0.035	-0.186	-0.213	-0.005	-0.009
4	-0.064	-0.063	0.035	0.032	-0.184	-0.212	-0.001	-0.010
5	-0.068	-0.064	0.038	0.033	-0.180	-0.210	-0.003	-0.006
6	-0.068	-0.060	0.040	0.033	-0.184	-0.205	-0.003	-0.004
7	-0.061	-0.059	0.038	0.027	-0.188	-0.211	-0.002	-0.003
8	-0.060	-0.058	0.040	0.028	-0.184	-0.211	0.000	-0.005
9	-0.053	-0.059	0.033	0.024	-0.181	-0.214	0.000	-0.006
10	-0.046	-0.059	0.025	0.022	-0.125	-0.200	0.000	-0.010

Average Elasticity by PUN Percentiles. 2011

Percentile	Mean		Stand. Dev		Min.		Max.	
	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak
1	-0.080	-0.062	0.031	0.023	-0.016	-0.010	-0.135	-0.103
2	-0.077	-0.057	0.031	0.026	-0.030	-0.004	-0.135	-0.109
3	-0.066	-0.057	0.023	0.026	-0.024	-0.003	-0.135	-0.109
4	-0.072	-0.056	0.029	0.025	-0.020	-0.006	-0.135	-0.112
5	-0.073	-0.057	0.029	0.027	-0.025	-0.010	-0.141	-0.112
6	-0.070	-0.058	0.030	0.024	-0.020	-0.008	-0.141	-0.095
7	-0.066	-0.061	0.028	0.022	-0.021	-0.012	-0.142	-0.107
8	-0.068	-0.059	0.031	0.024	-0.005	-0.006	-0.141	-0.120
9	-0.064	-0.059	0.030	0.025	-0.017	-0.009	-0.140	-0.111
10	-0.065	-0.059	0.026	0.026	-0.011	-0.009	-0.141	-0.115

Average Elasticity by PUN Percentiles. 2012

To summarize, elasticity is higher during off-peak hour and maximum segmentation division.

Chapter 5

The Heteroskedastic Model

In this chapter I postulate heterosedasticity assuming that hourly observations can have different price volatility and load variability.

Since observations refer to different hours, heteroskedasticity can be a plausible assumption to be explored: each hour is characterized by different price volatility and variability in the load which can be included in the model.

In the previous chapter we assumed the vector ε to be normal distributed with block diagonal covariance matrix having the same elements Σ. This statement is really a combination of a several assumptions, some of which may be relaxed. The assumption that the error terms have mean zero is an innocuos one since, if the model has non zero mean, this last can be incorporated into the intercept. A new model, identical to the old except for the intercept, can be created which have not mean zero errors. On the other hand, the assumption that $Var(\varepsilon) = \Sigma \otimes I_N$ is not innocuous in many applications. In this chapter we consider an empirical way of relaxing this assumption, in particular we assume:

1. For each equation the error terms ε_j, $j = 1, ..., m$ have a multivariate normal distribution with zero mean and covariance matrix Σ_j that is a positive definite matrix.

2. All elements of X remain fixed (i.e. are not random variables).

Heteroskedasticity refers to a model where the covariance matrix of the error terms are different across equations, that is, in the block diagnal matrix Ω, the non-null matrix are different from each other:

$$Var(\varepsilon) = \Omega = \begin{bmatrix} \Sigma_1 & 0 & ... & 0 \\ 0 & \Sigma_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \Sigma_m \end{bmatrix} = \begin{bmatrix} h^{-1} \times \Omega_1 & 0 & ... & 0 \\ 0 & h^{-1} \times \Omega_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & h^{-1} \times \Omega_m \end{bmatrix} \tag{5.1}$$

The same matrix can be written in terms of precison, substituting H_j^{-1} for Σ_j.

In this model we substitute for the covariance matrix Σ the precision h (the inverse of the variance $\sigma^2 = h^{-1}$ such that: $h^{-1} \times \Omega_j = \Sigma_j$ for all $j = 1, ..., m$. Moreover, we manage a hierarchical model, since we assume that we do not know the values assumed by the elements of the Ω_i matrixes. This model allows to free up the normality assumption, since unknown heteroskedasticity is equivalent to a linear regression model with Student-t errors.

5.1 The trasformed model

Before discussing the likelihood function, prior and posterior and computational issues, general results of the model are presented. Since Ω is a positive definite matrix, Cholesky decomposition can be applied, then it exist a $(Nm \times Nm)$ matrix P with the property that $P\Omega P' = I_{Nm}$.

Given the model:

$$y = X\beta + \varepsilon \tag{5.2}$$

with $\varepsilon_i|\Omega \sim N(0, h^{-1} \times \Omega_i)$

if we multiply both sides of the previous equation by P, we obtain the trasformed model

$$y^* = X^*\beta + \varepsilon^* \tag{5.3}$$

where $y^* = Py$, $X^* = PX$ and $\varepsilon^* = P\varepsilon$. It can be verified that $\varepsilon^* \sim N(0, I_{Nm})$. Hence the trasformed model falls again into the standard Normal linear regression model. There are two important implications to be discussed. If Ω is known Bayesian analysis is straightforward, on the other hand, if Ω is unknown, a computational issue arises since we need to derive its posterior density.

5.2 The Likelihood Function

Using the properties of the multivariate Normal distribution, the likelihood function of transformed model can be seen to be:

$$\begin{aligned} p(y|\beta, h, \Lambda) &= \frac{h^{\frac{Nm}{2}}}{(2\pi)^{\frac{N \cdot m}{2}}} |\Lambda|^{\frac{1}{2}} \exp\left[-\frac{h}{2}(y - X\beta)'\Omega^{-1}(y - X\beta)\right] = \\ & \frac{h^{\frac{Nm}{2}}}{(2\pi)^{\frac{N \cdot m}{2}}} \exp\left[-\frac{h}{2}(y^* - X^*\beta)'(y^* - X^*\beta)\right] \end{aligned} \tag{5.4}$$

We use indipendent Normal Gamma prior for β and h:

$$p(\beta) = N(\beta, V) \tag{5.5}$$

$$p(h) = G(\nu_0, s_0^{-2}) \tag{5.6}$$

Moreover, we assume that β and h have distributions indipendent on Ω, whose prior will be defined later.

$$p(\beta, h, \Omega) = p(\beta)p(h)p(\Omega) \tag{5.7}$$

Comparing to the previous homoskedastic model, the first difference has been to replace the covariance matrix with the scalar parameter $h^{-1} = \sigma^2$ multiplied by Ω_j, then the multivariate prior has changed in the corresponding scalar version: the Gamma distribution.

As in the previuos chapter, the prior hyperparameter elicitation comes from the previous empirical study (see Bigerna & Bollino [13]).

