Prediction of Market Value of Used Commercial Aircraft

Qian Feng

Risk and Environmental modeling group
Department of Applied Mathematics
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
Prediction of Market Value of Used Commercial Aircraft

MASTER OF SCIENCE THESIS

For the degree of Master of Science in Risk and Environmental Modeling Group (R&EM) at Department of Applied Mathematics at Delft University of Technology

QIAN FENG

13th August, 2012

Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
Delft, The Netherlands
The work in this thesis was supported by PROJECT of Risk and Environmental Modeling Group. Their cooperation is hereby gratefully acknowledged.

All rights reserved.
Copyright © Risk and Environmental Modeling Group
Applied Mathematics Department
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
Delft, the Netherlands
Prediction of Market Value of Used Commercial Aircraft

A Thesis submitted to the Faculty of Electrical Engineering, Mathematics and Computer Science

By

Qian Feng

in partial fulfillment of the requirements for the degree of Master of Science.

Daily Supervisor
Dr. Dorota Kurowicka

Responsible professor
Prof.dr. F.H.J. Redig

Committee member
Prof.dr. F.H.J. Redig
Dr. Dorota Kurowicka
Dr. F.H. van der Meulen
Dr. Anca Hanea
Bjorn Batenburg
Abstract

Aviation financers are interested in the current/future market value of used commercial aircraft, as this information is precious knowledge for them to support collateral position in the aircraft loan. In this paper, variables from general economy, airline industry and aviation fleet are explored to find out the factors predictive for used aircraft market. Two statistical methods - principal component regression and copula - are applied for building the prediction model of an aircraft.
# Table of Contents

List of Table ........................................................................................................................................... i
List of Figure ............................................................................................................................................... iii

Chapter 1 Introduction ............................................................................................................................ 1

Chapter 2 Analysis of available data ...................................................................................................... 7

Section 2.1 Available data ..................................................................................................................... 7
  Section 2.1.1 Specification of aircraft type .......................................................................................... 7
  Section 2.1.2 Historical market values .............................................................................................. 8
  Section 2.1.3 Data of airline industry indicators .............................................................................. 10
  Section 2.1.4 Data of global economy indicators ............................................................................ 11
  Section 2.1.5 Data of Airline fleet indicators .................................................................................. 13
  Section 2.1.6 Transformation of data .............................................................................................. 14

Section 2.2 Data exploration .................................................................................................................. 15
  Section 2.2.1 Market value versus type and age ............................................................................. 15
  Section 2.2.2 Market value and global economy .......................................................................... 18
  Section 2.2.3 Market Value growth versus type and age ............................................................. 19
  Section 2.2.4 Market value growth versus global economy and airline industry ...................... 24
  Section 2.2.6 Market value growth versus fleet ........................................................................... 28

Section 2.3 Conclusion ........................................................................................................................... 33

Chapter 3 Regression model of the market value growth ..................................................................... 35

Section 3.1 Classical linear regression ................................................................................................. 35
Section 3.2 Principal component regression ......................................................................................... 39
Section 3.3 Test of PCR model ............................................................................................................. 44
Section 3.4 PCR model - results .......................................................................................................... 48
Sections 3.5 Incorporating the age in the model .................................................................................. 50
Section 3.5.1 Varying coefficient PCR model .................................................................................... 50
Section 3.5.2 PCR model for clustered aircraft group ....................................................................... 54
Section 3.6 Conclusion ................................................................................................................. 55

Chapter 4 Model the market value growth with copula .......................................................... 57
  Section 4.1 Copula model ......................................................................................................... 57
  Section 4.2 Copula model-results ........................................................................................... 67
  Section 4.3 Incorporating the age in the model ...................................................................... 70
  Section 4.4 Influences between aircraft types ....................................................................... 73
  Section 4.4 Conclusion ............................................................................................................. 77

Chapter 5 Model performance ............................................................................................... 79
  Section 5.1 Comparison between PCR and copula model ..................................................... 79
  Section 5.2 Error analysis with age, year and aircraft type ..................................................... 82
  Section 5.3 Market value ......................................................................................................... 87
  Section 5.3 Conclusion ............................................................................................................ 89

Chapter 6 Conclusions ........................................................................................................... 91

Bibliography ............................................................................................................................... 93

Appendix A ................................................................................................................................. 95
  B.1 predicted market value based on PCR model ................................................................. 103
  B.2 predicted market value based on copula model ............................................................. 105
  B.3 polynomial function of coefficients of PCR with respect to age ................................. 107
  B.3 polynomial function of coefficients of copula with respect to age .............................. 108
# List of Table

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>the form of data of current market value</td>
<td>8</td>
</tr>
<tr>
<td>Table 2</td>
<td>aircraft selected [9]</td>
<td>9</td>
</tr>
<tr>
<td>Table 3</td>
<td>variables for describing the airline industry expansion</td>
<td>10</td>
</tr>
<tr>
<td>Table 4</td>
<td>variables for describing the world economy</td>
<td>12</td>
</tr>
<tr>
<td>Table 5</td>
<td>Data of Fleet of global airline industry, annual</td>
<td>13</td>
</tr>
<tr>
<td>Table 6</td>
<td>Data of manufacture annual</td>
<td>14</td>
</tr>
<tr>
<td>Table 7</td>
<td>New constructed data contain information of the second half year and the first half year</td>
<td>15</td>
</tr>
<tr>
<td>Table 8</td>
<td>aircraft groups</td>
<td>23</td>
</tr>
<tr>
<td>Table 9</td>
<td>Correlation coefficient between MV growth and number of aircraft in storage, for new Airbus A320-200</td>
<td>29</td>
</tr>
<tr>
<td>Table 10</td>
<td>indicative variables for model construction</td>
<td>34</td>
</tr>
<tr>
<td>Table 11</td>
<td>standard description for regression model</td>
<td>36</td>
</tr>
<tr>
<td>Table 12</td>
<td>coefficients and their statistics</td>
<td>36</td>
</tr>
<tr>
<td>Table 13</td>
<td>Correlation matrix and significance test</td>
<td>37</td>
</tr>
<tr>
<td>Table 14</td>
<td>95% confidence interval for coefficients</td>
<td>38</td>
</tr>
<tr>
<td>Table 15</td>
<td>KMO and Bartlett's Test</td>
<td>40</td>
</tr>
<tr>
<td>Table 16</td>
<td>Principal component analysis</td>
<td>41</td>
</tr>
<tr>
<td>Table 17</td>
<td>Loadings (principal component coefficients)</td>
<td>42</td>
</tr>
<tr>
<td>Table 18</td>
<td>coefficients for PRC</td>
<td>44</td>
</tr>
<tr>
<td>Table 19</td>
<td>Determining coefficient for models of new aircraft</td>
<td>44</td>
</tr>
<tr>
<td>Table 20</td>
<td>error comparisons between model of classical linear regression and PCR model for new Boeing 737-800</td>
<td>49</td>
</tr>
<tr>
<td>Table 21</td>
<td>Average error comparisons between model of classical linear regression and PCR model for all aircraft</td>
<td>49</td>
</tr>
<tr>
<td>Table 22</td>
<td>comparison of errors between three methods: PCR, varying coefficient PCR, PCR for clustered aircraft groups</td>
<td>49</td>
</tr>
<tr>
<td>Table 23</td>
<td>comparison of errors for model of new Boeing 737-800</td>
<td>54</td>
</tr>
<tr>
<td>Table 24</td>
<td>comparison of average errors for all 61 aircraft</td>
<td>68</td>
</tr>
<tr>
<td>Table 25</td>
<td>Boeing 767-300ER, comparison between Basic copula model and Complete copula model, applying normal distribution modeling the marginal functions</td>
<td>72</td>
</tr>
</tbody>
</table>
Table 26 for Boeing 767-300ER, comparison between Basic copula model and Complete copula model, applying kernel density estimation modeling the marginal functions................................. 72
Table 27 for Boeing 767-300ER, comparison between Basic copula model and clustered aircraft group, applying kernel density estimation modeling the marginal functions. ...................... 72
Table 28 direct competitors from two manufacturers......................................................................... 74
Table 29 members in one family in the same manufacturer ............................................................... 75
Table 30 aircraft produced by the same manufacturer, but with quite different size ................ 76
Table 31 the same aircraft type (Airbus A320-200) at different age ..................................................... 77
Table 32 Comparison of errors between the model of PCR and Copula for aircraft Boeing 737-800 over year 2001 to year 2010 ............................................................................................................ 80
Table 33 Comparison of errors between the model of PCR and Copula for all aircraft being modeled over year 2001 to year 2010 ......................................................................................................................... 81
Table 34 comparison of calculation time .............................................................................................. 82
Table 35 MAE of model for aircraft at different ages, PCR and Copula ........................................... 83
Table 36 error in year 2011, model of new aircraft ............................................................................... 86
Table 37 Money loss for Boeing 737-800, over year 2001 to year 2010 ........................................... 88
Table 38 Money loss in 2011 prediction for new aircraft ..................................................................... 88
Table 39 Comparison of money loss, model in [8], PCR model and Copula model, over year 2001 to year 2011......................................................................................................................... 89
List of Figure

Figure 1 difference of narrow-body and wide-body aircraft ............................................................. 8
Figure 2 Market value vs. seat capacity ............................................................................................... 16
Figure 3 Market Value vs. age, Airbus A320-200 ........................................................................... 17
Figure 4 Historical trends of Global GDP and aircraft market ......................................................... 18
Figure 5 MV growth patterns for new aircraft: Airbus A319-100, Boeing 777-200 ......................... 20
Figure 6 Difference between MV growth patterns of new Boeing 767-300ER, Airbus A330-200 and
Airbus A320-200 ............................................................................................................................... 21
Figure 7 MV growth patterns for Airbus A320-200 at age 0, age 3, age 11 and age 12 ..................... 22
Figure 8 Growth patterns of global economy, airline industry revenue and MV (Airbus A320-200) .................................................................................................................................................. 25
Figure 9 Scatter plots, global economy vs. MV growth of new Airbus A320-200 ......................... 26
Figure 10 Scatter plots, airline industry vs. MV growth of new Airbus A320-200 ......................... 27
Figure 11 Historical trends of fleet & manufacture, Airbus A320-200 ........................................... 28
Figure 12 scatter plots of growth of fleet & manufacture vs. MV growth, For Airbus A320-200.... 30
Figure 13 new Boeing 767-300ER; scatter plots of fleet & manufacture growth vs. MV growth... 31
Figure 14 14-year-old Boeing 767-300ER; scatter plots of growth of fleet & manufacture vs.
MV growth ......................................................................................................................................... 32
Figure 15 Box plot for the five variables .......................................................................................... 41
Figure 16 Scree plot .......................................................................................................................... 43
Figure 17 Standardized residuals versus standardized predicted values, Boeing 737-800 .......... 45
Figure 18 Lag plot, model of Boeing 737-800 .................................................................................. 46
Figure 19 Q-Q plot of residuals ........................................................................................................ 47
Figure 20 PCR model versus truth, new Boeing 737-800................................................................. 48
Figure 21 Relationship between coefficients of PCR model and age .............................................. 51
Figure 22 MV growth versus age at different year, for Boeing 767-300E ..................................... 52
Figure 23 coefficients versus age for Airbus A330-300 ..................................................................... 53
Figure 24 histograms for six variables .............................................................................................. 59
Figure 25 Empirical distribution vs. normal distribution fitting for GDP growth ....................... 60
Figure 26 Empirical distribution vs. kernel estimation for GDP growth ............................................. 61
Figure 27 scatter plot, probability integral transfrom of GDP growth verse probablity integral
transfrom of MV growth for new Boeing 737-800 ........................................................................... 63
Figure 28 scatter plots of 100 samples, generated from the bivariate distribution of variable for
GDP growth and variable for MV growth of new Boeing 737-800 ................................................. 64
Figure 29 Conditional distribution of MV growth in year 2006, for new Boeing 737-800, marginal functions are modelled by kernel density estimation ................................................................. 66
Figure 30 model versus truth, Copula model of new Boeing 737-800, modeling marginal function applying kernel density estimation ................................................................. 67
Figure 31 the histogram of simulated samples from the conditional distribution of MV growth in year 2002, copula model of new Boeing 737-800 ................................................................. 69
Figure 32 age vs. correlation between MV growth and economy factors .............................................. 70
Figure 33 Relationship between parameters of normal distribution fitting and age ............................. 71
Figure 34 Conditional distribution of MV growths for Airbus A320-200 and Boeing 737-800 in year 2002 and year 2005, size of the simulated samples =100 ............................................................ 74
Figure 35 the conditional distribution of MV growths for Airbus A330-200 and Airbus A330-300 in year 2002 and year 2005, size of the simulated samples =100 ............................................................ 76
Figure 36 comparison between PCR and Copula, prediction versus time .......................................... 80
Figure 37 MAE of model versus age ................................................................................................. 82
Figure 38 errors for all aircraft over year 2001 to year 2010, left: PCR; right: Copula ......................... 84
Figure 39 comparison between errors of new aircraft model, left: MAE, right: MAEX ......................... 85
Figure 40 model and truth versus time, new Boeing 767-300ER .............................................................. 86
Figure 41 true and modeled market values versus time ........................................................................ 87
Chapter 1 Introduction

Proper estimation of current and future aircraft value is an important knowledge for financer to support the collateral position of the aircraft. Commercial aircrafts are expensive assets worth of millions of US dollar; careless investment without evaluation of aircraft can put the aircraft bankers at risk. Bankers always prefer to provide loans for aircraft with strong market position, as they need to be sure that the aircraft reposed can cover the loss when put in the used-aircraft market again.

