Iraqi Statisticians Journal

https://isj.edu.iq/index.php/isj

ISSN: 3007-1658 (Online)

A Monte Carlo Simulation Study on the Performance of Linear and Quadratic Programming Estimators in Fuzzy ARDL Models

Sulaiman Hussien Jawad[1] and Munaf Yousif Hmood[2]

[1,2]Department of Statistics, College of Administration and Economics, University of Baghdad Iraq, Baghdad.

ARTICLE INFO

Article history:
Received 10 March 2026
Revised 11 March 2026
Accepted 23 March 2026
Available online 24 March 2026

Keywords:
Fuzzy ARDL,
Linear Programming,
Quadratic Programming,
Monte Carlo Simulation,
Estimation Performance

ABSTRACT

This study analyzes the relative effectiveness of the Linear Programming (LP) and Quadratic Programming (QP) techniques to estimate the Fuzzy Autoregressive Distributed Lag (FARDL) model, which can be regarded as the generalization of the classical ARDL model and is used to estimate fuzzy dynamic relationships among variables. In this model parameters are represented as fuzzy numbers defined in terms of central value and spread, hence allowing the direct adoption of structural uncertainty as being a part of the model architecture instead of using a conventional stochastic error. A Monte Carlo simulation experiment was implemented to meet the objectives of the study with different sample size and the complexity of the lag structure with the objective of testing the performances of the two estimation methods in different data settings. The comparative evaluation was based on two performance measures: the Root Mean Square Error (RMSE) which is an indicator of the accuracy of the estimation of central parameters and the Fuzzy Degree (FD) which is an indicator that quantifies the amount of uncertainty in the estimates. The results of the simulations show that QP method is always better than the LP method in all the sample sizes and lag settings with lower values in RMSE and FD respectively in most cases. This advantage is very strong especially in small samples and it persists as the sample size increases.

1. Introduction

The Autoregressive Distributed Lag (ARDL) is considered to be one of the most commonly used standard frameworks of studying dynamic interrelationships between the variables in time-series research. It allows investigation of both the short-run and long-run interactions to be considered in a common structure, despite the constituent variables representing heterogeneous orders of integration. The model has received a lot of academic interest in a wide range of empirical research because of its abilities to allow various lag settings, as well as its ability to admit variables possessing varying integration characteristics, specifically in the areas of macroeconomic and financial studies [1] [2].

Corresponding author E-mail: suliaman.hussien83@gmail.com
https://doi.org/10.62933/wbf13z44
This work is an open-access article distributed under a CC BY License (Creative Commons Attribution 4.0 International) under https://creativecommons.org/licenses/by-nc-sa/4.0/

The recent econometric literature has seen the significant increase in the application of fuzzy models as an alternative analytical framework to the traditional probabilistic models, especially in those cases when the structural uncertainty is involved that is difficult to bear with the help of probability distribution assumptions only [3]. Early papers in fuzzy regression, some of which are the book by Tanaka et al., (1982) have helped to develop this method by modeling the parameters in terms of fuzzy numbers that reflect levels of vagueness and imprecision [4].

In this context, the Fuzzy Autoregressive Distributed Lag (FARDL) model has emerged as an extension of the conventional ARDL model. It integrates fuzzy logic theory into the model structure by representing its parameters as fuzzy numbers consisting of a central value and a spread, thereby allowing uncertainty to be implicitly incorporated within the parameter structure rather than being confined to the random error term. Consequently, it provides a more flexible representation of dynamic relationships in general, and economic relationships in particular [5] [6].

Despite the notable development in the fuzzy regression literature and its widespread application in modeling relationships among variables, most empirical studies have primarily focused on static fuzzy regression models rather than dynamic time series frameworks. Consequently, this methodology has not been adequately integrated into dynamic time series modeling, particularly within the framework of the Fuzzy Autoregressive Distributed Lag (FARDL) model [7].

Moreover, studies on the use of mathematical programming models, specifically, the Linear Programming (LP) and the Quadratic Programming (QP), to estimate fuzzy parameters have tended to study these approaches either alone or in an applied perspective. As such, a study that takes a systematic comparative evaluation of their performance in different data conditions is still lacking. Consequently, the comparative advantages between the LP and QP are not decisively defined in the spheres of the econometric corpus especially taking into consideration the variability of sample sizes, the complexity of the model structure and the subtle complexities that are involved in the lag arrangements[8].

