

Citation: Özari Ç. & Ulusoy V. (2017), Estimation Of Bankruptcy Probabilities By Using Fuzzy Logic And Merton Model: An Application On Usa Companies, BMIJ, (2017), 5(4): 211-234 doi: <http://dx.doi.org/10.15295/bmij.v5i4.168>

ESTIMATION OF BANKRUPTCY PROBABILITIES BY USING FUZZY LOGIC AND MERTON MODEL: AN APPLICATION ON USA COMPANIES ¹

Çiğdem ÖZARİ²

Veysel ULUSOY³

Received Date (Başvuru Tarihi): 31/10/2017

Accepted Date (Kabul Tarihi): 04/01/2018

Published Date (Yayın Tarihi): 07/01/2018

ABSTRACT

In this study, we have worked on developing a brand-new index called Fuzzy-bankruptcy index. The aim of this index is to find out the default probability of any company X, independent from the sector it belongs. Fuzzy logic is used to state the financial ratio interruption change related with time and inside different sectors, the new index is created to eliminate the number of the relativity of financial ratios. The four input variables inside the five main input variables used for the fuzzy process, are chosen from both factor analysis and clustering and the last input variable calculated from Merton Model. As we analyze in the past cases of the default history of companies, one could explore different reasons such as managerial arrogance, fraud and managerial mistakes, that are responsible for the very poor endings of prestigious companies like Enron, K-Mart. Because of these kind of situations, we try to design a model which one could be able to get a better view of a company's financial position, and it could be prevent credit loan companies from investing in the wrong company and possibly from losing all investments using our Fuzzy-bankruptcy index.

Keywords: Merton Model, Factor Analysis, Clustering, Fuzzy Logic

JEL Codes: D74, F65, G32

İFLAS ETME OLASILIKLARINI BULANIK MANTIK VE MERTON MODEL KULLANARAK TAHMİN ETME: ABD ŞİRKETLERİ ÜZERİNE BİR UYGULAMA

ÖZ

Bu çalışmada bulanık-iflas endeksi altında yeni bir endeks oluşturduk. Endeksin amacı ait olduğu sektörden bağımsız herhangi bir X şirketinin varsayılan iflas etme olasılığını bulmaktır. Bulanık-iflas mantık endeksi finansal oran kesişim değişikliğini zamandan ve sektörden bağımsız olarak yorumlar. Yeni endeks finansal rasyoların farklı karar verme durumunu ortadan kaldırmak için de oluşturduk. Oluşturulan yeni değişkenin dört tanesi kümeleme ve faktör analizi sonuçlarından elde edilmiş olup, diğer değişken ise Merton modelinden elde edilmiştir. İflas etmiş şirketlerin geçmiş tarihteki olayları analiz edilirken; dolandırıcılık ve yönetim hataları gibi farklı birçok neden ile karşılaşılır. Buna örnek olarak Enron ve K-Mart gibi prestijli firmalar gösterilir. Bu tür durumlardan dolayı bu çalışmada herhangi bir şirketin finansal durumunun dahi iyi anlaşılacağı bir model tasarlanmayı

¹ Estimating Bankruptcy Probability Using Fuzzy Logic: An Application to a Panel of US and Turkish Sectors adlı doktora tezinden türetilmiştir. Kadir Has University, 2012.

² Assist. Prof. Dr., İstanbul Aydın University, cigdemozari@aydin.edu.tr

<http://orcid.org/0000-0002-3850-818X>

³ Prof. Dr. Veysel Ulusoy, Yeditepe University, vulusoy@yeditepe.edu.tr

<http://orcid.org/0000-0001-7227-894X>

hedeflemektedir. Bu model kredi yatırım şirketlerinin yanlış şirkete yatırım yapmalarını ve muhtemelen tüm yatırımlarını kaybetmelerini önleyebilir.

Anahtar Kelimeler: *Merton Model, Faktor Analizi, Kümeleme, Bulanık Mantık*

JEL Kodları: *D74, F65, G32*

1. INTRODUCTION

The main issue discussed in this article is how the Merton Model approach can be used to model a new measure of company's bankruptcy probability, which is also independent from its sectors. The original Merton model consists of some simple assumptions about the financial structure of a typical company. The event of default is determined by the market value of the company's assets in conjunction with the liability structure of the company (Tudela & Young, 2003). The most popular approach for estimating the default probability using market information is the Merton model. The Merton model assumes that a company has equity and certain amount of zero coupon debt that will become due at a future time. Much of the literature follows Merton (1974) by explicitly linking the risk of a company's bankruptcy to the variability in the company's asset value and viewing the market value of company's equity as the standard call option (Hull, 2004). The market value of company's asset with strike price equals to the promised payment of corporate debts (Wang et al., 2009). Merton notes that the shareholder's position can be thought as if the company was to buy a call option on their assets and that the price they would exercise the option with, was equal to the book value of company's debt due for payment in the defined time horizon. In this way, Merton became the first person to show that the default choice of a company could be modeled in accordance with the assumptions of Black and Scholes (1973).

Financial ratios are significant ratios that provide an understanding of the bankruptcy situation of any company and therefore must be used in the creation of the new index. Since there are many financial ratios, factor analysis is used to eliminate some of them. Factor analysis is one of the branches of statistical science, but due to extensions in psychology, this technique is mistakenly regarded as a psychological theory. The use of factor analysis in areas other than psychology has become very popular after 1950s. These fields include various disciplines such as economics, sociology, medical, taxonomy and political science.

The second step for constructing new index is to cluster sectors with variables derived from factor analysis. The sectors between clusters can be obtained from results of cluster analysis, and

this knowledge indicates that they are more volatile than the other sectors. The idea in here is to discover identical properties (if exists) of sectors which are between clusters. When one desires to construct or develop a new index, the objective should be to decrease number of input variables in order to make the new index more accurate, efficient and quick. However, in this papers' case study, correlation analysis was employed to see the relation between the cluster variables as clustering results did not provide enough information to reduce the number of financial ratios. Interpreting financial ratios may signal different meanings for different sectors due to the levels required or identical reasons. To cope with this difficulty, this study uses fuzzy process. For this process, main input variable is Merton default probability (MPD) and the other variables are chosen by the help of factor and cluster analysis. Factor analysis was used once and cluster analysis was used twice to remove financial ratios. To explain the work step by step: Factor Analysis, Cluster Analysis, Correlation Analysis, Cluster Analysis, Calculating MPD Values, Fuzzy Logic Process.