One may find it complicated when trying to understand aircraft market. Usually, the buyers of commercial aircraft are aircraft operators; however, it is difficult to tell the identity of the potential seller. It can be an aircraft manufacturer such as Boeing and Airbus, a banker with reposed aircraft from broken airline companies, airline companies who need to reduce their capacity or to get rid of old fleet, and so on. Aircraft being sold can be brand new, one year to more than ten years old. What’s more, the airline companies can make different choice of specification (such as engine type, avionic system) of the ordered aircraft within the options provided by the manufacturers.

These complexities make it a very difficult process to estimate the value of aircraft even under known economic environment. In aircraft market, the agreement on price is done by negotiation between buyers and sellers instead of a fixed money tag; after the agreement on price, the buyer will seek for a financial loan from aircraft financers, who will decide if to provide the loan based on all the information collected [1]. Generally, each party in the sale gets an evaluation of the aircraft being sold/ purchased independently, and thus consulting service of aircraft values becomes a very important business.

There are three levels of consulting services for estimation of current/future market value of an aircraft:

- Appraisal by experienced appraiser
An experienced appraiser will acquire all the facts about the aircraft in examination. He/she must inspect the aircraft personally to check its current physical condition as well as the maintenance and damage records. Together with estimate of the value the client will get a detailed report and the comprehensive assessment of the aircraft’s condition. For a specified aircraft, appraisal given by expert is the most reliable and thus most expensive option. Quality of the expert will be of crucial importance, however [2].

- Valuation online
  Appraiser will not visit the aircraft for checking. The client is required to provide all information needed by filling a questionnaire online. Then appraisal is made with software specially design for the purpose.
  Appraisal online is cheaper, and the client can still get a detailed analysis about the market value of the aircraft concerned [3].

- Books of aircraft value published by appraisal institutes
  Some aviation appraisal institutes publish books of aircraft values. One can find historical market values of aircraft for various models at each age. Market values presented in these books do not take into account specific conditions (hours flown, maintenance records etc.) of an aircraft of interest. They specify market value of the aircraft type of certain age in an ‘average’ condition. Usually this price is adjusted later to include extra information about the aircraft in question.
  Aircraft value book is the cheapest option which gives a quite rough estimation of aircraft value. The most commonly used ones among aircraft appraisers and brokers are the four ‘Aircraft Blue Book’ published in USA [4]. However, one may find lots of options for a book on aircraft values, as many professional aircraft appraisal institutes publish such a book as a profitable product.

Aircraft book has simplified greatly the estimation of an aircraft value; details such as maintenance document or damage history are not considered. Some books neglect even the small difference in detailed specifications of an aircraft under the same model (such as engine type), as some people only care about the history trends of various aircraft types in the market, instead of the detailed prices. The market value given in aircraft book is the conclusion made by experts based on experience and historical transaction information.
In this research, an aircraft values do not depend on aircraft’s physical conditions. This is also the basic assumption in the data source\(^1\) of aircraft market value used in this thesis. This assumption motivated by available data simplifies the task of building a quantitative model for estimation of aircraft value.

Within aircraft value consulting services, a distinction needs to be made between a ‘base value’ and a ‘market value’ when talking about aircraft value. The base value is a theoretical value that implies the underlying economic value of an aircraft in a hypothetical balanced supply and demand, while the market value is an estimation of most likely trading price under market condition at the time of sale. Both ‘base value’ and ‘market value’ change from year to year as they are influenced by the external economic environment [5]. Under the equilibrium of supply and demand, ‘base value’ and ‘market value’ are equal.

Equilibrium of supply and demand is a rather theoretical concept. Like any other physical asset, a crucial element determining the value of an aircraft is the relationship between supply and demand. One may explain or estimate the potential economic value of an aircraft as the revenue-generating ability from the perspective of an aircraft operator. However, we are concerned about how much an aircraft is worth in current market, hence only market values of aircrafts are analyzed in this thesis.

Vasigh and Erfani (2004) present internal and external factors influencing market value of an aircraft in [6]. Internal factors refer to aircraft specific characteristics, such as age and technical configurations of the type, which largely determine its value. External factors can be any factors that influence the aircraft market. The performance of general economy, the development of airline industry, the aircraft technology progress, and even some policies such as environmental regulations, all have some impacts on the valuation of commercial aircraft [6].

In [6], it is emphasized that it is a complex process to estimate the market value of a specific aircraft. Even more difficult is to predict an aircraft’s future value. Assuming that the current market value of an aircraft is known, its future market value is still not clear because of the uncertainty in external factors. Assumptions of future market demand needs to be made for prediction of aircraft markets. The future demand for an aircraft is estimated by forecasting the economic indicators (such as GDP growth, air traffic growth).

\(^1\) The data source is the aircraft value book published by ASCEND, which is an international company focusing on providing aviation information and consulting service.
Vasigh and Erfani in [6] developed a model to estimate the base value of an aircraft (the potential economic value) from perspective of finance theory. In this model, aircraft is considered as a machine generating revenue for aircraft operators. According to financial theory, the value of an asset (the aircraft) reflects the net present value (NPV) of the revenue-generating capacity over the economic life of the asset. When the operating costs exceed the revenue generating capacity, the economic life terminates. The model thus calculates the NPV of the aircraft according to the revenue generated and cost (capital cost, operational cost) over the economic life of the aircraft; and the future revenue and cost need to be estimated as well as they change with the general economy (such as oil prices, passenger numbers) as well.

Kelly\(^2\) presents a way to understand estimation of market values of used aircraft from a perspective an appraiser [7]. All other things being equal, aircraft values depreciate over time due to ageing structures that require an increasing amount of maintenance and increased oil consumption. The percentage of original value of a used aircraft can indicate depreciation value. Based on the AVITAS transaction database, Kelly states that age alone can explain 66% of the variance of the market value. Experienced experts classify aircraft into ten groups according to their market strengths, and model the relationship between percentage of original value and age for each group by polynomial fitting with historical data. Market strength of aircraft type is mainly determined by aircraft size and technology, and also influenced by some market forecasting and manufacturer status [7]. In this way, as long as we know the market value of a brand new aircraft, we are able to estimate market values of used aircrafts of all ages for the same type.

The model in [7] is used for estimation of current market value of used aircraft. Kelly points out that in order to forecast future market value, additional variables concerning economy must be introduced into the model. The effect of the demand cycle should be incorporated, as aircraft operators need to replace their fleet after a certain time. Only by introducing variables concerning economy, it is possible to see how far the values will fall when market changes significantly such as after 9/11.

Both research results focus on estimation of base value of an aircraft. The model in [6] explains the values of a used aircraft showing the potential economic value that can be generated by the aircraft, while the model in [7] describes the depreciation value of the used aircraft with respect to the new one.

\(^2\) an experienced expert of aircraft appraisal in AVITAS aviation consulting company,
Efforts to describe the market value change can be found in [8]. Ajang constructed a model to show the relationship between the departure of market value from base value and indicators of demand and supply. Year 2009 is defined as the base year, and the market values in 2009 are the base values. Four variables are considered as indicative factors for aircraft market values according to the analysis in [8]. These are: jet fuel price, backlog\(^3\) of orders, the total number of comparable aircraft types in storage\(^4\) and the ratio of new orders to current total number of this aircraft type. These four variables act as the explanatory variables in the linear regression model of departure of market value from base value.

Ajang goes further with including age as another parameter in the model. For a given type, the coefficients have to be recalculated for aircraft at each age level. Ajang modeled the relationship between coefficients in the regression model and age by polynomial fitting.

The idea in [8] is really inspiring. We see that the difference between base value and market value can be described by the factors from the economic environment of the aircraft market. As long as the values of the four variables are known, one can estimate the corresponding market value based on the model.

However, the analysis of the results shows that there is an inconsistency of the contributions of the variables to aircraft market. For example, the contribution of jet fuel price to model of Airbus A319-100 is positive, while the contribution of the same variable to Airbus A320-200 is negative; Airbus A319-100 and Airbus A320-200 both belong to Airbus A320 family, and one can expect that they have similar behavior in the market. In [8], the discussion in conclusion admits the inconsistent contributions of the same variable to models of similar aircraft, and mentioned that this is one problem in model validation\(^5\).

Another questionable decision in [8] is that, year 2009 is picked as the base year. Market values in other years are compared with values in 2009 to measure the change of the market. However, the whole world suffered from the worst economic recession in year 2009, and thus the aircraft market was severely impacted. The aircraft market value in this year does not represent the ‘real’ potential economic value of an aircraft as they are underestimated greatly.

---

\(^3\) Order: the number of new aircraft orders from the manufacturers; Backlog: the cumulative number of orders that are not delivered

\(^4\) Storage: the number of aircraft the operators put into storage

\(^5\) This problem has been discussed in the internship report of this project. The reason for this inconsistency is the high correlation among the explanatory variables. When two explanatory variables are highly correlated, they are conveying essentially the same information. So when both variables are included in the model, the coefficients cannot represent the real influence of the explanatory factors to the dependent variable anymore.
Despite these deficiencies, the model in [8] is the only complete model one may find in the literature for modeling the market value of aircraft with changing economic environment. This thesis is motivated by the research presented in [8]. Thinking over the idea in [8] we attempt to model the complicated relationships between aircraft market and all the potential influential factors in the general economic environment.

The main goal of this research is to build a prediction model of aircraft market values based on available data, and to understand the interrelationship between the variables in the general economy and aircraft market. We consider two predictive models for market value and study their strengths and weaknesses. The thesis consists of four main parts.

First, we make a first round of data selection to detect the useful information from the collected data. We will find out the influential factors for aircraft values from the internal specific characteristics of the aircraft and the external economic environment. Data exploration is made to check all the possible correlations between aircraft values and economic variables. In fact, it is found that instead of market value, it is better to model the market value growth.

The first model for market value growth is build in Chapter 3. First we use classical regression to model the relationship between market value growth and the explanatory variables; Principal component analysis is applied to handle the high dependence among the explanatory variables; the information in the explanatory variables is combined in one principal component. Comparison between classical linear regression and principal component analysis, and discussion concerning the sensitivity of the market value growth of an aircraft at different age is discussed.

The second model is an application of copula. We make a model of the joint distribution of all concerned variables. Kernel density estimation and normal distribution fitting can be used to model the margins separately; and we use normal copula to describe the whole dependence structure among Market value growth and the explanatory variables. Using copula model we are not only able to predict market value growth, but also to discuss the interactive influence between different aircraft types.

We compare the performance of these two models, and make detailed analysis concerning the results. One can see how the models response to the practical events in Chapter 5.
Chapter 2 Analysis of available data

It is quite challenging to identify the influential factors of the aircraft values. In [6] and [7], we can see that the internal characteristics of an aircraft and the general economic environment mainly determine the market value of an aircraft. In order to find out proper indicators, we will make an exploration of data in this chapter.