5.3 Posterior Distributions

The joint posterior distribution of all parameters is, as always, the likelihood function times the priors:

$$
\begin{aligned}
p(\beta, h, \Omega|y) \propto\ & p(\Omega) \times \\
& \exp\left\{\left[-\frac{h}{2}(y^* - X^*\beta)'(y^* - X^*\beta)\right]\right\} \\
& \times \exp\left\{-\frac{1}{2}(\beta - \beta_0)'V_0^{-1}(\beta - \beta_0)\right\} \\
& \times h^{\frac{1}{2}(Nm+\nu_0-2)} \exp\left[-\frac{h\nu_0}{2s_0^{-2}}\right]
\end{aligned} \tag{5.8}
$$

This posterior is based on the notation related to the transformed model; without writing the posterior distributions referring to the initial model, we can notice that this joint posterior does not take any well-known form and posterior inference needs the use of simulation in order to derive some useful summary statistics related to the parameters.

As in the previous chapter, the full conditional posterior distributions are derived.

The full conditional for β is a Normal:

$$\beta|y, h, \Omega \sim N(\beta_n, V_n) \tag{5.9}$$

where:

$$V_n = (V_0^{-1} + hX\Omega X)^{-1} \tag{5.10}$$

and

$$\beta_n = V_n(V_0^{-1}\beta_0 + hX\Omega X\widehat{\beta}(\Omega)) \tag{5.11}$$

with

$$\widehat{\beta}(\Omega) = (X^{*\prime}X^*)^{-1}X^{*\prime}y^* = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y \tag{5.12}$$

The posterior distribution of h conditional on the other parameters in the model is a Gamma:

$$h|y, \beta, \Lambda \sim G(s_n^{-2}, \nu_n) \tag{5.13}$$

where:

$$\nu_n = Nm + \nu_0 \tag{5.14}$$

and

$$s_n^2 = \frac{(y - X\beta)'\Lambda(y - X\beta) + \nu_0 s_0^2}{\nu_n} \tag{5.15}$$

Conditioning on Ω, the two full conditional distributions for β and h combine data and prior information.

Given Ω, the full conditional distributions of β and h are ascribable to a well-known analytical form, as in the traditional linear model, the Gibbs Sampling algorithm can exploit this two densities and constructs a stationary Markov Chain. The first part of simulation procedure has been defined. However, the full conditional posterior of Ω does not take any recognizable form:

$$p(\Omega|\beta, h, y) \propto p(\Omega)|\Omega|^{\frac{1}{2}} \left\{ \exp\left[-\frac{h}{2}(y - X\beta)'\Omega^{-1}(y - X\beta) \right] \right\} \tag{5.16}$$

The joint posterior distribution $p(\beta, h, \Omega|y)$ keeps remaining untractable since $p(\beta, h, \Omega|y) \neq p(\beta, h|\Omega, y) \cdot p(\Omega|\beta, h, y)$ and the inference related to the random parameters β, h and Ω has not been possible yet: The prior distribution for Ω needs to be deeplier investigated in order to design the second component of similation.

5.4 Heteroskedasticity of unknown form

When Ω is an unknown parameter, the elements of the matrix Ω in (5.1) become random variables. The treatment of heteroskedasticity of unknown form is a challenging task and involves the use of a hierarcical prior. Hierarchical priors have played a big role in many recent developments in Bayesian statistical theory and gradually have been becoming popular in econometrics as well. They are commonly used as a way of making parameter-rich models flexible to statistical analysis.

As it has already been done with the variance, that it has been replaced by its precision h, also the different matrixes Ω_j can be expressed in terms of precision substituting for $\{\Omega_j\}_{j=1}^m$ their inverse $\{\Lambda_j\}_{j=1}^m \equiv \left\{\Omega_j^{-1}\right\}_{j=1}^m$ for all $j = 1, ..m$.

Introducing unknown heteroskedasticity increases the number of parameters to be estimated becoming $k+1+N*\frac{m^2+1}{2}$; if we treat $\Lambda_1,...,\Lambda_m$ as completely indipendent and unrestricted matrixes, we would not have enough observations to estimate each one of them. For this reason the exchangebility of Λ_i becomes an assumption essential to deal with this high dimensional model. The prior for Λ becomes:

$$p(\Lambda)=\prod_{j=1}^{m} f_W(\Lambda_j|\Lambda_0,\nu_\lambda) \tag{5.17}$$

which states that Λ_js are different from one another but they are *i.i.d.* draws from the same Wishart distribution (the hierarchical prior). The hierarchical prior imposes a structure to the model that preserves flexibility and makes estimation be possible.

Some comments about the specification of the matrix Λ need to be done. Previous litterature about heteroskedasticity has defined Λ as a simple diagonal matrix, without considering correlation across observations pertaining the same equations. Given the exchangebility assumption, the non-null elements of Λ have been considered scalars drawn from a gamma distribution. Then, in the traditional heteroskedastic model, the prior is an univariate density. Instead, as it is shown above, in the current model matrix Ω_m are "full" since it is assumed correlation across observations of the same equation. Then, I adopted a multivariate density as hierarchical prior and that increases the dimension of the inferential problem as the computational difficulty.

Moreover, the use of a Wishart prior distribution (or equivantely assuming that Ω_j are drawn from an inverse Wishart with scale equal to Λ_0^{-1}) allows to turn out a linear regression model with *i.i.d.* Student-t error terms with $\nu_\lambda > m$ degrees of freedom.

In other words :

$$\varepsilon_i|\Lambda_i^{-1} \sim N(0,\sigma_i^2\Lambda_i^{-1}) \tag{5.18}$$

$$\Lambda_i \sim W(\Lambda_0,\nu_\lambda) \tag{5.19}$$

$$\varepsilon_i \sim t(0,\sigma^2,\nu_\lambda) \tag{5.20}$$

The model becomes more flexible since Student-t distribution that is a more general class of distributions that includes Normal density as a special case (occurring when the degrees of freedom ν_λ tent to infinity).

Our treatment of unknown heteroskedasticity is equivalent to a scale mixture of Normal. The error tems ε_i are disributed according to a mixture of m different normal distributions. That is:

$$\varepsilon_i = \sum_j e_{ij} \left(\alpha_j + (H_j)^{-1/2} \eta_{ij} \right) \tag{5.21}$$

where η_{ij} is i.i.d $N(0, I_m)$ for $i = 1, ..., N$, $j = 1...J$ and e_{ij}, α_j and H_j are all parameters. The e_{ij} is a dicotomous random variable and indicates which distribution component in the mixture the ith error is drawn from.

$$e_{ij} = \begin{cases} 1 & \text{if } \varepsilon_j \sim N(\alpha_j, H_j) \\ 0 & \text{otherwise} \end{cases} \tag{5.22}$$

Since it is unknown which component the ith error is drawn from, we define $p_j = P(e_{ij} = 1)$ for $j = 1, ..., m$ the probability of the error being drawn from the jth component in the mixture. Formally it means that e_{ij} are i.i.d draws from a Multinomial distribution

$$e_i \sim M(1, p) \tag{5.23}$$

where $p = (p_1 ... p_m)'$ is the probability vector p.