2. DATA AND METHODOLOGY

This part of the study consists of two sections, the first section describes the structure of the sample and the second section describes the construction of the brand-new fuzzy bankruptcy index.

2.1 Data

In this study, 3574 companies have been analyzed including foreign and non-foreign ones, resided in the USA. These companies belong to 78 different sectors such as medical supplies, banking, drug, internet, railroad, energy, **foreign electronics**, healthcare, automobile & truck, chemical, semiconductor, etc.

Table 1 illustrates the calculated and/or collected financial ratios that are most frequently used and significant in the literature for each company. usa.gov and data.worldbank.org web sites has been used gather annually observations for the period of 4 years.

Table 1: Financial Ratios

Total Debt	Return on Capital	Value Line Beta	EBITDA	EBIT
Enterprise Value/Sales	Revenues Last Year	EBIT(t-1)	Capital Expenditures	Enterprise Value/Trailing Sales
Cash as % Revenues	Invested Capital	Market Correlation	Change in Noncash Working Capital	Institutional Holdings
EV/EBITDA	Return on Equity	EV/EBIT	Trailing Revenues	Insider Holdings
Stock Price	Depreciation	Net Margin	Forward EPS	Trailing PE
Number of Shares Outstanding	Non-cash WC as % of Revenues	Intangible Assets /Total Assets	Non-Cash Working Capital	Current Price Earnings
Trailing 12- Revenues	Cash as % Total Assets	Free Cash Flow for the Firm	Three Year Regression Beta	Growth in Revenue (Last Year)
PBV Ratio	Market Cap	Hilo Risk	SG&A Expenses	Price Sales Ratio
Trading Volume	Dividends	Book Value of Assets	Market Debt to Capital	Market Cap
Three Year Standard Deviation	Cash as % Firm Value	Firm Value/Book Value of Capital	Fixed Assets / Total Assets	Expected Growth in EPS
Firm Value	Forward PE	Reinvestment	Book Debt to Capital	Dividend Yield

2.2 Methodology

In the first application, some basic financial ratios, as shown in Table 1, are calculated to find out the financial ratios which are appropriate for all sectors. Since there are so many financial ratios, factor analysis has been used to reduce the number of financial ratios that can explain the identical information and give approximately the same results.

This study obtained 11 financial ratios from the results of factor analysis. Among those 11, a financial ratio which is highly positively correlated with one financial ratio but not positively correlated with the remaining ones (Appendix A) is eliminated from the model. The names and abbreviations of these variables are shown in Table 2.

Table 2: Financial Ratios Derived from Factor Analysis and Correlation

Name	Abbreviation	Name	Abbreviation
Cash as Firm Value	CF	Stock Price	SP
Earnings Before Interest Taxes	EBIT	Trading Volume	TV
Expected Growth in EPS	EG	Total Debt	TD
Number of Shares Outstanding	NSh	Trailing Sales	TS
Pre-tax Operating Margin	PTOM	Value Line Beta	BVL

For the next step, the sectors have been clustered according to the variables, which were obtained from the factor analysis. One of the most fascinating result of cluster analysis was that there are always two particular clusters (Cluster-1 and Cluster-2) appeared each year. These two clusters (sets or groups) were formed considering these ten financial ratios as the resemblance feature.

When the results of cluster analysis is examined, it is observed that some sectors are located in two clusters at the same time, which means that they are between clusters. In other words, it was determined that some sectors belong to different clusters with different percentages. These sectors might have high volatilities. Participating in different clusters at the same time indicates that they are more volatile than the sectors that belong to a single cluster. Since it is desirable to create an index that can explain the possibility of bankruptcy, the volatile sectors will be extremely important data. With this information, similar and dissimilar features of these sectors can be easily found, and in relation to these dissimilar properties, variables that can explain the default probability can be captured.

Table 3 shows the sectors participating in different clusters (i.e. between clusters) with their belonging percentage and shows the sectors that only belong to Cluster-2. The sectors that are not included in Table 3 belongs to Cluster-1 with 100%.

Table3: Results of The First Year Clustering Analysis

	Cluster-1	Cluster-2
Automobile& truck	0.25	0.75
Drug	0.63	0.37
Food Processing	0.00	1.00
Foreign Electronics	0.00	1.00
Foreign Telecom Services	0.00	1.00
Internet	0.90	0.10
Medical Supplies	0.40	0.60
Petroleum	0.17	0.83
Semiconductor	0.87	0.13
Tobacco	0.50	0.50

As the information given in Table 3, the percentage of drug sector found in Cluster-1 is 63% and the percentage of Cluster-2 is 37%. The results of the first-year analysis shows that the sectors; automobile & truck, drug, internet, medical supplies, semiconductor, tobacco, and petroleum have the highest probability of bankruptcy among all sectors in the sample. The fact that any sector does not belong to only one cluster indicates that the sector is too volatile which means that the sector is likely to sink. Realize that, they have different percentages. For example, it can be said that the tobacco sector is more volatile than the semiconductor sector ($0.87 > 0.5$). Since the tobacco sector is among the clusters with 0.5 percentage; this sector equally belongs to Cluster-1 and Cluster-2. However, semiconductor companies belong to Cluster-1 with 0.87 percentage which is greater than 0.5. It can be assumed that the semiconductor companies are almost certainly belongs to Cluster-1. From these results, if the sectors tobacco and medical supplies are compared in terms of risk; the tobacco sector is riskier than the sector medical supplies. As the percentage of belonging to a cluster increases, the variability decreases. From this sentence, you can think that these variables are negatively correlated. But it is not true because variability is related to the fact that the percentage of belonging to a cluster is close to 0,5.

Table 4 illustrates the results of the same type of cluster analysis for the second year only. In addition, the tables that illustrates the results of cluster analysis do not provide information on Cluster-1 since the non-Cluster-2 sectors belong to Cluster-1. As the information given in Table 4, the percentage of internet sector found in Cluster-1 is 82% and the percentage of Cluster-2 is 18%.