Section 2.1 Available data

In this section, available data which is necessary in the following chapters to build a predictive model of the market value is presented.

Section 2.1.1 Specification of aircraft type

The aircraft type specifies basic properties of an aircraft given by manufacturer. Generally, the manufacturer may provide several options of the type with some extra equipment such as engine and avionic system. In this research, we will take the most commonly used specification for a type. In other words, aircrafts of a specific type produced in the same year are considered as identical.

Aviation Research (AR) department of DVB Bank presents an overview of modern western-built aircraft in [9]. The aircraft types that are included are the main western-built airplanes that are currently in operation. Older types, such as DC-9, B707, B727, have been omitted as they are less relevant for the Bank’s day-to-day business. Based on information provided by this book, one can consider the following properties of an aircraft:

- **Size**
  Seat capacity indicates the size of the aircraft type.

- **Technology**
  Year of first flight is the indication of age of the technology.

- **Class**
  Class of an aircraft is determined by its size and number of aisles it can contain in basic arrangement. In this respect aircrafts are divided onto narrow-body and wide-body also known as the single-aisle and the twin-aisle aircrafts, respectively (see Figure 1). Typically
the wide body aircrafts have larger maximum range (more than 10000 km) and can be used for longer flights (e.g. B777-300: 11029 km, A330-200: 12500km). The narrow-body aircraft typically have maximum range of about 6000 km (e.g. B737-700: 6230 km, A320: 5700 km).

Section 2.1.2 Historical market values

Data of past market values represent the historical trend of aircraft market. It is the most important data set for this research.

We use the data sheet of market values coming from ASCEND company in this report. ASCEND (Aerospace Information Redefined) is an international company providing aerospace information, and consultancy services on valuations and appraisals. It publishes a book of market values of aircrafts and updates the data every year (updates are done every July). The same data has been used by Ajang in [8].

Market values of aircrafts combine information of experts’ experience as well as some real transactions. This is a big asset of ASCEND and the methods used to produce these data are not public.

ASCEND data contains information about current market value of an aircraft of given type and age in an ‘average’ condition. Data of current market value has the form shown in Table 1:

<table>
<thead>
<tr>
<th>Type</th>
<th>Year of Manufacture</th>
<th>Year of Valuation</th>
<th>Current market value</th>
<th>Age</th>
</tr>
</thead>
</table>
Because of the properties of the available data we will predict the market value of an aircraft of certain age in average condition. Hence, maintenance or damage information is assumed not to influence the market value in our model. Moreover the aircraft is in standard configuration which means that if there are several options for engines, difference will not be considered [4].

Though there are over one hundred of aircraft types in original data, only a small part can be used for data analysis. First of all we focus only on passenger aircrafts. Moreover we will model aircraft types that are still in production and ones that are produced longer than five years.

Since the main goal of this report is to find indicative factors for used-aircraft market value we need to see the historic trends of market values of aircraft types at various ages. Hence, we will select for the analysis passenger aircraft types that have been in production before 2001 and are still in production in 2009. With this criterion only eight types are included for the analysis (see Table 2).

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Type</th>
<th>Class</th>
<th>Seat capacity</th>
<th>First flight year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airbus</td>
<td>A319-100</td>
<td>Medium narrow-body</td>
<td>124 (2 class), 134 (1 class)</td>
<td>1995</td>
</tr>
<tr>
<td>Airbus</td>
<td>A320-200</td>
<td>Medium narrow-body</td>
<td>150 (2 class), 164 (1 class)</td>
<td>1988</td>
</tr>
<tr>
<td>Airbus</td>
<td>A321-200</td>
<td>Medium narrow-body</td>
<td>185 (2 class), 199 (1 class)</td>
<td>1997</td>
</tr>
<tr>
<td>Boeing</td>
<td>737-800</td>
<td>Medium narrow-body</td>
<td>162 (2 class), 175 (1 class)</td>
<td>1997</td>
</tr>
<tr>
<td>Airbus</td>
<td>A330-200</td>
<td>Medium wide-body</td>
<td>253 (3 class), 193 (2 class)</td>
<td>1997</td>
</tr>
<tr>
<td>Airbus</td>
<td>A330-300</td>
<td>Medium wide-body</td>
<td>295 (3 class), 335 (2 class)</td>
<td>1992</td>
</tr>
<tr>
<td>Boeing</td>
<td>767-300ER</td>
<td>Medium wide-body</td>
<td>218 (3 class), 269 (2 class)</td>
<td>1986</td>
</tr>
<tr>
<td>Boeing</td>
<td>777-200ER</td>
<td>Large wide-body</td>
<td>301 (3 class), 375 (2 class)</td>
<td>1994</td>
</tr>
</tbody>
</table>

A319, A320 and A321 belong to Airbus family. A319 is the smallest, A320 medium and A321 the largest of the three. They are medium-range, narrow-body, commercial passenger jet airliners manufactured by Airbus. A320 family was developed to compete against the Boeing 737-Classics (-300/-400/-500), and has since faced challenges from the Boeing 737 Next Generation (-600/-700/-800/-900).
Boeing 767-300ER is the extended version of Boeing 767-300, which is the first stretched version of family Boeing 767. The main competitor of Boeing 767-300ER is the Airbus A330-200. Boeing 777-200 is a long-range, wide-body, twin-engine jet airliner which competes with A330-300 [14].

Section 2.1.3 Data of airline industry indicators

In [6] and [7], it has been emphasized that the demand from the airline industry is the driven factor for commercial aircraft market. We will use airline revenue and traffic demand as the indicator of the health of airline industry.

Data source of airline industry comes from internal Air Transport Association (IATA) data base. The IATA represents some 240 airlines comprising 84% of total air traffic [10]. It is a reputable organization over the whole airline industry. IATA provides the data representing the performance of airline industry over 30 years.

However, because this is a profitable service, only data from 2000 to 2011 can be accessed for free. Annual data is updated by the end of year and monthly data by the end of month. Here we will also give the reason for the selection of airline revenue and traffic demand as the indicator for airline industry [11].

- Airline revenue
  
  Airline revenue is the total income of airline industry. It represents the cash flow of the airline industry. We will take it as an important measure for the capital ability of the airline industry.
  
  Revenue growth is defined as the percentage change of revenue per year.

- Traffic demand
  
  There are two measures for passenger traffic. One is the conventional passenger number, and the other is revenue passenger kilometers (RPKs). RPKs is defined as:
  
  \[ \text{Revenue Passenger Kilometer} = \text{number of paying passengers} \times \text{kilometers flown} \]

  Detailed descriptions of the data available are presented in Table 3.

<table>
<thead>
<tr>
<th>Name</th>
<th>Unit</th>
<th>Indication for</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airline industry</td>
<td>Billion (USD)</td>
<td>Airline industry</td>
<td>The total money earned by the airlines</td>
</tr>
</tbody>
</table>

http://www.iata.org/whatwedo/economics/pages/traffic_analysis.aspx
Section 2.1.4 Data of global economy indicators

Global economy provides the basic trading environment for aircraft market. The following factors indicate the state of global economy, and are related to airline industry.

- **GDP**
  
  Global Domestic Product (GDP) is the most common measure of economy development [12]. It refers to the market value of all officially recognized final goods and services produced within a country in a given period.

- **Fuel price**
  
  Fuel price constitutes the main operating cost for aircraft operators. It is believed that when jet fuel price is high, operators prefer fuel-efficient aircraft; in other words, younger aircraft may be more popular. Crude oil is cheaper than jet fuel, but it can represent the market trend of jet fuel as well.

- **Interest rate**

---


10 Data source of RPK growth (monthly year-on-year): IATA monthly RPK monitor, tracked and provided by Aviation research department of DVB Bank.
Real interest rate represents the capital cost for the money borrowers. As few aircraft operators can do their payments by cash, capital cost is an important part in operating cost.

- Inflation rate

Market value may need to be adjusted by inflation. A consumer price index (CPI) measures changes in the price level of consumer goods and services purchased by households, and would be used as an index for inflation rate in this report.

Detailed introduction to variables describing the world economy is listed in Table 4.

<table>
<thead>
<tr>
<th>Name</th>
<th>Unit</th>
<th>Indication for</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual World GDP</td>
<td>Billion (USD)</td>
<td>World economy</td>
<td>World Global Domestic Product</td>
</tr>
<tr>
<td>Quarterly World GDP</td>
<td>Billion (USD)</td>
<td>World economy</td>
<td>World Global Domestic Product</td>
</tr>
<tr>
<td>World GDP growth</td>
<td>%</td>
<td>World economy</td>
<td>Percentage change of GDP in market exchange rates</td>
</tr>
<tr>
<td>Crude oil price</td>
<td>USD/ Barrel</td>
<td>Fuel price</td>
<td>Europe Brent Spot Price</td>
</tr>
<tr>
<td>Real interest rate in</td>
<td>%</td>
<td>Capital cost</td>
<td>The lending interest rate adjusted for inflation as measured by the GDP deflator in USA.</td>
</tr>
<tr>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPI</td>
<td></td>
<td>Inflation</td>
<td>Consumer Price Index</td>
</tr>
</tbody>
</table>

11 Data source of GDP: International Monetary Fund, World Economic Outlook Database.
13 Data source of GDP / GDP growth (annual): International Monetary Fund, World Economic Outlook Database.
14 Data source of Crude oil price (annual, monthly): U.S. Energy Information Administration
15 Data source of Real interest rate in USA (annual): The World Bank, Catalog Source World Development indicators
Section 2.1.5 Data of Airline fleet indicators

Airline fleet indicators give information about the makeup of a current fleet. This data is believed to describe preferences of airlines for certain aircraft types. Data of Airline fleet is also provided by ASCENT Company.

- Number in storage
  The aircraft operator sometimes put some aircraft in storage to deal with the reduced traffic. They will return to operation when traffic grows. However, some aircraft are put in storage only because they are old and are being replaced by newer ones. Sometimes the important parts are removed from stored aircrafts and they might never return to operations anymore.

- Number in service
  Number of aircraft in service can show the traffic demand and popularity of a specified aircraft type among the operators.

- Number of operators
  Number of operators owning a certain aircraft type might indicate the popularity of this aircraft. It is believed that the high number of operators owning the aircraft type influences positively its market value.

<table>
<thead>
<tr>
<th>Name</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number in storage</td>
<td>The total number of this aircraft type put in storage</td>
</tr>
<tr>
<td>Number in Service</td>
<td>The total number of this aircraft type in service</td>
</tr>
<tr>
<td>Operator number</td>
<td>The total number of aircraft operators who use this aircraft type</td>
</tr>
</tbody>
</table>

Another data set of the popularity of aircraft is data from manufacturers. Data of manufacturers is included in the data set in [8]. Net order can present the demand for the aircraft type. Generally, it takes two to three years for the manufacturer to deliver the aircraft after it has been ordered. With new orders coming and long delivery time it is common to see that there are backlogs for some popular aircraft types.

- Order backlogs

---

The firm order backlog indicates the potential for short-term deliveries; the higher the backlog, the higher the future deliveries. However too high backlog could have an adverse impact on orders, as the long waiting-time will discourage potential customers.

- Net order

Net orders can represent the current demand for the specified aircraft type in the market.

### Table 6 Data of manufacture annual 18

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order backlog</td>
<td>Cumulative amount of an aircraft type has been ordered but not delivered yet</td>
</tr>
<tr>
<td>Net Order</td>
<td>Total amount of orders placed for an aircraft type minus cancelation made per year</td>
</tr>
</tbody>
</table>

### Section 2.1.6 Transformation of data

In this thesis, one of the important goals is to investigate the relationship between aircraft values and indicators in general economy and airline industry. However, the difference in updating time of different data sets may cause some problem in the analysis’ conclusion.

As mentioned before, the data of market values is updated annually every July. However, data of general economy and airline industry is updated by the end of year. The difference in updating times may blur the relationships between the variables. Thus we will apply the quarterly or monthly data to make a new ‘annual’ data, which can combine the information in the previous second half year and current first half year. For example, a new GDP index can be constructed by adding the quarter GDP in previous second half year and the current first half year.