The assumption that Λ_i follows a Wishart distribution and that, given Λ_i, the errors are indipendent Normal $(0, h^{-1}\Lambda_i^{-1})$ is equivalent to the assumption that the distribution of error term ε is a weighted average of Normals having different variances but the same means (i.e. all errors have mean equal to zero). When we mix the error terms' normal distributions using $f_W(\Lambda_i|\Lambda_0, \nu_\lambda)$, they end up to be equal to the Student-t distribution. Intuitively, assuming that a Normal model is too restrictive, a more flexible distribution taking a mixture (the weighted average) of Normals can be created. As more and more Normals are mixed, as the distribution becomes more and more flexible and can approximate any distribution with high degree of freedom. Mixtures of Normal are powerful tool to be used when economic theory does not suggest any particular form of likelihood function and you wish to be more flexible.

However, this model use a finite mixture of Normal and it cannot be considered non-parametric in the sense that it can not accomodate any distribution, it is "just an extremely flexible modeling strategy" (Koop [47]).

Parameter ν_λ is not known and Bayesian framework imposes to define a prior distribution $p(\nu_\lambda)$. The prior of λ is specified in two steps, firstly we specify $p(\Lambda|\nu_\lambda) = \prod_{i=1}^{N} f_W(\Lambda_i|\Lambda_0, \nu_\lambda)$, secondly we define $p(\nu_\lambda)$; in this way these two steps refer to a hierarchical prior model. $p(\Lambda|\nu_\lambda)$ and $p(\nu_\lambda)$ are the features necessary to design the second part of the simulation procedure.

It must be mentioned the risk concerning the use for the degree of freedom ν_λ an imoroper prior which allocates same probability to all values between zero and infinite. For degrees of freedom greater than 100, where Student-t distribution approach the Normal distribution, this prior, far from being noninformative, states that the error terms are normally distributed.

5.5 Bayesian Computation

Now, let it focus on $p(\Lambda|y, \beta, h, \nu_\lambda)$ and $p(\nu_\lambda|y, \beta, h, \Lambda)$.

$$p(\Lambda|y, \beta, h, \nu_\lambda) = \prod_{i=1}^{N} p(\Lambda_i|y, \beta, h, \nu_\lambda) \tag{5.24}$$

where

$$p(\Lambda_i|y, \beta, h, \nu_\lambda) = W\left((\nu_\lambda + m)\left[h(\varepsilon_i \varepsilon_i')\right]^{-1} + \nu_\lambda, \nu_\lambda + m\right) \tag{5.25}$$

Conditional on β, ε_i can be calculated and hence also the parameters of the Gamma density can be sampled within the Gibbs sampler.

Problems arise in the derivation of full conditional posterior for ν_λ. Since ν_λ is positive we assume as a prior an exponential distribution that is a gamma with two degree of freedom:

$$p(\nu_\lambda) = G(\nu_0, 2) \tag{5.26}$$

Then, the full conditional posterior is:

$$p(\nu_\lambda|y, \beta, h, \Lambda) = p(\nu_\lambda|\Lambda) \propto p(\Lambda|\nu_\lambda)p(\nu_\lambda)^1 \tag{5.27}$$

The kernel of the posterior conditional of ν_λ is simply (5.24) times (5.26):

[1] Since ν_λ does not eneter in the likelihood $p(\nu_\lambda|y, \beta, h, \lambda) = p(\nu_\lambda|\lambda)$.

$$p(\nu_\lambda|\Lambda) \propto p(\Lambda|\nu_\lambda)p(\nu_\lambda)$$
$$\propto \left(\frac{\nu_\lambda}{2}\right)^{\frac{N\nu_\lambda}{2}} \Gamma\left(\frac{\nu_\lambda}{2}\right)^{-N} \exp(-\eta\nu_\lambda) \tag{5.28}$$

where

$$\eta = \frac{1}{\nu_0} + \frac{1}{2}\sum_{i=1}^{N}\left[\ln|\Lambda|_i^{-1}) + tr|\Lambda_0^{-1}\Lambda_i\right] \tag{5.29}$$

The density derived in (5.28) is not again a standard one, so another algorithm for the posterior simulation of the degree of freedom need to be performed. The simulation strategy is the following:

- Find an algorithm which simulates a sample from $p(\nu_\lambda|\Lambda)$.
- Given ν_λ, run Gibbs Sampling simulating $p(\beta, h, \Lambda, \nu_\lambda|y)$ using the $\{\nu_\lambda^t\}$ sample, $p(\beta|y, h, \Lambda)$ in (5.9) and $p(h|y, \beta, \Lambda)$ in (5.13).

Formally, the full conditional to be used in the Gibbs Sampler should have been $p(\beta|y, h, \Lambda, \nu_\lambda)$ and $p(h|y, \beta, \Lambda, \nu_\lambda)$, but condional on Λ, ν_λ adds no new information and thus $p(\beta|y, h, \Lambda, \nu_\lambda) = p(\beta|y, h, \Lambda)$ and $p(h|y, \beta, \Lambda, \nu_\lambda) = p(h|y, \beta, \Lambda)$.

When simulation concerns an univariate and bounded distribution, Geweke [38] recommends the use of Importamce Sampling techniques for the simulation of ν_λ from $p(\nu_\lambda)$. Here, I used the Random Walk Metropolis-Hastings Algorithm.

The candidate generating function is $q(\nu_\lambda^{(s-1)}; \nu_\lambda^*) = N([\nu_\lambda^* - \nu_\lambda^{(s-1)}], 0.2)$ and number of replications are set equal to 11000. After discarding the first 1000 realizations, the chain $\left\{\beta^i, h^i, \nu_\lambda^i\right\}_{i=1001}^{11000}$. The sequence simulate a sample from the posterior $p(\beta, h, \Lambda, \nu_\lambda|y)$.

5.6 Empirical Results

The statistical model has been applied to data referring the whole 2011. After relaxing homooskedastic assumption, average elsaticicities have increased compared with their analogous in previous model. Moreover, allowing the variance to differ across observations of the same equations has led the peak hour elasticities to be, on average, higher than off-peak ones. The introduction of new parameters (the heteroskedastic terms Ω_i in the covariance matrix) allows to process more information which determined the rise of peak hours elasticities. Hourly average elasticities range from -0.1434, recorded in September to -0.0360 recorded in November. Comparing to the off-peak estimates, peak hour elasticities show higher variability, going from -0.1434 to -0.0484, while the off-peak ones vary between -0.070 and -0.0359. It means that the application of heteroskedastic model affects essentially the variance of peak estimates.

