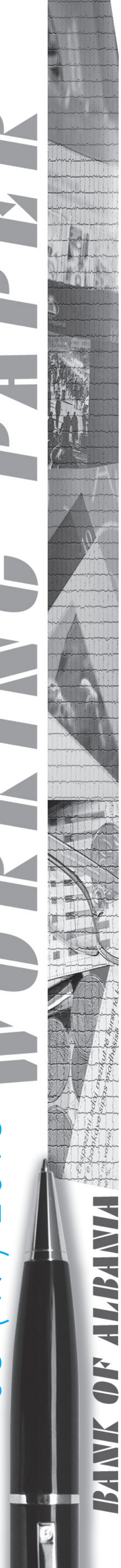
WORKING PAPER



08 (47) 2013

DENSITY ESTIMATION FOR ECONOMIC VARIABLES – A GENUINE APPLICATION

Altin Tanku
Kliti Ceca*

BANK OF ALBANIA

** Altin Tanku, Kliti Ceca Research Department, Bank of Albania*
e-mail: atanku@bankofalbania.org, kceca@bankofalbania.org

The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Bank of Albania.

CONTENTS

"If I understand something, I can simulate it". J. Gentle

SUMMARY

The paper studies and proposes the density estimation methodology as an alternative of modelling the economy as a set of economic variables. In the study, we pose the density estimates for economic variables in one and multidimensional case. This is done using several kernels, which are selected as probability density functions. In the one dimensional case, Gaussian, triangle, rectangle and Epanechnikov kernels are used. In the multidimensional case, Gaussian and Epanechnikov kernel are used. For the first time, we give the form of the densities (density estimates) of economic variables in their general form, based on density estimation methods. The methodology is illustrated on a set of economic data in the case of the Albanian economy. Graphical representations of the density estimates are given for selected scalar variables and two dimensional vectors. The paper also includes the algorithms for calculating the density estimates given in the paper. General discussions on the methodology together with clear references on the literature are given.

The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Bank of Albania.

JEL classification: C18, C51, C81, C82.

Keywords: empirical investigation, economic modelling, density estimation, joint density of economic variables, bandwidth, kernel function, distribution function estimation, mean square error, mean integrated square error.

1. INTRODUCTION

Understanding economic relationships and accurately modelling them play an important role from the analysis, policy making and forecasting point of view. These relationships are summarised in economic models both theoretic and empiric, single or multidimensional ones. The tools that the economic science uses are the hypothesis testing and the statistical analysis of the history of the random economic variables (time series analysis). Both approaches are building upon the mathematic fundamentals, like the theory of random processes, solution of the systems of simultaneous difference equations, and a large set of reasonable and "compulsory" assumptions regarding behaviour of the economy and the statistical properties of the data. This direction has defined the development of economics during the last 40 years. Yet despite such advances, currently, economics is suffering a setback for its unrealistic representation of the real economic phenomenon. An important element of the above mentioned critique focuses on the inability of the models to embody the stochastic nature of the real life.

This paper presents and discusses a new alternative econometric (statistic) method, which relies on the estimation of joint density probability function of the random economic variables. The general objective of this paper it to pose density estimation method as a toolkit to help understand and possibly replicate the random process that generates the vectors of economic data of interest at each moment in time or for a given time interval. That means to uniquely portray the density estimation of the set of variables that describe the economy and potentially identify the model that yields the consecutive value or the set of future values when the past and current values of the related variables and the own past values are known.

The uses of the density estimation methods have aroused increasing interest in many fields where further details on the distribution of the random variables have to be studied. For many years, it has become a deep research in the evaluation of estimates to better align with the functions coming from a real phenomenon, without making undue assumptions on specific elements in advance, which joint together can change the behaviour of the model as a whole.

In this project, we aimed among others, to bring this method to the attention of methodical statistical specialists in the field of applied economics, based on the rigorous mathematical definitions. In that way, we tend to create a clear perception, not only as a set of tools for theoretical applications, but also formally applying it in a multiple dimensional space, based on the perception of the economy as a space generated from a number of random variables. We discuss the potential of this new method and apply it in the case of the Albanian economy to provide a discussion on the monetary policy.

Current models that are used for forecasts and analysis have come under fierce critique from the public, politicians, and business. As a matter of fact, what is called "the failure of the current models" was so to say predicted by several economists, expressed in its best form by Lucas (with Lucas critique). Another prominent economist Solow (1985) states that "…economic theory learns nothing from economic history, and economic history is as much corrupted as enriched by economic theory".[1] Solo did return to the same theme in 2010 with a similar critique of the macro econometric models pointing to the failure of current models to predict the last crisis and yield reliable policy solutions for the financial and economic mess that followed. The main and most important part of this critique emphasizes three important facts. First, models are way too rigid and too stylized to describe the stochastic nature and the number of deterministic trends of real life. Second, life is too dynamic to yield long enough stationary series of random variables or support hypothesis testing approach. And third, a little cleverness and data mining can yield the desired yet good enough to be accepted statistical results.

As this is not enough, there exist however differences in the way theoretical economists and econometricians see the world, as it

[1] Solow R. (1985), Economic History and Economics, "The American Economic Review", Vol. 75, No. 2, May 1985.

is elaborately shown by Juselius and Franchi (2007). The former would start from a specific theory to build the model based on the particular believed theoretic knowledge upon which the entire model is constrained. In other words, the model is built to fit a predesigned theoretic structure. The latter observe that this structure is way too rigid to describe the real life data. Their approach let the data build the model, albeit adding particular structure of their own in the identification process.[2]

Both methods constrain reality with their structure and assumption to make it adoptable to the available toolkits. In this process, both methods transform reality to take it to the laboratory. The alternative would be to study the reality, in its natural form outside the framework of parameterised standard models.

Meanwhile, it is known that the probability density is the framework that fully explains probability behaviour of a variable or phenomenon. Therefore, the attention on the applied economics and finance should be directed not only to the study of the densities of the random variables, but also to the investigation of the probabilities of the random events. The mutual implications that arise from the interaction among these random variables define the behaviour of random phenomena that we study and can be useful in their modelling. Economy is a good example of this multidimensional space.

Currently, there exist advanced techniques that show how to construct a density estimate, with a good approximation to the real one. These methods are used in different fields of application where density estimation provides the statistical framework to interpret the probability density of random events and make accurate assumptions from these interpretations regarding future developments (i.e. the behaviour of this probabilistic phenomenon in the future). This framework could be also useful in the discussion of different episodes and shock scenarios exploring what would happen to the probability density of a scalar or vectorial random variable and what the consequences of potential changes in the probability densities are. This new form of the conditional densities of random

[2] These assumptions will be discussed in details in chapter 2

phenomena, represented by some fixed conditional variables, represents an alternative method of empirical investigation that is worth exploring to measure and forecast the impact of certain fixed conditions in the performance of the densities of other variables in a random phenomenon.

Also, the progress of the probability density over time could be another alternative method to see the development of the real phenomena across time.

The rest of this working paper is organized as follows:

The second chapter discusses the definition and representation of economy as a multidimensional phenomenon and, in this way, opens the road to the discussion of the multidimensional density estimation as an appropriate framework for its empirical investigation.

The third chapter describes the methodology of density estimates. The main applied methods are given together with the description of the main literature used.

The fourth chapter provides the concrete form of the density estimates of one and d dimensional economic variables. In addition, it offers a broad discussion on the parameters, measures of discrepancy and efficiency issues related with the method.

The fifth chapter discusses the intuition, results and data.

The sixth chapter provides concrete applications of the density estimation with the economic data from the Albanian economy. Density estimates are given on one and two dimensional variables. Also, graphs and comments on graphical representations are given as concrete applications of the methods.

The seventh chapter concludes the study.

The eighth chapter provides general discussions on the methodology and some directions for our future research.

The last part includes the literature used for the study.

In the appendices, there are graphics representations and relating comments and corresponding algorithms to construct the density estimates we applied.

The main symbols used in this paper are as follows

Small letters x_i are used for the numerical values of the sample.
Capital letters X_i are used for the sample, as random variables.
The underscore sign on each letter, i.e. $\underline{t}$ or $\underline{x}$, means that this is a vector.
h is the smoothing parameter, or the bandwidth – relating with the histogram concept.
$f(.)$ is the density of the phenomena (of the random variable) that we will estimate.
$\hat{f}(.)$ is the density estimate.
π is the mathematical constant, calculated as $\pi \approx 3.14159$.
PDF stands *for probability density functions*.
JPDF stands *for joint probability distributions*.
PDA stands *for probability distribution analysis*.
MSE means squared *error at a point*.
MISE means integrated *squared error*.
DGP means data *generating process*.
p.i.t. means probability integral transforms.
ALL means Albanian Lek (currency).

2. PORTRAYING THE ECONOMY AS A RANDOM MULTIDIMENSIONAL EVENT

Social events like economics represent multidimensional simultaneous random events that result from the interaction of many variables, each embodied with a certain probability density function. We study this social phenomenon along its many dimensions, which we perceive like economic variables. This view of the economy as a social phenomenon is discussed by Hendry and Richard (1983). In their definition, economy is the interaction among economic variables that we observe as a set of economic data, hence the data generating process (DGP). The analysis of the economy is related to this definition of the random economic event as a data generation process (DGP), meaning the natural process that generates the particular behaviour of an experiment through time and space. Economy as random event and the DGP are later reviewed by Ericsson, Hendry and Mizon (1998), who formalise the definition of the economy or economic behaviour in the form of a multiple dimensional random process. The authors describe the DGP in the form of a probability space $[\Omega, \mathcal{F}, \mathrm{P}(.)]$, where Ω represents a sample space of the vectors of d variables, denoted vector x_t, that describes the process (in the study of economics that would describe x_t as a vector whose elements represent the measurements at time period $t \in T$ where $T = \{1, 2, ..., t_{last}\}$ for variables like inflation, GDP, unemployment, money demand, interest rates, etc.), $\mathcal{F}$ is the event space (e.g. the event that all economic variables mentioned above take a particular value simultaneously), and $\mathrm{P}(\cdot)$ is the probability measure for the events in $\mathcal{F}$ (e.g. the probability that a particular combination of variables materialises at time t).

The DGP is formally expressed as the joint density function of the initial conditions or vector X_0, a vector of parameters ζ, and all the subsequent vectors x_t for all $t \in \{1, 2, ..., t_{last}-1\}$ as follows:

$$D_X(X_T \mid X_0,\zeta) = \prod_{(t=1)}^{T} D_x\,(x_t \mid X_{t-1},\zeta_t), \qquad 2.1.$$

where X_{t-1} represents the stochastic process $(X_0, x_1,\dots,x_{t-1})$ and ζ_t represents a subset of the parameters set $\zeta = (\zeta_t, ..., \zeta_T)$, whose value is assumed to be known to the researcher or the policy making authorities (or it is the object of inquiry of the research). The set of parameters is the mechanism that relates all random variables together and for that reason estimation of $\zeta = (\zeta_t, ..., \zeta_T)$ is the focus of the empiric research. The estimation of the parameters is necessary for forecasts and policy analysis and renders this probabilistic representation of the economy useful to authorities. This is the time series method at its best and more complete analysis with the Sims (1980) VAR approach and its later developments including structural VARs and VECM. Referring to Hendry and Richard (1983), Juselius (2006) describes the VAR analysis as a very convenient way to explain economic behaviour of the rational agents with partial information. Therefore it becomes a useful tool in the hands of the policy makers. The time series representation of the DGP process has the formulation conditional process of marginal events given the knowledge of initial conditions X_0 and the parameter matrix that describes the stochastic process in the vector of the random variables of the random events. In principal this is done by consequently decomposing the joint probability into a conditional probability and marginal probability for each $t \in T$ repeating the process until we reach t_0 as in eq. 2.1. above. Therefore the VAR process describes the conditional process $\{x_t | X^0_{(t-1)}\} \sim N(\mu, \Omega)$.

The method has been very well accepted because it does a good job in describing the economy.[3] *The framework is formally based on the time series and linear spaces theory and the solution comes in the form of a set of parameters that depict the relationships among the present and past values of the variables of interest. Such coefficients would allow the policymaker to predict the jump in the variable x_t along with the accompanied i.i.d. vector of errors, which accounts for the errors of estimation and "fits" the model to the stochastic real world. However, the estimation of the parameters requires the following assumptions:*

- That all x_t have the same distribution which is approximately normal with known first and second moment.
- Constant mean and variance

3 See Stock and Watson (2001) for a detailed analysis of the VAR and its performance.

Both assumptions impose very strong restrictions on the data and the random process X_t. Moreover, several other long and short term restrictions of assumed theoretic relationships and identifications are imposed and required to generate a uniquely identified solution (referring to the decomposition or ordering of endogenous shocks), enforcing additional structure in the assumed data generating process.[4] The fundamental question is whether these assumptions and the imposed structure are taking the econometric model away from the DGP? One can't help asking whether this approach takes the assumed model away from the real DGP. Hendry (2011) summarises this discussion in six problems related to the knowledge about DGP and the estimated parameters:

- Specification of the set of relevant variables,
- Measurement of $x - s$,
- Formulation of the $D_{X_T}(\cdot)$
- Modeling of the relationships,
- Estimation of parameters,
- Properties of $D_{X_T}(\cdot)$, which determine the intrinsic uncertainty.

Among these problems, we would like to address the first, mix the third with the fourth and eliminate the fifth with a new non parametric pack of methods. In general we propose to study the DGP without related assumptions and the imposed theoretic structure. We instead focus on the study of the joint PDF and CDF of the DGP for all the variables.

This alternative situation observes and studies the DGP as a random event in the form of a joint probability density function rather than in the form of a system of difference equations (or rather than seeking the solution of a system of difference equations). The joint PDF that we propose preserves the simultaneity of the events despite dropping the time index, in similar fashion as time series does. Then, conditioning the dataset to reflect researchers' expectations regarding one or more of the random variables can generate a conditional forecast or analysis for the random event that reflects those expected developments.

---

[4] See Juselius and Franchi (2007) for a detailed discussion of the implication of the imposed theoretic restrictions and their implications.

In principal, this nonparametric analysis would eliminate the need for the marginalization and the conditioning assumptions of constant mean and variance required by the current empiric estimation methods. It is an alternative way of estimation of the economic phenomenon in its original DGP form. Considering the DGP that results from the random event we call economy, one can reasonably define the economy as a multidimensional space generated by d different random variables (where d corresponds to the number of chosen variables), in which each m (where $m < d$) dimensional space represents a subspace of the entire space of our random event. If we were to scale all these possible ξ_m spaces in their relationship, we will get the follows: $\xi_1 \subseteq \xi_2 \subseteq \cdots \subseteq \xi_m \subseteq \cdots \subseteq \xi_d$. It would mean that starting from ξ_m, each following event would represent a new possible subspace of the original d dimensional event. It would simultaneously mean that in substance we can represent the economy as an expanding sequence of spaces.

Assuming that all added variables are linearly dependent on the first chosen variables means that the new space will carry itself and so preserve the same DGP along fewer dimensions. If one or more of the new added variables are independent from the rest (or exogenous), then the distribution of the new space will also 'carry' this attribute to its distribution. These are all the conditions that we have to judge through joint densities of the variables creating those spaces. The proposed investigation method is developed upon the framework of density estimation of the DGP. Therefore, the following chapter provides an extended discussion of the density estimation of the single and multidimensional random events.

3. A DESCRIPTION OF THE METHODS OF DENSITY ESTIMATION

Probability density function is called a nonnegative function $f(.)$, that satisfies the condition:

$$\forall a, b \in R: P(a < X < b) = \int_a^b f(x)dx. \quad 3.1.$$

Knowing or estimating the probability distribution is an essential purpose in probability, based on the fact that: all the information of a random variable is contained by the values of this function and its probability distribution. The first part of this fundamental concept, consists in finding the values of a function (random variable in our case). On the other hand, it takes much effort to estimate the probability distribution.

Generally, the evaluation of the distributions (and the densities in particular) develops in two directions. The first study case, probably more explored in the literature, is the case of parametric distributions. It means that the sample data is assumed that comes from a random variable with known parameterized probability distribution, and what should be achieved is the estimation of the parameters.

The second study case is the estimation of non-parametric distributions – and it includes the situations where no preliminary information on distributions is available. Estimation of densities, in the frame used in this article, will focus on this second study case, and will also assume that the sample is made from random variable with a probability distribution function f.

In substance, density estimation is a natural generalization of the histogram of the sample from a random variable or vector – setting aside the histogram as a naive estimator. While the "kernel density estimation", as a part of density estimation, can be seen as analog transformation of Fourier transformation of a periodic function, or Taylor expansions of a continuous function.[5]

---

[5] The density estimation theory is described in several monographies and other literature mentioned in the article.

In that concept, let us mention a simple generalization of the histogram, in its classic form, as treated in the corresponding literature. Thus, assume that we have a sample $x_1, x_2, ..., x_n$, from the random variable X, let x_0 be a starting point and h the parameter of the bin width. The h parameter acts as a smoothing parameter, in the meaning that increasing the value of h suppresses the statistical noise and gradually wipes the statistical significance of the curve while, while decreasing the value of h increases the statistical noise and gradually makes the statistical unreadable.[6]

In the right-closed intervals $[x_0 + mh, x_0 + (m+1)h[$[7] build verticals equal with absolute or relative quantity of the number of x_i that fall in the same interval. Functional form of the histogram can be given:

$$\hat{f}(t) = \frac{1}{nh} \cdot (\text{nr.of } x_i \text{ in the same interval with } t) \qquad 3.2.$$

If the random variable X has the density $f(.)$, it is true that:

$$f(t) = \lim_{h \to 0} \frac{1}{2h} P(t\text{-}h < X < t+h) \qquad 3.3.$$

Then, as a natural estimator of the density $f(.)$, we can have:

$$\hat{f}(t) = \frac{1}{2nh} \cdot (\text{nr. of } x_i \in]t\text{-}h, t+h[\,) \qquad 3.4.$$

In a more formal way, the above estimator can be written:

$$\hat{f}(t) = \frac{1}{nh} \cdot \sum_{i=1}^{n} w\left(\frac{t\text{-}x_i}{h}\right), \text{ where } w(x) = \begin{cases} \frac{1}{2}, \text{ for } |x|<1 \\ 0, \ x \text{ other} \end{cases} \qquad 3.5.$$

The above estimator in the literature is known as "naive estimator" and has two main features: first, it is a direct generalization of the histogram, and second, its canonical form allows the further generalization into the "kernel density estimation".[8]

The estimation using a kernel is a natural generalization of the expression given in 3.5. Assume that $K(x)$ is a density probability

[6] Discussions of the smoothing parameter h are given in the eighth chapter.