New GDP (year<sub>n</sub>) = GDP in 3rd quarter (year<sub>n−1</sub>) + GDP in 4th quarter(year<sub>n−1</sub>)
+GDP in 1st quarter (year<sub>n</sub>) + GDP in 2nd quarter(year<sub>n</sub>)

In this way, the new GDP index can contain the information that corresponds to the data of market value. New GDP growth can also be obtained by calculating the percentage change of new GDP indices.

---

For factors that only annual data are available, the new data can be estimated by taking the average values in precious and current year. For example, new revenue growth can be constructed by

\[
new \text{ revenue growth (}\text{year}_n\text{)} = \frac{1}{2} (\text{revenue growth (}\text{year}_{n-1}\text{)} + \text{revenue growth (}\text{year}_n\text{)})
\]

This leads to the new variables that are summarized in Table 7:

Table 7 New constructed data contain information of the second half year and the first half year

<table>
<thead>
<tr>
<th>Name</th>
<th>Unit</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>New world GDP</td>
<td>Billion (USD)</td>
<td>Average of quarterly GDP</td>
</tr>
<tr>
<td>New GDP growth</td>
<td>%</td>
<td>Percentage change of new GDP indices</td>
</tr>
<tr>
<td>New crude oil price</td>
<td>USD/ Barrel</td>
<td>Average of monthly oil price in the second quarter</td>
</tr>
<tr>
<td>New interest rate</td>
<td>%</td>
<td>Average of annual real interest rate in USA</td>
</tr>
<tr>
<td>New RPK growth</td>
<td>%</td>
<td>Average of monthly RPK growth</td>
</tr>
<tr>
<td>New revenue growth</td>
<td>%</td>
<td>Average of annual revenue growth</td>
</tr>
<tr>
<td>New passenger growth</td>
<td>%</td>
<td>Average of annual passenger growth</td>
</tr>
<tr>
<td>New order</td>
<td>%</td>
<td>Average of annual order</td>
</tr>
</tbody>
</table>

Section 2.2 Data exploration

We have collected all the data in section 2.1 which we think may be of importance to aircraft market value. In [7], it has been mentioned that the relationships between aircraft market, global economy and airline industry are quite complicated. In order to find out the influential factors from this complicated background, we will make an exploration of available data to see the relationship between them.

Section 2.2.1 Market value versus type and age

It is natural that market values would be related to the specification of aircraft (type), as manufacturers make the price of a product according to the configuration of the aircraft. The most important characteristic of a model / type is its size. In Figure 2, we see the relationship
between the size of an aircraft (measured by seat capacity in typical configuration) and its market value.

Figure 2 shows that larger aircrafts are more expensive. The relationship between seat capacity and market value can be described well by a quadratic polynomial trend line. One can find the seat capacity of the eight aircraft types considered in this thesis in Table 2.

![Figure 2 Market value vs. seat capacity](image)

In Figure 2, seat capacity has explained most of the variance of the market value, but we can also see that some points (such as Airbus A330-300) deviate from the curve. The rest of the variance can be e.g. explained by the difference of the technology applied in each aircraft. Technology is another important characteristic of an aircraft type. It includes the core technology applied in the engine, avionics systems, and so on.

It is difficult to measure the technology of an aircraft. In [9], the year of first flight is used as a guide to the age of the technology. However, technology cannot be represented completely by the first flight year; the aircraft manufacturer sometimes uses the same technology in new products.

Despite of this, technology helps with explaining the variance in Figure 2 greatly. For example, market value of Boeing 777-200 is much higher than the trend line because of the advanced technology that it contains. Boeing 777 family has computer mediated controls; it is also
the first entirely computer-designed commercial aircraft; airlines have described this type as a comparatively fuel-efficient alternative to other wide-body jets.

It is quite natural that the older aircrafts are cheaper than new ones (all other things being equal). Older aircraft require an increasing amount of maintenance and consumes more fuel and thus has higher operational costs. When an aircraft cannot generate enough revenue to cover its operating cost, its economic life has come to the end [6].

Relation between age of an aircraft market value for A320-200 aircraft can be seen in Figure 3. Notice that smaller amount of data are available for older aircrafts, as it takes some time for an aircraft to ‘grow’ old after the beginning of its production. For example, in year 2001 only market values of A320-200 from age 0 to age 13 are available; while in year 2009 one may find the market value for aircraft which is 21 years of age.

![Figure 3 Market Value vs. age, Airbus A320-200](image)

In the same year for a given aircraft type, there is an apparently negative relationship between age and aircraft value. In Figure 3, the trend lines (2nd order polynomials) between markers fit the set of data quite well. We see that the determining coefficients $R^2$ are all higher than 0.99, which indicates a very good fit. For other aircraft types the similar figures can be obtained, indicating that age is another determining factor for market value of an aircraft. This conclusion supports the results given by Kelly in [6].
Another interesting feature can be observed in Figure 3. We see that the influence of age on aircraft market varies with years. The trend lines in year 2009 (yellow diamonds) is much lower than the one in year 2001 (green squares); in other words, compared to year 2001, older aircraft are relatively cheaper in year 2009, though the market value for new ones are similar in these two years. This phenomenon is caused by the different market situation in each year. It has been pointed out in [15] that when aircraft operators need to replace or expand their fleet in recession years of aircraft market, they are prone to buy newer aircraft instead of older ones.

Section 2.2.2 Market value and global economy

Aircraft values also depend on global economy. Vasigh and Erfani showed in [7] that GDP is an important indicator for the traffic demand, and thus important for aircraft market. In Figure 4 we compare historical trends of Global GDP and aircraft market values for two aircraft types.

![Figure 4 Historical trends of Global GDP and aircraft market](image)

The grey bars show global GDP (left axis), red solid line presents market value of Airbus A320-200 and blue dashed line the market value of Boeing 737-800. Market values of both aircraft types fluctuate in quite similar way: they suffered from a sharp down turn in year 2002, grew up steadily during year 2004 to year 2008, reached peak in year 2008, and fell down again in year 2009.
The economic developments can explain the pattern of market values well. The drop in year 2002 is caused by the 9-11 attacks in year 2001. As data of market value is updated in July, the impacts of 9-11 on aircraft market are observed only in year 2002. Year 2004 to year 2007 is a boosting period for the global economy, and thus the aircraft market can enjoy growth as well. The recession started in the end of year 2008, and the slowing down of the world economy in 2009 put the aircraft market into another low point.

Though we can see the influence of economy on aircraft market value, its relationship with GDP is not so strong. The linear correlation coefficient between GDP and market value of Airbus A320-200 is only $0.1546$ while the spearman’s correlation is $0.0287$. However as discussed before the annual GDP might not be the best indicator for the market values which are updated in July. If we check the correlation of new GDP (GDP in the second half of the previous year and the first half of this year), the results are quite disappointing as well with Pearson correlation coefficient equal to $0.071$ and Spearman’s correlation of $0.1071$. With the amount of data available the correlations are statistically insignificant.

Other variables in airline industry such as revenue, RPKs and passenger growth, though expected to have a strong correlation with aircraft market value, do not show a statistically significant correlation with market value either.

The reason for this is that external economic variables in global economy and airline industry influence the change of aircraft market value instead of market value itself. Essentially, it is the internal factors -age and type- have determined the underlying market value of an aircraft. When we compare aircraft market value with economic variables directly, the influence is covered greatly by the characteristic of the aircraft.

In Figure 4 we can see that aircraft market drops when GDP growth slows down. This observation motivates us to look at the growth of market values versus the economy growth rather than at market values versus economic indicators itself.

In the following section, we will focus on the market value growth to see interrelations between aircraft market, global economy and airline industry.

**Section 2.2.3 Market Value growth versus type and age**

Market Value growth (MV growth) is defined the percentage change of market value of an aircraft from year to year.

$$\text{MV growth(year } n\text{)} = \frac{\text{MV(year } n\text{)} - \text{MV(year } n-1\text{)}}{\text{MV(year } n-1\text{)}} \times 100$$ (2.1)
In this section, we will discuss the influence of internal and external factors on MV growth. Internal factors include type and age, while external factors refer to variables in the economic environment of aircraft market.

In Figures 5 and 6 the MV growth for different aircraft types are shown. In Figure 5, we can see the MV growth of two aircrafts with very different sizes. Airbus A319-100 is the smallest model while Boeing 777-200 is the largest one among the eight aircraft types we selected. The plots are quite similar which indicates that in contrast to MV the size is not a very influential factor for MV growth. The differences between MV growths of these aircrafts can be due to size or technology or other factors.

It looks like the variation of MV growth for both aircraft types can be attributed to the economic environment for aircraft market.

Figure 5  MV growth patterns for new aircraft: Airbus A319-100, Boeing 777-200
In Figure 6, we see MV growths of three aircraft types: Airbus A330-200, Boeing 767-300ER and Airbus A320-200. They are all comparable in sizes and in the technology (measured in the year where they were flown first time). However the market value of new Boeing 767-300ER behaves quite differently from the other two. It declined more than 15% in year 2002, and reached peak in year 2006; while the other two suffered smaller losses in year 2002 and reached peak in year 2005. It becomes difficult to explain this behavior and find out reasons for the differences of MV growth patterns.

From Figure 6 it becomes also apparent that aircraft market values in 1990s and in 2000s are very different. Aircraft market in 2000s suffered much bigger fluctuations. This conclusion is supported by research in [15] which confirms a cyclical behavior of business. Additionally aviation industry has changed greatly since year 2001 (9.11 attacks lead to many changes in the aviation industry [16]).

Since the aircraft market in 2000s is different, and that data set of airline industry is only available from year 2001 in the following analysis we only take into account data from 2001. This decision might be controversial (as we are left with a very small data history) but we believe that the behavior of the market in the near future will be more comparable to years 2000s than to
Analysis of available data

If long time predictions are needed one might consider re-quantifying the model to take into account also more stable years for market values.

Next we will check the influence of age on MV growth. As Airbus A320-200 is produced in year 1986, more data of older aircrafts are available. We will take this type as the example to compare MV growth patterns of aircraft at different ages.

![Figure 7 MV growth patterns for Airbus A320-200 at age 0, age 3, age 11 and age 12](image)

As we can see in Figure 7, the differences in MV growth patterns of age 0 and age 3 are quite small; also the differences between MV growth of age 11 and age 12 are not very significant. Hence for aircrafts of ‘similar’ ages MV growth patterns are comparable. We observe the same phenomena for other aircraft types as well.

As we can see that the difference of MV growth is quite small for aircraft at a similar age level, we can consider to put the aircraft which have similar MV growth pattern in the same group; the average values of MV growth of the group present for the MV growth level for all aircraft in this groups. In other words, aircraft in the same group are assumed to have the same market behaviors, the difference among them is caused by some random factors.

For a given aircraft type, the clustering procedures are as following: first we assume that aircraft 0 and aircraft 1 can be put in the same group; then for each year, calculate the average of
the MV growth of aircraft at age 0 and age 1, and consequently we can get the difference terms for each year for two aircraft. According to the assumption, this difference terms should come from some random factors. So we will consider that they follow a normal distribution with mean 0 and a constant variance.

In order to support the assumption, a test of the difference term need to be made. We will test if the difference term follows a normal distribution, and if the mean and the variance of the difference term is ‘small’ enough that we can accept that there is no big difference among the aircraft being analyzed.

Kolmogorov-Smirnov test is used for the test of normality of the distribution. We define the ‘small’ difference term according to the difference in the money: the mean of the difference term should be smaller than 0.01, and the difference in money caused by taking the average of the group should be smaller than 1 million with a probability 95%. With the normal distribution assumption and the std. deviation of the difference term, the 95% confidence interval for the difference term is calculated as following:

\[ \text{mean} - 1.96 \times \text{std deviation}, \text{mean} + 1.96 \times \text{std deviation} \]

In combination of the definition of MV growth, we can thus calculate the 95% confidence interval for difference in money. As long as the money difference is within 1 million, we will accept that the difference amount the MV growth of the aircraft is small.