This issue arises from the nature of fuzzy estimation itself, as it is not limited to estimating the central values of the parameters but also extends to estimating their degree of spread, which adds an additional dimension to the comparison between estimation methods. Linear Programming seeks to minimize the total spread of the fuzzy parameters under data containment constraints, whereas Quadratic Programming allows the quadratic structure of the objective function to be taken into account. This may be reflected in the accuracy and stability of the estimates, particularly in small samples or in environments characterized by high levels of uncertainty[9].

This problem is increased in the context of economic studies in developing countries, where data often have a small sample size or not all observations are available. As a result, its performance of estimators in small sample is both a methodological and practical imperative. Although a few studies have conducted comparative studies on the use of fuzzy estimates strategies [10], the issue of sensitivity of these strategies to small sample conditioning in the context of dynamic model representation is limited in the existing literature.

On this basis, the current research paper will attempt to use a Monte Carlo simulation to compare the relative performance of Linear Programming and Quadratic Programming in estimating the FARDL model under different conditions with respect to the size of the sample and the complexity of the lag structure. Comparison is drawn between the Root Mean Square Error (RMSE) objective measure of determining the accuracy of the central values of the parameters, as well as the Fuzzy Distance (FD) measure, which is used to quantify the degree of uncertainty regarding the dispersion of the parameters. The research also focuses on investigating the performance of estimators in the small-sample setting hence presenting the data constraints that are often experienced in most developing economies.

Accordingly, this study contributes to filling a methodological gap in the fuzzy econometrics literature by providing a systematic and comparative evaluation of parameter estimation methods for the FARDL model using a simulation approach. It focuses on the effects of sample size and model complexity, thereby enhancing the reliability of fuzzy model results and deepening the understanding of their behavioral properties under diverse data environments.

2. Theoretical Background of Fuzzy Modeling

2.1 Fuzzy Logic

Fuzzy logic, firstly developed by Zadeh (1965) [11], provides a strict mathematical system of defining non-probabilistic uncertainty by giving variables the degree of membership to objects in the closed unit interval [0, 1], thus avoiding the traditional binary categorization based on full-membership or total non-membership. The construct allows the representation of a situation that could not be accurately delineated using deterministic numerical values, through the characterization of variables or model parameters as fuzzy sets that represent a range of allowable values, instead of a point estimate. This concept constitutes the theoretical foundation upon which fuzzy regression models are based, where uncertainty is incorporated directly into the structure of the model rather than being confined to a statistical error term[12].

2.2 Fuzzy Regression Analysis

The concept of fuzzy regression was first introduced in 1982 by Tanaka et al. [3] as an extension of traditional regression models to address situations in which the relationship between variables is characterized by a degree of vagueness or structural instability [13]. Fuzzy regression is also considered an appropriate alternative in cases where some of the assumptions or properties of conventional models are not satisfied, or when sufficient data are not available to construct a reliable statistical model [14].

Unlike probabilistic regression models which only allow uncertainty to be constrained to random disturbances, as represented by an error term, fuzzy regression allows the precise modeling of structural imprecision in the model parameters. In this case, the coefficients are estimated to be fuzzy numbers and thus representing a range of possible values instead of a point estimate [6]. Such a methodological framework is especially appropriate in situations where stability of structural relationships is hard to postulate or where structural relationships themselves can be well described using conventional probabilistic models, particularly when dealing with a situation that is of vagueness or structural uncertainty. There are several formulas in making the fuzzy regression models depending on the subject of the variables and parameters to be used [15]. The particular case of crisp variables and fuzzy coefficients is taken over to the current study.

3. Fuzzy Autoregressive Distributed Lag Model – FARDL

The steps for constructing the FARDL model are represented in the following stages:

3.1 Model Specification

The functional specification of the FARDL model could be expressed as a fuzzy linear expression integrated into a linear system associated with one dependent variable(Y_t), and one explanatory variable(X_t). For simplicity of exposition and methodological analysis, the present study considers the case of a single explanatory variable. Both variables are crisp, and their respective lagged values are also crisp and precisely determined. The model also includes a constant term and a deterministic trend component. In the case of a sample space with (T) observations, the model is as follows.

$$Y_t = \alpha_0 + \psi t + \sum_{j=1}^{p} \Phi_j Y_{t-j} + \sum_{i=0}^{q} \theta_i X_{t-i}$$

$$\rightarrow Y_t = \alpha_0 + \psi t + \Phi_1 Y_{t-1} + \ldots + \Phi_p Y_{t-p} + \theta_0 X_t + \theta_1 X_{t-1} + \ldots + \theta_q X_{t-q} = AZ$$

$$, t = 1, 2, \ldots. T \qquad (1)$$

$$, Y, X \text{ Integrated } I(0) \, or \, I(1)$$

$$j = 1, 2, \ldots\ldots, p \qquad , i = 0, 1, \ldots\ldots, q$$

Where:

(j) : represents the lagged values of the dependent variable and (i) represents the lagged values of the independent variable.