Table 4: Results of The Second Year Clustering Analysis

	Cluster-1	Cluster-2		Cluster-1	Cluster-2
Automobile& truck	0.25	0.75	Petroleum	0.86	0.14
Chemical (Diversified)	0.00	1.00	Petroleum (Integrated)	0.00	1.00
Drug	0.56	0.44	Precious Metals	0.75	0.25
Food Processing	0.33	0.67	Semiconductor	0.91	0.09
Foreign Electronics	0.00	1.00	Telecom Equipment	0.50	0.50
Internet	0.82	0.18	Telecom Services	0.56	0.44
Medical Supplies	0.86	0.14	Tobacco	0.00	1.00
Oilfield Services & Equipment's	0.00	1.00			

From the second-year cluster analysis results, only the tobacco companies now belong to only Cluster-2 that means belong to Cluster-2 with 100%. Since in the first-year, the tobacco sector was between clusters with 50%, it can be said that tobacco companies have stabilized according to the clustering variables. Also, some sectors start to be between clusters and some sectors start to be in Cluster-2 only. In other words, new sectors included to Cluster-1 and Cluster-2 in the second year with different percentages and new sectors included only to Cluster-2. The included sectors are food processing, precious metals, telecom services and telecom equipment. In addition, they are the sectors which become more volatile than the first year. For the second-year analysis results, since the percentages of belonging one of the clusters are approximately 0.5, the sectors of drug, telecom equipment and telecom services are the most volatile sectors. To be aware of the percentage in one of the clusters is very important because the volatility of sectors is measured with this identification. In other words, we measure volatility of companies with respect to being in different clusters. The consideration of the difference between being one cluster with 50% and 90% is very important. Thus, the issue that a sector belonging to a cluster with 50% is more volatile than a sector belonging a cluster with 90%, is one of the most striking findings of this study. As can be seen from Table 4, that the drug, telecom services and telecom equipment sectors are more volatile than the internet, medical supplies and precious metals sectors.

Surprisingly, the precious metals and automobile& truck sectors have same percentages in different clusters. In cluster analysis, it was not possible to say that one cluster is better than the other one. And, this result can be considered as the proof of them being in the same position and

in the same risk group. In addition, the chemical and oilfield services & equipment's sectors start to belong Cluster 2 with 100%.

Table 5 illustrates the results of third-year cluster analysis. In addition, it illustrates the sectors which belong to Cluster-2 with 100% and illustrates the sectors between clusters with percentage information. For instance, food processing sectors belonging Cluster-1 and Cluster-2 with 50%. As the information given in Table 5 the percentage of automobile & truck sector found in Cluster-1 is 40% and the percentage of Cluster-2 is 60%.

Table 5: Results of The Third Year Clustering Analysis

	Cluster-1	Cluster-2
Automobile& truck	0.40	0.60
Drug	0.92	0.08
Entertainment	0.83	0.17
Food Processing	0.50	0.50
Internet	0.89	0.11
Metals & Mining	0.50	0.50
Petroleum	0.00	1.00
Precious Metals	0.80	0.20
Semiconductor	0.92	0.08
Telecom Equipment	0.75	0.25
Telecom Services	0.85	0.15

From the third-year cluster analysis results; the sectors drug, semiconductor, internet and automobile& truck are still between clusters with different percentages. Now, petroleum is the only sector that belongs to Cluster-2 with 100% and now the tobacco sector belongs to Cluster-1 with 100%.

If we consider the three-year analysis for the tobacco sector only; the first year it was between clusters, the second year it belongs to Cluster-2 with 100% and the third year it belongs to Cluster-1 with 100%. Table 6 illustrates the results of fourth-year cluster analysis.

Table 6: Results of The Fourth Year Clustering Analysis

	Cluster-1	Cluster-2		Cluster-1	Cluster-2
Aerospace/Defense	0.75	0.25	Internet	0.80	0.20
Automobile& truck	0.40	0.60	Maritime	0.00	1.00
Biotechnology	0.00	1.00	Medical Supplies	0.75	0.25
Cable	0.00	1.00	<i>Metals & Mining</i>	0.50	0.50
Chemical	0.67	0.33	Petroleum	0.00	1.00
Chemical	0.00	1.00	Power	0.00	1.00
Diversified Co.	0.00	1.00	Precious Metals	0.00	1.00
Drug	0.90	0.10	Railroad	0.00	1.00
E-Commerce	0.67	0.33	Semiconductor	0.90	0.10
Entertainment	0.00	1.00	<i>Telecom Equipment</i>	0.57	0.43
<i>Food Processing</i>	0.50	0.50	Telecom Services	0.83	0.17
Hotel & Gaming	0.00	1.00	Utility	0.00	1.00

As the information given in Table 6, the percentage of chemical sector found in Cluster-1 is 67% and the percentage of Cluster-2 is 33%. When the analysis of the fourth year is examined, it is observed that there are more members of Cluster-2. For instance; the sectors entertainment, hotel & gaming and railroad are now the sectors which belong to Cluster-2. From Table 6, it appears that the most volatile sectors are food processing, metals & mining. Telecom equipment and automobile & truck sectors can be seen as also volatile sectors than the others but not as high as food processing and metal & mining sectors.

In addition, the semiconductor and drug sectors belong to Cluster-1 with 90% in the fourth-year cluster analysis results. These sectors practically belong to Cluster-1. From Table 6, the sectors chemical and e-commerce are between clusters with same percentage, which means that they have same type of risk. Medical supplies and aerospace / defense industry sectors are among the clusters and the percentages of being in the clusters are the same. Table 7 illustrates the sectors between clusters for each year. The semiconductors, drug, internet and automobile &

truck sectors were found to be among the clusters each year according to the cluster analysis results made every year for 4 years.

Table7: Sectors Between Clusters

For all years	Automobile& truck, Drug, Internet and Semiconductor
First two years	Petroleum
First, second and last years	Medical supplies
Last three years	Food Processing, Telecom Equipment and Telecom Services
Last two years	Metals & mining
Second and third years	Precious Metals
First year only	Tobacco
Third year only	Entertainment
Last year only	Aerospace/Defense, Chemical and E-Commerce

The cluster analysis shows that food processing, medical supplies, telecom equipment and telecom services are sectors between clusters for three times for different time-periods. Since four-year analysis is examined, they are between clusters with probability 0.75.

In addition; the sectors metals&mining, petroleum and precious metals are the sectors between clusters for two times for different time-periods. In other words, they are between clusters with probability 0,5. The sectors such as tobacco, entertainment, chemical, e-commerce and aerospace/defense are the sectors between clusters for only one year, that means they are between clusters with probability 0.25. The other sectors that are not mentioned in Table 7 are the sectors belongs to Cluster-1 or Cluster-2 with 100%.