If the test show that the assumption is accepted, then we will put the aircraft at age 0 and aircraft at age 1 in the same group, and consider if aircraft at age 2 can be included as well; if the test does not support the assumption, we will take put aircraft at age 0 in the original group, and take aircraft at age 1 as a the first member of a new group, and then consider if aircraft at age 2 can be put in this new group.

In this way, we get clustered groups in Table 8.

### Table 8 aircraft groups

<table>
<thead>
<tr>
<th>Name</th>
<th>Group</th>
<th>Variance of difference term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airbus A319-100</td>
<td>Group 1: age 0 to age 5</td>
<td>1.1162</td>
</tr>
<tr>
<td>Airbus A320-200</td>
<td>Group 1: age 0 to age 4</td>
<td>1.1435</td>
</tr>
<tr>
<td></td>
<td>Group 2: age 5 to age 10</td>
<td>1.4999</td>
</tr>
<tr>
<td></td>
<td>Group 3: age 11 to age 12</td>
<td>1.9554</td>
</tr>
<tr>
<td>Airplane Model</td>
<td>Group 1: age 0 to age 4</td>
<td>Group 1: age 0 to age 3</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td><strong>Boeing 737-800</strong></td>
<td>Group 1: age 0 to age 3</td>
<td></td>
</tr>
<tr>
<td><strong>Airbus A330-200</strong></td>
<td>Group 1: age 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 2: age 1 to age 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 3: age 3</td>
<td></td>
</tr>
<tr>
<td><strong>Airbus A330-300</strong></td>
<td>Group 1: age 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 2: age 1 to age 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 3: age 3 to age 5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 4: age 6 to age 8</td>
<td></td>
</tr>
<tr>
<td><strong>Boeing 767-300ER</strong></td>
<td>Group 1: age 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 2: age 1 to age 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 3: age 3 to age 7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 4: age 8 to age 13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 5: age 14</td>
<td></td>
</tr>
<tr>
<td><strong>Boeing 777-200ER</strong></td>
<td>Group 1: age 0 to age 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Group 2: age 2 to age 4</td>
<td></td>
</tr>
</tbody>
</table>

Notice that for the more expensive aircraft, the number of ages that can be bundled together is smaller. Groping aircrafts allows us to simplify the modeling of MV growth of aircraft in the following chapters.

**Section 2.2.4 Market value growth versus global economy and airline industry**

In Figure 8 we observe that there is a close relationship between global economy, airline industry and aircraft market. Vasigh and Erfani have discussed in [7] that, like any other physical asset, a crucial element determining the value of an aircraft is the relationship between supply and demand. Airline industry is the customer of aircraft market, and global economy provides the trading environment. Variables in these two fields are important predictive factors for aircraft market value.
We consider different indicators for the global economy: GDP growth, oil price change, CPI (inflation rate) and interest rate. We will firstly check which variables influence the MVG the most.

Scatter plots in Figure 9 show the relationships between MVG of new Airbus A320-200 and the four considered variables as well as the relationships between these variables. We can see that, there are strong positive relationship between MVG and GDP growth as well as oil price change; while the correlations between MV growth and CPI and interest rate are very week.
Figure 9 Scatter plots, global economy vs. MV growth of new Airbus A320-200

It is understandable there is a close relationship between GDP growth and MV growth. Economy is the main driving factor for air traffic demand. The rapid growth of world trade and international investment has led to growth in business travel, and growth in family incomes has contributed to an increase in tourism.

Slightly surprising can be to see that oil price changes also have a positive relationship with the aircraft market. One reason is the close relation between oil price and world economy, which we can also see from the scatter plots of oil price change vs. GDP growth in Figure 9. Strong correlation does not indicate a causal relationship, but we can still take oil price change as an important indicator for MV growth.

CPI measures the changes in price level of consumer goods and services by households. Probably due to aircraft costs (millions of dollars), the price of aircraft is not affected by inflation rate very much. 1 US dollar in year 1998 has the same buying power as 1.38 US dollar in year 2008. However, a new Boeing 767-300ER was worth 79 million (US dollar) in year 1998 and only 69.9 in year 2008. The inflation has little influence on the aircraft values.

Interest rate indicates the lending cost of capital. When buying an aircraft, aircraft operators need to balance the capital cost and operational costs. Older aircraft are cheaper,

---

19 MVG: market value growth; IR: interest rate; GDP refers to GDP growth and Oil refers to oil price change.
implying smaller money burden, but they lead to higher operational cost due to more maintenance and higher fuel consumption. However, the MV growth of an aircraft, as we can see in scatter plots in Figure 9, does not show the strong relationship with interest rate.

Airline industry is the customer of aircraft market [11]. The health of the commercial aircraft market is highly dependent on the state of the airline industry. Next we will consider variables in airline industry, including revenue growth of airline industry, RPK passenger growth and passenger growth. They are selected as they can indicate the expansion of airline industry [7].

The first row/column in the scatter matrix in Figure 10 presents the relationship between MV growth of new Airbus A320-200 and other variables. As we can expect, all these three variables show a strong positive correlation with MV growth. When airline industry is expanding fast, the demand for aircraft grows.

Revenue growth shows the capital flow condition for airline industry. We can expect that when there is more money available, aircraft operators are more willing to expand/replace their fleet. It is also easier for them to borrow money from banks.

Both RPK and passenger growth are measures of traffic demand. We can see that MV growth has the strongest relationship with these two variables, i.e. the air traffic demand.

Figure 10 Scatter plots, airline industry vs. MV growth of new Airbus A320-200

20 In Figure 10, each variable refers to the corresponding growth, e.g. revenue indicates revenue growth; MVG refers to market value growth
The picture presented so far for A320-200 are also made for other types of aircraft at each age and it turns out these five variables have a high correlation with MV growths. Thus we can be sure that variables in Table 8 will be important indicative factors for aircraft market in general.

**Section 2.2.6 Market value growth versus fleet**

Data of fleet may contain the information about the preference and popularity of an aircraft. We will still take Airbus A320-200 as the example to check the underlying relationships. Again, we will try to extract some information from the raw data of the fleet.

In Figure 11, we can see the historical trends of fleet and manufacture. From upper to lower, left to right, the variables for Y-axis are number in service, number in storage, operator number, backlog, net order and MV growth of Airbus A320-200 separately.

![Figure 11 Historical trends of fleet & manufacture, Airbus A320-200](image)

As we can see in the first plots in Figure 11 the number of aircraft in service (the number of aircraft flying) has been increasing over the past 10 years. It is natural to expect that, as air travel has been a large and growing industry for more than twenty years. In the past decade, air travel has grown by 7% per year [10]. We can expect that the fleet of airline industry will still
expand in the following years. But comparing it with the last plots representing MV growth, we can see that it is not a proper variable as an indicator for MV growth.

The second plot shows the number of aircraft in storage varies over years. Compare the pattern with MV growth in plot 6, it looks like that there is some negative relationship between number in storage and MV growth. In year 2007, when aircraft market was enjoying a good time, the number of aircraft put in storage declined greatly. It looks that the availability of used aircraft (stored) would have a negative impact on aircraft market.

In Table 9 we show product moment and rank correlation coefficients between MV growth and stored aircraft. We need to remember that the size of the data is only 10; p-values for testing the hypothesis of no correlation against the alternative that there is a nonzero correlation are present in Table 8 as well. If p-value is small, say less than a significance level 0.05, then the correlation is significantly different from zero.

In Table 9, we see that indeed the correlations are negative but given amount of data they are not statistically significant with the data size of 10.

Table 9 Correlation coefficient between MV growth and number of aircraft in storage, for new Airbus A320-200

<table>
<thead>
<tr>
<th>Type</th>
<th>correlation coefficient</th>
<th>p-value of test against hypothesis of zero correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson</td>
<td>-0.4667</td>
<td>0.2125</td>
</tr>
<tr>
<td>Spearman</td>
<td>-0.4461</td>
<td>0.2287</td>
</tr>
</tbody>
</table>

The number of operators keeps growing as well. This is also the result of continuously expanding airline industry.

Backlog grew fast during year 2004 to year 2008, and slowed down in year 2009, as net orders declined greatly in year 2009. Comparing the shape of historical trend of backlog/ net order with pattern of MV growth, we can hardly see any direct close relationship.

Though we are expecting that there should be a positive relationship between net order and MV growth, it is very disappointing to see that MV growth is not correlated with net order very much. Pearson’s coefficient of net order and MV growth is only 0.3648 (p-value of Pearson test equal to 0.3344) for Airbus A320-200. And Boeing 737-800, which is a newer and larger type, the Pearson’s coefficient between net order and MV growth is 0.2832 with p-value of significance test equal to 0.4602.
It seems that there is little information we can get from the data of fleet. In order to investigate more, we will check the correlation between MV growth and the change of fleet, i.e. the percentage change of each factor.

Figure 12 shows the scatter plots of MV growth with respect to the percentage change of each variable. The last plot in Figure 6, a variables used in [8] is also checked, which is defined as the ratio of order/current fleet number (current fleet number is the total of number in service and number in storage).

![Figure 12 scatter plots of growth of fleet & manufacture vs. MV growth, For Airbus A320-200](image)

Among the six scatter plots, plots of MV growth vs. change of number in storage, MV growth vs. backlog growth and MV growth vs. order growth seem to show some correlations. In other plots, the points look quite random.

In order to see if the similar conclusions can be drawn for other aircraft types, we will also check Boeing 767-300ER, which is a bigger aircraft model. The scatter plots are show in Figure 13.
The results are quite disappointing: there seems to be no correlation at all in plots in Figure 13.

We also think about the possibilities that older aircraft may have a stronger correlation with the condition of fleet, so the same observation in Figure 13 is made again between MV growth of 14-year-old aircraft and fleet & manufacture growth; the results are present in Figure 14.

As we can see in Figure 14, the scatter plots still indicates that there is hardly any relationship between variables of fleet and MV growth of 14-year-old Boeing 737-300ER. The scatter looks quite random and are shown in Figure 14.
In fact, after a long and careful exploration of data of fleet, a disappointing conclusion is that one can hardly find a common variable which can indicate the influence of fleet of an aircraft type and its market value for all aircraft types.

For example, for Boeing 767-300ER, there is a strong negative correlation between number in storage and MV growth with Pearson’s coefficient -0.6628; while for Boeing 737-800, this coefficient is only -0.1730. Considering that Boeing 767-300ER is an earlier model, the availability of more old aircraft in service may have a bigger impact on its aircraft market; however, for Airbus A320-200, which is an earlier model than Boeing 767-300ER, the p-value of testing null hypothesis of zero correlation between number in storage and MV growth is as high as 0.2125 (see Table 10).

Reasons for the storage of aircrafts are discussed in [15]. Some aircraft are stored and will be available for sale, some will be put into use again (in better times), and some will retire forever because their maintenance cost surpasses the revenue they can generate. The aircrafts that reached the end of their economic life cannot be sold or used anymore. Hence it is quite difficult to see how many of aircrafts in storage still hold economic values. Information contained in the raw data is not sufficient for use in building a model.
In conclusion, it is difficult to find any indicative factor for MV growth from fleet data.

Section 2.3 Conclusion

In this chapter information contained in the available data has been analyzed. The following conclusions can be drawn:

The determining factor for aircraft market value is its internal characteristic—type and age. Generally, the bigger and newer the aircraft is, the more expensive it is. Since aircraft is regarded as a ‘machine’ to generate revenue for airline operators, its value is based on the underlying economic value it may produce. Specification of aircraft type implies its ability to carry passengers, while age is related with the operational cost. Bigger aircraft are more expensive as they can carry more passengers at one time. Moreover due to landing capacities of airports large aircraft are more economical. Older aircraft are cheaper as their operational cost is higher.

Aircraft values also depend on the economic environment in the time of sale. Aircraft market in 2000s is different from the one in 1900s. After the September 11 attacks in 2001, significant change has happened within airline industry and impacted aircraft market. For the analysis we will use data from year 2001 to year 2010.

It is the growth of market value that is strongly correlated with the economic variables. As the internal properties of an aircraft have determined its market value largely, volatility in economic environment causes the fluctuation in aircraft market. Both GDP growth and change in oil price have a strong positive relation with MV growth. Aircraft market suffered recession when world economy growth slows down.