$Z = [1 \quad t \quad Y_{t-j} \quad X_{t-i}\,]^t$: represents the system inputs and is a non-fuzzy (crisp) vector.

$\tilde{A} = [\widetilde{\alpha_0} \quad \tilde{\psi} \quad \widetilde{\phi_j} \quad \widetilde{\theta_i}\,]^t$: represents the system parameter vector, which is expressed in the form of symmetric triangular fuzzy numbers.

3.2 Estimation of the FARDL Model Parameters

3.2.1 Estimation Using Quadratic Programming

The fuzzy parameters estimation is carried out by applying a quadratic programming paradigm as described by Tanaka and Lee [16]. The given methodology is based on mathematical optimization strategies that are aimed at defining a range of parameters which best fit observed data and its estimations in a fuzzy analytical model. What can be of interest, is the ability of the quadratic programming method to combine the property of central tendency of least-squares estimation with the feasibility property of fuzzy regression to strike a balance between the accuracy of the estimation and the level of fuzzy uncertainty of the model.

As a result, the quadratic programming problem can be expressed with symmetrical triangular fuzzy coefficients. It will have an objective that will reduce cumulative spread of the model, and therefore it should be formulated as follows[8]:

The quadratic *programing model is*

$$\min J = \left(c_{\alpha_0}\left|X_0\right|\right)^2 + \left(c_{\psi}\left|t\right|\right)^2 + \sum_{j=1}^{p}\left(c_{\Phi_j}\left|Y_{t-j}\right|\right)^2 + \sum_{k=1}^{K}\sum_{i=0}^{q_k}\left(c_{k\theta_{ki}}\left|X_{k,t-i}\right|\right)^2 + \xi\alpha'\alpha$$

$$\rightarrow \min J = c'\left|Z\right|\left|Z\right|' c + \xi\alpha'\alpha$$

Subject to (2)

$$Y_t \geq \alpha' Z_t - (1-h)c'\left|Z_t\right|$$

$$Y_t \leq \alpha' Z_t + (1-h)c'\left|Z_t\right|$$

$$Z_t = \left[\,1, t, Y_{t-1}, \ldots. Y_{t-p}, X_t, X_{t-1}, \ldots. X_{t-q}\right]'$$

$$, c' \geq 0$$

Here (ξ) represents a positive number with a very small value used to transform the objective function into a quadratic form. The objective function in equation (2) contributes to generating a spread vector that is more flexible and less sharp for the model coefficients. In essence, it seeks to reduce the degrees of spread of the fuzzy parameters, which leads to lowering the overall ambiguity measure of the model. At the same time, it is required that the membership degree of each observation (Y_t) be greater than or equal to the value of the$(\alpha - cut)$, such that:

$$\mu_{Y*}(Y_t) \geq h \qquad , h = (\alpha - cut) \qquad for\ all\ t = 1, 2, \ldots. T \tag{3}$$

This criterion simply expresses the fact that the fuzzy output (response variable) of the model must cover all observed values with a specified degree of the$(\alpha - cut)$. This requirement is referred to as the threshold condition, and the inequality representing the threshold condition can be reformulated as follows:

$$(1-h)c'\left|Z\right| - \left|Y - Z'\alpha\right| \geq 0 \quad , Z \neq 0 \tag{4}$$

3.2.2 Estimation Using Linear Programming-LP

To estimate the parameters of the FARDL model using the Linear Programming method, the estimation problem is transformed into a Linear Programming (LP) problem under an objective function that aims to minimize the total spread of the fuzzy parameters in the model, as follows[17]:

The linear programing model is

$$\min J = c_{\alpha_0}\left|X_0\right| + c_{\psi}\left|t\right| + \sum_{j=1}^{p} c_{\Phi_j}\left|Y_{t-j}\right| + \sum_{k=1}^{K}\sum_{i=0}^{q_k} c_{k\theta_{ki}}\left|X_{k,t-i}\right|$$

$$\rightarrow \min J = c'\left|Z\right|$$

$$\text{Subject to} \qquad (5)$$

$$Y_t \geq \alpha' Z_t - (1-h)\,c'\left|Z_t\right|$$

$$Y_t \leq \alpha' Z_t + (1-h)\,c'\left|Z_t\right|$$

$$Z_t = \left[\, 1, t, Y_{t-1}, \ldots . Y_{t-p}, X_t, X_{t-1}, \ldots . X_{t-q} \right]'$$

$$, c' \geq 0$$

It is necessary to mention that the reason behind the use of triangular functions can be associated with the simplicity of their computations and the lack of information concerning the linguistic variable (the model parameters). When a specific linguistic term needs to be encoded as a part of a fuzzy logic system, the most basic acceptable form is that which attaches a prototypical value (complete membership) with an upper and lower spread limits. Moreover, the balance in the entropy, a measure of the level of ambiguity of a fuzzy set, is obtained by the application of triangular functions during the modelling of linguistic variables [18].

The membership function of the fuzzy parameters ($\tilde{A}$) is given by the following expression [19]:

$$\mu_{A_i}(a_i) = \begin{cases} 1 - \dfrac{\alpha_i - a_i}{c_i} & , \alpha_i - c_i < a_i < \alpha_i + c_i \\ 0, & \text{other wise} \end{cases} \qquad (6)$$

$$i = 1, 2, \ldots, k$$

$$c_i > 0$$

Where:

a_i : It represents the fuzzy parameter, as illustrated in Figure 1.

c_i : represents the spread of the fuzzy parameter.

α_i: It is the center of the fuzzy set.

k: represents the number of model parameters.

The effectiveness of the model presented can be measured using the Root Mean Square Error (RMSE), which is an indicator of the difference between the actual values and the computed ones based on the central parameters of the fuzzy model. Moreover, single measure of the extent of fuzziness of the model is also measured. What is necessary here is to make a distinction that the error term here is used to refer to the difference between the observed data and the estimates made by the fuzzy model, not the probabilistic estimation errors. The comparative criteria can be obtained with the help of the following expressions [20][4] .

1- Root Mean Square Error-RMSE:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}\left[Y_t - Y_t^c\right]^2}{T}} \quad (7)$$

2- The degree of the fuzziness:

$$F(M^h) = \sqrt{(c_{\alpha_0}{}^h)^2 + \quad (c_{\psi}{}^h)^2 + \quad (c_{\Phi_j}{}^h)^2 + \quad (c_{\theta_i}{}^h)^2} \quad (8)$$

$$, j = 1,2,\ldots p \ and \ \ i = 0,1,\ldots q$$

4. Experimental results

4.1 Data Generating Process

The process of data generation in the simulation model was worked out on the basis of the real quarterly data of the Iraqi economy on the fiscal surplus/deficit balance as the dependent variable and the public revenues as the independent variable within the period of 2013Q1 to 2023Q4. These data have been chosen because of the fact that macroeconomic variables have high level of fuzziness and uncertainty, which varies with time. The empirical model estimates were used to build the Data Generating Process (DGP) so the series generated accurately capture the dynamic properties, both in the lag structure and time dependence between the variables, whilst the distributional assumptions were simplified on purpose to aid the process of simulation.

4.2 Design and Implementation of the Simulation Program:

To build a Monte Carlo simulation that was supposed to recreate the FARDL model and test the effectiveness of Linear Programming and Quadratic Programming models in estimating model parameters, the programming language R (version 4.4.2) was utilized. The simulation architecture would be divided into four main steps, each step would be dedicated to the model estimation and further analysis of the results, as explained below:

Stage One: Model Specification and Construction, which consists of the following:

1- Model Identification:

The models were defined in the Fuzzy Autoregressive Distributed Lag (FARDL) framework in the empirical component in which a group of specific structures were developed, which vary in the lag orders of the dependent variable (p) and independent variable(q). Also the deterministic elements were changed and this was the constant term and the general trend. This variety of structures is aimed to generalize a variety of dynamic situations that reflect the behavior of time series that is being studied, and as a result, provide a complete and accurate empirical analysis that is worthy of the desired sample sizes. The following table summarizes the lag orders and deterministic terms of each of the models.

Table (1) presents the structures of the proposed FARDL models.