	Hour	January	February	March	April	May	June	July	August	September	October	November	December
Peak	9	-0.0756	-0.0699	-0.0869	-0.0676	-0.0853	-0.0495	-0.0506	-0.0730	-0.1388	-0.0579	-0.0603	-0.0599
	10	-0.0764	-0.0696	-0.0864	-0.0679	-0.0855	-0.0492	-0.0509	-0.0731	-0.1321	-0.0582	-0.0598	-0.0600
	11	-0.0771	-0.0695	-0.0867	-0.0682	-0.0859	-0.0484	-0.0510	-0.0769	-0.1387	-0.0582	-0.0599	-0.0579
	12	-0.0771	-0.0705	-0.0872	-0.0682	-0.0863	-0.0485	-0.0504	-0.0728	-0.1342	-0.0578	-0.0600	-0.0556
	13	-0.0769	-0.0707	-0.0877	-0.0677	-0.0868	-0.0490	-0.0505	-0.0754	-0.1355	-0.0568	-0.0599	-0.0571
	14	-0.0761	-0.0701	-0.0876	-0.0686	-0.0863	-0.0493	-0.0503	-0.0777	-0.1416	-0.0569	-0.0603	-0.0579
	15	-0.0759	-0.0701	-0.0878	-0.0677	-0.0859	-0.0498	-0.0505	-0.0758	-0.1374	-0.0571	-0.0607	-0.0615
	16	-0.0758	-0.0706	-0.0875	-0.0682	-0.0852	-0.0496	-0.0508	-0.0756	-0.1398	-0.0578	-0.0602	-0.0588
	17	-0.0751	-0.0707	-0.0872	-0.0678	-0.0853	-0.0488	-0.0512	-0.0757	-0.1371	-0.0579	-0.0605	-0.0610
	18	-0.0751	-0.0706	-0.0868	-0.0676	-0.0859	-0.0486	-0.0514	-0.0767	-0.1369	-0.0578	-0.0611	-0.0612
	19	-0.0759	-0.0703	-0.0869	-0.0679	-0.0856	-0.0493	-0.0520	-0.0795	-0.1387	-0.0586	-0.0610	-0.0613
	20	-0.0753	-0.0710	-0.0869	-0.0688	-0.0856	-0.0499	-0.0516	-0.0782	-0.1434	-0.0582	-0.0611	-0.0581
Off-Peak	21	-0.0571	-0.0528	-0.0452	-0.0580	-0.0630	-0.0438	-0.0594	-0.0577	-0.0498	-0.0707	-0.0395	-0.0607
	22	-0.0572	-0.0541	-0.0456	-0.0592	-0.0657	-0.0434	-0.0588	-0.0576	-0.0487	-0.0706	-0.0419	-0.0612
	23	-0.0573	-0.0556	-0.0489	-0.0580	-0.0629	-0.0432	-0.0592	-0.0576	-0.0495	-0.0676	-0.0447	-0.0596
	24	-0.0576	-0.0522	-0.0459	-0.0613	-0.0662	-0.0427	-0.0588	-0.0577	-0.0498	-0.0664	-0.0418	-0.0595
	1	-0.0576	-0.0553	-0.0455	-0.0595	-0.0650	-0.0432	-0.0587	-0.0578	-0.0496	-0.0701	-0.0422	-0.0601
	2	-0.0581	-0.0557	-0.0470	-0.0603	-0.0679	-0.0439	-0.0588	-0.0578	-0.0503	-0.0675	-0.0460	-0.0599
	3	-0.0580	-0.0559	-0.0504	-0.0619	-0.0632	-0.0438	-0.0588	-0.0577	-0.0514	-0.0645	-0.0372	-0.0600
	4	-0.0579	-0.0556	-0.0521	-0.0641	-0.0642	-0.0439	-0.0590	-0.0577	-0.0506	-0.0668	-0.0365	-0.0604
	5	-0.0573	-0.0532	-0.0526	-0.0594	-0.0665	-0.0429	-0.0593	-0.0569	-0.0504	-0.0622	-0.0359	-0.0607
	6	-0.0567	-0.0519	-0.0517	-0.0584	-0.0635	-0.0429	-0.0591	-0.0569	-0.0510	-0.0675	-0.0362	-0.0603
	7	-0.0570	-0.0531	-0.0500	-0.0582	-0.0666	-0.0435	-0.0593	-0.0569	-0.0509	-0.0689	-0.0360	-0.0609
	8	-0.0572	-0.0526	-0.0475	-0.0624	-0.0676	-0.0441	-0.0591	-0.0565	-0.0513	-0.0675	-0.0366	-0.0611

Hourly Average Elasticity. 2011

Average Elasticity	2011			
	I Quarter	II Quarter	III Qaurter	IV Quarter
Peak	-0.0778	-0.0677	-0.0882	-0.0591
Off-Peak	-0.0533	-0.0562	-0.0556	-0.0558

Peak/Off-Peak Average Elasticity by Quarter. 2011.

Estimates, aggregated by zone segmentation, confirm the results derived in the previous chapter. During Peak hours, higher elasticities has infact been recorded when the single market occurred. When the transmission constraints are violated, elasticity becomes lower and this is particular evident during the peak hours. As we said before, high levels of expenditure availability affect demand elasticity reducing the responsiveness to change in price. Moreover, frequent congestions during peak hour may suggest that electricity is an essential commodity whose demand is stiff and whose consumption can not be postponed.

Off-peak estimates instead, do not show a well defined behaviour, particularly in the first semester; in the second semester instead elasticitity records its highest average value when there was maximum segmentation.

Average Elasticity												
January	February	March	April	May	June	July	August	September	October	November	December	Year
-0.0667	-0.0622	-0.0678	-0.0640	-0.0755	-0.0463	-0.0550	-0.0666	-0.0941	-0.0626	-0.0500	-0.0598	-0.0642

Monthly Average Elasticities. 2011.

Zone	2011-First Semester			2011-Second Semester		
	Average	Peak	Off-Peak	Average	Peak	Off-Peak
5	0	0	0	0	0	0
4	-0.062	-0.068	-0.056	-0.061	-0.057	-0.065
3	-0.060	-0.067	-0.053	-0.067	-0.076	-0.058
2	-0.063	-0.069	-0.056	-0.065	-0.071	-0.059
1	-0.062	-0.070	-0.054	-0.062	-0.070	-0.055

Average Elasticity by Zone Segmentation. 2011.

	January		February		March		April		May		June	
	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.
5	0	.	0	.	0	.	0	.	0	.	0	.
4	5	-0.077	11	-0.061	4	-0.091	1	-0.065	16	-0.079	17	-0.035
3	91	-0.077	146	-0.061	125	-0.085	110	-0.067	107	-0.084	133	-0.029
2	269	-0.075	178	-0.060	226	-0.088	215	-0.068	225	-0.088	205	-0.034
1	7	-0.077	1	-0.059	17	-0.090	34	-0.071	24	-0.081	5	-0.043

Peak Average Elasticity by Zone Segmentation. Jan-Jun. 2011.

	July		August		September		October		November		December	
	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.
5	0	.	0	.	0	.	0	.	0	.	0	.
4	22	-0.065	0	.	10	-0.047	7	-0.056	1	-0.061	2	-0.054
3	152	-0.060	52	-0.067	221	-0.147	146	-0.060	163	-0.061	137	-0.060
2	196	-0.062	311	-0.060	124	-0.132	202	-0.056	156	-0.060	202	-0.059
1	2	-0.076	9	-0.088	5	-0.074	17	-0.064	40	-0.059	31	-0.057

Peak Average Elasticity by Zone Segmentation. Jul-Dec. 2011.