Airline industry is the customer of aircraft market. Naturally, variables in airline industry are all influential factors for aircraft market. We can see that revenue growth and traffic growth are highly correlated with MV growth. They are all important factors when modeling growth of market value.

Information provided by data of fleet is very vague. We can hardly find a common factor to explain the variance of MV growth for all aircraft types. So in the end, none of variables in the data of fleet can be used.

In Table 10 we show the most influential variables for MV growth together with the Pearson correlation they have with MV growth as well as the p-value of the test of significance for the correlation.
Table 10 indicative variables for model construction

<table>
<thead>
<tr>
<th>Field</th>
<th>Name of variables</th>
<th>Pearson’s coefficient with MV growth (^{21})</th>
<th>p-value against zero correlation hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>General economy</td>
<td>GDP growth</td>
<td>0.7556</td>
<td>0.0115</td>
</tr>
<tr>
<td></td>
<td>Oil price change</td>
<td>0.7803</td>
<td>0.0077</td>
</tr>
<tr>
<td>Airline industry</td>
<td>Revenue growth</td>
<td>0.8724</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>RPK growth</td>
<td>0.8727</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>Passenger growth</td>
<td>0.8778</td>
<td>0.0008</td>
</tr>
</tbody>
</table>

We can see that, these variables in general economy and airline industry in fact contain a similar information of demand for aircraft in the current economic environment. The GDP growth is the driven factor for air traffic, and the revenue growth and traffic growth (passenger growth and RPK growth) indicates the health of the airline industry, which is the main consumer of commercial aircraft.

The variables in Table 10 show that there are strong positive correlations between these variables and MV growths of all aircraft regardless of type and age. It indicates that the increased demand will always boost the market value growth for all aircraft, and the demand is the main factor for the change of market value of commercial aircraft, as we have not found another variable which has a strong correlation with all aircraft.

\(^{21}\) Here we use MV growth of new Airbus A320-200
Chapter 3  Regression model of the market value growth

In the scatter plots in section 2.2, we can see linear relationships between MV growth and the variables in Table 9. It is quite natural to apply ordinary linear regression (OLS) to build a model of MV growth. We first discuss the classical linear regression model for MV growth and its deficiency. Then the principle component regression is applied to circumvent problems of highly correlated explanatory variables in the classical regression model.

We follow the idea presented in [8] and build one regression model for the market value of a given aircraft type with coefficients that depend on the age of the aircraft.

Section 3.1 Classical linear regression

We will model the MV growth of new Boeing 737-800 with the indicative variables in Table 10. Considering the strong correlation between MV growth and the five indicative variables, it is quite natural to apply multiple linear regressions to describe the relationship between the explanatory variables (variables in Table 9) [13] and the response variable (MV growth of new Airbus A320-200) by fitting a linear equation to observed data. We will use the following notations:

\[ Y \text{ -- MV growth of new Boeing 737-800; } \\
X_1 \text{ -- GDP growth; } \\
X_2 \text{ -- Oil price change; } \\
X_3 \text{ -- Revenue growth; } \\
X_4 \text{ -- RPK growth; } \\
X_5 \text{ -- Passenger growth. } \\
\]

It is assumed that the model has the following form

\[ Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon \tag{3.1} \]

where \( \varepsilon \) is normally distributed with mean zero and standard deviation \( \sigma \) (denoted as \( \varepsilon \sim N(0, \sigma^2) \)) and \( \beta_i \quad i=1,\ldots,5 \) are parameters in the model.
The size of observed data set is 10 (year 2001 to year 2010). By applying OLS method, we can get the estimated values of the coefficients of the best-fitting line, which is calculated by minimizing the sum of the squares of the vertical deviations from each data point to the line (for example, if a point lies on the fitted line exactly, then its vertical deviation is 0).

### Table 11 standard description for regression model

<table>
<thead>
<tr>
<th>Model</th>
<th>$R^2$</th>
<th>Adjusted $R^2$</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>F-Statistics</td>
</tr>
<tr>
<td>0.901</td>
<td>0.778</td>
<td>7.302</td>
<td>0.039</td>
</tr>
</tbody>
</table>

Table 11 shows the goodness of fit of the regression model. The determining coefficient $R^2$ is a common measure of how well a regression model fits the data. The closer $R^2$ is to 1, the better the model fits. For model of MV growth of new Boeing 737-800, $R^2$ is equal to 0.901, which indicates that over 90% of the variance in MV growth is explained by the explanatory variables in the model.

The F-statistics for the null hypothesis that $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$ follows distribution $F(5,5)$. F-statistics is equal to 7.302, and $Pr(>7.302)=0.039$. With significance level 0.05, the p-value for the F-test provides strong evidence against the null hypothesis. In other words, at least one of the explanatory variables is linearly related to MV growth, which supports that this linear model can be applied to fit the observed data.

We also need to test the properties of the residuals to see if there is independence and normality in the residuals. Here we use Durbin-Watson statistics and Kolmogorov-Smirnov test to test the independence and normality and it turns out that both assumptions are accepted for this model.

Next we want to check if the contribution of each variable to the model are statistically significant. We will do this by check the significance test for each variable in Table 12.

### Table 12 coefficients and their statistics

<table>
<thead>
<tr>
<th>Model</th>
<th>B</th>
<th>Std. Error</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Constant)</td>
<td>-6.736</td>
<td>2.461</td>
<td>-2.738</td>
<td>.052</td>
</tr>
<tr>
<td>$X_1$</td>
<td>-.053</td>
<td>.794</td>
<td>-.067</td>
<td>.950</td>
</tr>
<tr>
<td>$X_2$</td>
<td>.056</td>
<td>.050</td>
<td>1.127</td>
<td>.323</td>
</tr>
<tr>
<td>$X_3$</td>
<td>.188</td>
<td>.448</td>
<td>.419</td>
<td>.696</td>
</tr>
</tbody>
</table>
In Table 12, we can see the estimated values and statistical tests for the coefficients. Each significance test is to test against the null hypothesis $\beta_i = 0, i = 1,2, \ldots, 5$ separately. According to the test results in Table 12, none of the five explanatory variables has a significant contribution to the model.

The results in Table 11 and Table 12 seem unintuitive. The results says that the whole model fits the data well, even though none of the explanatory variables has a statistically significant impact on $Y$. The reason for this is that the explanatory variables selected in Chapter 2 are all highly correlated with each other, which can be seen in Table 13.

### Table 13 Correlation matrix and significance test

<table>
<thead>
<tr>
<th></th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1.000</td>
<td>0.582</td>
<td>0.757$^*$</td>
<td>0.626$^*$</td>
<td>0.749$^{**}$</td>
</tr>
<tr>
<td>$X_2$</td>
<td>0.582</td>
<td>1.000</td>
<td>0.690$^*$</td>
<td>0.708$^*$</td>
<td>0.624$^*$</td>
</tr>
<tr>
<td>$X_3$</td>
<td>0.757$^{**}$</td>
<td>0.690$^*$</td>
<td>1.000</td>
<td>0.828$^{**}$</td>
<td>0.896$^{**}$</td>
</tr>
<tr>
<td>$X_4$</td>
<td>0.626$^*$</td>
<td>0.708$^*$</td>
<td>0.828$^{**}$</td>
<td>1.000</td>
<td>0.875$^{**}$</td>
</tr>
<tr>
<td>$X_5$</td>
<td>0.749$^{**}$</td>
<td>0.624$^*$</td>
<td>0.896$^{**}$</td>
<td>0.875$^{**}$</td>
<td>1.000</td>
</tr>
</tbody>
</table>

$^{**}$. Correlation is significant at the 0.01 level (2-tailed).

$^*$. Correlation is significant at the 0.05 level (2-tailed).

In Table 13, we can see that almost all the variables are have statistically significant correlations (with 0.05 significance level) with each other (except for correlation between variable $X_1$ and $X_2$). They all convey essentially the same information. In this case, none may contribute significantly to the model after others are included. But together they contribute a lot.

This co-linearity among the $X$ variables may be not a big problem when the model is used just for prediction. However, when we try to understand how each $X$ variable impact $Y$, this causes some trouble. First the individual p-value is quite misleading; you cannot judge the importance of the variables from the p-value now as the p-value can be high even though the variable is important. Second, the confidence interval on the regression coefficients will be very wide, and even include zero, which means that one cannot be confident whether an increase in the $X$ variable is associated with an increase or a decrease in $Y$. When excluding a variable (or
adding a new one) we can change the coefficients dramatically, and even change their signs because of the wide confidence interval [14].

<table>
<thead>
<tr>
<th>Model</th>
<th>95.0% Confidence Interval for coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
</tr>
<tr>
<td>1 (Constant)</td>
<td>-13.568</td>
</tr>
<tr>
<td>$X_1$</td>
<td>-2.257</td>
</tr>
<tr>
<td>$X_2$</td>
<td>-.083</td>
</tr>
<tr>
<td>$X_3$</td>
<td>-1.055</td>
</tr>
<tr>
<td>$X_4$</td>
<td>-.973</td>
</tr>
<tr>
<td>$X_5$</td>
<td>-1.517</td>
</tr>
</tbody>
</table>

The confidence interval in Table 14 provide evidence for discussion about problem caused by co-linearity. Zero is included in confidence interval for each variable. The sign of the coefficient is not reliable for analysis of the impact of the explanatory variables to the model now.

As we can see from the analysis above, the classical analysis is not appropriate for available data.

A common solution to co-linearity is to remove some of the correlated variables as they measure essentially the same thing. It is difficult to make the decision of selection just by analyzing the logical relationships of the explanatory variables to the model, as they are all important (based on the analysis in Chapter 2); and as they are highly correlated with each other, just removing one or two cannot solve the problem.

One may think of stepwise regression [15] for the selection of appropriate variables for model prediction. Applying stepwise regression (with criteria of probability-of-F-to-Enter <=0.050, probability-of-F-to-remove >= 0.100) only passenger growth is included and the determining coefficient of the model is 0.842, lower but still high.

However, stepwise regression does not provide a good solution to our problem. There are 8 aircraft types at different ages, and stepwise selection have to be made for model of each aircraft type at a certain age. Unfortunately, generally, the selections of predictive variables are different. For example, the stepwise regression picks the passenger growth for model of new Airbus A320-200 and revenue growth for Boeing 777-200 (with criteria of probability-of-F-to-Enter <=0.050, probability-of-F-to-remove >= 0.100). This will make it quite confusing when
interpreting the contribution of the variables to the model; we hope that there are some common variables that can be used for all aircrafts.

Another way to solve co-linearity is to find a way to combine the variables. Principal component analysis thus can be applied to handle this problem.

### Section 3.2 Principal component regression

In statistics, principal component regression (PCR) is a regression analysis that uses principal component analysis (PCA) when estimating regression coefficients [16]. It can be used to handle the problem when the explanatory variables are close to being collinear.

In PCR instead of regressing the dependent variable on the independent variables directly, the principal components of the independent variables are used. Though the number of principal components can be equal to the number of original variables, generally only the components with the highest variance or big eigenvalues are selected.

Following this idea, we can see that there are three steps in PCR [17]:

1. The first step is to run a principal component analysis on the table of the explanatory variables;
2. Applying the ordinary least squares (OLS) regression to get parameters for the selected components;
3. The parameters of the model are computed for the explanatory variables.

As long as we have observed values for the selected component (calculated from the observed value of the explanatory variables), we are able to make prediction for the dependent variables (MV growth). The idea of PCR model can be seen from the following chart:

![Diagram](image-url)
We are going to find out the underlying principal component by applying principal component analysis first and then get the final linear model based on selected components.

1. Generating principal components

Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett’s test of sphericity are two tests for the suitability of data for principal components.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy is a statistic that indicate the proportion of variance in our variables that might be caused by underlying factors. It is a index between 0 and 1, and high value (close to 1) generally indicate that PCA can be used; if the results is less than 0.50, then PCA may not be helpful very much [18].

Bartlett’s test of sphericity tests the hypothesis that your correlation matrix is an identity matrix, which would indicate that your variables are unrelated and therefore unsuitable for structure detection [18].