[7] The choice of x_0 may have no restrictions in the classical way of building the histogram.

[8] Other methods are in place and they are not the focus of this paper.

function. In the terminology of "density estimation" such a function used hereinafter will be called kernel function (or simply Kernel). If the function $w(x)$ is substituted with kernel $K(x)$ in the expression 3.5., the general form of density kernel estimation takes the form:

$$\hat{f}(t) = \frac{1}{nh} \cdot \sum_{i=1}^{n} K\left(\frac{t - x_i}{h}\right), \text{ where } h \text{ is smoothing parameter.} \qquad 3.6.$$

As it is described in detail in the following chapter, in the case of random vectors the main form of the kernel density estimation is the same. The methods of density estimation are described in some main monographies and textbooks.

Description of the main literature:

The literature on density estimation is extensive. First, basic ideas come with Fix and Hodges (1951) and Akaike (1954). Influential papers of Rosenblatt (1956) and Parzen (1962) initiated the mathematical theory and stimulated further interest in the subject. For an overview of the literature on kernel density estimation, we refer to the books of Devroye and Gyorfi (1985), Silverman (1986), Devroye (1987), Wand and Jones (1995), and Devroye and Lugosi (2000).

The derivation of the Epanechnikov kernel from optimization arguments is due to Bartlett (1963) and Epanechnikov (1969). Hodges and Lehmann (1956) did it even earlier, although not in the context of density estimation. A short proof implying that the Epanechnikov Kernel is the optimum estimator in the meaning of the mean square error is given, e.g., by Devroye and Gyorfi (1985).

Wand and Jones (1995), "Kernel Smoothing", Chapman Hall elaborates the estimations with several dimensions. To reduce the level of the error on approximations, the authors reduce the dimensions of the space they work. They also give, estimations on statistical error of approximations.

Gentle (2009), "Computational Statistics", Springer, explores the estimation process of the densities on a kernel based, together with other techniques of smoothing on a topological space.

Tsybakov (2009), "Introduction to Nonparametric Estimation", Springer, is playing in the nonparametric sphere, and explores the estimates in a one dimensional case. There are treated also, methods of construction of the estimators, their statistical properties (like convergence and rates of convergence), optimality of the estimators, etc.

Silverman (1986) "Density Estimation for Statistics and Data Analysis" Chapman Hall, describes also the details of the estimation process of a density distribution. It is probably the most comprehensive text, which explores the development of the concept and the main documents on the literature of that time.

4. CONCRETE FORM OF DENSITY ESTIMATES, ONE AND MULTIDIMENSIONAL CASES

In this chapter, we portray the methodology for the estimated densities for one and more dimensional cases, for some different kernels selected in advance. Enriching the list will be the focus of future research. In fact, the most important applications of the density estimates are in the multivariate space. Since the multivariate applications stand as the generalization of univariate applications, we are starting with one dimensional density estimators.

Densities that are modelled after some of the necessary transformations are given in their canonical form. Calculating and giving density estimations for some economic series represents the first overall outcome of this article. Those estimations and the actual analytical expression of the densities can be useful in the situations where the variables are used for policy analysis. In order to enable a wider range of involvement of the same densities, we are in the process of enhancing the list of kernels. Different considerations on which kernel can be used are described in the literature (we referred to some of them in the last chapter of this paper). Annex 3 provides the compiled algorithms for calculations on estimators of densities.

4.1 ECONOMIC VARIABLES, ONE DIMENSIONAL CASE

For the estimation of densities in one dimensional variable we consider four kernels, as described in the following table. Efficiencies calculations are given by Silverman (1986).

Table 4.1.1.: Kernel densities and their efficiencies, one dimensional case

	Kernel	Density	Efficiencies
1.	Gaussian	$K(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}, t \in R.$	$\sqrt{\frac{36\pi}{125}} \approx 0.9512$
2.	Triangle	$K(t) = \begin{cases} 1-\lvert t\rvert, & \text{for } \lvert t\rvert<1 \\ 0, & t \text{ others} \end{cases}$	$\sqrt{\frac{243}{250}} \approx 0.9859$
3.	Rectangle	$K(t) = \begin{cases} \frac{1}{2}, & \text{for } \lvert t\rvert<1 \\ 0, & t \text{ others} \end{cases}$	$\sqrt{\frac{108}{125}} \approx 0.9295$
4.	Epanechnikov	$K(t) = \begin{cases} \frac{3}{4\sqrt{5}}\left(1-\frac{1}{5}t^2\right), & \text{for } \lvert t\rvert<\sqrt{5} \\ 0, & t \text{ others} \end{cases}$	1

Presentation of the density estimation for each of the kernels is made using the formula 3.6. mentioned above.

In practice, to estimate the density of the distribution of an economic random variable we make the following assumptions.

Let's be X a random variable with the density $f(.)$, which will be estimated. Assume also that we have n realizations of this variable, respectively $x_1, x_2, ..., x_n$.

To estimate the density $f(.)$, using the kernels listed in the table above, we made substitutions and corresponding transformations and final functional forms are given below:

Density estimation with Gaussian kernel:

$$\hat{f}(t) = \frac{1}{nh} \cdot \frac{1}{\sqrt{2\pi}} \cdot \sum_{i=1}^{n} \exp\left[\frac{-(t-x_i)^2}{2h^2}\right], \qquad 4.1.1.$$

where $t \in R$ and h smoothing parameter.

Density estimation with Triangle kernel:

$$\hat{f}(t) = \frac{1}{nh^2} \sum_{i=1}^{n} a_i, \qquad 4.1.2.$$

where $a_i = \begin{cases} (h-\lvert t - x_i\rvert), & \text{for } (h-\lvert t - x_i\rvert) > 0 \\ 0, & \text{for } (h-\lvert t - x_i\rvert) \leq 0' \end{cases} \quad t \in R.$

Density estimation with Rectangular kernel:

$$\hat{f}(t) = \frac{1}{2nh}\sum_{i=1}^{n} a_i, \qquad 4.1.3.$$

where $a_i = \begin{cases} 1, & \text{for } (h-|t - x_i|) > 0 \\ 0, & \text{for } (h-|t - x_i|) \leq 0' \end{cases} \qquad t \in R.$

Density estimation with Epanechnikov kernel:

$$\hat{f}(t) = \frac{3}{4\sqrt{5}} \cdot \frac{1}{nh}\sum_{i=1}^{n}\left[1 - \frac{1}{5}\left(\frac{t - x_i}{h}\right)^2\right], \qquad 4.1.4.$$

where $\frac{|t - x_i|}{h} < \sqrt{5}, t \in R.$

Practically, the above functional forms, given as functions of the variable t, serve as density estimates of density $f(.)$ of random variable X. Discussions regarding the selection of kernel and parameter h, are given in the eighth chapter of this paper.

Those functional forms can be easily used in the case of the need for the explicit form of the density of an economic variable.

The probability density function (PDF) theoretically represents the fullest stochastic information of an economic variable (or any random variable for that matter). The PDF based analysis can describe the variable and its random nature accurately, yet this information is not very informative in the day-to-day policy decision-making despite revealing the true probability nature during the entire known history of the variable. In the case of the socio economic developments, the researcher is interested in knowing the probability that a particular phenomenon materializes in the next s periods. Let us take inflation for example and assume that we know its mean and variance accurately; however, for the inflation targeting central bank, this information is useless unless it forecasts with accuracy the probability of inflation hitting the target value 12 months from now or 24 months from now. In this respect, the PDA would tell the bank the entire distribution of inflation without providing information where in this distribution inflation will be in the next periods. The central bank's credibility and success depends on the predictability of this figure at these particular moments in time, not on the mean of the inflation series, which stays the same regardless of the realization of the last inflation (experiment). In the same fashion, markets and policy makers are interested in knowing with predicted accuracy the future value of their final objectives and intermediate targets, and predict with accepted accuracy the effects of prescribed policies. In this respect, the probability density functions of each individual random variable would not be very helpful in terms of the dynamic developments of the multidimensional random events in the real word. The reason why the probability distribution analysis (PDA) of single variables does not satisfy these needs (therefore, it is not used despite its informative nature) derives from two important facts:

First, individual probability density functions (PDF) strips each realization from its time indexation. Therefore, in the context of PDA, the probability of the variable hitting a particular value is invariant to the time of investigation.

Second, losing its time indexation the PDA loses the existing interrelations and relationships among individual variables of a multiple dimensional random event. Therefore, if one were to adopt the vision of the economy as a d dimensional random phenomena presented above focusing only on the analysis of the PDF of the economic variable x, is not very informative. This is due to the fact that probability distribution by itself neither preserves nor conveys information of relationship (dependence causality, co-movement) between the expected value of variable x with the other $d-1$ variables in the economy or its own previous values due to the absence of time or any other form of indexation.

From the point of view of the economists and policymakers, this 1 dimensional nonparametric method could appear complementary of the existing methods and not sufficient. It can help in the observation of the stochastic feature of the density of the variable.

Literature acknowledges this problem and proposes the joint or multivariate density forecasts as a potential method to solve inter and intra-temporal dynamics in the contest of the PDF analysis. The method is based on the factorization of the random event of interest into conditional and marginal probabilities, and the calculation of the probability integral transforms (p.i.t.) values for the conditional and marginal process separately.[9] This method is discussed by Clements and Smith (2002)[10], who find that the p.i.t. method lacks the power whenever correlation structure is miss-specified. The failure to preserve the temporal pairing and potential autocorrelation versus independence of the conditional and marginal process are potential sources of misspecification.

In this respect, our proposal broadens the above analysis with the introduction of the joint density probability function of the random event by estimating the JPDF of the d dimensional random event. The method has been successfully tested and applied in other fields of applied sciences. We here intend to introduce it in the study of economics with the general objective to understand and possibly replicate the random process that generates the vectors of economic data of interest at each moment in time or for a given time interval. That means to uniquely identify the "function" that yields the consecutive value or the set of future values when the past and current values of the related variables and the own past values are known. Despite "removing" the time indexation the method (by its design shown below) preserves the simultaneity of developments among the random variables of interest. The following section proceeds with the presentation of the d dimensional PDF of a random event.

4.2. ECONOMIC VARIABLES D – DIMENSIONAL CASE

Like in the one dimensional case, for d dimensional vectors of data the density estimates are given. For simplicity, we initially selected two densities in the role of the kernel, which are more useful in the relevant literature. Those densities are given in the table below.

Assume that $\underline{t} = (t_1, t_2, \ldots, t_d)'$ is the variable (argument) from R^d.

---

[9] See Diebold et al. (1998) for a description.

[10] Clements, M.P., and Smith, J., 2002, Evaluating Multivariate Forecast Densities: a Comparison of Two Approaches, International journal of Forecasting, 18, 2002, pp. 397-407.

Table 4.2.1.: Kernel densities, d dimensional case

	Kernel	Density
1.	Gaussian	$K(\underline{t}) = \frac{1}{(2\pi)^{d/2}} \cdot \exp(-\frac{1}{2}\underline{t}'\underline{t})$, where $\underline{t} \in R^d$, $K(\underline{t}) \in R$
2.	Epanechnikov	$K_e(\underline{t}) = \begin{cases} \frac{1}{2C_d}(d+2)(1-\underline{t}'\underline{t}), & \text{if } \underline{t}'\underline{t} < 1 \\ 0, & \text{for other } \underline{t} \end{cases}$, where C_d is the volume of unit d dimensional sphere $\underline{t} \in R^d$, $K_e(\underline{t}) \in R$

To provide kernel density estimation of the density $f(.)$, first define variables of the sample and the variables of the density that we are trying to estimate.

Let us assume that there have been n realizations of the d dimensional variable $\underline{X}$, respectively $\underline{x_1}$, $\underline{x_2}$,..., $\underline{x_n}$. As vectors, the sample can be written:

$$\begin{cases} \underline{x_1} = (x_1^1, x_1^2, \ldots, x_1^d) \\ \underline{x_2} = (x_2^1, x_2^2, \ldots, x_2^d)' \\ \underline{x_n} = (x_n^1, x_n^2, \ldots, x_n^d) \end{cases} \qquad 4.2.1.$$

Let us assume that $t = (t_1, t_2, \ldots, t_d)'$ element of R^d, is the variable (argument) of the density that will be estimated.

As an analogy with the one dimensional case, the expression of the kernel density estimation $K(\underline{t})$, for each of the kernels, as it is mentioned in Silverman (1986) and Wand and Jones (1995), is given as follows:

$$\hat{f}(\underline{t}) = \frac{1}{nh} \cdot \sum_{i=1}^{n} K\left(\frac{1}{h}(\underline{t} - \underline{x_i})\right), \qquad 4.2.2.$$

where: $t \in R^d$; $\underline{x_1}$, $\underline{x_2}$,...,$\underline{x_n}$ are the elements of sample from R^d; and $\hat{f}(\underline{t}) \in R$.

To estimate the density $f(.)$, using the kernels mentioned in the table above, we made substitutions and corresponding transformations, and the final functional forms are given below:

Density estimation, d – dimensional case, Gaussian kernel:

$$\hat{f}(t)=\frac{1}{nh^d\ (2\pi)^{d/2}}\cdot\sum_{i=1}^{n}\exp\{-\tfrac{1}{2h^2}[(t_1-x_i^1)^2+(t_2-x_i^2)^2+\cdots+(t_d-x_i^d)^2]\}, \quad 4.2.3.$$

where: $\underline{t}\in R$, h is smoothing parameter.

Density estimation, d – dimensional case, Epanechnikov kernel:

$$\hat{f}(t)=\frac{1}{nh^d}\frac{d+2}{2\cdot C_d}\cdot\sum_{i=1}^{n}a_i, \quad 4.2.4.$$

where: $a_i=\begin{cases}\{1-\frac{1}{h^2}[(t_1-x_i^1)^2+(t_2-x_i^2)^2+\cdots+(t_d-x_i^d)^2]\}, & \text{if } \{\ldots\}>0\\ 0 & \text{, if } \{\ldots\}>0'\end{cases}$

and $C_d=\begin{cases}\frac{\pi^{d/2}}{(\frac{d}{2})!}, & \text{d-even}\\ \frac{2^{\frac{d+1}{2}}\cdot\pi^{\frac{d-1}{2}}}{d!!}, & \text{d-odd}\end{cases}$

Like in the one dimensional case, the functional forms above, which are given as functions of the variable $\underline{t}$, serve as kernel density estimates for the density $f(.)$ of the random vector X. Discussions relating with the choice of the kernel and the smoothing parameter h are given the eighth chapter of this paper.

These functional forms can be easily used in the case of need of the concrete density of a random vector.

Also, the list of potentially useful kernels could be extended based on the concrete needs of the research. The fix point is always that a kernel is density probability function. Other considerations on that choice relate to the efficiency of the kernel and the advantages that a kernel may represent towards the others. We discuss these issues in the following section.

4.3. MEASURES OF DISCREPANCY, THE BEST CHOICE OF THE SMOOTHING PARAMETER H AND THE EFFICIENCY OF THE KERNELS – GENERAL CONSIDERATIONS

In the case of univariate data, various measures have been studied relating to the discrepancy of the estimate $\hat{f}(.)$ from the density we have to estimate $f(.)$. The choice for one or another kernel sometimes is strongly related to the discrepancy measure that those functions ($\hat{f}(.)$ and $f(.)$) have between them. In further more complicated techniques, just because of the "measures of discrepancy" issues, the choice of the kernel is even "indifferent" for the main concept of the given definition of the kernel – being a probability density, and getting other types of functions. Such conditions are given in further mathematical studies of the natures of the estimates.

As a general estimator of the measure of the discrepancy, the literature considers a natural measure: the mean square error. Furthermore on that estimator, it has two appearances:

- as a local measure of the discrepancy in a single point t, shortly mentioned as MSE, defined by: $MSE_t\,\hat{f} = E[\hat{f}(t) - f(t)]^2$,

- as a global measure of discrepancy in the zone where the distributions are defined, shortly mentioned as MISE, defined by: $MISE\,(\hat{f}) = E\int[\hat{f}(t) - f(t)]^2\,dt$.

As the mean square error is common in the classical parametric statistics to measure the closeness of an estimate of a parameter with that parameter, this concept is quite known. This measure is also used, in the base literature which explores the density estimation.

Detailed considerations of the measures of the discrepancy are given in Wand and Jones (1995) for one and multi – dimensional cases. Also, Gentle (2009) treats them to evaluate this measure of discrepancy for function estimates in general. Tsybakov (2009) gives technical details and corresponding calculations on both types of measures. The transformations are given in details and in a formal way. Silverman (1986) also describes the estimates of both types of errors in one and multidimensional cases.

Other authors try to give more specific forms to avoid the fact that after the calculations of both errors, it is still a function of smoothing parameter h, or how to calculate it exactly, etc., such as the efforts given by Marron and Wand (1992).