The sectors between clusters are important sectors and data for us to determine our bankruptcy index. To do this, one have to find similar and identical properties of these sectors according to the clustering variables. For these sectors, basic statistical information is used, in other words the mean and standard deviation of the clustering variables are calculated. Table 8 shows the mean and standard deviation of the clustering variables of all companies for each year. In addition, the mean and standard deviations (yearly) for the sectors between clusters (semiconductor, internet, drug and automobile& truck) are calculated. The values are given in Appendix B.

Table8: Descriptive Statistics of All Sectors

	SP	TV	NSh	TD	PTOM	TS	EG	VLB	EBIT	CF
1st YearMean	45.61	2099800.68	129.23	866.28	-1.1	18.06	0.18	0.61	819.22	0.14
1st YearSt. Dev.	534.69	53541582.78	466.02	4524.82	5.54	163.01	0.12	0.56	3307.06	0.34
2nd YearMean	50.43	1970577.11	144.23	906.41	-1.47	23.82	0.17	0.88	705.35	0.15
2nd YearSt. Dev.	611.97	49117175.39	468.47	4557.92	7.16	262.36	0.12	0.42	3196.35	0.39
3rd YearMean	49.35	2976954.13	154.83	1326.63	-1.79	11.05	0.16	0.9	859.79	0.18
3rd YearSt. Dev.	612.69	64667284.05	485.6	10374.06	8.72	75.41	0.09	0.38	3890.77	0.33
4th YearMean	30.86	2429465.97	175.99	1523.87	-1.28	3.58	0.14	0.99	1052.35	0.43
4th YearSt. Dev.	405.08	55590958.26	521.34	10558	6.03	29.18	0.12	0.42	4568.08	1.07

The average of SP and TV of these sectors are smaller than the average of all sectors. In addition, the average of TD and EBIT are also smaller for the sectors drug semiconductor and internet. Only for the automobile& truck sector, the average of TD and EBIT are higher than the sector averages. For this reason, it is not possible to generalize that the sectors in the both clusters (i.e. between clusters) have lower TD and EBIT values.

After that, the correlation matrix of the clustering variables and results are shown in Table9 and the bold ones are positively correlated ones. Since the correlation matrix is symmetric, the other values are not included.

Table 9: Correlation Matrix of the Financial Ratios Derived from Factor Analysis

	NSh	CF	TS	EG	PTOM	SP	TD	TV	VLB	EBIT
NSh	1									
CF	-0.28	1								
TS	-0.03	-0.33	1							
EG	-0.36	0.15	0.39	1						
PTOM	-0.31	-0.32	0.03	-0.12	1					
SP	0.04	-0.22	0.48	0.24	0.19	1				
TD	0.41	-0.20	-0.16	-0.34	0.03	0.11	1			
TV	0.45	-0.18	0.35	0.02	0.21	0.29	-0.56	1		
VLB	-0.08	0.06	0.05	0.76	-0.01	-0.21	-0.13	0.01	1	
EBIT	0.65	-0.27	-0.13	-0.40	0.16	0.20	0.89	0.07	-0.20	1

From Table 9, the highest correlation coefficient is 0.89 and between clustering variables EBIT and TD. This means that before the sectors between clusters have similar properties. They

have both smaller TD and EBIT values and EG has highest correlation with EBIT. To make distinct clusters, the variable EG among these three variables will be chosen.

To select other clustering variables, one needs to find variables that have no relationship or almost no relationship. Because of this reason, we look at the smallest correlation coefficient between clustering variables and choose the variable TV, BVL, CF and PTOM.

We again cluster the sectors for each year with respect to these new variables (EG, TV, BVL, CF and PTOM) and find out that there are no sectors between clusters. In other words, the sectors are between clusters due to other clustering variables (EBIT, TD, SP, TS and NSH). From the second cluster analysis result gives us an information that these financial ratios (ratios make the analysis between clusters) measure the volatility. In other words, these financial ratios are better observers of volatility than other financial ratios. Since EBIT and TD has highest correlation, TD was also eliminated from the fuzzy process.

To summarize, there are four variables for constructing a new index through factor analysis and clustering analysis. For the last important variable, the MPD of all sectors must be calculated. Next section explains how the MPD values of any company and MPD values of each sector can be calculated.

2.3 Merton Probability of Default Model

The probability of default can be measured with a model that uses company's financial information to obtain an indication of how likely the firm is to enter distress in the near future (Saretto, 2004). This financial information, in the Merton Model, consists of several variables. The volatility of the assets, the risk-free rate, debt of the company and the timing of the action are the input variables of the Merton Model. The volatility of the assets has been determined by analyzing the stock data return of the last three years. To measure the risk-free rate, a one-year constant maturity rate was analyzed. As trying to provide a based model, the time period is taken as one year and the balance sheets of the companies are used to figure the model.

With the parameters mentioned above, one could say that related with the Merton Model, there are 2 equations with 2 unknown parameters which are current value of the assets and volatility of the assets. As soon as possible we discover the un-known parameters, we could easily calculate the MPD of each company. The connection between the asset volatility and equity volatility defined by Ito's Lemma. Black-Scholes is used to value the firm's equity as a function of its asset value and its assets volatility. To find these unknown parameters, we need to find values for them that would satisfy both equations. By the solution of these equations;

MPD values of each company for each year are calculated. Sector MPD values are calculated by taking the average of MPD values of all companies in each sector.

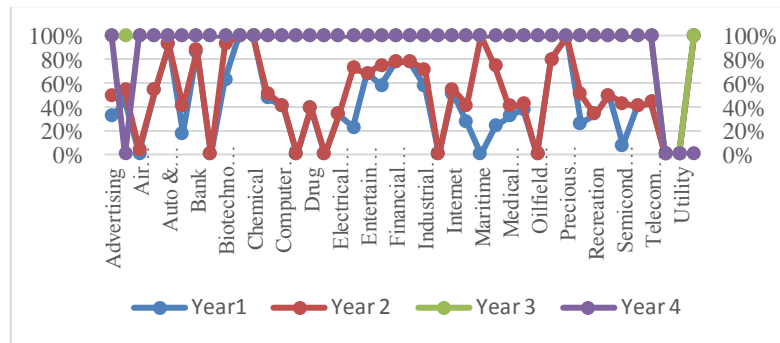
As mentioned before, sectors which are between clusters are more volatile than the sectors which belong to only one cluster. However, when we calculate the MPD values of these sectors, for each year none of them have higher probability than 0.5. Only the tobacco sector has probability 0.5 at the first year. This also indicates the necessity of building a default bankruptcy probability index. Table 10 illustrates MPD values of the sectors, which are between clusters.