<table>
<thead>
<tr>
<th></th>
<th>Kaiser-Meyer-Olkin Measure of Sampling Adequacy</th>
<th>Bartlett's Test of Sphericity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.822</td>
<td>Approx. Chi-Square</td>
</tr>
<tr>
<td></td>
<td></td>
<td>df</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sig.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>36.798</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.000</td>
</tr>
</tbody>
</table>

As we can see in Table 14, results of both tests support that PCA can be used for this problem. Statistics of KMO is equal to 0.888, close to 1; p-value of Bartlett’s Test is smaller than 0.001, indicating that the variables are not uncorrelated.

PCA is based on the covariance matrix and thus not invariant to the scale of the variables. The first element will be dominated by the the variable with highest variance. So first we would like to see the scale of the five predictive variables. In Figure 15, we can see the variability of the variables from the box plot.

For each box, the central mark is the median, the edge of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually as ‘+’. The bigger box indicates a larger variability of the variable.
It is easy to see that there is substantially more variability in the oil price change than in GDP growth. The length of the 50% percentile box for oil price change is around 40 while for GDP growth it is less than 10. In this case, we would consider that using correlation matrix is more proper for our problem. Correlation matrix in fact is the covariance matrix of standarized variable, where the standarized variable is caculated by

\[ ZX_i = \frac{X_i - \bar{X}_i}{\sigma_{X_i}} \]  

where \( X_i \) is the mean of \( X_i \) and \( \sigma_{X_i} \) is the std. variance of \( X_i, i = 1, \ldots, 5. \)

The eigenvalues of the correlation matrix of the five variables can be seen in the second column in Table 15.

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial Eigenvalues</th>
<th>Extraction Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>% of Variance</td>
</tr>
<tr>
<td>1</td>
<td>3.948</td>
<td>78.969</td>
</tr>
<tr>
<td>2</td>
<td>0.459</td>
<td>9.189</td>
</tr>
<tr>
<td>3</td>
<td>0.373</td>
<td>7.467</td>
</tr>
<tr>
<td>4</td>
<td>0.142</td>
<td>2.831</td>
</tr>
<tr>
<td>5</td>
<td>0.077</td>
<td>1.544</td>
</tr>
</tbody>
</table>
In Table 16, we can also see the percentage of the variance explained by each component. The first component can explain to 78.969% of the total variance, while the first one and second one together can explain 88.158% of the variance. The rest of the components explained little. The form for the principal component is as following.

<table>
<thead>
<tr>
<th>Component</th>
<th>ZX₁</th>
<th>ZX₂</th>
<th>ZX₃</th>
<th>ZX₄</th>
<th>ZX₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component1</td>
<td>0.4193</td>
<td>0.4042</td>
<td>0.4751</td>
<td>0.4598</td>
<td>0.4728</td>
</tr>
<tr>
<td>Component2</td>
<td>-0.5202</td>
<td>0.7886</td>
<td>-0.1339</td>
<td>0.1716</td>
<td>-0.2452</td>
</tr>
<tr>
<td>Component3</td>
<td>0.6690</td>
<td>0.4201</td>
<td>-0.1326</td>
<td>-0.4922</td>
<td>-0.3406</td>
</tr>
<tr>
<td>Component4</td>
<td>-0.3024</td>
<td>0.0594</td>
<td>0.7573</td>
<td>-0.5756</td>
<td>0.0162</td>
</tr>
<tr>
<td>Component5</td>
<td>-0.1201</td>
<td>0.1863</td>
<td>-0.4066</td>
<td>-0.4307</td>
<td>0.7746</td>
</tr>
</tbody>
</table>

X₁,X₂,...X₅ is positively correlated with each other. As a result, we can see that component 1 is some kind of average index of X₁,X₂,...X₅, as the loadings are quite similar to each other. Component 2 is mainly about the shortage of oil and low state of economy. Component 3 presents the different states of economy and airline industry. The remaining components are not as easily identified.

2. Make a selection of the components

A Scree plot can make it more clear for this observation in Figure 16. The eigenvalue of each component in the initial solution is plotted in their decreasing order in the scree plot, and we can see that the plot looks like the side of a mountain in Figure 16.
For a regression model of prediction, we need to make a selection of the components which have a significant contribution to the model. We have seen that the eigenvalues of the first component is as high as 3.946 and for other components the eigenvalues are lower than 1 from the information in Table 16 and Figure 15. The first component can explain to 78.969% of the total variance, and the other components can explain less than any of the original variable. As the main purpose is to build a prediction model first, we will only select the first component as the independent variable for the prediction model.

3. Building the regression model with respect to the selected component

The principal component regression has the following form:

$$ Y = \gamma_0 + \gamma_1 \times \text{component} + \epsilon $$ \hspace{1cm} (3.2)

where $\epsilon \sim N(0, \sigma^2)$.

The parameter $\gamma_0, \gamma_1$ can estimated applying OLS method.

With the principal component coefficients in Table 16 and the calculation of standardized variable, we can get the final form of the regression model as following

$$ Y = \gamma_0 + \gamma_1 \times (c_1 \times ZX_1 + c_2 \times ZX_2 + c_3 \times ZX_3 + c_4 \times ZX_4 + c_5 \times Z_5) + \epsilon $$ \hspace{1cm} (3.3)
where \( \epsilon \sim N(0, \sigma^2) \), \( c_i \) is the loading coefficient (Table 17) of the first component for \( zX_i \), \( zX_1 \) is the standardized variable for \( X_1 \).

### Section 3.3 Test of PCR model

We need to see how well the model fits the data, which is mainly checked by the determining coefficient \( R^2 \). On the other hand, it is necessary to see if the assumption of the regression model is satisfied by the data. The most important assumption about the regression model is about the noise term \( \epsilon \), which should follow a normal distribution with mean 0 and constant variance \( \sigma^2 \). In order to validate our model, we need to analyze the residuals. We show as an example the model of MV growth for new aircraft Boeing 737-800.

**Table 18 coefficients for PRC**

<table>
<thead>
<tr>
<th>Model of MV growth of Boeing 737-800</th>
<th>( \gamma_0 )</th>
<th>( \gamma_1 )</th>
<th>( R^2 )</th>
<th>P-value of F-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.323</td>
<td>7.373</td>
<td>0.890</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

The result in Table 17 show that 88% of the variance in the dependent variable can be explained by the model. The p-value of F-statistics for the model provides a strong evidence that the model fits the data well. From the sign of the estimated value of \( \gamma_1 (7.373) \), we know that the component has a positive contribution to the model, which is quite reasonable from the interpretation of the first component. The first component is considered as an average index of the five explanatory variables, and all the five have a positive correlation with MV growth.

**Table 19 Determining coefficient for models of new aircraft**

<table>
<thead>
<tr>
<th>New aircraft</th>
<th>( R^2 )</th>
<th>New aircraft</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airbus A319-100</td>
<td>0.8006</td>
<td>Airbus A330-200</td>
<td>0.8023</td>
</tr>
<tr>
<td>Airbus A320-200</td>
<td>0.8800</td>
<td>Airbus A330-300</td>
<td>0.8309</td>
</tr>
<tr>
<td>Airbus A321-200</td>
<td>0.8483</td>
<td>Boeing 767-300ER</td>
<td>0.7582</td>
</tr>
<tr>
<td>Boeing 737-800</td>
<td>0.8900</td>
<td>Boeing 777-200</td>
<td>0.8293</td>
</tr>
</tbody>
</table>

In Table 18, we can see that the determining coefficients \( R^2 \) are all above 0.8 except for Boeing 767-300ER, which indicates that the demand index fits the data well, but cannot explain the market value growth of Boeing 767-300ER quite well.
Next we will make the residual analysis to see if the three assumptions of the regression model can be accepted based on the data [19].

First we would check the constant variation across the data. Figure 17 presents the scatter plots of standardized residuals versus standardized predicted values for model of Boeing 737-800. One can hardly see any pattern from the 10 points. In other words, we will accept that the error term in the model has a constant variance not related to predicted values.

![Figure 17 Standardized residuals versus standardized predicted values, Boeing 737-800](image)

The same test is performed for models of other aircraft as well. And it turns out that for all of them, the constant-variance assumptions can be accepted.
Figure 18 tests the independence of the residuals. Xlabel shows the residuals of year \(_t\) while ylabel shows the residuals of year \(_{t+1}\). From the plots we can see that the points are scattered randomly. We can take this as a strong indication that residuals are independent of each other.

We also use Durbin-Watson statistics to test the independence of residuals. The Durbin-Watson test is used to test if the residuals are uncorrelated, against the alternative that there is autocorrelation among them. Small values of P value indicate that the residuals are correlated. For model of Boeing 737-800, the Durbin-Watson statistics is equal 2.78, which support the conclusion in Figure 17 that the the residuals are uncorrelated.

However, when we perform the Durbin-Watson statistics for all the models, there is a small parts of them do not accept the hypothesis of non-correlation. These models are for Airbus A320-200 from age 0 to age 4, and Airbus A330-200 from age 0 to age 2, 8 out of the 61 aircraft modelled.

It does not imply that the linear model is wrong. First it is really difficult to tell if the auto correlation conclusion is a truth or just a coincidence because of the small size of the training data. For Airbus A320-200, not only its direct competitor Boeing 737-800, but also its families A319-100 and A321-200 do not suffer from the problem of the Durbin-Watson test.
Of course we can still handle this auto correlation problem with a autoregresssive model of the residuals. However, we need to keep it in mind that only ten years’ data is available, and this problem is only for a small part of the 61 aircrafts analyzed. So instead of persisting with the theory, here we would tolerate this flaw in our model.

Figure 19 is the Q-Q plot of the residuals. It is a quantile-quatile plot of the residuals versus theoretical quantiles from a normal distribution, used to test the normality of the residuals. As we can see, the points almost follow the straight line well, indicating that the residuals follow a normal distribution.

![Q-Q Plot of Sample Data versus Standard Normal](image)

**Figure 19 Q-Q plot of residuals**

We also performs a Kolmogorov- Smirnov test of the default null hypothesis that the residual comes from a distribution in the normal family, against the alternative that it does not come from a normal distribution. It turns out that the null hypothesis is accepted at 5% significance level, which is consistent with the Q-Q plot results.

Kolmogorov- Smirnov test is performed for all the models, and for all the aircraft, it is accepted that the residuals follow a normal distribution.

Based on the residual analysis, we can accept that the model is validated (despite of some flaws of the autocorrelation for some aircrafts) we will use this model for data fitting and prediction.
Section 3.4 PCR model - results

In this section we will check in details how the model fits the data for each aircraft. We will first take Boeing 737-800 as an example.

In Figure 20, one can see the results of PCR model versus truth for new Boeing 737-800 from year 2001 to 2011. The blue spots represent the data (truth), and the solid green line shows the fitting of the model; and the dashed red lines gives the 95% confidence intervals of the prediction. One need to take care that data from year 2010 to year 2011 is used for obtaining the parameters of the model, and data in year 2011 is for prediction.

Figure 20 PCR model versus truth, new Boeing 737-800

As we can see, the model can follow the trend of truth, but does not predict well in some years (e.g. year 2002).

Confidence interval gives us more confidence about the estimation. For example, in year 2009, the estimation of the market value of Boeing 737-800 is $-3.2811$, and we are 95% confident that the true value will fall in the interval $[-20.6937, -5.2413]$. As we can see, all points of truth fall in the 95% confidence interval of predicted values; in other words, there is no outliers for model of Boeing 737-800.

However, the confidence interval is quite wide; the length of the intervals is between $[13,6689, 15,4524]$. For a new Boeing 737-800 which is worth at least 34 million USD (in year
2002), this indicates at least 4.6474 million uncertainty in the estimation of market values. The uncertainty in the estimation in the model is quite big.

Here we can also compare the results between PCR model and classical linear model. We will use two kinds of error measure to check the fitness of the model prediction. One is mean absolute value (MAE), which measure the average of the forecast errors; the other one is maximum absolute value (MAXE), which test the large errors in the forecasting.