Regarding the smoothing parameter (or window width), there is a certain level of research trying to find optimized values of it, or, as it is sometimes called in the literature, "the ideal window". The optimization procedure is normally based on minimizing the MISE mentioned above. The transformations of this optimization are given by "Parzen Lemma" (1962), and are also mentioned by Silverman (1986), where the optimum h is given by:

$$h_{opt} = k_2^{-2/5} \{\int_R K(t)dt\}^{1/5} \{\int_R f^{(2)}(t)^2 dt\}^{-1/5} n^{-1/5}, \qquad 4.3.1.$$

where:

$$k_2 = \int_R t^2 K(t)dt.$$

Further improvement of this base formula of h optimum can be found in the literature. So, Wand and Jones (1995) and Silverman (1986) give details in pure mathematical optimization procedures finding h, for different situations. Gentle (2009) also gives details and approximations of the h optimum. Additional comments on the issue are discussed in the eighth chapter of this paper "General discussion".

On the other side, the optimized kernel (with the condition of being density probability, as mentioned in this paper), solved by Hodges and Lehmann (1956) and by Silverman (1986), gives the kernel $K_e(.)$, as follows, and also suggest the density estimation techniques by Epanechnikov:

$$K_e(t) = \begin{cases} \frac{3}{4\sqrt{5}}(1-\frac{t^2}{5}), & -\sqrt{5} < t < \sqrt{5} \\ 0, & t \text{ other} \end{cases} \qquad 4.3.2.$$

The idea of efficiency arose by comparing any kernel with the above mentioned one – Epanechnikov Kernel. The definition of efficiency of a kernel, as described by Silverman (1986) is:

$$\mathrm{eff}(K) = \frac{3}{5\sqrt{5}} \left\{\int_R t^2\, K(t)dt\right\}^{-1/2} \cdot \left\{\int_R K(t)^2\, dt\right\}^{-1}, \qquad 4.3.3.$$

Relating to the concept of efficiency, all other literature is referring to the same concept mentioned above. Choosing a high efficiency kernel is quite needed, as it is perfect legitimate to choose a kernel that fulfills other requirements, such as degrees of differentiability that can be necessary in some computational efforts.

5. INTUITION, RESULT INTERPRETATION AND DATA

Time series has been the workhorse of empirical analysis in the study of economic phenomenon. In principle time series methods calculate "the walk" of the variable around its mean along its distribution under the directions given by the matrix of estimated parameters, along with a i.i.d. set of errors. These errors measure the deviation in the location of the current realization of the variable at period t from the forecasted location based on the estimated parameters values, which are estimated from the covariance matrix.

As described in chapter 2, time series methods build upon the probability theory to generate estimates of the parameter values of the random event. These probability based models are used to perform four major tasks for the benefit of the policymakers: first, provide a coherent and credible description of the economic data; second, provide reliable forecasts for economic variables; third, provide hints of the structural inference; and last but not the least provide good foundation or policy analysis. Stock and Watson (2001)[11] provide a good discussion of the role and the performance of VAR in each of the four areas above. They also highlight the fact that certain aspects of the VAR methodology are more useful than the estimated coefficients themselves depending on the needs of the researcher or policymaker. Though the focus of the time series method is on the estimation of the parameters matrix, frequently coefficients may go unreported and the analysis relies more on exogeneity, structural form, variance decomposition, impulse response functions and the properties of the estimated matrix.

As an alternative and complementary to time series analysis, we propose the calculation of d dimensional joint density probability functions as anew and alternative nonparametric method of representation, analysis and forecast of the DGP of ad dimensional random variable in R^d. Therefore, in order to be an alternative, the new method needs to supply similar information and reliability in the description of the DGP like the rest of the existing methods.

[11] Stock, J.H., Watson M.W. (2001), Vector Autoregression, The Journal of Economic Perspectives, Vol. 15, No. 4. (Autumn, 2001), pp. 101-115.

Despite its nonparametric nature, this new method is capable to provide similar information on policy and forecast analysis based on the likelihood of simultaneous realization of two or more particular random variables via conditioning of the estimated (calculated) d dimensional joint density probability distribution. The obtained conditional probabilities are based on one particular or a set of expected values of the event rather than on the marginal probabilities of the previous events.

Therefore depending on the interest of the research, policy analysis, or forecast one can calculate the JPDF of an m dimensional random event and report the DGP in the form of the joint conditional as $D_{x_m}(\underline{t}_d, \underline{x}_d, \hat{f}(\cdot) \,|\, \underline{t}_{d-m})$, where x_m represents a m dimensional vector, $m \in (1, 2, \ldots, d-1)$, or unconditional process as $D_{x_d}\left(\underline{t}_d, \underline{x}_d, \hat{f}(\cdot)\right)$.

Based on the above representation of the DGP, the new method can perform similar tasks with econometrics methods for a scalar time series in all those cases where $m=1$ yielding forecasts based on the data in levels, as well as structural and shock analysis based on the first difference and on the percentage changes of the same data, giving this information in the form of the probability of distribution of x_m– the variable of interest for each particular value or a set or vector(representing a set of values) of the other $d - m$ random variables.

Generally, the analysis for each variable of interest x_m is based on the comparison and the evolution of the moments of conditional distributions of x_m. The method could also be used on differenced data (in order to remove potential build in common time or stochastic trends as is commonly the practice in time series) or percentage changes in variables with respective distributions potentially revealing information regarding structural relationships or the contemporaneous links among variables and finally growth rates, respectively.

The JPDA has the potential to provide information regarding another important issue, the constancy of the estimated parameters and potential structural breaks in the data, e.g. the existence of a multimodal relationship could potentially imply the presence of structural breaks in the data or shifting relationships in the differenced or elasticity in percentage changes representation of the data.[12] In this respect, the information coming out of the JPDF can potentially provide the researcher with the equivalent of causality, structural analysis, and impulse response and variance decomposition, which is the set of information provided by the current empiric investigation methods.

In conclusion, calculation of the JPDF of a d dimensional random event opens the possibility to use this d dimensional JPDF to display in graphic representation and analytical form the JPDF of the random variables or subspaces of interest within this d dimensional random event given the known or forecasted values (restrictions) of the other $d-1$ random variables of our event. The resulting conditioned JPDF will provide at least the same or improved (in fact much more informative and real) information than the parametric method with regard to the four tasks above. Moreover, it does so without the setback required by the parametric methods.

Meaning, one does not need to make assumptions regarding constant mean and variance, or the normality of the errors (distribution) including all other assumptions regarding the ordering (the endogeneity) of the variables and the structural form of relationships. These characteristics relax the functional and statistical needs for log linearization in the data. Therefore the data can be modelled in their natural form. Additional benefits derive from the fact that time series empirical analysis, and other parametric methods for that matter, are traditionally modeled and conducted on the log and log differences of observed variables. This log transformation of the original data is a necessity given the non-linear functional form of the economic relationships and/or due to the particular interest in growth rate equations (or law of motions), which are commonly approximated by changes in the log of a variable. In addition log linearization brings additional statistical benefits related to the homogenous variance and the normality of the residual associated with the log models (See Bardsen and Lutkepohl (2011)).

[12] It remains to be tested and will be the target in our future research.

Finally, due to the marginalization and conditioning process, which renders DGP workable for policy-related analysis as depicted by equation 2.1., the errors that are made in early periods in the process of forecast or impulse responses analysis do carry over and accumulate in the next periods. This accumulation results in the expansion of the confidence interval for each consecutive period of forecast or impulse responses analysis eventually leading to reduced predictability and accuracy in results. The representation of the DGP in the form 5.1 eliminates this error accumulation.

At the current stage, the analysis of the densities based on a dataset of the Albanian economy is given as an illustration and is based mostly on the graphical presentation. As such, the analysis is limited to the two dimensional JPDF, due to lack of visualizing apparatus and intangible nature of d dimensional vector constrained by the three dimensional space. The results and their interpretation are given in the seventh chapter and in the first and second annexes. And this is one of the main constrains of the method mentioned in the base literature (see 4.5. at Silverman (1986)). Nonetheless, we have some more concrete results (still draft) that try to bypass such difficulty, which will be part of our future publication on this issue.

6. APPLICATION OF THE DENSITY ESTIMATION METHOD, STUDY OF VARIABLES, GRAPHS PRESENTATIONS AND INTERPRETING THE RESULTS

Density estimation represents a complete (comprehensive) method for estimation, analysis and simulation of a random phenomenon and, because of these attributes has found various uses in different fields of study. Our paper proposes the application of this method for economic and financial variables. In this chapter, we illustrate the use of those methods for a set of economic variables in the case of Albania. As mentioned above, the analysis of selected variables is based on graphical representation of the estimated one and two dimensional densities.

In the case of this article, we will look at some key aspects (introductory) based on the application of the "density estimation". The selection of the variables we study and methods are presented below.

The data is used in the scalar form – per each of the variables the density estimates are evaluated, and vectorial form – density estimates of the random vectors are evaluated. Density estimates are given for several kernels.

6.1. DATA DESCRIPTION, MODELING AND GRAPHICS

In this section, we will illustrate the use of density estimation as a method of economic analysis investigating the relationship between money, inflation, exchange rate, GDP and the interest rate for the period 1998-2011 in the Albanian economy. Despite being very important elements of policy design and decision making for the Bank of Albania, the set of variables and their relationship is the objective of several research papers produced by the Bank of Albania staff, which provide abundant material to match and compare against our results. The dataset represents quarterly observations for the period Q1, 1998 to Q4, 2011, and are reported in levels, first difference and as percentage changes (with the exception of the interest rate, which is reported only in its level and first difference form).

Money is represented by the stock of monetary aggregate M2, as reported by the Bank of Albania. Exchange rate represents the nominal effective exchange rate (as estimated by the Bank of Albania)[13], with positive changes representing ALL depreciation. GDP represents annualized gross domestic product statistics measured by INSTAT. Quarterly observations for the period 2005 -11 are reported by INSTAT, the rest of the period represents disaggregated quarterly GDP data of the annual ones reported by INSTAT[14]. CPI represents the consumer price index measured and reported by INSTAT. Interest rate variable represents the real interest rate of 12-month deposits as reported by the Bank of Albania. Money and GDP data are expressed in millions of Albanian ALL.

The modelling of the selected data consists in the following: we construct density estimates for the selected macroeconomic variables, one and multidimensional variables. To give the idea of allocation of probability one and two dimensional variables, we provide the graphical representation of these evaluators.

Graphic analysis is used as the first tool to explore the behavior of probability densities one or two dimensional. A challenge in that case, quite differently from the parametrical case where we faced well-known distributions, is to see behind the mixed shapes and colours of quite natural distributions – creating experience is a challenge that we suggest.

Based on canonical concrete expressions of the density estimates mentioned in the chapter 4.1. and 4.2., we contributed to programming the software made for that case, as the readymade software in that field of research is missing. Firstly build in PHP

[13] See Vika (2006) for a description of the methodology. Vika, I., 2006, "Kursi real efektiv i këmbimit në Shqipëri - konceptet dhe matja e tij", Bank of Albania WP series 2(18), ISBN 99943-864-1-7

[14] Statistics and methodology come from the dataset of the Bank of Albania macro-econometric model (MEAM)

programming language; we plan to contribute to completing it with other techniques in density estimation in our future research on that field. The software actually creates graphical representations of one and two dimensional estimates, and also gives the graphical representations of d dimensional estimated density with $d-2$ and $d-1$ elementary conditions.[15]

6.2. ONE DIMENSIONAL DENSITY ESTIMATES OF SOME MACROECONOMIC DATA OF THE ALBANIAN ECONOMY

For the variables described in the following table, the density estimates are represented in graphics, using Gaussian and Triangle kernels. The formulas used are given in 4.1.

Table 6.2.1.: Variable name and description, one dimensional case

Variable description	Variable name	Time period
M2 real (billions ALL)	RM2-bln	[Q1, 1996-Q4, 2011]
M3 real (billions ALL)	RM3-bln	[Q1, 1996-Q4, 2011]
CPI Index	CPI	[Q1, 1996-Q4, 2011]
Interest rate Euribor (percentages)	ER_3M	[Q1, 1996-Q4, 2011]
Interest rates 12 months deposits	R-12M-Dep	[Q1, 1996-Q4, 2011]
GDP real (billions ALL)	Y-REAL-bln	[Q1, 1996-Q4, 2011]
GDP nominal (millions ALL)	YN	[Q1, 1996-Q4, 2011]
Nominal Effective Exchange Rate, quarterly (ALL/Eur)	NEER-3M	[Q1, 1996-Q4, 2011]

The graphs of the density estimates are presented in Annex A1.

6.3. TWO DIMENSIONAL DENSITY ESTIMATES OF SOME MACROECONOMIC DATA OF THE ALBANIAN ECONOMY

The data used are in level, in first difference and relative difference.

[15] "Elementary conditions", as it is used in our paper (as analogy with "elementary event"), means that the probability condition is a single point, respectively in d-2 and d-1 dimensions.

For those variables, the density estimates are represented in graphics, using Gaussian kernel. The formula used is given in 4.2. Surface and contour graphics representations are used.

Table 6.3.1.: Variable name and description, two dimensional case

	Variable description	Variable name	Variable name	Variable description	Time period
Data in levels	M2 real (billions ALL)	RM2-bln	CPI	CPI Index	[Q1, 1996-Q4, 2011]
	M3 real (billions ALL)	RM3-bln	CPI	CPI Index	[Q1, 1996-Q4, 2011]
	Interest rate	R-12M-Dep	CPI	CPI Index	[Q1, 1996-Q4, 2011]
	GDP real (billions ALL)	Y-REAL-bln	CPI	CPI Index	[Q1, 1996-Q4, 2011]
	Nominal Effective Exchange Rate, quarterly	NEER-3M	CPI	CPI Index	[Q1, 1996-Q4, 2011]
Data in first difference	M2 real (billions ALL) (First difference)	dRM2-bln	dCPI	CPI Index (First difference)	[Q1, 1997-Q4, 2011]
	M3 real (billions ALL) (First difference)	dRM3-bln	dCPI	CPI Index (First difference)	[Q1, 1997-Q4, 2011]
	Interest rate (First difference)	dR-12M-12-Dep	dCPI	CPI Index (First difference)	[Q1, 1997-Q4, 2011]
	GDP real (billions ALL) (First difference)	dY-REAL-bln	dCPI	CPI Index (First difference)	[Q1,1997-Q4, 2011]
	Nominal Effective Exchange Rate, quarterly (First difference)	dNEER-3M	dCPI	CPI Index (First difference)	[Q1, 1997-Q4, 2011]
Data in relative difference (%)	M2 real (billions ALL)	pRM2-bln	INF	Inflation	[Q1, 1997-Q4, 2011]
	M3 real (billions ALL)	pRM3-bln	INF	Inflation	[Q1, 1997-Q4, 2011]
	Interest rate	dR-12M-12-Dep	INF	Inflation	[Q1, 1997-Q4, 2011]
	GDP real (billions ALL)	pY-REAL-3M	INF	Inflation	[Q1, 1997-Q4, 2011]
	Nominal Effective Exchange Rate, quarterly	pNEER-3M	INF	Inflation	[Q1, 1997-Q4, 2011]

The graphs of the density estimates are presented in Annex A2.

Below we present the graphical representation of the density estimations for the variables described in Table 6.2.1. The kernels used in the density estimation are Gaussian and Triangle kernels, as mentioned in the label of the vertical axes in each graph.

As an interpretation: these are the forms of distribution of the Albanian economic variables represented.

7. CONCLUSIONS

7.1. Parametric methods of estimation have played an important role in the study of economics. Among them, time series analysis has become the fundamental tool of research in economics. However, the practical implementation of this theoretic framework is rendered possible only by a set of rigid restrictions and constraints. In this respect, the study of economics based on this method portrays the economy in a very unrealistic way and transforms the data to provide a different fit for these constrains. The purpose of this paper is to present and propose density estimates tools, in the framework of multidimensional probability analysis, as an alternative tool of empiric investigation of the economic phenomenon, without the restrictions imposed by time series and other parametric estimation methods. In principle, we propose defining the economy as a social phenomenon in the form of a random event in a $\mathbf{d}$ dimensional space. We use this definition to study the economy (the data generating process) in the form of a random event as a joint density function departing from the traditional difference equation framework. This interpretation comes after constructing the density estimates in the $\mathbf{d}$ – dimensional space, based on the mathematical postulate that knowing the values and the distribution of a random variable means that we know all the information about this variable.

7.2. On this ground, the paper proposes the use of density estimation as a method of approximating through decomposition of the unknown density with known kernels. In that way, considering decomposing the stochastic information of our economic phenomenon in the form of weighted sum of known probability densities, we keep most of the stochastic information of a random phenomenon. In this respect we describe the methodology of density estimation in one and multidimensional cases for several selected kernels. The material provides the analytical functional forms of the one and multidimensional density function together with a broad discussion of the literature as well as a technical discussion on the choice of the kernel and the smoothing parameter. The study of economics through this alternative methodology leaves behind the restricting assumptions employed in the current empirical methods including the reliance on normal distribution family.

7.3. The fundamental change proposed by the density estimation method is that the correlations and comovement of the economic variables are defined, in substance, from the probability measures rather than from the values of those variables. The advantage of the proposed methodology over other nonparametric probability based inference methods is that multidimensional density function preserves the simultaneity of singular event occurrence across all dimensions of the random event during the time of investigation. Therefore conditioning the process along one or more (up to $d-1$) dimensions provides the conditional density estimates of occurrence for the variable(s) of interest with the moments of resulting distribution yielding similar information to the other currently used empiric investigation methods. The conditional analysis can be applied with scalar or vector condition (representing dynamic movement in the known variable), with the results representing the dynamic movements of the conditional distribution of the random event. Even further probabilistic analysis may give more detailed results based on analysis of marginal distributions.