Table 10: MPD of the Sectors which are Between Clusters

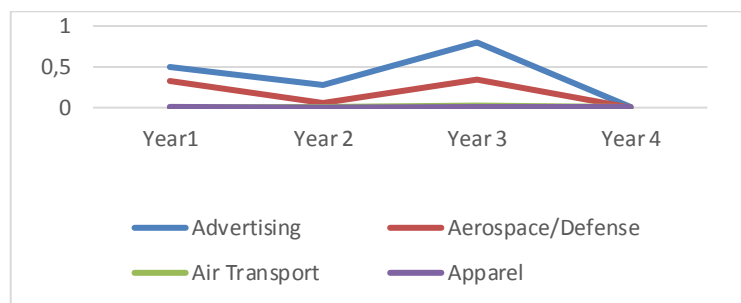
Sector	MPD-1 st Year	MPD-2 nd Year	MPD 3 rd Year	MPD 4 th Year
Aerospace/Defense	-	-	-	0.000
Automobile& truck	0.2500	0.0000	0.02135	0.0077
Chemical	-	-	-	0.0024
Drug	0.0736	0.0003	0.1151	0.0626
E – Commerce	-	-	-	0.016
Entertainment	-	-	0.1684	-
Food Processing	-	0.0004	0.0548	0.0079
Internet	0.2741	0.0203	0.2441	0.1117
Medical Supplies	0.1562	0.04	-	-
Metals & Mining	-	-	0.3702	0.052
Petroleum	0.1288	0.00023	-	-
Precious Metals	-	0.0003	0.0024	-
Semiconductor	0.0001	0.000451	0.078	0.0127
Telecom Equipment	-	0.00033	0.3025	0.0506
Telecom Services	-	0.00024	0.2873	0.0599
Tobacco	0.5000	-	-	-

From Table 10, one can easily observe that in the first year the tobacco sector has a greater MPD value (0.5) than the other sectors and this sector is also between clusters in the same year. This sector is the riskiest sector than the other sectors in the first year. Also in this year, the second risky sector is the internet. In addition, in the second year the medical supplies sector has a greater MPD value (0.04) than the other sectors and again this sector is between clusters. From the third and last year analysis results, the sector metals & mining and internet have greater MPD values than the other sectors and they are also between clusters at these years. Also from Table 10, if MPD values are considered only for the sectors which are between clusters for all four years, big changes are observed. This is another reason that why we need another index to construct, which is a modified version of MPD and Graph 1 and Graph

2independently illustrate all these problems that MPD values have (quickly change, near values to each other)



Graph 1: MPD Values of Selected Sectors



Graph 2: MPD Values of Selected Sectors

2.4 Fuzzy Model for Bankruptcy Probability

In this section of the study, firstly, emergence of fuzzy logic and its application areas are explained, then explain why and how we make use of the fuzzy logic and how we construct the model to create the new index.

The “Fuzzy Sets” discovered by Lotfi A. Zadeh at 1965 in his well-known paper called Information and Control. As Mr. Zadeh is a prestigious scholar working on control-theory, the first steps of fuzzy logic are represented by himself. The concept of ‘state’ which designed by Zadeh figure out the basis of the modern control theory. In the early 60’s, he thought that classical control theory had put too much emphasis on precision therefore could not handle the complex systems. In the beginning of 1962, he studies to handle biological systems “we need a radically different kind of mathematics, the mathematics of fuzzy or cloudy quantities which are not describable in terms of probability distribution (Zadeh, 1962). At the end of these

studies, he states his ideas in his paper Fuzzy Sets. Later he formalized the ideas into his paper Fuzzy Sets. Mr. Zadeh has proposed many of the fundamental concepts in fuzzy process in end of 60's and beginning of 70's.

As far as Zadeh presenting the fuzzy sets in 1965, he proposed the fuzzy algorithm in 1968, fuzzy decision making in 1970 (Bellman & Zadeh), and fuzzy ordering in 1971 (Zadeh). On the other hand, he has published a seminal article called "Outline of a new approach to the analysis of complex systems and decision process" (Zadeh, [1973]), that is also known as the first starting point of fuzzy control in 1973. In this article, he introduced concept of linguistic variables and proposed to use if-then rules to formulate human knowledge. Since then, fuzzy set theory has been rapidly developed by Zadeh himself and a large number of researchers, and in a wide range of unexpected fields, successful applications of this theory have begun to emerge

The economics and finance areas are one of the areas that fuzzy logic can be applied. There are also many optional subjects which are suitable for fuzzy logic process; such as modeling complex sales and trading systems, cost benefit analysis, investment appraisal and portfolio analysis. Also, in banking system fuzzy logic approach is used for the assessment of credit demand or rating credits. Fuzzy systems are used mostly for estimating, decision-making and mechanical control systems.

In daily facts, we are using fuzzy logic for comparative concepts that can be changed related with the persons, cultures, time and with similar concepts. As an example, a person could feel cold in a place when the other one sitting nearby him could feel the opposite. Another example from finance, the percentage of the increase of one stock is makes much more profit than another one with the same increase percentage.

In our fuzzy model, 5 fuzzy variables are defined as inputs and for each input, 5 membership functions that 2 of them are trapezoids and 3 of them are triangles. With the need to find the border points of each membership function, first we calculate the basic statistics such as mean, maximum, minimum and standard deviation of each input. In the light of these data and expert opinion, we decide the borders of each membership function as seen in Table 13, Table 14 and Table 15.

Table 13: Range of the Input Variables

	SP	EBIT	NSh	TS	MPD	Output
Min	-1	-1	-1	-10	0	0
Max	20	15	15	15	1	1
Mean	-0.0321	-0.0158	0.0344	-0.0227	0.022	0.4853
St. Dev	0.7641	0.9372	11.141	0.8423	0.2883	0.0366

Table 14 shows the borders of triangular membership functions and Table 15 shows the borders of trapezoid membership functions of the input variables of the fuzzy process.