Model	1	2	3	4	5	6	7	8	9	10
p	1	2	1	2	2	3	3	4	4	4
q	1	2	2	1	3	2	1	3	0	4
Const?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trend?	No	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes

2- Determining the initial values:

Having identified the different model structures, the initial values of the proposed models were established through the estimation of the models using actual data under study and the ARDL model. The adoption of these values is used to integrate them as a starting point in the construction of the parameters of the fuzzy FARDL model in the simulation experiments, thus making the synthetic data real as much as possible and approximates their values to the real statistical characteristics of the time series. The estimated value of the parameters is given in Table (2).

Table (2) shows the initial values of the parameters.

Model	1	2	3	4	5	6	7	8	9	10
α_0	-6.039	-10.3	-8.617	-6.971	-11.05	-15.04	-8.932	-18.91	-7.042	-12.6
ψ	-	-0.115	-	-0.092	-0.144	-	-0.124	-0.302	-0.196	-0.28
ϕ_1	-0.131	-0.217	-0.17	-0.169	-0.235	-0.259	-0.209	-0.38	-0.17	-0.35
ϕ_2	-	-0.144	-	-0.094	-0.138	-0.14	-0.106	-0.284	-0.092	-0.23
ϕ_3	-	-	-	-	-	-0.407	-0.228	-0.436	-0.167	-0.3
ϕ_4	-	-	-	-	-	-	-	0.033	0.316	0.162
θ_0	0.323	0.431	0.345	0.392	0.454	0.342	0.401	0.531	0.516	0.693
θ_1	-0.019	-0.142	-0.203	0.039	-0.145	-0.085	0.158	0.058	-	-0.01
θ_2	-	0.309	0.273	-	0.271	0.472	-	0.412	-	0.284
θ_3	-	-	-	-	0.073	-	-	0.166	-	0.51
θ_4	-	-	-	-	-	-	-	-	-	-0.61

Stage Two: Data Generation

In this stage, the data required to simulate the FARDL model are generated in order to evaluate its performance and to compare the Linear Programming and Quadratic Programming methods used in parameter estimation across six sample sizes: (10,15, 20, 30, 50, and 100), with each simulation experiment repeated (1000) times for each sample size.

The sample sizes were chosen with varying purposely to test the performance of the estimation methods in heterogeneous informational conditions, that is, in situations that are typified by small sample sizes. This purposeful choice is associated with economic environments in which paucity of data dominate thus increasing the uncertainty in estimations and reducing the strength of dynamic relationships. Besides, the sample size heterogeneity allows questioning the stability and the efficiency of estimators as a function of information density.

Descriptive statistics and inferential statistics showed that the original time series is not normally distributed because the mean is 25.027 and the standard deviation is 9.167. However, the construction of the simulation experiments in the current study is not intended to achieve the replication of the distributional properties of the original data; instead, it is aimed at the evaluation of the FARDL model and its results. The explanatory variables were therefore produced with an assumption of normality with equal mean and equal standard deviation which was an approximation but was only used to generate the data as opposed to being the true representation of the economic data.

The process of data generating was based on a Data Generating Process (DGP) constructed in the manner of specification on the Autoregressive Distributed Lag (ARDL) model. Under this model, the dependent variable was synthesized with using only the lagged observations of the independent variable along with the lagged observations of the dependent variable combined with a stochastic error component. The analysis also assumes that the error term follows a standard normal distribution and thus to make certain that the errors produced is independent and identically distributed over time, it would mean that the analysis assumes that the error term follows a normal distribution, and therefore the error is independently and identically distributed across the time.

To ensure that the generated series are consistent with the requirements of the conventional ARDL model, the data-generating process was designed in a way that the independent variable in the data process is stationary in levels, being generated with a constant mean and constant variance. The dependent variable on the other hand was obtained by an ARDL specification based on its own lagged values and those of the independent variable and hence captured the dynamic interaction between the two variables . Though the independent variable is integrated into the system of data generation, the autoregressive part, which is associated with the lagged values of the dependent variable, dominates the stability of the model. In this regard, the parameters of this component were selected to meet the requirement of dynamic stability.

In order to confirm this property, stability of the model that was estimated, which was used in the process of data generation, was tested. The analysis of the roots of the characteristic polynomial indicating the autoregressive component showed that all the roots have inverse roots within the unit circle and this means that the time series of the dependent variable is stationary in the levels i.e. $I(0)$.