7.4. The purpose of the paper was to present the density function as an alternative method to describe and visualize the data generating process and as an alternative mean of empirical investigation. For illustration purposes, we have applied the proposed method to the study of relationship among inflation, money, income, interest rate and exchange rate in the case of the Albanian economy. The results in this paper are presented in graphical form, which limits the analysis to two dimensional case. The paper reports for the first time density estimate functions for inflation with each of the other variables.

Looking at the graphs, it is interesting to observe the shapes and generally the forms of densities. So normally we interpreted them in the following way:

- We observed some multimodal distribution located near the same linear path, which we interpreted as an indication of the variables sharing the similar stochastic or time trends.
- In the most dominated unimodal distribution we accept them in some cases as an uncorrelated bi-dimensional distribution.

Anyway, the correlation could and must be calculated in some cases, especially when the form of the distribution is oval shaped.

- The same conclusion is drawn even in the cases when many bell shapes are irregularly distributed around the horizontal plane.
- Oval shape probabilistic allocations, which are located almost orthogonal with one of horizontal axes, are in fact low variance on the axe of orthogonality relatively to the variance of the other axe. We interpret this observation as a sign of independency on those variables.

Further analysis could be based on the visual interpretation of density estimates. Also, getting a deeper interpretation is straight by interpreting the moments of conditional density functions. However, we acknowledge that this process might yield the challenge of dealing with “strange” distribution from the reality.

7.5. The illustration engages macroeconomic variables but the density estimation is useful and can be straight forward extended to the study of phenomenon that is described in the form of a random event, especially in the cases of large data sets where the marginalization of conditional probabilities renders the parametric estimation methods difficult due to the large number of calculations and low degrees of freedom.

This work is extended in several directions, which will be the focus of our future research.

8. GENERAL DISCUSSIONS FOR THE METHODOLOGY AND FUTURE RESEARCH

The methods we propose in this paper have the potential to provide a lot of answers relating to the history of the economic variables and their relationship and, on the other side raise a lot of questions. In this chapter, we would like to bring to the reader's attention and discuss few important issues that emerge from the base literature and also from our experience in dealing with the application of the methodology in the estimation of densities.

The choice of the kernel:

- Some of the questions that may rise on the choice of the kernel that we use on the density that can be used. Practically, based on our attempts, the general considerations of the choice of the kernel relate to the aim of the use of the density estimation. In general, if we look for a smoother curve or surface, we have to choose a "more smoothed" kernel. However, the choice of the kernel will also depend on the expression of the density in the analytical transformations, as we have to choose a kernel that is possible to be easily transformed from the mathematical operators we plan to apply, i.e. such as degrees of differentiability that can be necessary in some computational efforts. Anyway, as we mentioned in this paper, we suggest the "smoothing parameter" h as an alternative (or better way) to achieve desired smoothness in respective one and multidimensional density estimates. A very important point that we have to consider is that: a discontinuity estimator, as the histogram itself is, gives extreme difficulty in a process where estimates of the derivatives are required.

Further in that framework, in the literature there are different ways of modifying the kernel method into similar methods that are suggested for different situations.

So, *the nearest neighbour class of estimates* represents the effort to arrange the amount of smoothing to the local density of data. In that framework, they are normally used in the study of the tails and is unlikely to be appropriate to study the density as a whole, since this estimate is not itself a density. *The adaptive kernel method*, as is shown by Silverman (1986), has certain potential practical advantages over both the kernel and the nearest neighbour methods as a method for smoothing long tail distribution. Other methods come directly as "smoothing of a function" techniques, like *orthogonal series estimates*. They are derivative of all orders and can be efficient in the cases similar to their application in the other types of functions than densities.

- Based on our experience, all the methods used have their own pro and cons elements. However, despite all arguments, the use of kernel methods is the best and most practical choice; especially in the case when one studies the density as a whole. If the situation calls for the study of local features of the density curve, other local estimates can be chosen.

The *efficiency* of the kernel:

- Another feature on the choice of the kernels is *the efficiency*. Optimization process of minimizing measures of discrepancy between a density that must be estimated and the estimate gives a solution of the *most efficient* kernel, which is Epanechnikov kernel in one dimensional case, as is mentioned in 4.3. Details on the techniques of the optimization are given in the main literature of the field mentioned in this paper.

It is also worth mentioning that any other one dimensional kernel used has *the efficiency* less than 1, as it is defined here.

- Finally, approaching an answer to the main question "which kernel to use and when", what we can add to the findings of the main literature, based on our empirical work, partly published in this paper, is: to know the main features of a concrete distribution, try with the main known distribution (as in chapter 4) and keep in control the efficiency and the smoothing parameter.

The choice of the smoothing parameter h.

- Other questions relate to *the choice of the smoothing parameter h*. Literature suggests easily practical guidance on the optimal choice of the smoothing parameter h. The experienced user can try different values of h considering the general rule that: small values of h may increase 'statistical noise' of the density estimates and large values may cause the loose of the statistical attributes. As we mentioned in this paper, there exist research relating with the optimization of the smoothing parameter. Silverman (1986), Wand and Jones (1995), Gentle (2009) etc., described techniques in ideal window width and kernel.

Based on our experience, we suggest higher values of h in the study of the whole density and smaller values of h whenever the focus of investigation falls on particular local topography of the density. To be mentioned is that based on our experience, we can suggest optimized values of h if the interest is on viewing the density estimates as a whole and smaller values of h if the interest is in a narrow local zone of the density. In the procedures that need to produce a large number of graphs, and because of this reason need to be automatised, the literature suggests approximations of h. The relating literature is mentioned in 4.3. in more details.

- The choice of the smoothing parameter in the multidimensional spaces in principle is "liberalized" in the meaning that it is not obligatory to choose the same unified smoothing parameter in all directions (dimensions) of the space R^d of the variables of the density estimates. So, in that way, the needed modifications of the formula 4.2.2. can be made. This is considered in the literature as *the more sophisticated ways of choosing the windows width* and in substance is advised to be applied relating with the densities, which appear radically not symmetric.

- What we can suggest is that "playing" with different values of smoothing parameter h is a necessity. Empirically, we may find values of h, which optimized the shape of the density estimate.

In terms of potential application and our future research, we will continue to apply this method in the exploration of the particular features of economic variables and their interaction respectively via the investigation of one and multidimensional, density estimation techniques proposed in this paper.

9. LITERATURE

Bardsen, G., Lutkepohl, H., (2011) Forecasting levels of log variables in vector autoregressions, International Journal of Forecasting, Vol. 27, issue 4, Oct. 2011, pp. 1108-1115

Clements, M.P. and Smith, J.(2002), Evaluating Multivariate Forecast Densities: a Comparison of Two Approaches, International journal of Forecasting, 18, 2002, pp. 397-407.

Devroye L., Györfi L., (1985), Nonparametric density estimation, the L1 view. John Wiley & Sons, Inc.

Devroye, L., Lugosi, G., (2000), Combinatorial Methods in Density Estimation. Springer- Verlag, New York.

Diebold, F. X., Gunther, T. A., & Tay, A. S. (1998).Evaluating density forecasts with applications to financial risk management. International Economic Review, 39, 863– 883.

Epanechnikov, V. (1969), Nonparametric estimates of a multivariate probability density, Theory of Probability and its Applications, 14, 153-8.

Ericsson, N., R., Hendry, D., F., Mizon, G., E., (1998), Exogeneity, Cointegration, and Economic Policy Analysis, Journal of Business & Economic Statistics, Vol. 16, No. 4 (Oct., 1998), pp. 370-387

Gentle, J. E. (2009), "Computational Statistics", Springer.

Hendry, D. F. and J.-F.Richard (1983). The econometric analysis of economic time series (with discussion). International Statistical Review 51, 111 - 163.

Hendry D., F., (2011) Mathematical Models and Economic Forecasting: Some Uses and Mis-uses of Mathematics in Economics, University of Oxford, Department of Economics Discussion papers, No. 530, ISSN 1471-0498, Feb. 2011

Hodges J.L., Lehmann, E.L. (1956), The efficiency of some nonparametric competitors of the t-test. The Annals of Mathematical Statistics, 27.

Juselius, K., Franchi, M., (2007), "Taking a DSGE model to the data meaningfully, Economics, the open-access, open-assessment'', E-Journal, No. 2007-4, Jun. 2007.

Juselius, K.,(2006), The Cointegrated VAR Model: Methodology and Applications, Oxford University Press, ISBN 0-19-928566-7.

Marron J.S., Wand M.P. (1992), "Exact mean integrated squared error", The annals of Statistics, 1992, Vol.20, No.2, 712-736.

Parzen E., (1962), On Estimation of a Probability Density Function and Mode, Source: Ann. Math. Statist. Volume 33, Number 3 (1962), 1065-1076.

Rosenblatt M., (1956), Remarks on some nonparametric estimates of a density function, Source: Annals of Mathematical Statistics, Volume 27, Issue 3 (Sep.,156) , 832-837.

Silverman B.W., (1986), "Density Estimation for Statistics and Data Analysis" Chapman Hall.

Silverman, B. W., Jones, M. C. (1989), "E. Fix and J.L. Hodges (1951): An Important Contribution to Nonparametric Discriminant Analysis and Density Estimation: Commentary on Fix and Hodges (1951)", International Statistical Review / Revue Internationale de Statistique, Vol. 57, No. 3 (Dec., 1989), pp. 233-238.

Sims, C., A., (1980), Macroeconomics and Reality, Econometrica, Vol. 48, No. 1 (Jan., 1980), pp. 1-48.

Solow R., (1985), Economic History and Economics, "The American Economic Review", Vol. 75, No.2, May 1985.

Solow, R., (2010), "Building a Science of Economics for the Real World", Prepared Statement for Congressional Testimony, House Committee on Science and Technology Subcommittee on Investigations and Oversight, July 20, 2010.

Stock, J.H., Watson M.W., (2001), Vector Autoregressions, The Journal of Economic Perspectives, Vol. 15, No. 4. (Autumn, 2001), pp. 101-115.

Tsybakov (2009), "Introduction to Nonparametric Estimation", Springer.

Vika, I.(2006), "Kursi real efektiv i këmbimit në Shqipëri–konceptet dhe matja e tij", Bank of Albania WP series 2(18), ISBN 99943-864-1-7.

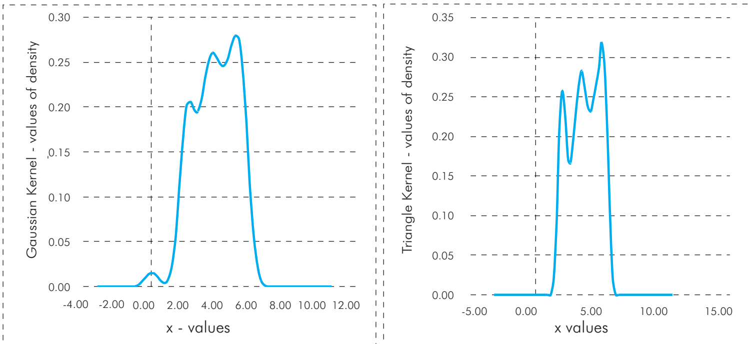
Wand M.P., Jones M.C., (1995), "Kernel Smoothing", Chapman Hall.

ANNEX 1: ONE DIMENSIONAL DENSITY ESTIMATES OF SOME MACROECONOMIC DATA OF THE ALBANIAN ECONOMY – GRAPHICAL REPRESENTATION OF DENSITY ESTIMATES.

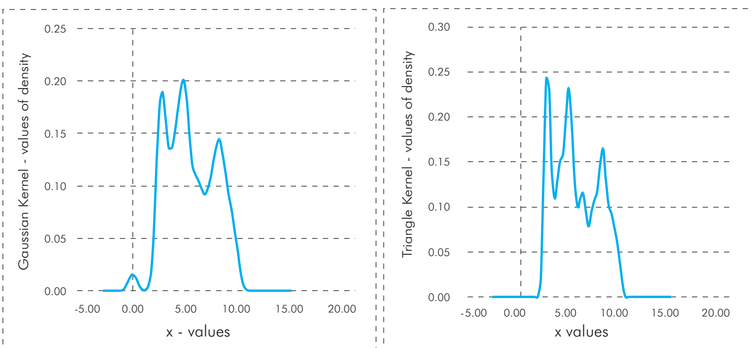
Below we present the graphical representation of the density estimates for the variables described in Table 6.2.1. The kernels used in the density estimates are Gaussian and Triangle kernels, as mentioned in the label of the vertical axes in each graph.

As an interpretation: these are the forms of distribution of the Albanian economic variables represented.

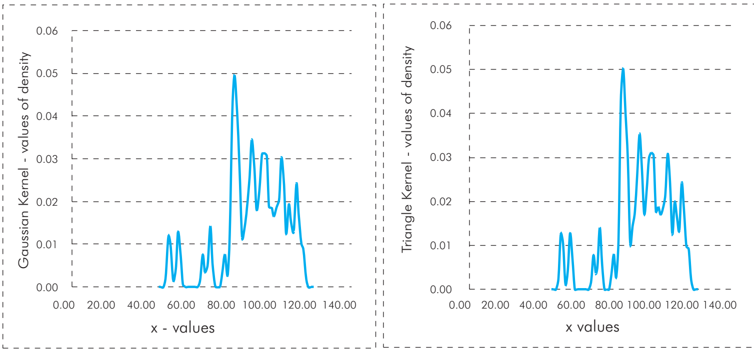
Graph A1.1.: Density estimates for the variable M2 real (bln ALL), for the time period [Q1, 1996 – Q4, 2011]



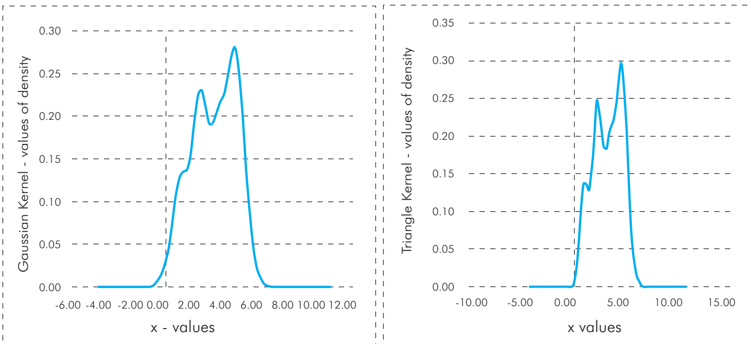
Graph A1.2.: Density estimates for the variable M3 real (bln ALL), for the time period [Q1, 1996 – Q4, 2011]



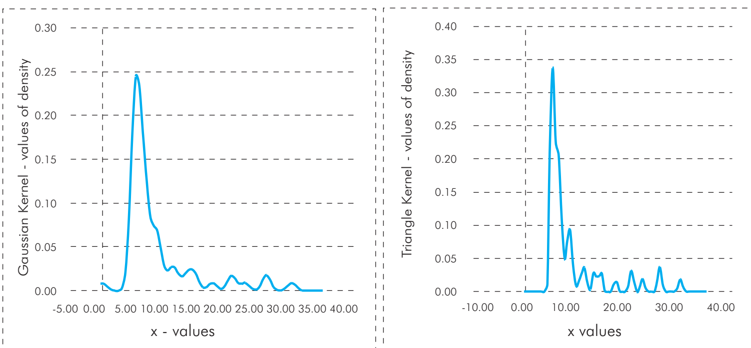
Graph A1.3.: Density estimates for the variable CPI Index, for the time period [Q1, 1996 – Q4, 2011]



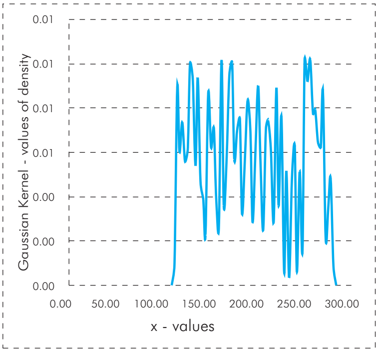
Graph A1.4.: Density estimates for the variable Interest rate Euribor (percentages), for the time period [Q1, 1996 – Q4, 2011]

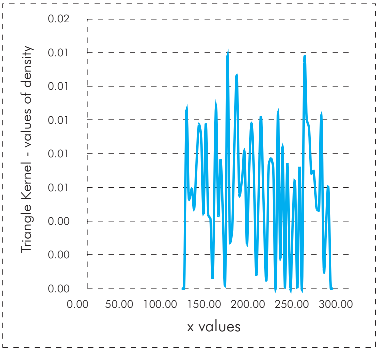


Graph A1.5.: Density estimates for the variable Interest rates deposits 12 months, for the time period [Q1, 1996 – Q4, 2011]

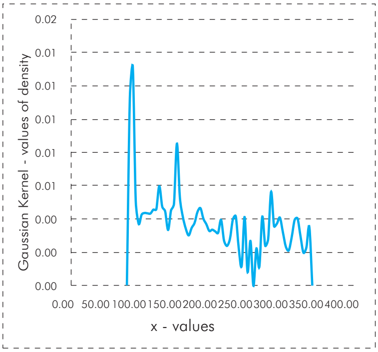


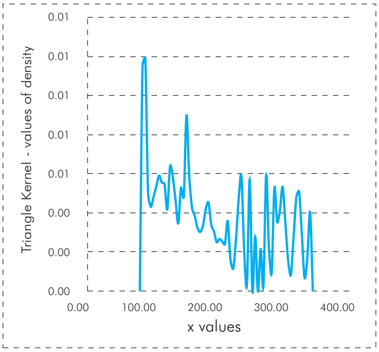
Graph A1.6.: Density estimates for the variable GDP real (bln ALL), for the time period [Q1, 1996 – Q4, 2011]