Table 14: Properties of Triangular Membership Functions: Fair, Good and Very Good

Fair	E	F	G	Good	O	P	Q	Very	H	I	J
SP	0.027	2.14	4.08	SP	1.08	4.47	6.08	SP	5.36	10.36	16.61
EBIT	1.79	3.82	6.34	EBIT	4.7	6.61	8.68	EBIT	6.38	9.16	13
NSh	2.89	4.73	7.13	NSh	4.47	6.53	9.09	NSh	8.17	9.65	10.99
TS	-7.78	-	-1.99	TS	-3.16	-1.11	0.68	TS	-1.17	3.67	7.29
MPD	0.39	0.54	0.86	MPD	0.054	0.2579	0.521	MPD	0,009	0,036	0,059
Output	0.1	0.3	0.5	Output	0.41	0.5	0.59	Output	0.5	0.7	0.9

Table 15: Properties of Trapezoid Membership Functions: Excellent and Poor

Excellent	K	L	M	N	Poor	A	B	C	D
SP	14.25	17.9	22.1	38.9	SP	-1	-1	0.25	0.47
EBIT	8.68	10.65	15	15	EBIT	-1	-1	0.70	4.72
NSh	9.23	11.42	15	1	NSh	-1	-1	0.6	4.68
TS	2.70	8.62	15	15	TS	-10	-10	-7.5	-4.93
MPD	0	0	0.0025	0.033	MPD	0.52	0.82	1	1
Output	0.7	0.9	1	1	Output	0	0	0.1	0.3

In our fuzzy logic process, we use and/or method with taking minimum/maximum values of variables. For the implication process, we use minimum and for the aggregation process, we use maximum method. As we explained before, for the defuzzification, we use centroid method.

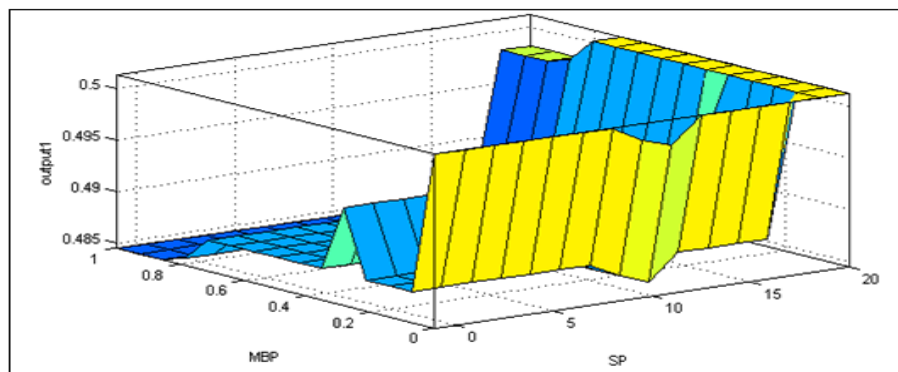
The membership functions are defined as “Poor”, “Fair”, “Good”, “Very Good”, and “Excellent”. In this study, for the membership functions we decided to use piecewise linear functions such as triangular and trapezoidal. The triangular membership functions are “Fair”,

“Good”, “Very Good”. “Poor” and “Excellent” has the shape of trapezoidal because of their over situation.

To sum up; 5 input variables and one output variable, which we defined as fuzzy bankruptcy index and denoted FBI, which was produced from these 5 variables. One of the input variables is the default probability that we calculated from the Merton Model, which is denoted by MPD. The other input variables are the variables from the results of factor and cluster analysis (SP, NSh, EBIT and TS).

For the next step of modeling, the rules of fuzzy process need to define. Notice that, if there are only 2 variables with 5 membership functions for each, the total number of rules should be 25. Since there are 5 input variables with 5 membership functions are defined, the number of rules that must be applied is 5^5 . According to average and standard deviation applications, rulings have been decided for indicators. For the “and” rules, if all inputs are same, the same linguistic variable for the output is assigned. For the “or” rules, if at least three of them are same, the most used linguistic variable for the output is assigned. In other words, if all input variables are member of poor membership function, trivially the output must be the member of poor membership function. After determining the membership functions and rules of the fuzzy process, Mamdani fuzzy inference system is employed to create output for our model. Graph 3 illustrates the relationship between two of the inputs of the model with the output.

Graph 3: Surface that Shows Relationship between MPD, SP and Output



Using Graph 3, the relation of MPD and SP can be easily seen with the value of output. So, after investigating the graph, you can change up your mind and decide to use another indicator the best fits the output and Table 16 illustrates the FBI and MPD values of randomly selected sectors.

Table 16: FBI and MPD Values

Sector	Index	Year 1	Year 2	Year 3	Year 4
Utility	FBI	0.3814	0.4231	0.3814	0.3979
	MPD	0.0000	0.0000	0.0000	0.0000
Wireless Networking	FBI	0.5000	0.3814	0.3817	0.2315
	MPD	0.1429	0.0000	0.0010	0.0000
Advertising	FBI	0.3814	0.5000	0.3814	0.5000
	MPD	0.5000	0.2875	0.7983	0.0082
Telecom Services	FBI	0.4981	0.5000	0.3961	0.5000
	MPD	0.2257	0.0002	0.2874	0.0599

To sum up the steps of fuzzy process are:define the data, define the borders of the membership functions, define the model that develop the output, define the rules, fuzzification process, defuzzification process and constructing the output FBI.

3. CONCLUSION

In this study by the help of fuzzy process and Merton Model, a new default parameter named fuzzy-bankruptcy index, which is denoted by FBI, is constructed. The main aim is also to find an index, which is nearly acute to all securities. To do this, our veri consists of 78 different sectors stated in USA. MPD values of these companies were calculated and by the help of these values, the MPD values of any sector by taking the average MPD values of the companies that belong to this specific sectorwere also calculated. To determine other input variables of our fuzzy logic model, factor analysis and cluster analysis are employed. First,factor analysis is used to reduce the number of financial ratios andin the analysis, elevenfactors were found with the identical information. One particular financial ratio from the result of factor analysis that have positively correlated to one of the others and weak correlation with the restwas also eliminated from the index.

Since we decided to use fuzzy process, we know that must define membership functions for each variable. The minimum number of membership functions which is suitable to our study can be 3, such as fair, good and very good. However, we want to measure the sensitivity also, because of this, we define 5 membership functions such as poor, fair, good, very good and excellent. The problem starts at here. For the fuzzy process,10 input variables with 5 membership functions are too many to be significant. In this model, defining too many

membership functions indicates too many fuzzy rules, which gives too much initiative to the researcher. In this case, two options exists, one is to reduce the number of membership functions to 3(Fair, Good, Very Good) or reduce the number of input variables. Because of the sensitivity, we decided to reduce the number of input variables and use cluster analysis again. Realize that we decided to reduce the number of membership functions,we cannot cope with the difficulty of explaining all sectors. Figure 5 illustrates the main steps of the study.