Accordingly, the generated series satisfy the standard conditions of the ARDL model, in that they do not contain unit roots and do not exceed integration of order one. The independent variable was also made stationary, thus avoiding passing non stationary dynamic to the dependent variable as it is described in the following specification:

$$Y_t = \alpha_0 + \sum_{j=1}^{p} \phi_j Y_{t-j} + \sum_{i=0}^{q} \theta_i X_{t-i} + u_t \qquad , u_t \sim N(0,1) \qquad (9)$$

<u>Stage Three:</u> Estimation

This stage is devoted to estimating the parameters and constructing the FARDL model, based on the following two estimation methods:

1- Quadratic Programming-QP Method.

2- Linear Programming-LP Method.

<u>Stage Four:</u> Results and Comparison

The two metrics used to measure the efficiency of the estimation techniques applied in the simulation framework of the Fuzzy Autoregressive Distributed Lag (FARDL) model were the Root Mean Square Error (RMSE) and the Fuzzy Degree (FD), to compare the estimation methods used and determine the most efficient one. In this regard, RMSE is treated as a distance value between the measured values and the estimates generated by the fuzzy model on the basis of the central values of the estimated fuzzy parameters, and not as a distance value within a more traditional probabilistic scheme; and the FD criterion measures the amount of fuzzy uncertainty that exists in the estimated parameters.

The presentation of the simulation results in this section is limited to the values of the RMSE and FD criteria for each of the proposed models according to the different sample sizes, with the aim of focusing on the core comparison indicators between the estimation methods without expanding on other quantitative details.

Tables (3) and (4) present the results of the RMSE and FD criteria for the proposed models according to the variation in sample sizes and estimation methods.

Table (3): RMSE Criterion Results for the Proposed Models According to Sample Sizes and Estimation Methods

Models	Sample size						Method
	10	15	20	30	50	100	
1	0.945	0.781	0.672	0.559	0.435	0.315	LP
	0.237	0.216	0.198	0.311	0.134	0.098	QP
2	1.664	1.347	1.176	0.948	0.716	0.503	LP
	0.261	0.184	0.177	0.159	0.131	0.096	QP
3	1.493	1.276	1.106	0.911	0.733	0.521	LP
	0.211	0.205	0.190	0.165	0.134	0.097	QP
4	1.146	0.925	0.780	0.634	0.476	0.335	LP
	0.288	0.215	0.183	0.163	0.132	0.096	QP
5	1.658	1.336	1.156	0.927	0.712	0.498	LP
	0.214	0.167	0.161	0.155	0.129	0.096	QP
6	3.031	2.228	1.832	1.386	1.010	0.669	LP
	0.195	0.184	0.176	0.158	0.130	0.096	QP
7	1.712	1.307	1.081	0.857	0.641	0.443	LP
	0.194	0.182	0.176	0.158	0.130	0.097	QP
8	3.051	2.252	1.853	1.390	1.003	0.671	LP
	0.275	0.142	0.134	0.117	0.106	0.094	QP
9	1.642	1.383	1.229	0.983	0.746	0.507	LP
	0.191	0.177	0.170	0.157	0.131	0.096	QP
10	3.394	2.721	2.292	1.817	1.373	0.965	LP
	0.200	0.168	0.159	0.142	0.124	0.095	QP

Table (4): FD Criterion Results for the Proposed Models According to Sample Sizes and Estimation Methods

Models	Sample size						Method
	10	15	20	30	50	100	
1	2.652	2.506	2.336	2.187	1.896	1.598	LP
	1.187	1.173	1.146	1.075	0.958	0.801	QP
2	4.054	3.642	3.440	3.050	2.547	2.132	LP
	1.036	1.079	1.051	0.969	0.856	0.758	QP
3	3.904	3.675	3.394	3.074	2.702	2.201	LP
	1.122	1.133	1.097	1.015	0.888	0.733	QP
4	2.967	2.713	2.491	2.203	1.876	1.649	LP
	1.057	1.077	1.041	0.972	0.878	0.802	QP
5	3.932	3.536	3.294	2.918	2.474	2.062	LP