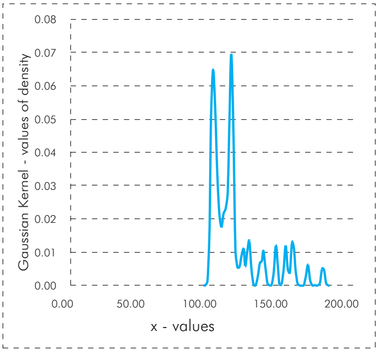


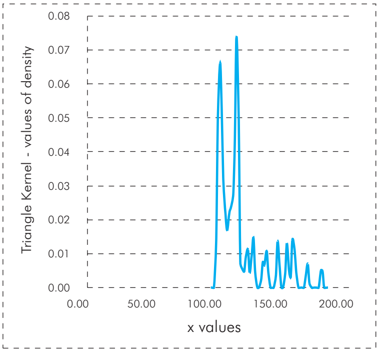
Graph A1.7.: Density estimates for the variable GDP nominal (mln ALL – devided by 000), for the time period [Q1, 1996 – Q4, 2011]





Graph A1.8.: Density estimates for the variable Nominal Effective Exchange Rate quarterly (ALL / EUR), for the time period [Q1, 1996 – Q4, 2011]





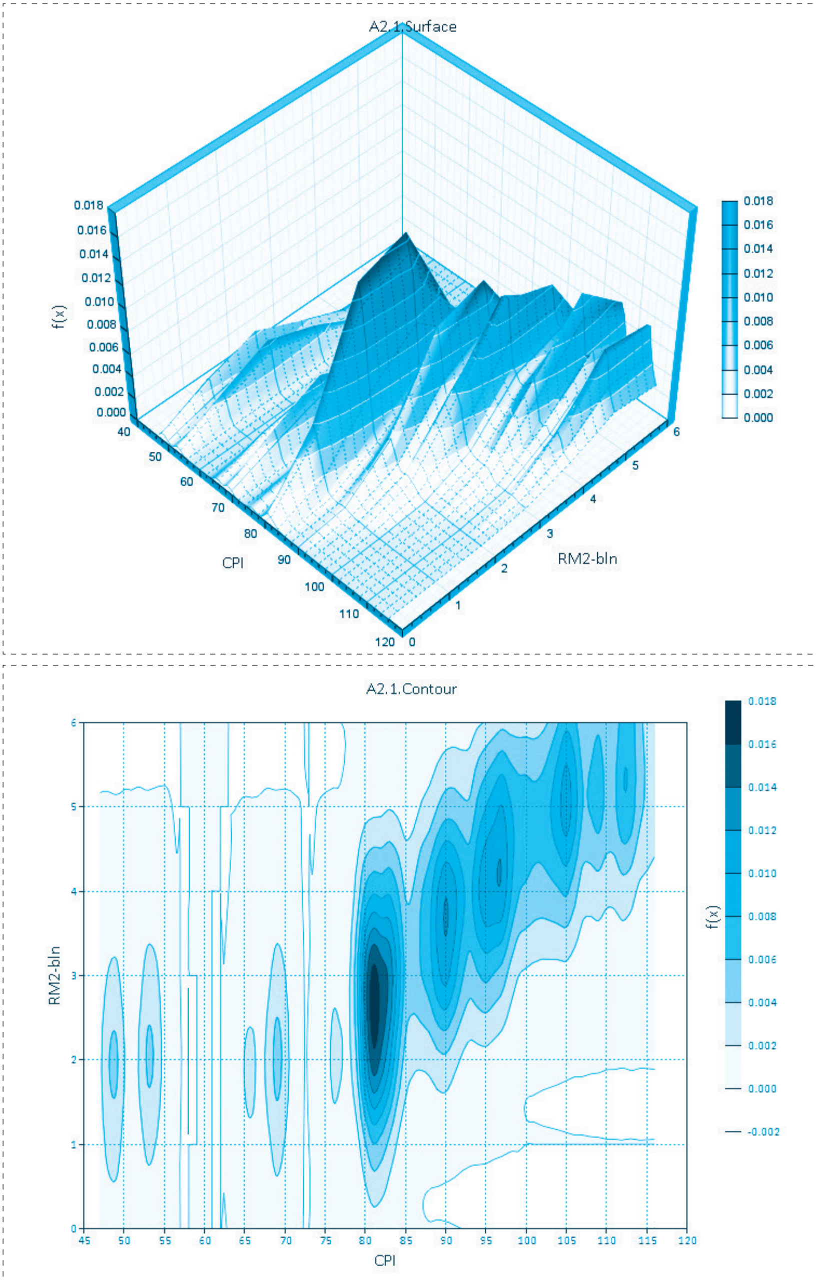
ANNEX 2: TWO DIMENSIONAL DENSITY ESTIMATES OF SOME MACROECONOMIC DATA OF THE ALBANIAN ECONOMY – GRAPHICAL REPRESENTATION OF DENSITY ESTIMATES.

Here we illustrate the method by applying it to explore the relationship among inflation, money, interest rate, exchange rate and aggregate demand (approximated by GDP). The data are described above in section 6.3. Here we discuss an intuitive interpretation of the two dimensional graphical representation of the data in level, difference and percentage changes (with the exception of the interest rate, which is only in its level and difference form). We have to note that, as mentioned in the paper, in the third dimension there are the values of the function of density estimates.

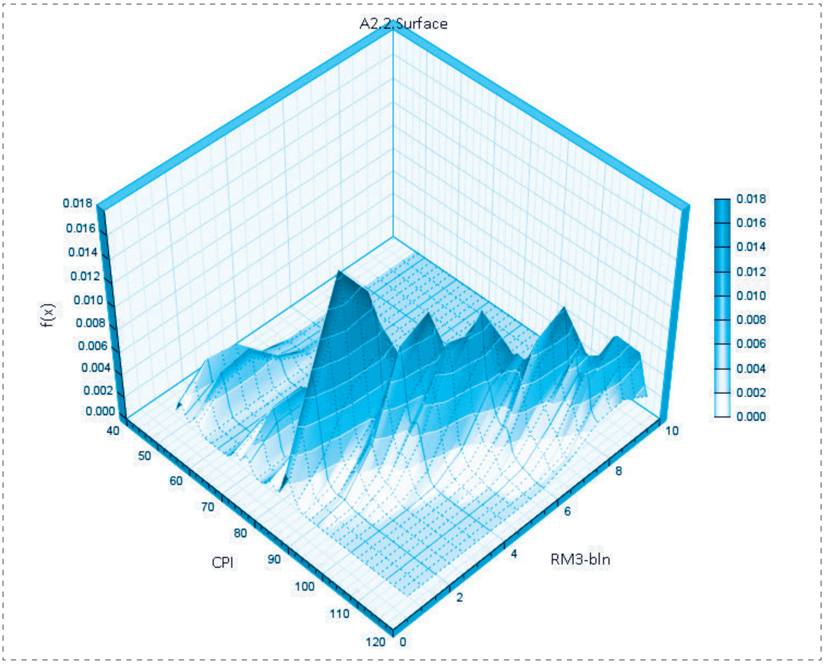
DATA IN LEVEL OF THE VARIABLES

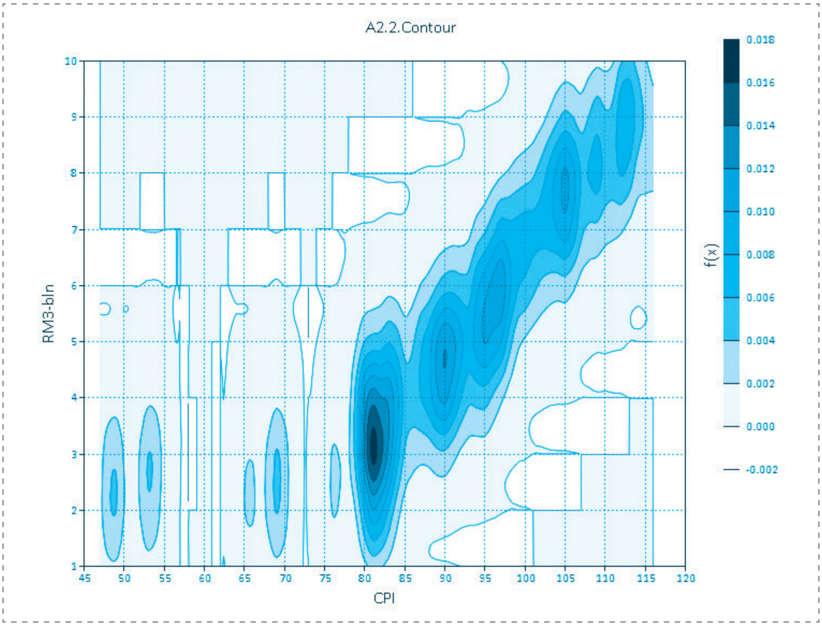
Graphs A2.1. – A2.5. represent density estimates of joint density distribution of two dimensional vector of inflation and all other economic variables, all in levels. The first interesting observation is that all graphs, with the exception of the interest rate and quarterly exchange rate, show that it is highly probable to find density estimates that are positioned along the diagonal indicating a positive relationship between the level of CPI and Money. Or the existence of a joint time trend in the data. There is no clear pattern of densities emerging in the case of interest rate and exchange rate. However, the presence of multiple bell shapes scattered along the graph indicates that the variables are not independent. Independence in the data in level would imply the existence of a single and regular bell shaped graph.

Graph A2.1.: Density estimation for the variables CPI and M2 real (bln ALL), Gaussian kernel, for the time period [Q1, 1997 – Q4, 2011]

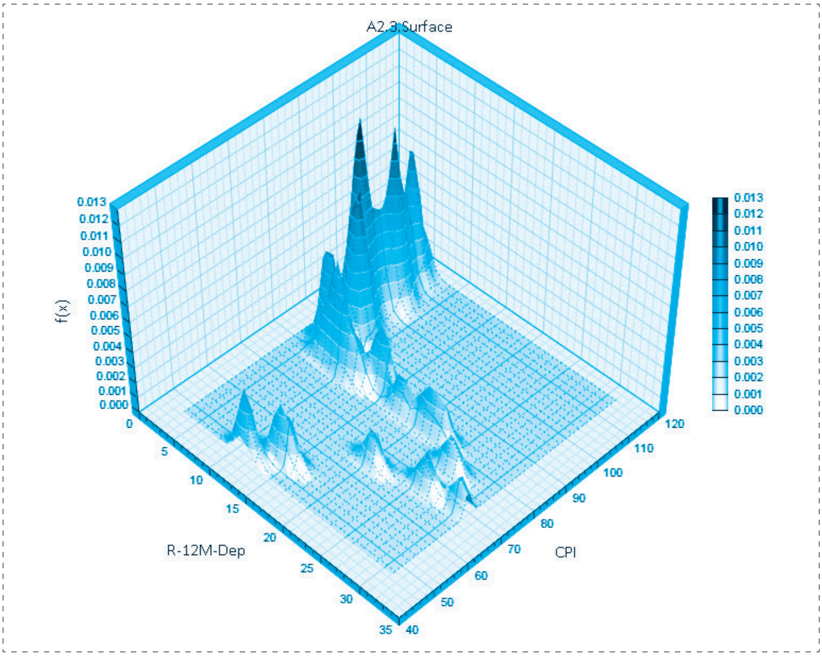


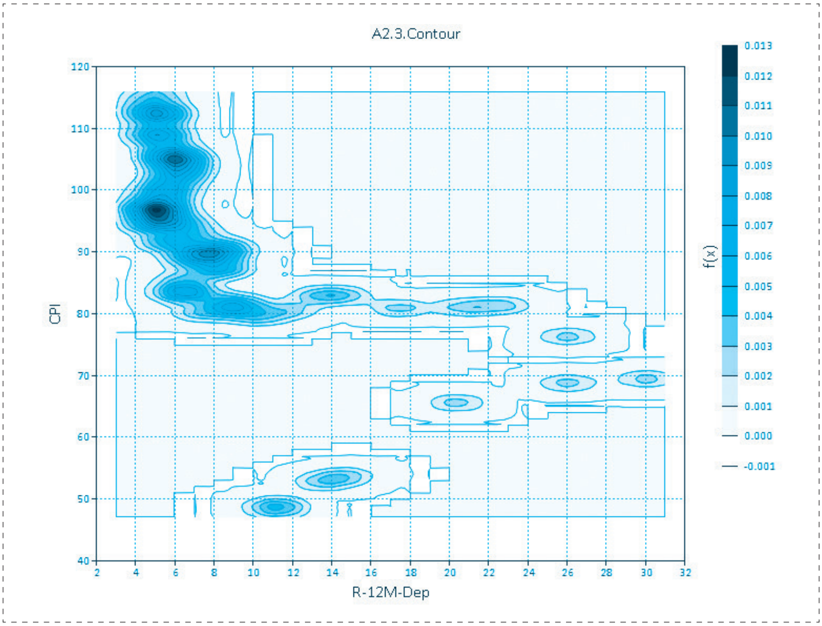
Graph A2.2.: Density estimates for the variables CPI and M3 real (bln ALL), Gaussian kernel, for the time period [Q1, 1997 – Q4, 2011]



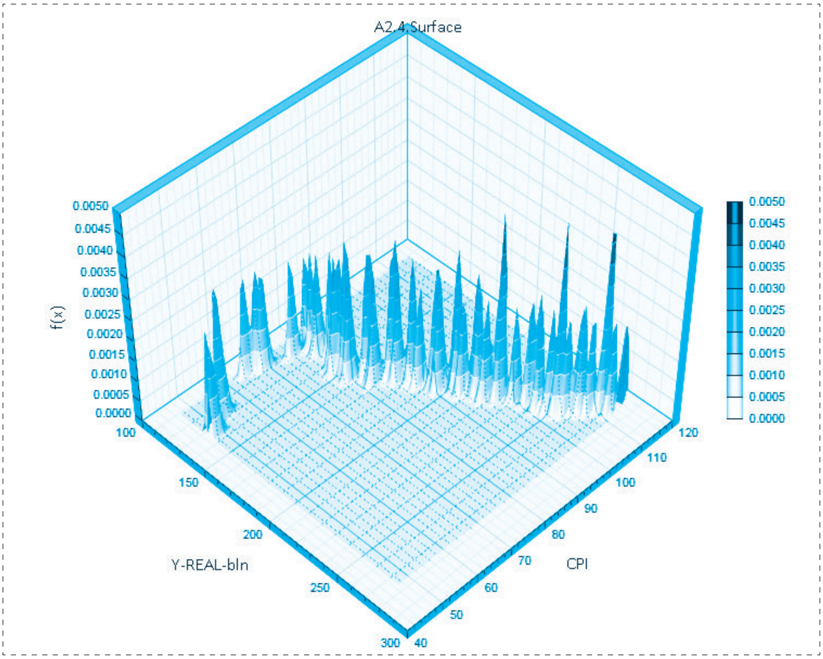


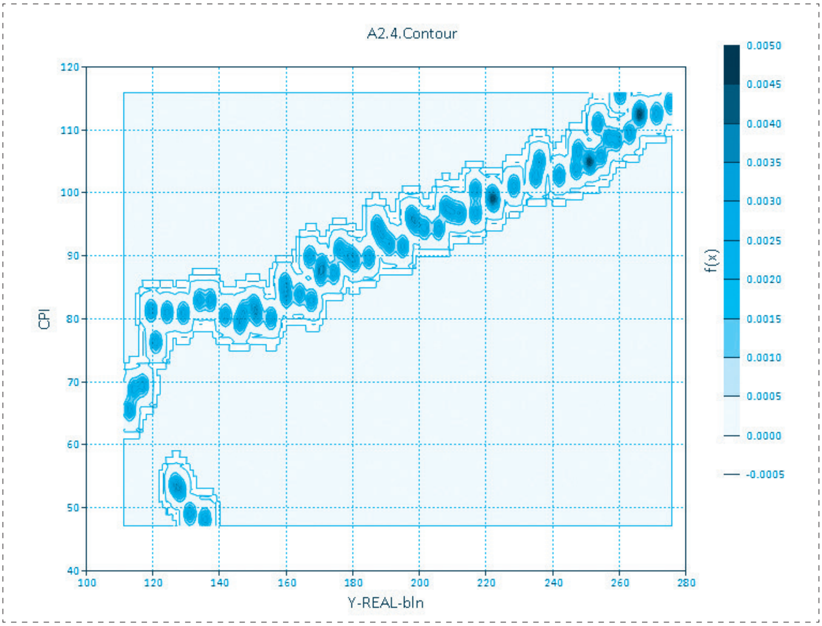
Graph A2.3.: Density estimates for the variables R and CPI, Gaussian kernel, for the time period [Q1, 1997 – Q4, 2011]



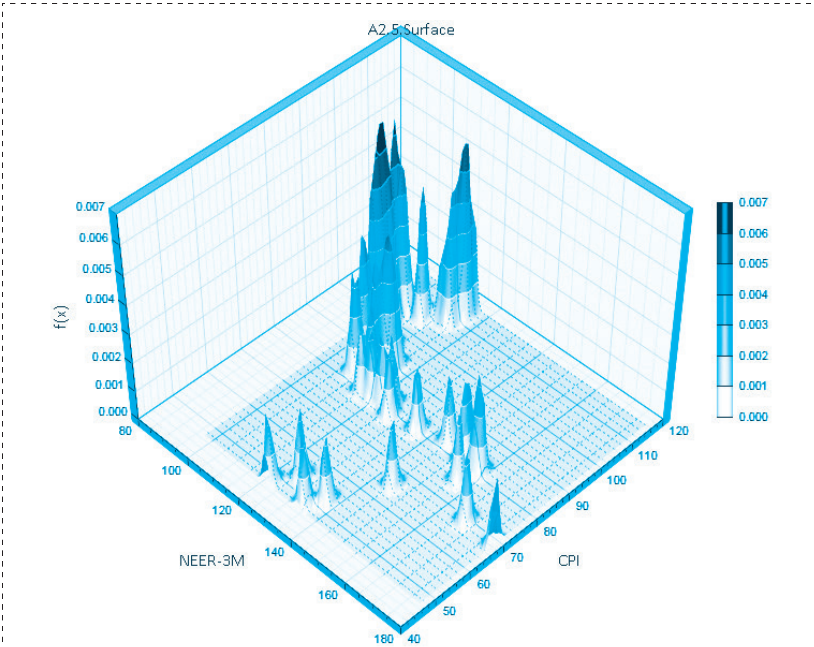


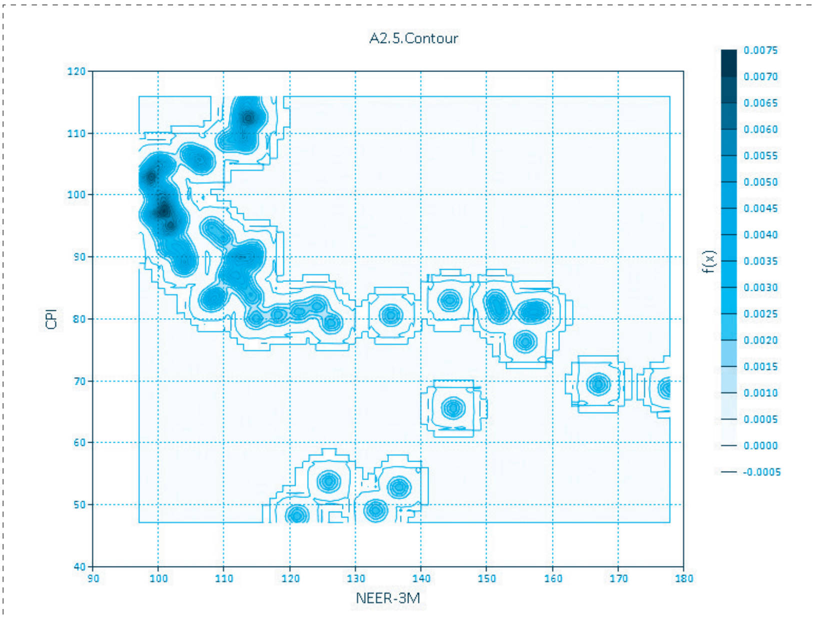
Graph A2.4.: Density estimates for the variables Ymld and CPI, Gaussian kernel, for the time period [Q1, 1997 – Q4, 2011]





Graph A2.5.: Density estimates for the variables NEER and CPI, Gaussian kernel, for the time period [Q1, 1997 – Q4, 2011]





DATA IN FIRST DIFFERENCE OF THE VARIABLES

Being aware of the fact that the relationship in levels could be driven by the presence of the time or other commonly shared stochastic trends, we continue the investigation by repeating the procedure with the differences and percentage changes.