	January		February		March		April		May		June	
	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.
5	0	.	0	.	0	.	0	.	0	.	0	.
4	7	-0.055	10	-0.048	3	-0.070	8	-0.058	15	-0.067	18	-0.040
3	117	-0.057	73	-0.052	105	-0.053	152	-0.060	162	-0.061	123	-0.035
2	121	-0.057	135	-0.057	122	-0.042	120	-0.061	165	-0.070	157	-0.051
1	127	-0.058	118	-0.053	141	-0.050	80	-0.060	30	-0.061	62	-0.043

Off-Peak Average Elasticity by Zone Segmentation.Jan-Jun. 2011.

	July		August		September		October		November		December	
	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av. Elast.	Freq.	Av Elast.
5	0	.	0	.	0	.	0	.	0	.	0	.
4	25	-0.057	5	-0.079	7	-0.059	7	-0.067	1	-0.079	4	-0.048
3	193	-0.058	162	-0.057	172	-0.054	93	-0.066	94	-0.057	67	-0.052
2	109	-0.060	128	-0.060	99	-0.047	155	-0.069	127	-0.060	153	-0.057
1	45	-0.062	77	-0.051	82	-0.046	117	-0.067	138	-0.051	148	-0.053

Off-Peak Average Elasticity by Zone Segmentation.Jul-Dec. 2011.

Elasticities aggregated by PUN percentiles show an opposite behaviour compared to their corresponding homoskedastic estimates. They are higher when higher levels of equilibrium price have been recorded. Since elasticities are higher during the peak periods when higher levels of equilibrium price are recorded, the aggregation by PUN percentile expresses this buyers' behaviour recording higher values when the equilibrium price are higher.

Percentile	January		February		March		April		May		Jun	
	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak
10	-0.0703	.	-0.0710	.	-0.0858	-0.0723	-0.0656	-0.0438	-0.0832	-0.0742	-0.0520	-0.0574
9	-0.0738	.	-0.0694	-0.0374	-0.0861	-0.0502	-0.0695	-0.0664	-0.0878	-0.0587	-0.0518	-0.0439
8	-0.0742	-0.0692	-0.0708	-0.0630	-0.0854	-0.0406	-0.0709	-0.0613	-0.0859	-0.0715	-0.0497	-0.0447
7	-0.0762	-0.0664	-0.0685	-0.0508	-0.0877	-0.0597	-0.0685	-0.0714	-0.0883	-0.0731	-0.0481	-0.0447
6	-0.0796	-0.0560	-0.0700	-0.0512	-0.0862	-0.0594	-0.0646	-0.0587	-0.0859	-0.0682	-0.0485	-0.0473
5	-0.0761	-0.0542	-0.0697	-0.0558	-0.0880	-0.0299	-0.0695	-0.0600	-0.0826	-0.0729	-0.0514	-0.0466
4	-0.0763	-0.0552	-0.0704	-0.0522	-0.0887	-0.0430	-0.0666	-0.0566	-0.0836	-0.0585	-0.0456	-0.0404
3	-0.0718	-0.0575	-0.0729	-0.0565	-0.0902	-0.0499	-0.0670	-0.0607	-0.0842	-0.0633	-0.0496	-0.0455
2	-0.0716	-0.0613	-0.0753	-0.0498	-0.0892	-0.0535	-0.0719	-0.0619	-0.0812	-0.0657	-0.0457	-0.0403
1	.	-0.0547	.	-0.0568	.	-0.0478	-0.0713	-0.0590	-0.0799	-0.0570	-0.0394	-0.0435

Monthly Average Elasticity By PUN Percentile. Jan-Jun. 2011.

Percentile	July		August		September		October		November		December	
	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak
10	-0.0505	-0.0648	-0.0840	-0.0480	-0.1652	-0.0498	-0.0577	-0.0715	-0.0609	-0.0384	-0.0620	-0.0618
9	-0.0531	-0.0685	-0.0840	-0.0612	-0.1211	-0.0501	-0.0570	-0.0614	-0.0608	-0.0449	-0.0600	-0.0620
8	-0.0490	-0.0604	-0.0721	-0.0588	-0.1273	-0.0552	-0.0586	-0.0766	-0.0589	-0.0305	-0.0536	-0.0583
7	-0.0522	-0.0578	-0.0749	-0.0565	-0.1157	-0.0483	-0.0574	-0.0651	-0.0608	-0.0470	-0.0481	-0.0612
6	-0.0513	-0.0613	-0.0732	-0.0533	-0.0808	-0.0443	-0.0571	-0.0696	-0.0588	-0.0412	-0.0630	-0.0566
5	-0.0511	-0.0654	-0.0736	-0.0563	-0.1743	-0.0480	-0.0601	-0.0668	-0.0585	-0.0426	-0.0641	-0.0644
4	-0.0494	-0.0603	-0.0715	-0.0618	-0.1503	-0.0507	-0.0593	-0.0638	-0.0621	-0.0385	-0.0636	-0.0699
3	-0.0516	-0.0581	-0.0777	-0.0556	-0.1614	-0.0511	-0.0564	-0.0562	-0.0617	-0.0517	-0.0633	-0.0663
2	-0.0468	-0.0605	-0.0668	-0.0636	-0.1293	-0.0564	-0.0612	-0.0666	-0.0632	-0.0434	-0.0917	-0.0533
1	-0.0552	-0.0526	-0.0692	-0.0599	.	-0.0541	-0.0630	-0.0759	-0.0640	-0.0299	.	-0.0608

Monthly Average Elasticity By PUN Percentile. Jul-Dec. 2011.

Percentile	Mean		Stand. Dev		Min.		Max.	
	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak
10	-0.084	-0.056	0.060	0.024	-0.309	-0.099	-0.002	0.000
9	-0.078	-0.057	0.049	0.024	-0.298	-0.119	-0.001	-0.001
8	-0.078	-0.056	0.045	0.024	-0.282	-0.124	-0.001	0.000
7	-0.073	-0.057	0.037	0.024	-0.252	-0.129	0.000	0.000
6	-0.070	-0.054	0.032	0.026	-0.291	-0.127	0.000	0.000
5	-0.070	-0.055	0.027	0.027	-0.292	-0.126	-0.005	0.000
4	-0.068	-0.059	0.021	0.024	-0.217	-0.132	-0.001	-0.002
3	-0.070	-0.053	0.022	0.025	-0.262	-0.125	-0.012	-0.001
2	-0.072	-0.052	0.019	0.027	-0.216	-0.130	-0.010	0.000
1	-0.070	-0.054	0.023	0.028	-0.224	-0.131	-0.010	0.000

Average Elasticity by Pun Percentiles. 2011

Chapter 6

Conclusion

This thesis tried to explain buyers' behaviour of Italian Wholesale Electricity Market through Bayesian method.