	0.942	1.029	1.024	0.959	0.841	0.718	QP
6	7.552	6.209	5.664	4.721	3.932	3.072	LP
	1.264	1.225	1.162	1.057	0.919	0.754	QP
7	4.123	3.532	3.176	2.810	2.348	1.941	LP
	1.027	1.058	1.024	0.946	0.852	0.801	QP
8	7.259	5.907	5.287	4.354	3.530	2.714	LP
	1.073	1.125	1.091	0.998	0.876	0.743	QP
9	3.714	3.408	3.196	2.868	2.667	2.770	LP
	1.164	1.131	1.092	0.997	0.912	0.882	QP
10	7.702	6.825	6.197	5.374	4.448	3.623	LP
	1.287	1.332	1.290	1.158	1.002	0.846	QP

As shown in Tables (3) and (4), the Quadratic Programming (QP) outperforms the Linear Programming (LP) method in the different models and sample sizes as indicated by the RMSE and FD values. Generally, (QP) has smaller values of both criteria, which indicates that it is closer to the actual parameters and results in fuzzy estimates of parameters with less uncertainty. As a result, it is possible to consider the (QP) approach as more effective and more reliable in estimating a parameter as compared to an (LP) counterpart.

A review of effects in sample size shows that small samples are utilized to highlight the differences between the two methodologies. In the limited sample settings, the (LP) method shows a greater degree of sensitivity; the root-mean-square error and fuzzy distance measures increase significantly, indicating a relative lack of taking advantage of the limited data and a subsequent lack of stability in the parameter estimates. However, the (QP) methodology maintains relatively smaller and more consistent values of metrics in similar conditions and, therefore, represents a better ability to address the data lack by reducing the gap between the true and approximated parameters and, at the same time, dampening the fuzziness of the approximated parameters.

In large samples, the results show that there is overall improvement in the performance of both the methods as the values of RMSE and FD continue to go down with an increase in sample size. This observation is consistent with the methodological assumption that, the bigger the sample, the better the estimation and reduce uncertainty.

This advantage, however, does not eliminate the differences between the two procedures the Quadratic Programming (QP) method retains the relative best performance even at $(T = 50, T = 100)$ but the difference in performance becomes less in comparison to the smaller sample conditions. Such findings suggest that excellence of (QP) cannot be refined only on the size of a sample but also on the inherent nature of all its estimation procedures.

It has been shown by analyses of the effect of the lag orders (p, q)، that models with a larger lag structure show a strong increase in both the root-mean-square error (RMSE) and the frequency dimension (FD) when the (LP) method is used, a fact that is a manifestation of the increased sensitivity of the method to the complexity involved in the model dynamic architecture as well as to the number of estimated parameters. However, an equivalent escalation is not evident in case of such an increase with the application of the (QP) method; the corresponding

RMSE and FD values are relatively smaller and more consistent even on models of higher lag orders. This behavior implies that the (QP) methodology has a high ability to support dynamic complexity.

5. Conclusion

The results of the simulation experiments show that Quadratic Programming (QP) method is superior to the Linear Programming (LP) method in estimating the parameters of Fuzzy Autoregressive Distributed Lag (FARDL) model based on the Fuzzy Degree (FD), and Root Mean Square Error (RMSE) metrics. QP approach was found to possess a better ability to reduce the gap between the actual parameters and the estimates aid in the generation of fuzzy estimates with a lesser level of uncertainty as compared to the estimates generated by the LP method. Moreover, the findings also show that an increase in the size of the sample leads to better performance of both the methods as it reduces the value of RMSE and FD. However, this improvement does not invalidate the discrepancies between methods as the QP method remains relatively superior in terms of sample sizes and lag structure. These results, therefore, suggest that Quadratic Programming in the FARDL model can be used to estimate parameters more accurately and consistently and where the dynamic complexity is high or the sample size is small.

References

[1] M. H. Pesaran, Y. Shin, and R. J. Smith, "Bounds testing approaches to the analysis of level relationships," *J. Appl. Econom.*, vol. 16, pp. 289–326, 2001, doi: 10.1002/jae.616.

[2] O. Aregbeyen and T. M. Ibrahim, "Testing the Revenue and Expenditure Nexus in Nigeria : An Application of the Bound Test Approach," *Eur. J. Soc. Sci.*, vol. 27, no. 3, pp. 374–380, 2012.

[3] H. Tanaka, S. Uejima, and K. Asai, "Linear Regression Analysis with Fuzzy Model," *IEEE Trans. Syst. Man. Cybern.*, vol. 12, no. 6, pp. 903–907, 1982, doi: 10.1109/TSMC.1982.4308925.