The results for the differenced series are presented in charts A2.6. – A2.11. In principle, assuming the existence of a relationship, one would expect that similar changes in one variable are matched by similarly scaled changes in the other variable, portrayed by the presence of a single bell shaped graph or a regularly shaped set of bells streamlined along the diagonal or any other regularly shaped curve.

The graph that emerges is different from the one in the levels. In the case of money M2 and M3, density estimates[16] are represented by many irregular bells with modes concentrated mainly in the d(CPI) interval 0 and 5. Changes in money M2, M3 and CPI are not uniform; surprisingly, in both cases, it seems that the probability of having a larger change in CPI increases with the lower or negative changes in money almost indicating a negative relationship between changes in money and changes in inflation. This is a surprising result from the theoretic point of view.

The density estimates for the vector [d(CPI), d(R-12M-3-Dep)] yielded more or less a regular bell, which can be interpreted as a sign of a single linear slope for the relationship between both random variables. The mode of the joint distribution is, however, located close to the coordinates [3, 0] in the plane [d(RM3-bln), d(CPI)]. Both scalar values are close to the deserved monetary policy equilibrium, with a lot more variance in the interest rate than in d(CPI).

The same can be said with regard to the density estimates of the vector [d(CPI), d(YN)]. The chart reveals a single bell in the interval 0-5 for d(CPI) component. Less probable events are located along

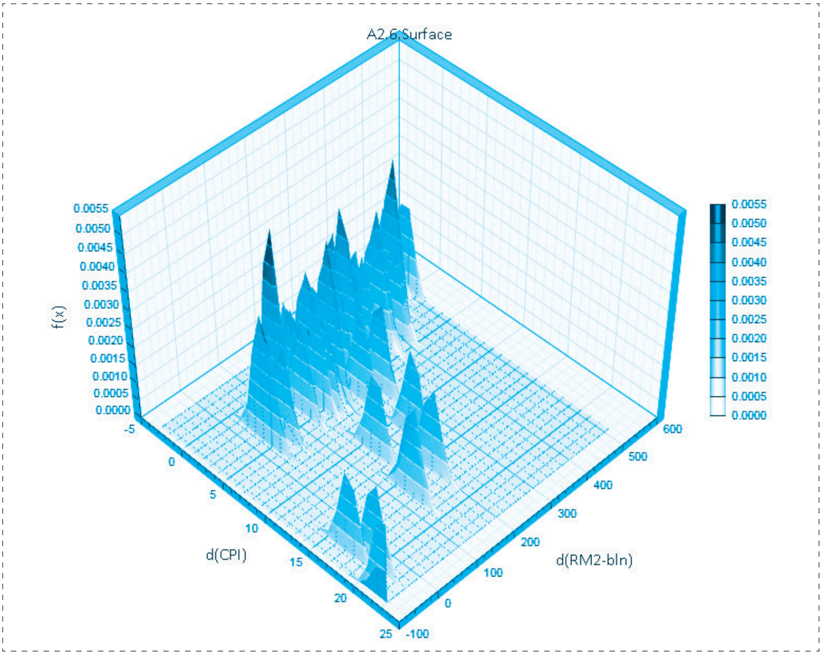
[16] In that part, we are referring to data on first difference, so we are not mentioning it specifically.

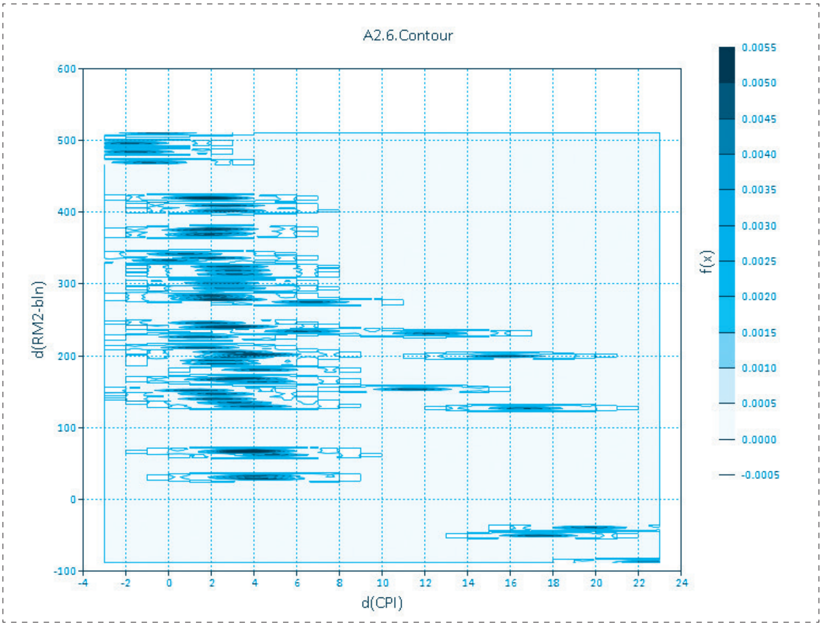
the northwest – northeast diagonal, providing evidence that joint events are observed along the diagonal, albeit with a very low probability. Most of the density estimates that the events are jointly distributed with the mode falling at around 10 and 2, for d(YN) and d(CPI), respectively.

In all cases above, it is interesting to observe that the bell shaped density estimates function tales a similar variance distribution along d(CPI) with the mode happening approximately at the middle of the interval. It is surprising to observe that density function along the d(CPI) dimension does not change depending on the paring random variable.

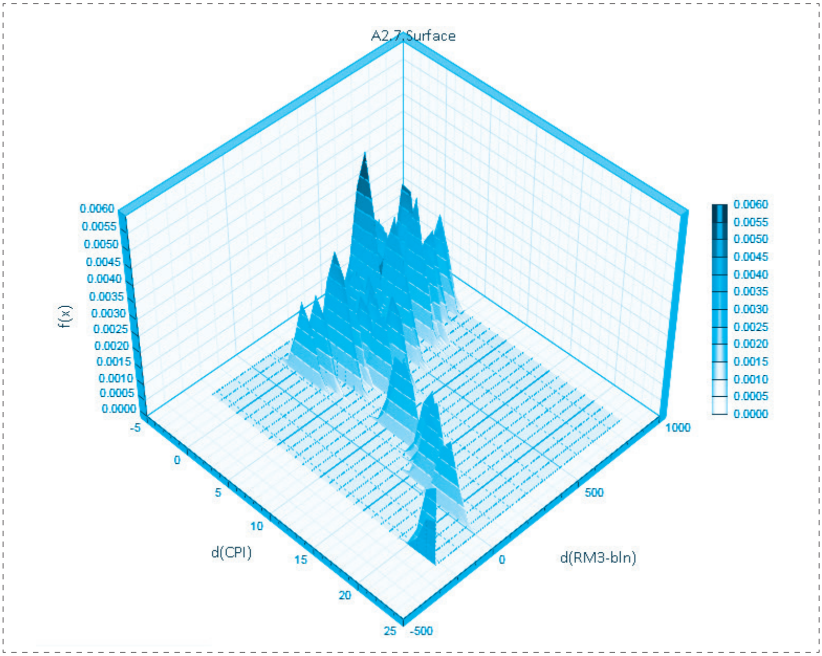
The graph is almost the same in the case of the d(CPI) with d(NEER-3M), with the only difference that the density estimation depicts more activity in the north-eastern corner of the graph A2.10. And the tail of the distribution of the main "bell" becomes fatter in the d(CPI) dimensions. Regardless of the shape of the density estimates function indicates independence between variables. Graph A2.11., which depicts the d(CPI) – d(NEER –12M) rate, presents a similar situation. However, there is a more pronounced evidence of densities located in the north-eastern corner. The presence of the multimodality in the Δ(NEER3m) happens along the 0 – 5 interval on the Δ(CPI), with all modes falling more or less in the middle of the segment 0 – 5. The overall graph would indicate independence; however, there is neglectable evidence that indicates a simultaneous increase in the volatility in Δ(CPI) and Δ(NEER3m).

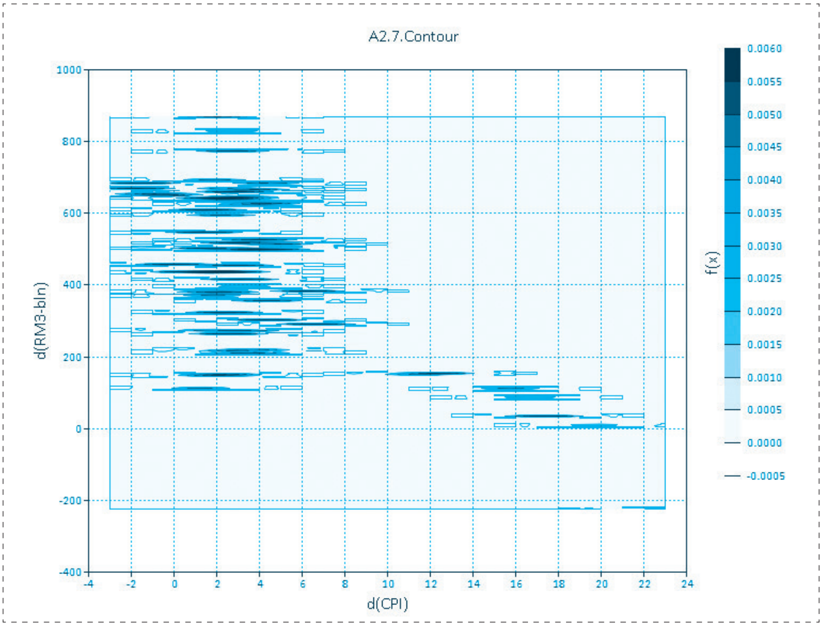
Graph A2.6.: Density estimates for the variables CPI and M2 (data on first difference), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011]



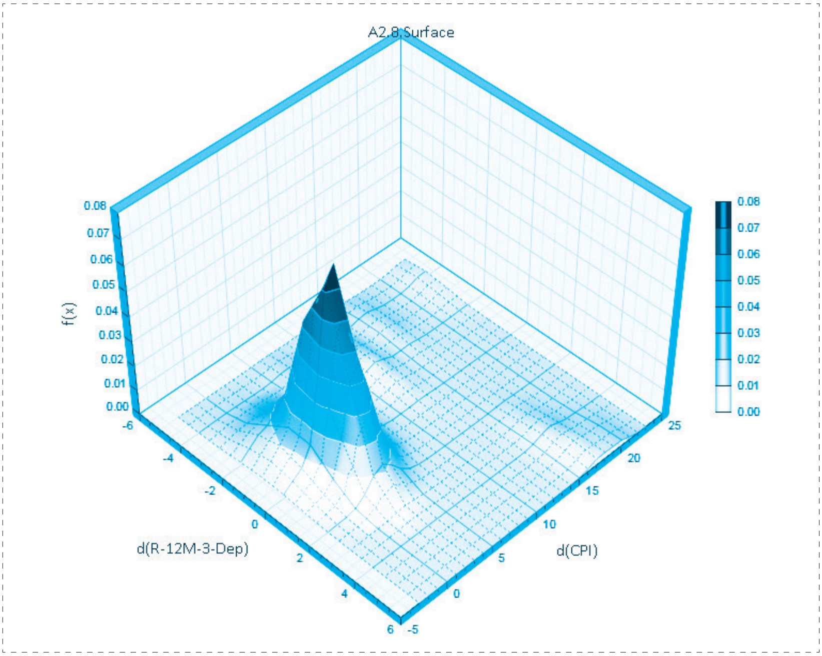


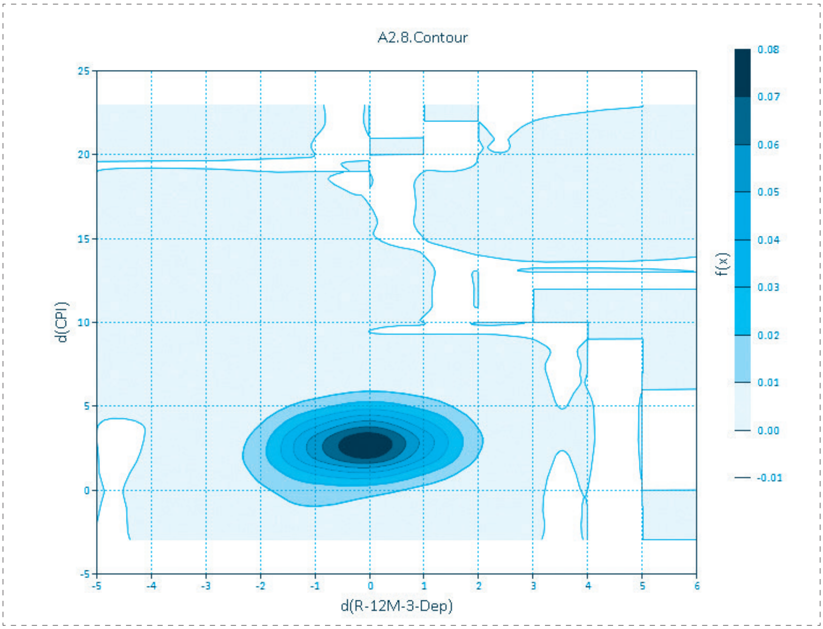
Graph A2.7.: Density estimates for the variables CPI and M3 (data on first difference), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011]



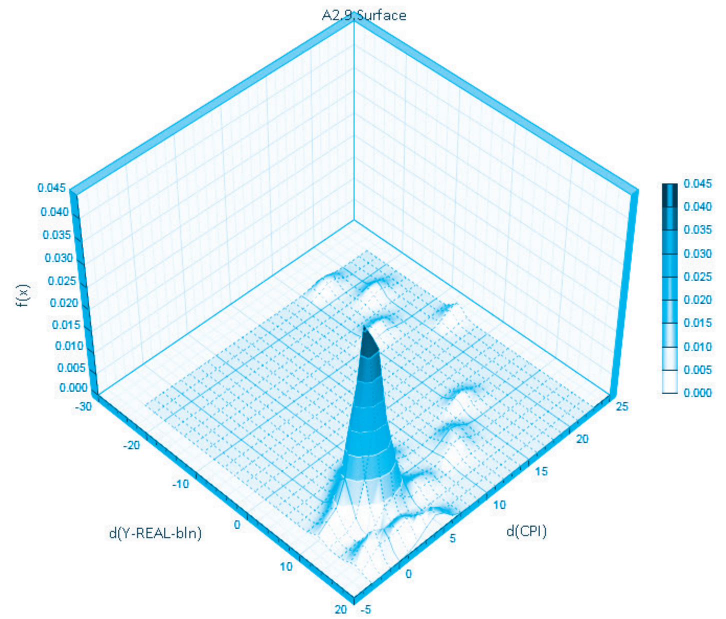


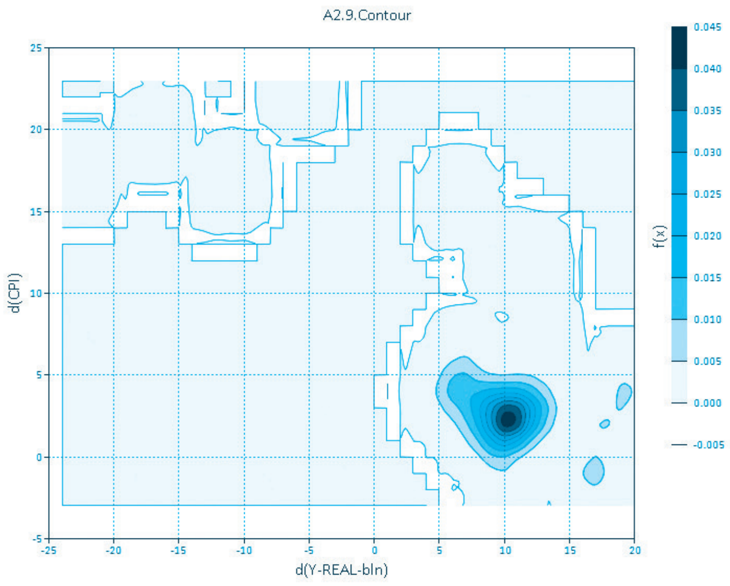
Graph A2.8.: Density estimates for the variables R and CPI (data on first difference), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011]