The first two chapters explain the reason of preferring the Bayesian method rather than the classical frequentist method and enlighten the main elements of Bayesian model.

The third chapter outlines the main features of the Italian Electricity Market after the deregulation process. Given the strategic relevance of electricity sector, the liberalization had been a challenging process which had to face two main issues. First, market structure had been defined in order to guide investments to enhance the trasmission grid or for more flexible and efficient generation plants. Secondly, market structure has to ensure the constant covering of demand profiles in an efficient and competitive way.

After giving summary statistics related to the main features of Italian Electricit market, the last two chapters provide two statistical models applied to derived electricity demand elaticity. The first model is a Seemingly Unrelated Regression equations where the hourly demand equations are supposed to be correlated within the peak/off-peak group of hours. In the Second model instead, I allowed the demand equations within the group of hours to differ in their variance covariance matrixes, introducing in this way Heteroskedasticity. Both the homoskedastic and heteroskedastic model highlight that buyers in the Italian Wholesale Electricity Market react to change in price since the estimated elasticities are different from zero.

Moreover, the estimates differ from one another during the day, on the strenght of the level of electricity loads, the market segmentation structure and the levels of PUN.

In the Homoskedastic Multivariate Linear Model, during the 2011, hourly elasticities were higher during the off-peak hours, when there was maximum segmentation of the market and lower levels of PUN. For the first semester of the 2012 instead, elasticity was higher during peak hours, but its behaviour did not change with respect to the market segmentation and the PUN level.

In the Heteroskedastic Multivariate Linear Model, the elasticities recorded during peak hours became higher than Off-Peak estimates. Moreover, as in the Homoskedastic model, buyers mantain their higher reactivity to changes in price when there is not congestion (particularly during the peak period) and when PUN records high values.

Further development of the research may be the application of the Heteroskedastic model to the data referring the 2012. Moreover, also the computational part can be implemented. Heteroskedastic model represents a novel in the empirical analyses, but the multi-dimension of the inferential problem made the construction of the algorithm the most challenging task of my thesis. The heteroskedastic model designs for empirical data a statistical framework rigorous and detailed. However, posterior simulation requires the discretional choice of the proposal density and the tuning of its parameters, first of all the variance, affecting the behaviour of the chain and the resulting posterior. Running the procedure with other parameters and alternative functional forms of the candidate generating density could be a possible development in order to make a comparison between the different derived estimates.

Appendix

A.1 Stationarity of Gibbs Sampling Algorithm

Let be θ the p-vector of parameters. A Markov Chain used for Monte Carlo simulation has to converge to the target distribution we want to simulate, such that $p_{t+1} = p = p_t$. Recalling the formula 2.10 in Chapter 2, the Gibbs sampling Algorithm will converge to the posterior distribution if there exists a solution a Markov Chain whose kernel satisfies the integral equation:

$$p(\theta') = \int_{\Theta} p(\theta) K(\theta, \theta') d\theta$$

Gibbs sampler partitions the parameter vector θ in $(\theta_{(1)}, \theta_{(2)}, ...\theta_{(B)})$ where $\theta_{(j)}$ for $j = 1, 2, ...B$ is a scalar or vector and B is the number of partitions.

Let demonstrate stationarity condition in the usual case in which $B = 2$. The transition kernel of Gibbs Sampler uses the full conditional distribution: $p(\theta^{(1)}|\theta^{(2)}, y)$ and $p(\theta^{(2)}|\theta^{(1)}, y)$ such that:

$$K(\theta^t, \theta^{t+1}) = p(\theta_{(1)}^{t+1}|\theta_{(2)}^{t}, y) \;\; p(\theta_{(2)}^{t+1}|\theta_{(1)}^{t+1}, y)$$

Let demonstrate that kernel satisfies the integral equation in (2.8):

$$
\begin{aligned}
\int K(\theta^t, \theta^{t+1}|y)p(\theta|y)d\theta &= \int p(\theta_{(1)}^{t+1}|\theta_{(2)}^t, y) \;\; p(\theta_{(2)}^{t+1}|\theta_{(1)}^{t+1}, y) \;\; p(\theta_{(1)}^t, \theta_{(2)}^t|y) \;\; d\theta_{(1)}^t d\theta_{(2)}^t \\
&= \int p(\theta_{(1)}^{t+1}|\theta_{(2)}^t, y) \;\; p(\theta_{(2)}^{t+1}|\theta_{(1)}^{t+1}, y) \;\; p(\theta_{(2)}^t|y) \;\; d\theta_{(2)}^t \\
&= p(\theta_{(2)}^{t+1}|\theta_{(1)}^{t+1}, y) \int p(\theta_{(1)}^{t+1}|\theta_{(2)}^t, y) \;\; p(\theta_{(2)}^t|y) \;\; d\theta_{(2)}^t \\
&= p(\theta_{(2)}^{t+1}|\theta_{(1)}^{t+1}, y) \int p(\theta_{(1)}^{t+1}, \theta_{(2)}^t, y) \;\; d\theta_{(2)}^t \\
&= p(\theta_{(2)}^{t+1}|\theta_{(1)}^{t+1}, y) \;\; p(\theta_{(1)}^{t+1}, y) = p(\theta_{(2)}^{t+1}, \theta_{(1)}^{t+1}, y) \\
&= p(\theta^{t+1}|y)
\end{aligned}
$$

Appendix

A.2 Stationarity of Metropolis Algorithm

First suppose that the the transition kernel of Metropolis Algorithm satifies the condition called *detailed balance condition:*

$$K(\theta, \theta')p(\theta) = K(\theta', \theta)p(\theta')$$

for all θ and $\theta \in \Theta$. Then p is the stationary distribution of the chain. Consider now θ' belong to any set $B \subset \Theta$ and note:

$$\begin{aligned}
\int_{\Theta} K(\theta, B)p(\theta)d\theta &= \int_{\Theta}\int_{B} K(\theta\prime, \theta) \;\; p(\theta'|\theta) \;\; d\theta\prime \;\; d\theta, \\
&\text{by the detailed balance condition} \\
&= \int_{\Theta}\int_{B} K(\theta, \theta') \;\; p(\theta|\theta') \;\; d\theta\prime \;\; d\theta, \\
&= \int_{B} p(\theta|\theta\prime) \;\; d\theta
\end{aligned}$$

because $\int_{\Theta} K(\theta, \theta')d\theta\prime = 1$. Thus $p(\cdot)$ is a stationary distribution of the chain.