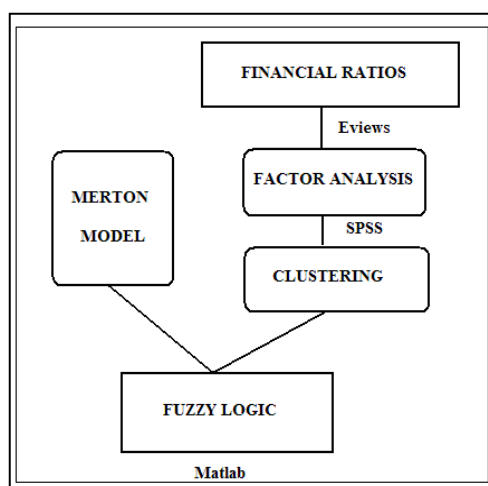


Figure 5: Main Steps of Study

Volatile sectors by the support of cluster analysis working like an exploratory data analysis tool for organizing observed data into meaningful groups. In this study, our first aim is to analyze the sectors which are between clusters. The companies between clusters means that they are the companies which are not stable or companies with higher bankruptcy probabilities. We try to find identical properties of these sectors as soon as we selected sectors between clusters. Searching for the similarities of the financial ratios of companies which are between clusters, we examine that they are the properties which makes companies between clusters. In addition, the correlation coefficients of the financial ratios which we found from factor analysis was also examined. With these information, we try to find specific financial ratios that only measures the volatility concept.

In the light of correlation analysis, we eliminate some financial ratios and cluster sectors again with these variables and find again two clusters. However, now the sectors belong to only one cluster. Thus, there were financial ratios that can measure bankruptcy sensitivity. The financial ratios which does not make sectors belong to only one cluster are the financial ratios we are looking for. Since two of them have a high correlation coefficient, one of them is not included to the analysis.

5 input variables were obtained to construct our new index which indicates bankruptcy probabilities. To use fuzzy logic, first we analyze the input variables by the help of basic statistics. In the light of these knowledge we can construct the membership functions and fuzzy rules between our input variables.

To check the effectiveness of our index, we compare the values of MPD and the new index FBI, and figure out that their correlation is positive. This means that these variables are acting on same direction which indicates that they both increase or decrease at the same time. Notice that we construct FBI because MPD takes values so close to 0 or takes to values so near to each other, which we cannot easily make estimation with this knowledge or we cannot see the distinguishes of default probabilities between years. Our results show that FBI is much more sensitive than the MPD. Since the main objective of this study is to improve the probability of bankruptcy of the Merton model and to complete the deficiencies, the new index is compared only with MPD values. Sensitivity and minor differences can be observed with the new variable, and major changes that should not have been observed have been removed. In other words, since the MPD values are input variables of FBI, we can not conclude a result that FBI calculation is easier than MPD. With the help of fuzzy process and the sector-independent financial ratios resulting from cluster and factor analysis an index was created which is more sensitive and does not show major decision differences in the short term.

REFERENCES

- Bandemer, H. and Gottwald, S. (1995). *Fuzzy Sets, Fuzzy Logic, Fuzzy Methods with Applications*, Wiley, New York, NY, USA.
- Benos A. and Papanastasopoulos G. (2005). Extending the Merton Model: A Hybrid Approach to Assessing Credit Quality. <http://econwpa.repec.org/eps/fin/papers/0505/0505020.pdf>
- Bharath S. T. and Shumway T. (2004). Forecasting Default with the KMV-Merton Model, Working Paper in University of Michigan.
- Black, F., and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81, 637-659.
- Crouhy. M. Galai. D., Mark. R. (2000). A Comparative Analysis of Current Credit Risk Models. *Journal of Banking and Finance*.
- D'amico, G. (2008). A Convergence Result in the Estimation of Markov Chains with Application to Compound Options. *Journal of Statistical Theory and Practice*, 2(4), 693-705.
- Duffie D. and . Leandro A, Saita. Ke Wang. (2006). Multi-period Corporate Default Prediction with Stochastic Covariates, *Journal of Financial Economics*. 59-117.
- Duffie, D., & Singleton, K. J. (2012). *Credit risk: pricing, measurement, and management*. Princeton University Press.
- Chen, D. H., Chou, H. C., Wang, D., & Zaabar, R. (2009). the Predictive Performance of a Barrier Option Credit Risk Model in an Emerging Market. Working Paper.
- Groenen, P. J., Kaymak, U., & van Rosmalen, J. (2007). Fuzzy clustering with Minkowski distance functions. *Fuzzy Clustering and its Applications*, Wiley, 53-68.
- Haack, S. (1979). Do we need "fuzzy logic"? *International journal of man-machine studies*, 11(4), 437-445.
- Hull J. C., *Options* (1997). *Futures and Other Derivatives*. Prentice Hall.
- Hull J., White A., Joseph L. Rotman. (2000). Valuing Credit Default Swaps I: No Counterparty Default Risk, April.
- Hull, J., Nelken, I., & White, A. (2004). Merton's model, credit risk, and volatility skews. *Journal of Credit Risk* Volume, 1(1), 05.
- Miyake, M., & Inoue, H. (2009). A default probability estimation model: An application to Japanese companies. *Journal of Uncertain Systems*, 3(3), 210-220.
- Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance*, 29(2), 449-470.
- Lin, S., Ansell, J., & Andreeva, G. (2007). *Merton models or credit scoring: modelling default of a small business*. University of Edinburgh Management School
- Saretto, A. (2005). Predicting and pricing the probability of default.
- Shnaider, E., & Kandel, A. (1989). The use of fuzzy set theory for forecasting corporate tax revenues. *Fuzzy Sets and Systems*, 31(2), 187-204.

Huschens, S., Vogl, K., & Wania, R. (2005). Estimation of default probabilities and default correlations. In *Risk Management* (pp. 239-258). Springer Berlin Heidelberg.

Tudela, M., & Young, G. (2005). A Merton-model approach to assessing the default risk of UK public companies. *International Journal of Theoretical and Applied Finance*, 8(06), 737-761.