[4] F. Gkountakou and B. Papadopoulos, "The use of fuzzy linear regression and ANFIS methods to predict the compressive strength of cement," *Symmetry (Basel).*, vol. 12, no. 8, 2020, doi: 10.3390/SYM12081295.

[5] M. Eren, "Fuzzy Autoregressive Distributed Lag model-based forecasting," *Fuzzy Sets Syst.*, vol. 459, pp. 82–94, 2023, doi: 10.1016/j.fss.2022.06.003.

[6] F.-M. Tseng, G.-H. Tzeng, H.-C. Yu, and B. J. C. Yuan, "Fuzzy ARIMA model for forecasting the foreign exchange market," *Fuzzy Sets Syst.*, vol. 118, pp. 9–19, 2001, doi: https://doi.org/10.1016/S0165-0114(98)00286-3.

[7] S. Eslamian and K. Ostad-ali-askari, "Evaporation and production efficiency modeling using classical least square regression and fuzzy linear recurrence," *SRN Prepr.*, pp. 1–30, 2021, [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4761502

[8] A. mutlag Abdulateef and E. H. Aboudi, "Estimating the Fuzzy Regression Panel Data Model Based on Approximate Bayesian Computation with Application," *Iraqi Stat. J.*, vol. 2, no. 1, pp. 163–176, 2025.

[9] S. Donoso, N. Marin, and M. A. Vila, "Quadratic Programming Models for Fuzzy Regression," in *International Conference on Mathematical and Statistical Modeling in Honor of Enrique Castillo*, Granada,Spain, 2006.

[10] M. Hojati and C. Bector, "A simple method for computation of fuzzy linear regression," *Eur. J. Oper. Res.*, vol. 166, pp. 172–184, 2005, doi: 10.1016/j.ejor.2004.01.039.

[11] L. A. Zadeh, "Fuzzy Sets," *Inf. Control*, vol. 8, pp. 338–353, 1965, doi: 10.1016/j.procs.2022.09.516.

[12] M. Eren and B. De Baets, "Forecasting Turkey ' s Primary Energy Demand Based on Fuzzy Auto-regressive Distributed Lag Models with Symmetric and Non-symmetric Triangular Coefficients," *Int. J. Fuzzy Syst.*, vol. 27, no. 1, pp. 237–249, 2025, doi: 10.1007/s40815-024-01773-5.

[13] P. Pandit and P. Dey, "Comparative Assessment of Multiple Linear Regression and Fuzzy Linear Regression Models," *SN Comput. Sci.*, vol. 2, no. 76, pp. 75–76, 2021, doi: 10.1007/s42979-021-00473-3.

[14] K. J. Kim, H. Moskowitz, and M. Koksalan,

"Fuzzy versus statistical linear regression," *Eur. J. Oper. Res.*, vol. 92, no. 2, pp. 417–434, 1996, doi: 10.1016/0377-2217(94)00352-1.

[15] A. Shapiro, "Fuzzy regression models," *Artic. Penn State Univ.*, pp. 1–17, 2005, [Online]. Available: http://alm.soa.org/library/research/actuarial-research-clearing-house/2006/january/arch06v40n1-ii.pdf

[16] H. Tanaka and H. Lee, "Interval regression analysis by quadratic programming approach," *IEEE Trans. Fuzzy Syst.*, vol. 6, no. 4, pp. 473–481, 1998, doi: 10.1109/91.728436.

[17] H. Tanaka, I. Hayashi, and J. Watada, "Possibilistic linear regression analysis for fuzzy data," *Eur. J. Oper. Res.*, vol. 40, no. 3, pp. 389–396, 1989, doi: 10.1016/0377-2217(89)90431-1.

[18] W. Pedrycz, "Why triangular membership functions?," *Fuzzy Sets Syst.*, vol. 64, no. 1, pp. 21–30, 1994, doi: 10.1016/0165-0114(94)90003-5.

[19] D. T. Redden and W. H. Woodall, "Properties of certain fuzzy linear regression methods," *Fuzzy Sets Syst.*, vol. 64, no. 3, pp. 361–375, 1994, doi: 10.1016/0165-0114(94)90159-7.

[20] K. Kılıç and D. Topuz, "The evaluation of potentially toxic elements using artificial neural networks and fuzzy linear regression analysis methods in cappadocian volcanic ash soils of Turkey," *Heliyon*, vol. 9, no. 9, 2023, doi: 10.1016/j.heliyon.2023.e19448.