It only remains to show that the Metropolis kernel satisfies detailed the balance condition. To see this, note that the transition kernel of Metropolis Algorithm can be written as:

$$K(\theta, \theta') = \rho(\theta, \theta')q(\theta'|\theta) + (1 - r(\theta))\delta_\theta(\theta')$$

where $\delta_\theta(\theta')$ is the Dirac delta function equal to one if $\theta' = \theta$ and zero otherwise. Here,

$$\rho(\theta, \theta') = \min\left(\frac{p(\theta')}{p(\theta)}, 1\right)$$

and

$$r(\theta) = \int \rho(\theta, \theta')q(\theta'|\theta)d\theta'$$

The explanation of transition kernel is that $q(\theta'|\theta)$ is the probability that θ' is produced and $\rho(\theta'|\theta)$ is the chance that it will be accepted so the first term of the sum is the probability that $\Theta = \theta$ is produced and accepted. $r(\theta)$ is the sum of these probability over θ' and so is the chance that the produced θ' is accepted. It follows that the final terms $(1-r(\theta)$ is the probability that θ' is produced and not accepted so the chain remains in θ. Finally, multiplying $K(\theta, \theta')$ by $p(\theta')$ it is straightforward to verify that the Metropolis kernel satisfies the detailed balance conditon and thus it has $p(\cdot)$ as its stationary distribution.

Bibliography

[1] Aalami, H.A., Parsa Moghaddam, M. and Yousefi, G.R. (2010). Modeling and prioritizing demand response programs in power markets, *Electric Power Systems Research,* **80**, pp 426–435.

[2] Albadi, M.H. and El-Saadany, E.F. (2008). A summary of demand response in electricity markets. *Electric Power Systems Research*, **78,** pp. 1989–1996.

[3] J M. Alston, J. A. Chalfant and Piggott, N. E. Estimating and testing the compensated Double-Log Demand Model. *Applied Economics*, **34**, pp 1177-1186.

[4] Anderson, R. D. J., Beker, R. and Osborn, D. R. (2009). Heteroskedasticity and Autocorrelaton Robust Inference for a System of Regression Equations. *Newcastle Discussion Papers in Economics*, **11**, pp 1-53.

[5] Arellano M. S. (2003). Diagnosing and Mitigating Market Power Chile's Electricity Industry. Working Papaer *Cambridge University Press.*

[6] Barnett, V. *Comparative Statistical Inference.* Wiley Series in Probability and Mathematical Statistics, 1982.

[7] Bask, M. and Widerberg, A. (2009). Market structure and the stability and volatility of electricity prices. *Energy Economics* **31**, pp 278–288.

[8] Bianco, V., Manca, O. and Nardini, S. (2009). Electricity consumption forecasting in Italy using linear regression models. *Energy*, **34**, pp 1413–1421.

[9] Bhattacharyya, S. C. and Timilsina, G. R. (2009). Energy Demand Models for Policy Formulation: A Comparative Study of Energy Demand Models. *World Bank Policy Research Working Paper Series,* 4866.

[10] Bernardo, J, M. *Bayesian theory,* Wiley & Sons, 1994,Chichester.

[11] Bernstein, M.A and Griffin, J. (2006).Regional Differences in the Price-Elasticity of Demand for Energy. *National Renewable Energy Report.*

[12] Brendt, E. R. and D. O. Wood. (1975) Technology, prices and the derived demand for energy, *The Review of Economics and Statistics*, **57**(3), pp. 259-268.

[13] Bigerna, S. and Bollino, C. A. (2014). Electricity Demand in Wholesale Italian Market. *The Energy Journal*, **35,** pp. 25-46.

[14] Borenstein, S., Bushnell, J. and Knittel, C. R. (1999). Market Power in Electricity Markets: Beyond Concentration Measures. Working Paper, PWP 059r, *Power.*

[15] Borenstein, S., Bushnell, J. and Wolak, F. (2002). Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market, Working Paper, CESM WP-102, *Center for the study of Energy Market.*

[16] Bushnell, J., Mansur, E.T. and Saravia, C. (2004). Market Structure and Competition: A Cross-Market Analysis of U.S. Electricity Deregulation, Working Paper, CESM WP-126, *Center for the study of Energy Market.*

[17] Borenstein, S. and Bushnell, J.An Empirical Analysis of the Potential for Market Power in California's Electricity Industry. Working Paper, PWP 044r, *Power.*

[18] Browne, W., (2002). MCMC Algorithm for costrained variance matrices. London: Institute of Education, University of London.

[19] Ciccarelli, M. and Rebucci, A. (2003). Measuring Contagion with a Bayesian Time-Varying Coefficient Model. Working Paper 2003/171. *International Monetary Fund*

[20] Chib, S., Greenber, E. (1994). Hierarchical analysis of SUR models with extension to correlated serial errors and time-vrying parameter models. *Journal of Econometrics*, **68**, pp. 339-360.

[21] Chib, S., Greenber, E. (1995). Understanding the Metropolis Algorithm. *The American Statistician,* **49,** pp. 327-335.

[22] Chib, S., Greenberg E. and Winkelmann R. (1998). Posterior simulation and Bayes factors in panel count data models. *Journal of Econometrics,* **86,** pp 33-54.

[23] Chib, S. and Hamilton, B. H. (2000). Bayesian analysis of cross-section and clustered data treatment models. *Journal of Econometrics,* **97**, pp 25-50.

[24] Cowles K. M. and Carlin, B. P.(1996). Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. *Journal of the American Statistical Association*, **91** (434), pp 883-904

[25] Creel, M. and Farell, M. (1996). SUR Estimation of multiple time-series models with heteroskedasticity and serial correlation of unknown form. *Economics Letters*, **53**, pp. 239-245.

[26] de Cristofaro, R. (2008). A new formulation of the Principle of Indifference. *Synthese*, **163**, pp. 329–339.

[27] de Cristofaro, R. (2003). L'inferenza statistica induttiva. *Statistica e Società,* **2**, pp. 15-20.

[28] de Cristofaro, R. (2003). The Analitical Solution to the problem of statistical induction. *Statistica*, **LXIII**, 2.

[29] De Finetti, B. *Probabilità e induzione,* Bologna 1993, CLUEB.

[30] De Finetti, B., *Teoria della Probabilità*, Torino 1970, 2vol.

[31] Frisch, R. (1959). A Complete Scheme for Computing All Direct and Cross Demand Elasticities in a Model with Many Sectors. *Econometrica*, **27** (2), pp. 177-196.

[32] Glachant, J. M. and Saguan, M.(2007). An Institutional Frame to Compare Alternative Market Designs in EU Electricity Balancing. Working Paper 07-001, Center for Energy and Envirnmental Policy Research.

[33] Gelfand, A. E., and Smith, A. F. M. (1990). Sampling-Based Approaches to Calculating Marginal Densities, *Journal of the American Statistical Association*, **85**, pp 398-409.

[34] Gelfand, A. E., (2000). Gibbs Sampling. *Journal of the American Statistical Association*, **95**, pp. 1300-1304.

[35] Gelman, A., Carlin, R., Stern, H. S., Dunson, D. B., Rubin, D. B. and Vehtari A. *Bayesian Data Analysis*, Third Edition, Chapman & Hall.

[36] Gelman, A., Roberts, G. O., Gilk, W. R., (1994). Efficient Metropolis Jumping Rule. BayesianStatistics, **5**, pp.599-607.