Zadeh L.A. (1965). Fuzzy sets. *Info. & Ctl*, Vol. 8. 338-353.

Zadeh L.A. (1968). Fuzzy Algorithms. *Info. & Ctl*, Vol. 12. 94-102.

Zadeh, L. A. (1984). Making computers think like people [fuzzy set theory]. *IEEE spectrum*, 21(8), 26-32

Appendix A: Results of Factor Analysis

Factor	Variance	Cumulative	Difference	Cumulative	Proportion
F1	12.66	12.66	6.96	0.38	0.38
F2	5.7	18.36	1.22	0.54	0.17
F3	4.49	22.85	2.02	0.68	0.13
F4	2.47	25.32	0.88	0.75	0.07
F5	1.59	26.91	0.14	0.8	0.05
F6	1.45	28.36	0.12	0.84	0.04
F7	1.32	29.68	0.24	0.88	0.04
F8	1.08	30.76	0.07	0.91	0.03
F9	1.01	31.77	0.02	0.94	0.03
F10	0.99	32.76	0.01	0.97	0.03
F11	0.97	33.73		1	0.03
Total	33.73	293.16			1

Appendix B: Descriptive Statistics of the Industry: Between Clusters

		SP	TV	NSh	TD	PTO	TS	EG	VL	EBIT	CF
Semiconductor	All Mean	9.36	725932.01	325.87	250.86	-0.5	10.22	0.2	0.98	491.44	0.7
	All St. Dev.	9.17	2396277.3	1081.6	568.5	4.54	67.3	0.1	0.52	1352.2	1.3
	1st Year Mean	10.2	294371.05	298.83	244.51	0.15	3.12	0.2	0.74	803.22	0.3
	1st Year St. Dev.	6.78	672097.1	1068.5	565.11	0.32	2.02	0.0	0.76	1674.7	0.3
	2nd Year Mean	10.9	533954.95	316.23	252.07	0.1	30.24	0.2	1.17	376.14	0.2
	2nd Year St.	9.56	998316.42	1097.2	568.75	0.4	124.2	0.1	0.43	1173.8	0.2
	3rd Year Mean	10.7	1620146.6	331.56	220.7	0.09	4.67	0.2	1.02	455.97	0.3
	3rd Year St.	11.5	4504342.8	1123.2	551.14	0.38	9.99	0.1	0.38	1424.3	0.3
	4th Year Mean	5.44	455255.43	356.86	286.16	-1.82	-1.29	0.2	1.04	493.75	1.6
	4th Year St.	7.53	970393.1	1116.2	626.37	8.08	7.88	0.1	0.29	1360.2	2.0
Internet	All Mean	10.2	294371.05	298.83	244.51	0.15	3.12	0.2	0.74	803.22	0.3
	All St. Dev.	6.78	672097.1	1068.5	565.11	0.32	2.02	0.0	0.76	1674.7	0.3
	1st Year Mean	10.9	533954.95	316.23	252.07	0.1	30.24	0.2	1.17	376.14	0.2
	1st Year St. Dev.	9.56	998316.42	1097.2	568.75	0.4	124.2	0.1	0.43	1173.8	0.2
	2nd Year Mean	10.7	1620146.6	331.56	220.7	0.09	4.67	0.2	1.02	455.97	0.3
	2nd Year St.	11.5	4504342.8	1123.2	551.14	0.38	9.99	0.1	0.38	1424.3	0.3
	3rd Year Mean	5.44	455255.43	356.86	286.16	-1.82	-1.29	0.2	1.04	493.75	1.6
	3rd Year St.	7.53	970393.1	1116.2	626.37	8.08	7.88	0.1	0.29	1360.2	2.0
	4th Year Mean	15.6	302994.6	52.38	9.79	-0.47	17	0.2	0.9	163.23	0.3
	4th Year St.	46.2	549909.1	54.12	20.55	2.7	37.84	0.1	0.4	827.34	0.7
Drug	All Mean	10.1	377235.58	188.37	660.38	-7.28	54.84	0.1	0.89	691.78	0.5
	All St. Dev.	15.2	899302.64	519.27	2675.7	14.32	500.2	0.0	0.53	2712.4	1.1
	1st Year Mean	11.8	318453.09	176.86	411.44	-7.27	19.45	0.1	0.52	695.41	0.2
	1st Year St. Dev.	14.3	957752.63	539.53	1527.3	13.81	30.65	0.0	0.55	2431.0	0.1
	2nd Year Mean	11.7	402909.59	186.86	656.19	-8.09	175.3	0.1	0.98	573.58	0.3
	2nd Year St.	16.7	877208.66	518.56	2885.7	14.76	967.8	0.0	0.41	2361.6	0.3
	3rd Year Mean	9.89	457125.18	194.22	581.87	-7.82	12.36	0.1	1.05	730	0.4
	3rd Year St.	16.3	1039490.1	523.18	2071.1	16.7	31.62	0.0	0.48	2916.3	0.5
	4th Year Mean	7.13	330454.45	195.56	992.02	-6.02	4.08	0.0	1.04	769.05	0.9
	4th Year St.	13.1	706295.52	509.24	3728	12	18.9	0.1	0.49	3073.7	2.1

Automobile& truck	All Mean	11.8	318453.09	176.86	411.44	-7.27	19.45	0.1	0.52	695.41	0.2
	All St. Dev.	14.3	957752.63	539.53	1527.3	13.81	30.65	0.0	0.55	2431.0	0.1
	1st Year Mean	11.7	402909.59	186.86	656.19	-8.09	175.3	0.1	0.98	573.58	0.3
	1st Year St. Dev.	16.7	877208.66	518.56	2885.7	14.76	967.8	0.0	0.41	2361.6	0.3
	2nd Year Mean	9.89	457125.18	194.22	581.87	-7.82	12.36	0.1	1.05	730	0.4
	2nd Year St.	16.3	1039490.1	523.18	2071.1	16.7	31.62	0.0	0.48	2916.3	0.5
	3rd Year Mean	7.13	330454.45	195.56	992.02	-6.02	4.08	0.0	1.04	769.05	0.9
	3rd Year St.	13.1	706295.52	509.24	3728	12	18.9	0.1	0.49	3073.7	2.1
	4th Year Mean	14.3	957752.63	539.53	1527.3	13.81	30.65	0.0	0.55	2431.0	0.1
	4th Year St.	11.7	402909.59	186.86	656.19	-8.09	175.3	0.1	0.98	573.58	0.3