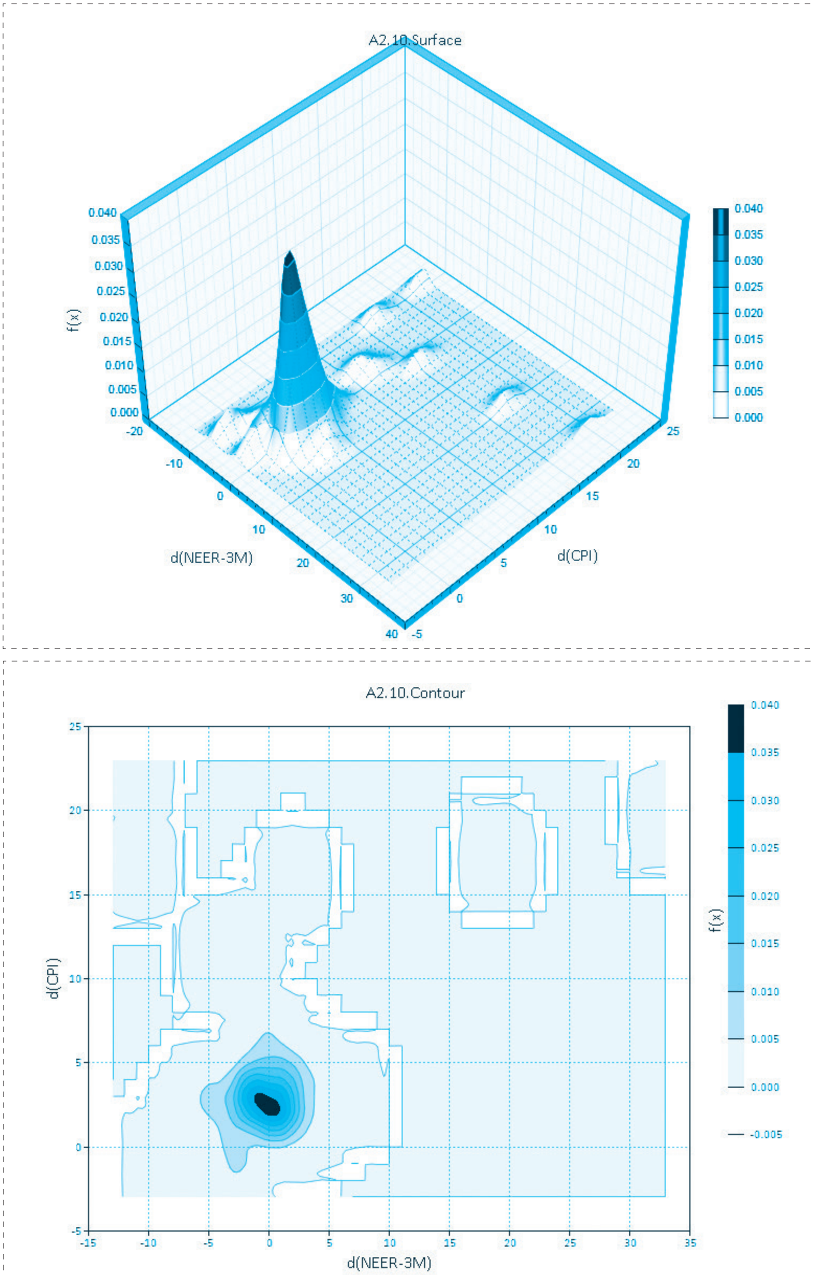


Graph A2.9.: Density estimates for the variables Y and CPI (data on first difference), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011]

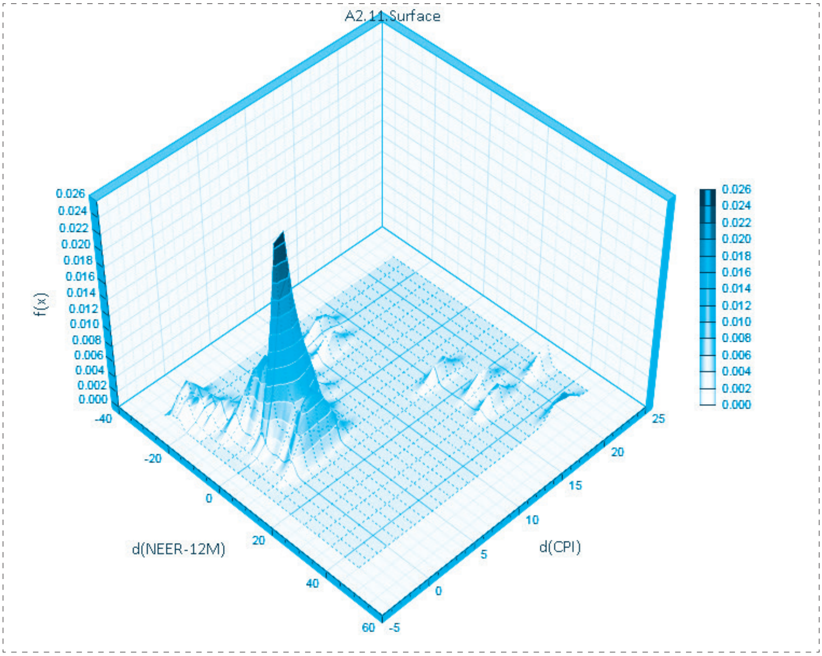


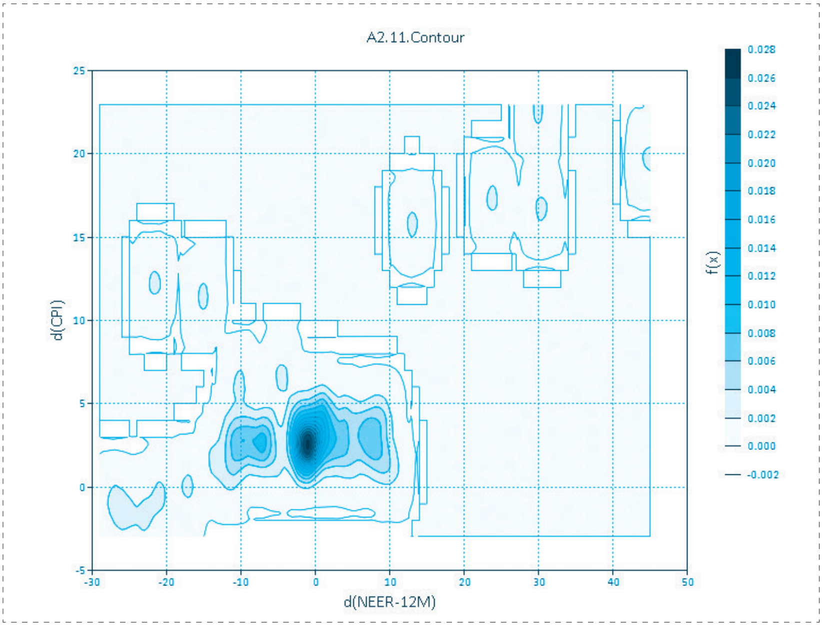


Graph A2.10.: Density estimates for the variables NEER 3m and CPI (data on first difference), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011]



Graph A2.11.: Density estimates for the variables NEER (annual) and CPI (data on first difference), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011]





DATA IN PERCENTAGE CHANGES

Finally, we repeat the same exercise on density estimation of the percentage changes of the variables. However, before starting the narrative, we would like to remind the reader that the percentage changes have transformed the CPI variable into inflation.

Assuming the existence of a relationship between inflation and our variables of interest, one would expect that percentage changes in one variable are matched by corresponding scaled percentage changes in the other variable, portrayed by the presence of a single bell shaped graph or a set of bells in the case of multimodal distributions.

The results of density estimation are shown in graphs A2.12. – A2.17. These graphs seem to depict a little more action in terms of densities.

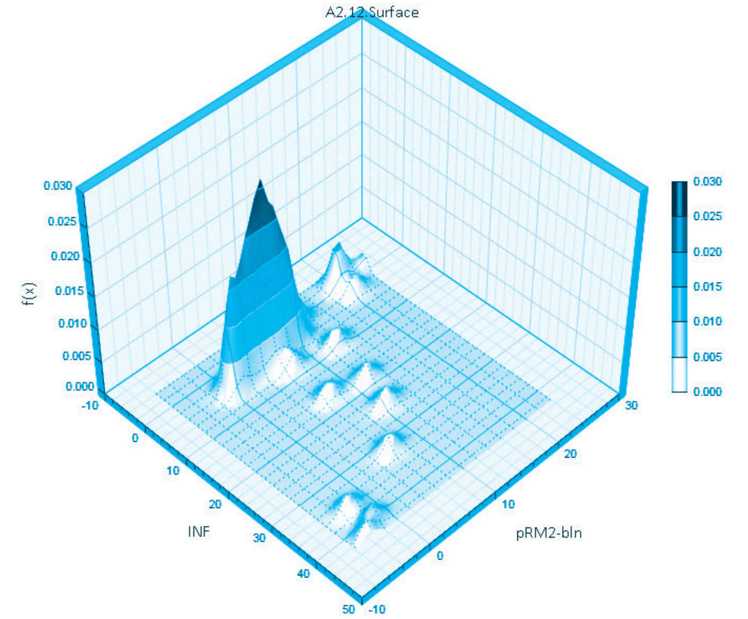
In both cases, M2 and M3 (in graphs A2.12. – A2.13.), density estimates are clearly depicted by a multimodal distribution along the money dimension, with dominant ones close to 3 percent of inflation. The variance of the money growth does not seem to affect the variance of inflation – the traces of the main bell that results from the intersection of itself with the plans orthogonal to horizontal plan, along money dimension, yields different curves with similar variance.

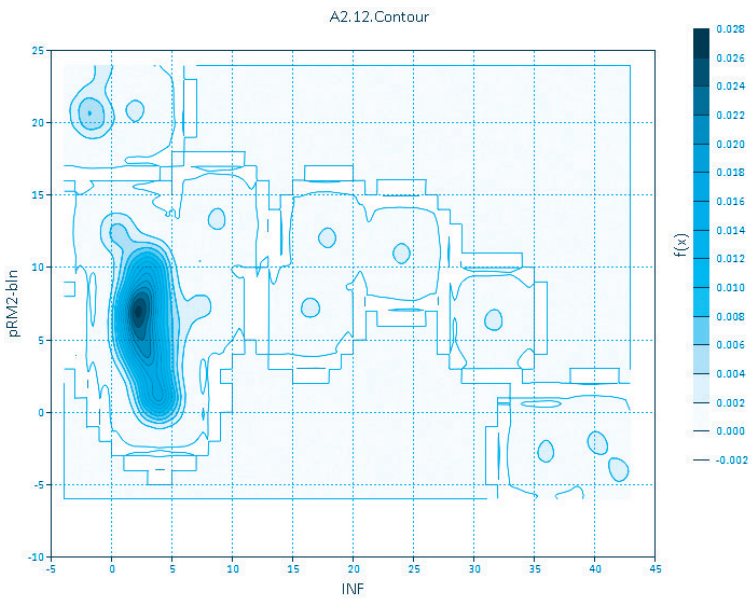
Despite this dominant characteristic in both cases (M2 and M3), there is a clear density pattern evolving along the diagonal indicating again a negative relationship between money and inflation.

Patterns are less pronounced in the INF – R and INF – Y (respectively in graphs A2.14. – A2.15.), to re-emerge in the INF – NEER relationship (in the graphs A2.16. – A2.17.). The latter depict increase in the probability of observing densities in the southwest-northeast diagonal.

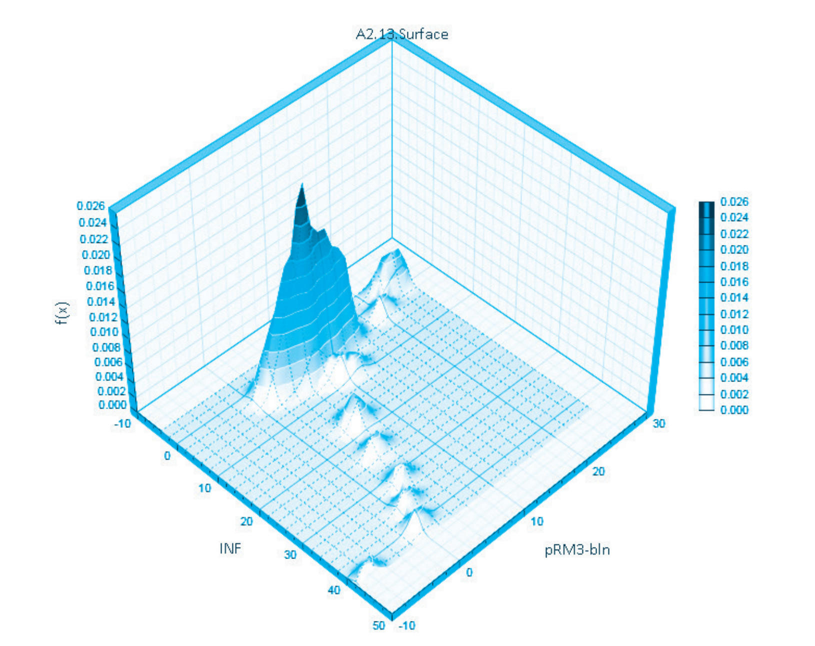
Overall, it is striking to observe that the variance along the ΔCPI and INF dimension stays within the same interval and is not clearly affected by any of the other random variables of interest. While we do not report estimates of expected values of inflation (this will be the focus of future research), the charts seem to imply that it is very likely (with a high probability) that CPI and inflation are not affected by simultaneous (within the same quarter) changes in money, economic activity, interest rate and exchange rate. Price developments seem to follow more or less their own trend, yielding inflation very close to the objective of the Bank of Albania, 3 ± 1 %, potentially interpreted as a sign of public's credibility in the Bank of Albania and its monetary policy.

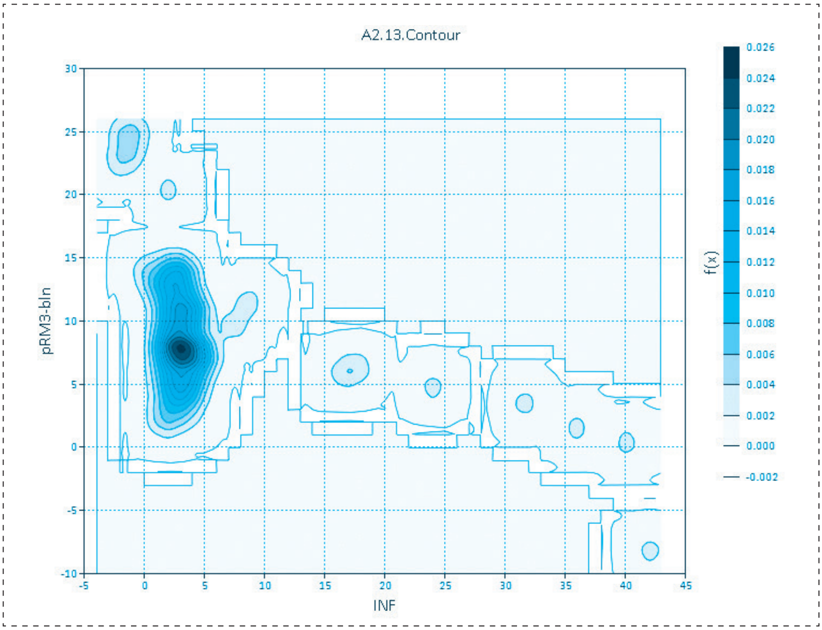
Graph A2.12.: Density estimates for the variables INF and M2 (data on percentage), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011]