[37] Gelman, A. and Rubin, D. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, **7**, pp. 457–511.

[38] Geweke, J. (1989). Bayesian Inference in Econometric Models using Monte Carlo Integration. *Econometrica*, **57**, pp. 1317–1340.

[39] Geweke, J. (1989). Exact Predictive Densities in Linear Models with ARCH Disturbances. *Journal of Econometrics*, **40**, pp. 63–86.

[40] Bernardo, J. M., Berger, J. O., Dawid, A. P., Geweke, J. and Smith, A. F. M. (1998). *Bayesian Statistics 6*-Simulation methods for model criticism and robustness analysis. Oxford University Press.

[41] Gilks, W. R., Richardson, S., Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice.* Chapman and Hall.

[42] Hasting, W.K. (1970). Monte Carlo Sampling Methods using Markov chain Monte Carlo and their applications. *Biometrika*, **57,** pp. 97-109.

[43] Hausmann, J. A., Kinnucan, M. and Mc Fadden, D. (1979).A two-level electricity demand model: Evaluation of the connecticut time-of-day pricing test. *Journal of Econometrics*, **10** (3), pp 263-289.

[44] Holt, M. T. and Goodwin, B. K. (2009): The Almost Ideal and Translog Demand Systems, in contributions to *Quantifying Consumer Preferences* (2009), Vol. 288, Emerald Group Publishing.

[45] Kahn, E. P., (1998). Numerical Techniques for Analyzing Market Power in Electricity. *The Electricity Journal*,

[46] Kirschen, D. S., Strbac, G., Cumperayot, P. and de Paiva Mendes, D.(2000). Factoring the Elasticity of Demand in ElectricityPrices. IEEE Transaction on Power System, **15** (2), pp 612-617.

[47] Koop, G. (2003). *Bayesian Econometrics*, Wiley.

[48] Koop, G., Poirier, D. J. and Tobas, J. L.(2007). *Bayesian Eocnometric Methods.* Cambridge University Press.

[49] Labandeira, X., Labeaga J. M., López-Otero, X. Estimation of Elasticity Price of Electricity with Incomplete Information. *Fedea,* WP 2009-18.

[50] Labandeira, X., Labeaga J. M., López-Otero, X. (2011) Energy Demand for Heating in Spain: An Empirical Analysis wiith Policy Purposes. *Economics for Energy*, WP 06/2011.

[51] Lancaster, T. (2004). *An introduction to moder Bayesian econometrics*, Blackwell Publishing.

[52] Lindley D. V. and Smith, A. F. (1973). Bayes estimates for the Linear Model. *Journal of the Royal Statistical Society*, **34**, pp. 1-41.

[53] Lehmann, E. L.(1995) Neyman's Statistical Philosophy, *Probability and Mathematical Statistics*, **15**, pp. 29-36.

[54] Mertens, J. F., Samuel-Cahn, E., Zamir, S. (1978). Necessary and sufficient conditions for recurrence and transience of Markov Chain, in terms of inequality. *Journal of applied Probability*, **15**, pp 848-851.

[55] Metropolis, N., Rosebluth, A.W., Rosebluth, M.N., Teller, A.H. and Teller, E. (1953). Equations of State Calculations by Fast Computing Machines. *Journal of Chemical Physics* **21**, pp. 1087-1092.

[56] Ng, S. (1995). Testing for Homogeneity in Demand System when the regressors are Nonstationary. *Jounal of Applied Econometrics,* **10** (2), pp147-163.

[57] Lazaridis, P. /(2004). Demand elasticities derived from consistent estimationof Heckman-type models, *Applied Economics Letters*, **11** (8), pp 523-527.

[58] Patrick, R. H. and Wolak, F. *Estimating the Customer-Level Demand for Electricity Under Real-Time Market Prices*, Preliminary Draft.

[59] Poirier, D. J. (1995). *Intermadiate Statistics and Econometrics. A comparative Approach.* MIT Press.

[60] Ravikumar, B., Surajit R. and Eugene Savin N.(2000). Robust Wald Tests in SUR System with Adding-up Restriction, *Econometrica*, **68**, pp 715-719.

[61] Robert, C. P. and Casella, G. (1999). *Monte Carlo Statistical Methods.* Springer-Verlag.

[62] Strevens, M. The Bayesian Approach to the Philosophy of Science. *Encyclopedia of Philosophy*, Macmillan.

[63] Smith, A. F. (1973). A General Bayesian Linear Model. *Journal of the Royal Statistical Society*, **35**, pp. 67-75.

[64] Tanner, M. A. and Wong, W. H.(1987), The calculation of posterior distribution by data augmentation. *Journal of the American Statistical Association*, **82**, pp 528-550.

[65] Taylor, L. D. (1975). The Demand for Electriciy: A Survey. *The Bell Journal of Eonomics,* **6 (**1), pp 74-110.

[66] Tierney, L. (1994). Markov Chain for Exploring Posterior Distributions. *Annals of Statistics*, **22**, pp 1701-1762.

[67] Tomohiro, A. (2011). Bayesian Variable Selection for the Seemingly Unrelated Regression Models with a large numberof predictors. *Japan Statics Society*, **41**, pp 187-203.

[68] Tomohiro, A. , Zellner, A. (2010). Hierarchical Bayesian Analysis of the Seemingly Unrelated Regression and Simulation Equatins Model using a Combination of Direct Montecarlo and Importance Sampling Techniques. *Bayesian Analysis*, **5**, pp. 65-96.

[69] Ventosa, M., Baıllo, A. , Ramos, A., Rivier, M.(2005). Electricity market modeling trends. *Energy Policy,* **33**, 897–913.

[70] Wang, J., S. Kennedy, and J. Kirtley. “A new wholesale bidding mechanism for enhanced demand response in smart grids. Institute of Electrical and Electronics Engineers Press, 2010.

[71] Western, B. and Jackman S. (1994). Bayesian Inference for comparative research. *The american political Science Rewiev*, **88** (2), pp 412-423.

[72] Willems, B., Rumiantseva I. and Weigt, H. Cournot versus Supply Functions: What does the data tell us?. *Energy Economics*, **31,** pp 38-47.

[73] Wiper, M.P., Girón, F.J. and Pewsey, (2005). A. Bayesian inference for the half-Normal and half distribution.Working Paper, *Statistics and Econometrics Series* 09.

[74] Wolak F.A and R.H. Patrick (1997). "Estimating the Consumer-Level Demand for Electricity Under Real-Time Market Prices," mimeo, Stanford University

[75] Zellner, A. and Tomohiro, A. (2010). A direct Monte Carlo approach for Bayesian analysis of the seemingly unrelated regression model. *Jounal of Econometrics,* **159**, pp. 33-45

[76] Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressin and tests of aggregation bias, *Journal of the American Statistical Association,* **57**, pp. 500-509.

[77] Zou, X. (2009).Double-sided auction mechanism design in electricity based on maximizing socialwelfare. *Energy Policy,* **37**, pp 4231–4239.