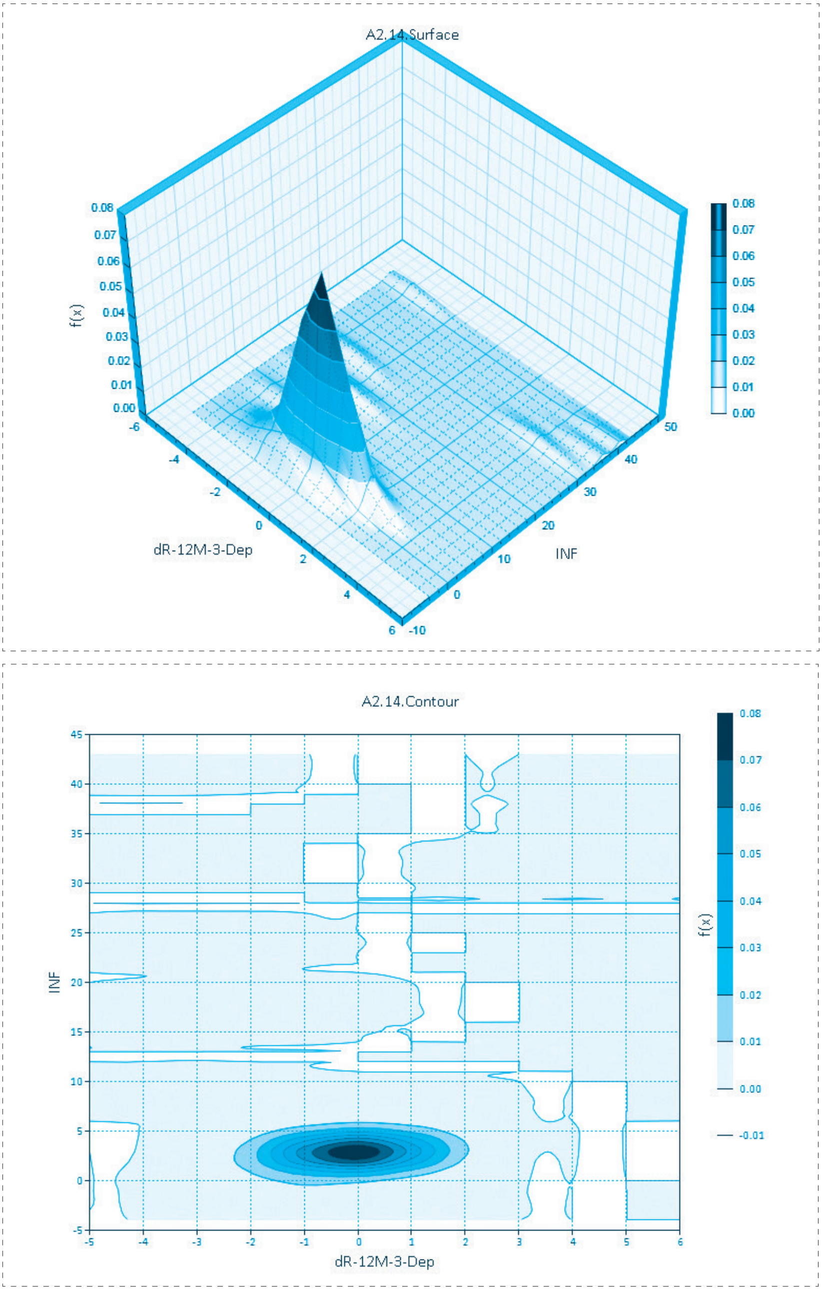


Graph A2.13.: Density estimates for the variables INF and M3 (data on percentage), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011]

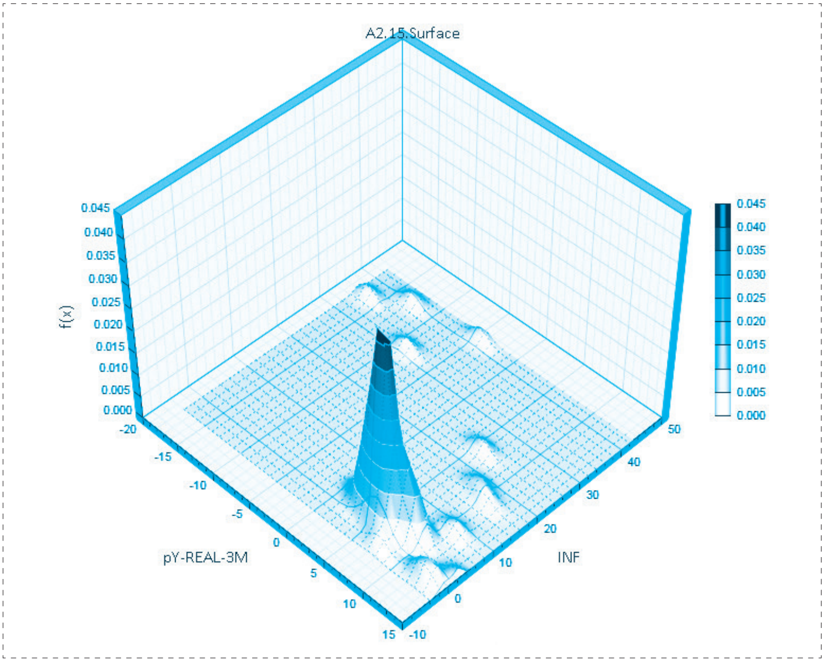


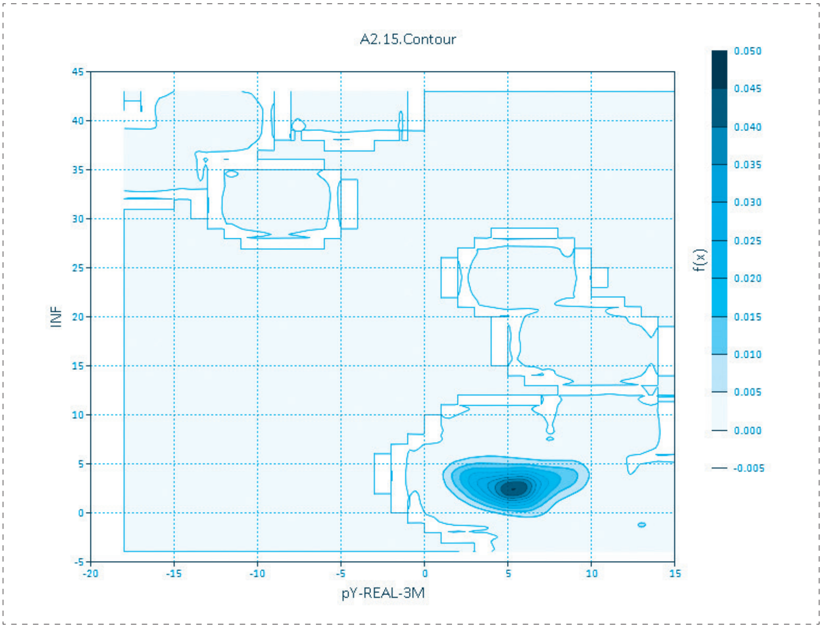


Graph A2.14.: Density estimates for the variables R and INF (data on percentage), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011]

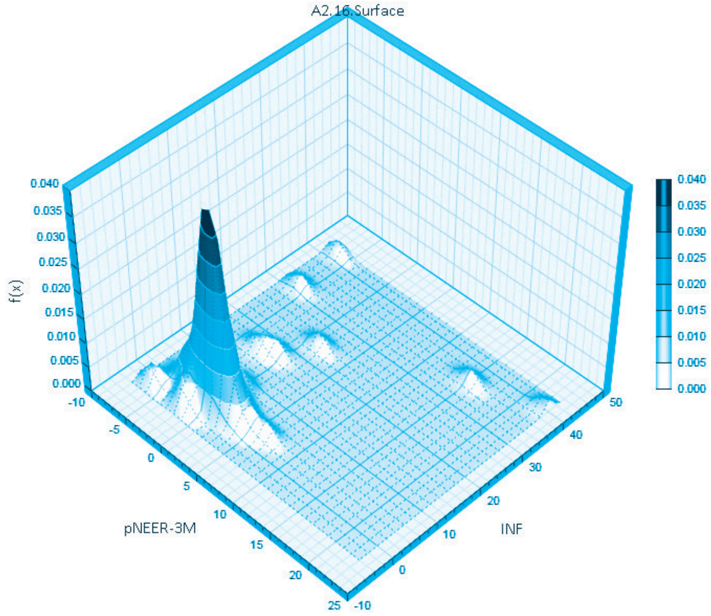


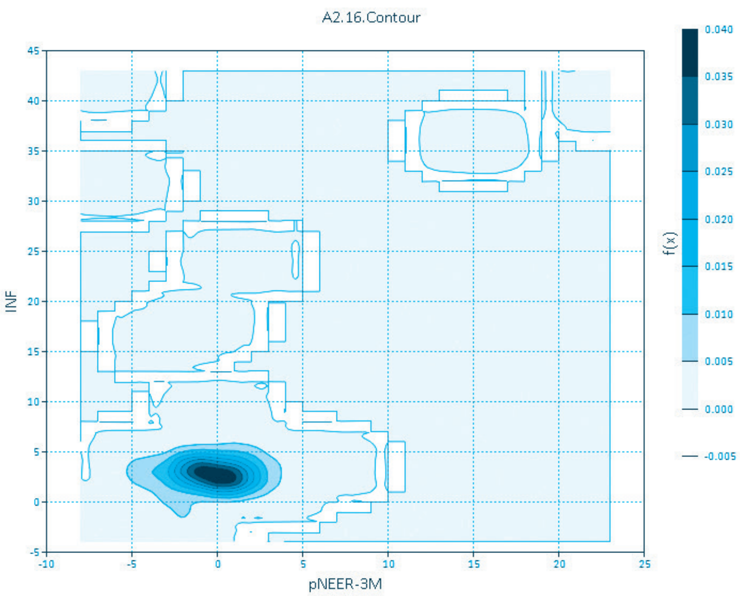
Graph A2.15.: Density estimates for the variables Y and INF (data on percentage), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011]



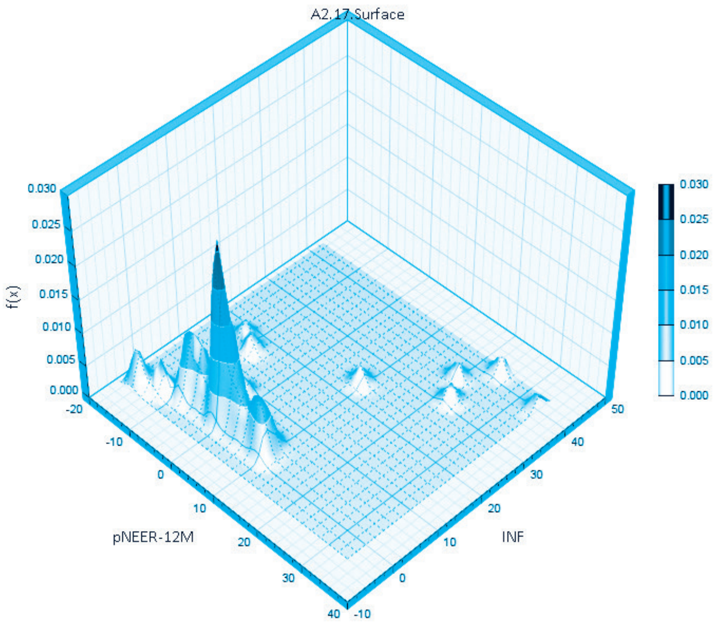


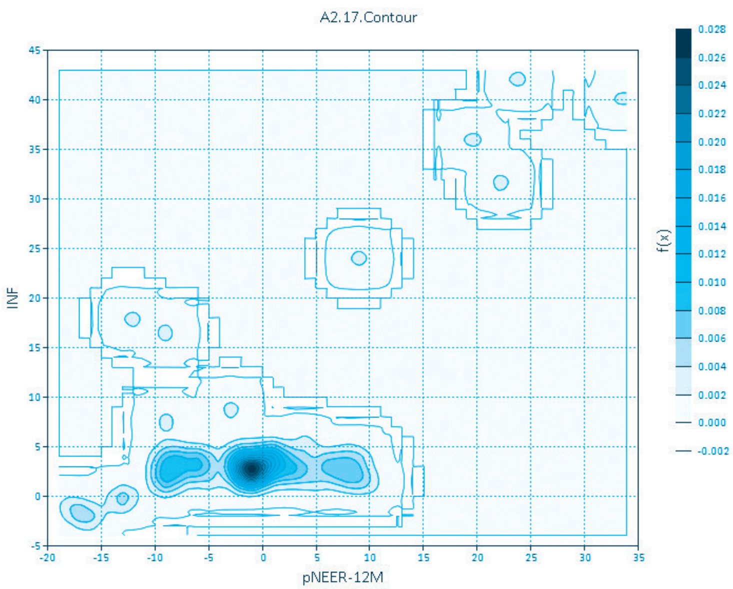
Graph A2.16.: Density estimates for the variables NEER 3m and INF (data on percentage), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011].





A2.17.: Density estimates for the variables NEER (annual) and INF (data on percentage), Gaussian kernel, for the time period [Q1, 1997-Q4, 2011]





ANNEX 3

Algorithms given in this chapter are a direct output of the formulas posed in the fourth chapter. They are constructed for one and multidimensional densities. In each of them, the functions of the densities are constructed and additional technical explanations are given on the graphical design techniques.

In the first block of algorithms: algorithm 1 – algorithm 4, we construct density estimates on a given sample $x_1, x_2, ..., x_n$ using a one dimensional kernel as they are given in 4.1.1. – 4.1.4. To see the graphical map of each density, plot a 2 dimensional graph $(t, \hat{f}(t)), \forall t \in R$. To judge on the variations of the density relating with the smoothing parameter h, repeat the algorithm for different values of h – also, optimized values of h have to be considered.

Algorithm 1: Density estimates and graphic map of the sample $x_1, x_2, ..., x_n$ using a normal kernel

Step 1: Fix a parameter $h > 0$.

Step 2: $\forall t \in R$, calculate $\hat{f}(t) = \frac{1}{nh} \cdot \frac{1}{\sqrt{2\pi}} \cdot \sum_{i=1}^{n} \exp\left[\frac{-(t - x_i)^2}{2h^2}\right]$.

Step 3: End.

Algorithm 2: Density estimates and graphic map of the sample $x_1, x_2, ..., x_n$ using a triangle kernel

Step 1: Fix a parameter $h > 0$.

Step 2: $a_i = \begin{cases} (h-|t - x_i|), & \text{for } (h-|t - x_i|) > 0 \\ 0, & \text{for } (h-|t - x_i|) \leq 0' \end{cases} \quad t \in R.$

Step 3: Calculate $\hat{f}(t) = \frac{1}{nh^2} \sum_{i=1}^{n} a_i$,

Step 4: End.

Algorithm 3: Density estimates and graphic map of the sample $x_1, x_2, \ldots, x_n$ using a rectangular kernel

Step 1: Fix a parameter $h > 0$.

Step 2: $a_i = \begin{cases} 1, & \text{for } (h-|t - x_i|) > 0 \\ 0, & \text{for } (h-|t - x_i|) \leq 0' \end{cases} \qquad t \in R.$

Step 3:Calculate $\hat{f}(t) = \frac{1}{2nh} \sum_{i=1}^{n} a_i$.

Step 4: End.

Algorithm 4: Density estimates and graphic map of the sample $x_1, x_2, \ldots, x_n$ using a Epanechnikov kernel

Step 1: Fix a parameter $h > 0$.

Step 2: $\forall t \in R$, calculate:

if $t \in R$, $\frac{|t - x_i|}{h} < \sqrt{5}$, then $\hat{f}(t) = \frac{3}{4\sqrt{5}} \cdot \frac{1}{nh} \sum_{i=1}^{n} \left[1 - \frac{1}{5}\left(\frac{t - x_i}{h}\right)^2\right]$.

Step 3: End.

In the second block of algorithms: algorithm 5 – algorithm 6, we construct density estimates on a given d dimensional sample of size n: $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ using a d dimensional kernel as they are given in 4.2.3. – 4.2.4. In this general case the points of the density are in d + 1 dimensional space: $(\underline{t}, \hat{f}(t)) \in R^{d+1}, \forall \underline{t} \in R^d$.

While the following algorithms are posed to construct the values of densities in d dimensional samples, the graphical representation can be done only if $d = 2$, means: $(\underline{t}, \hat{f}(t)) \in R^3, \forall \underline{t} \in R^2$.

Like in the one dimensional case, to judge on the variations of the density relating with the smoothing parameter h, repeat the algorithm for different values of h.

Algorithm 5: Density estimates of the d dimensional sample $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ using a normal kernel

Step 1: Fix a parameter $h > 0$.

Step 2: $\forall \underline{t} \in R^d$, calculate:

$$\hat{f}(t) = \frac{1}{nh^d\ (2\pi)^{d/2}} \cdot \sum_{i=1}^{n} \exp\left\{-\frac{1}{2h^2}\left[(t_1 - x_i^1)^2 + (t_2 - x_i^2)^2 + \cdots + (t_d - x_i^d)^2\right]\right\}.$$

Step 3: End.

Algorithm 6: Density estimates of the d dimensional sample $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ using Epanechnikov kernel

Step 1: Fix a parameter $h > 0$.

Step 2: Calculate: $C_d = \begin{cases} \frac{\pi^{d/2}}{\left(\frac{d}{2}\right)!}, & \text{d-even} \\ \frac{2^{\frac{d+1}{2}} \cdot \pi^{\frac{d-1}{2}}}{d!!}, & \text{d-odd} \end{cases}$

Step 3: $\forall \underline{t} \in R^d$, calculate:

$$a_i = \begin{cases} \left\{1 - \frac{1}{h^2}\left[(t_1 - x_i^1)^2 + (t_2 - x_i^2)^2 + \cdots + (t_d - x_i^d)^2\right]\right\}, & \text{if } \{\ldots\} > 0 \\ 0 & \text{, if } \{\ldots\} > 0' \end{cases}$$

Step 4: Calculate: $\hat{f}(t) = \frac{1}{nh^d} \frac{d+2}{2 \cdot C_d} \cdot \sum_{i=1}^{n} a_i$.

Step 5: End

CIP Katalogimi në botim BK Tiranë

Altin Tanku, Kliti Ceca
Density estimation for economic variables
– a genuine application - /
/ Tanku Altin, Ceca Kliti - Tiranë:
Banka e Shqipërisë, 2013

-76 f; 15.3 x 23 cm.

Bibliogr.
ISBN: 978-99956-42-92-6.

You may find this paper in the following address:

www.bankofalbania.org

If you want to receive a hard copy of this paper, write to us at:

Bank of Albania
Sheshi "Avni Rustemi", Nr. 24, Tiranë, Shqipëri
Tel.: + 355 4 2419301/2/3; + 355 4 2419409/10/11
Fax: + 355 4 2419408

or send an e-mail to:

public@bankofalbania.org

Printed in 380 copies