**Table 20 error comparisons between model of classical linear regression and PCR model for new Boeing 737-800**

<table>
<thead>
<tr>
<th>Method</th>
<th>$R^2$</th>
<th>MAE</th>
<th>MAXE</th>
<th>Error in 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical linear regression</td>
<td>0.9013</td>
<td>1.8924</td>
<td>4.2395</td>
<td>0.5715</td>
</tr>
<tr>
<td>PCR</td>
<td>0.8900</td>
<td>2.1928</td>
<td>3.7118</td>
<td>0.4586</td>
</tr>
</tbody>
</table>

Table 20 shows the comparison between model of classical linear regression and PCR model for new Boeing 737-800. We can see that the determining coefficient $R^2$ of the classical linear regression is a little bit higher than PCR model. Notice that there are five explanatory variables in classical linear regression, while there is only one explanatory variable (the first principal component) in PCR model. It indicates that the variable we used in PCR can almost represent the information as well as the five explanatory variables in classical linear regression.

The MAE indicates that the classical linear regression fits the data a little bit better than PCR, which we can expect from the higher $R^2$; the MAXE and the forecasting error in 2011 turns out that the PCR does a little bit better job.

In order to have a full view of the performance of these two methods, we will check the results for all aircraft. For each aircraft, one can obtain the MAE, MAXE and error in 2011 seperately for two methods. We want to check the average performance, and the most convenient way is to calculate the the average MAE, MAXE and absolute error in 2011 over the 61 aircraft we modelled.

**Table 21 Average error comparisons between model of classical linear regression and PCR model for all aircraft**

<table>
<thead>
<tr>
<th>Method</th>
<th>Average MAE</th>
<th>Average MAXE</th>
<th>Average absolute Error in 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical linear regression</td>
<td>2.9059</td>
<td>7.7888</td>
<td>2.2535</td>
</tr>
<tr>
<td>PCR</td>
<td>3.4943</td>
<td>8.3292</td>
<td>2.4485</td>
</tr>
</tbody>
</table>
The results in Table 21 again show that, the classical linear regression does a little bit estimation of the market value growth than PCR method in average. For each measure of the error, the classical linear regression has a smaller average error.

However, there is only one explanatory variables in PCR, and the results in Table 20 and Table 21 indicates that this variable can represent the information contained in the five explanatory variables very well. Though for prediction, there is no big problem with classical linear regression method, it causes lots of confusion when explaining the contribution of the variables to the model. For the principal component used for regression, however, there is a clear positive relationship between the component and the MV growth for all aircraft.

We can also see a practical meaning of the component from its form: it is the average of the five explanatory variables, indicating the demand level from the economic environment. Combining the positive correlation between the component and the MV growth for all models, it tells a basic fact of the aircraft market: the demand is the driven factor for the market value growth for all aircraft, regardless of age and type.

So despite that classical linear regression does a little bit better fitting for the data, we will still use PCR method for regression; the principal component we use is called the demand factor.

Section 3.5 Incorporating the age in the model

When applying PCR, one has to estimate the parameters $\gamma_0, \gamma_1$ separately for the 61 aircraft we have. We are wondering if there is some way to include more parameters about the characteristics of the aircraft in the model, and we can have a more completed model which can simplifying the process of calculating work.

In this section, we will show two ways to reducing the number of models we need to handle. One way is to model the relationship between coefficients $\gamma_0, \gamma_1$ and age; another way is to ignore the small difference between the similar aircraft and get a general estimation of the MV growth.

Section 3.5.1 Varying coefficient PCR model

Data of Boeing 767-300ER at age 0 to age 14 is available, which is quite sufficient for discussion about models for different ages. We will take Boeing 767-300ER as the example to show to the methods.
There are two parameters $\gamma_0, \gamma_1$ that we need to model. $\gamma_0$ is the coefficient for the constant, and $\gamma_1$ is the coefficient for the demand factor. We will observe the relationship between $\gamma_0, \gamma_1$ and age of the modeled aircraft from the scatter plots in Figure 21.

\[
\gamma_0(a) = -0.01a^2 - 0.058a - 1.82, \quad R^2 = 0.9972
\]

\[
\gamma_1(a) = 0.0034a^2 - 0.072a^2 + 0.68a + 8.02, \quad R^2 = 0.9945
\]

**Figure 21 Relationship between coefficients of PCR model and age**

As we can see in Figure 21, the coefficient of constant $\gamma_0$ decreases with respect with age (the left scatter plot), and the coefficient of demand factor $\gamma_1$ increases with respect with age (the right scatter plots). It conveys clear information that market value growth of older aircrafts are more sensitive to the market condition. In other words, we can expect a bigger fluctuation of the market value growth for older aircraft over the past years. The results of variance support this conjecture: from year 2001 to year 2010, the variance of MV growth for 0-year-old Boeing 737-600ER is 80.23 while the variance of MV growth for 14-year-old Boeing 737-800 is 271.85.

In other words, the young aircraft can maintain their value better. For investment, the older aircraft may be a better option as the values of older aircraft decline faster (in terms of percentage) in bad years, and increase more (percentagewise) in the good times of the aircraft market.

In Figure 21, we can see the sensitivity of the market value growth of older aircraft directly. Year 2002 and year 2009 is the recession time for aircraft market because of the 9-11
attack and 2009 economy recession; In year 2007, the growth of aircraft market value almost reaches the peak, while the year 2008, the growth of market value begin to slow down. Figure 22 presents the behavior of the market value growth for aircraft Boeing 767-300ER at each age in these four years through the scatter plots of MV growth and age.

As we can see, in year 2002 and year 2009, the market value of older aircraft declines much more than the young aircraft percentagewise. The new Boeing 767-300ER lost -18.4783% of the market value in year 2002, while the 14-year-old Boeing 767-300ER lost -28.6713%; in year 2007, the market value of new-Boeing 767-300ER increased by 2.7306%, and the the 14-year-old Boeing 767-300ER increased by 7.4205%. The market conditions have a stronger impact on the market value of older aircraft.

What’s more, both trends in Figure 21 are so clear, that we can use polyniminal fitting the model the relationships. In other words, the coefficient can be regarded as the function of age. As we can see from the polynomial form above the plots, the determining coefficient for both fitting is higher than 0.99.

We can even detect the change of the aircraft market change from the relationship between MV growth and age. As we can see in year 2008 (the left lower scatter plot), the aircraft values still kept growing, but the growth of older aircraft has slowed down greatly, and the 14-year-old aircraft value almost has stopped growth. It indicates that the market is going to
experiencing some turning point. The market value growth of older aircraft are more sensitive to the change of market condition as we can see.

However, Airbus A330-300 is a special case for this conclusion. The market value growth of older Airbus A330-300 is more stable than the younger ones. And as we can see in Figure 23, the correlation between coefficient of demand factor $\gamma_1$ and age is not positive. From model of aircraft at age 1, the value of coefficient $\gamma_1$ began to decrease with respect with age.

It is difficult to find out the reasons for this. Its similar family aircraft type Airbus A330-200, and its direct competitor Boeing 777-200, all follow the regular pattern that models of older aircraft have a bigger values of coefficient $\gamma_1$. We can only take this as an irregular phenomenon.

Figure 23 coefficients versus age for Airbus A330-300

Despite of this, it is not a problem for that we can model the coefficients $\gamma_0, \gamma_1$ with a polynominal function with respect to age. We have showed the function for coefficients $\gamma_0, \gamma_1$ for models of Boeing 767-300ER in Figure 20. We can thus include age into the original PCR model by having coefficients varying with age as following:

$$MV\ growth\ (a, x) = (-0.01a^2 - 0.058a - 1.82) + (0.0034a^3 - 0.072a^2 + 0.68a + 8.02)x + \varepsilon$$

where $a$ is the variable for the age of the Boeing 767-300ER modelled, $x$ is the variable of demand index of current year, $\varepsilon \sim N(0, \sigma^2)$. 
For each aircraft type, we can repeat the same procedures and get the functions for coefficients $\gamma_0, \gamma_1$. In this way, the characterisic properties of aircraft, age, is included into the model; instead of running 61 PCR models, we have 8 varying coefficient PCR models to predict the MV growth for each aircraft type at different ages (Appendix B).

Despite that for most of the aircraft, the determining coefficient $R^2$ of polynomial fitting is higher than 0.99, there is still some difference between the results of PCR model and varying coefficient PCR model. We will compare the difference after presenting another methods of reducing number of models.

**Section 3.5.2 PCR model for clustered aircraft group**

We have showed that there is some difference in reactions of aircraft of the same type at different ages: the market value of older aircraft are more sensitive to the market condition. However, at the same time, this difference in MV growth of aircraft of the same type at similar ages is quite small. In chapter 2, we have clustered the aircraft into 20 groups, and for one group, we would use the average MV growth to represent for the MV growth of all aircraft in it.

This is another way to reducing the number of models. Instead of describing carefully the relationship between coefficients $\gamma_0, \gamma_1$ and age, we can ignore the small difference between the aircraft in the same group, and treat them equally.

Of course, this simplification of the model will results in bigger errors. It is still a solution to reduce the number of model. Though it seems more accurate with the varying coefficient PCR model, it would cause some errors as well when use the fitted function to calculate the value of coefficients for models of aircraft at each age.

In Table 22, one can check the comparison of errors between three methods: PCR, the varying coefficient PCR, and the PCR for clustered aircraft groups.

**Table 22 comparison of errors between three methods: PCR, varying coefficient PCR, PCR for clustered aircraft groups**

<table>
<thead>
<tr>
<th>Method</th>
<th>Average MAE</th>
<th>Average MAXE</th>
<th>Average Error in year2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td>3.4943</td>
<td>8.3292</td>
<td>2.4485</td>
</tr>
<tr>
<td>Varying Coefficient PCR</td>
<td>4.8277</td>
<td>13.3842</td>
<td>3.6580</td>
</tr>
<tr>
<td>PCR for clustered aircraft groups</td>
<td>3.5294</td>
<td>8.3771</td>
<td>2.4472</td>
</tr>
</tbody>
</table>

It turns out that the varying coefficient model has the biggest error. The function of coefficient is obtained through the polynomial fitting of coefficients with age, but the results of
varying coefficient PCR faded away from the original model a lot. The errors in the estimation of coefficient $\gamma_0, \gamma_1$ are transferred to the results of the whole model, and are enlarged because of the important contributions of the coefficients to the whole model. Despite that there is a clear logic relationship between age and coefficient $\gamma_0, \gamma_1$, it may not be a good option when we want an more accurate model.

For method of PCR for clustered aircraft group, the results are just a little worse than the PCR model, and now we only need to calculate parameters for 20 models (as we have 20 groups in Table 9). The work is reduced greatly and we can still hold a good estimation of the aircraft value growth. However, one need to keep in mind that the difference between the aircraft in the same group does exist. In bad years, there is a tendency that we underestimate the MV growth of younger aircraft and overestimate the MV growth of older aircraft, and vice versa.

As we can see, the more factors we want to put into the model, the more complicated the model become, the less the accurate of the prediction is. However, it helps us understanding the inter-relationship between aircraft MV growth and age better.

**Section 3.6 Conclusion**

In this chapter, we present a regression model for MV growth. Applying classical linear regression, we are able to make a good estimation of MV growth based on the selected explanatory variables. However, because of the high dependence among the explanatory variables, we can not use the coefficients of the variables to explain their contribution to the model.

In order to handle the high correlation among the explanatory variables, principal component regression (PCR) is applied. The first principal component is selected as the predictor for the regression model; it is almost the average values of the five explanatory variables, indicating the demand level in the current economic environment. This demand indicator has a positive relationship with MV growth of all aircraft regardless of age and type. For prediction, Principal component regression can still do a job with only the demand factor.

We also discussed the relationship between age and the coefficients for model of Principal component analysis. The results see that the coefficient of the demand indicator (the first principal component ) has a negative correlation with age. It implies that the MV growth of older aircraft are more sensitive to the change in economic environment. In bad years such as year 2002 and year 2009, the values of older aircraft delined much more than the younger ones.
Applying polynomial fitting, we are able to describe the two parameters of Principal component regression (the coefficient of constant $\gamma_0$ and the coefficient of the demand indicator $\gamma_1$) using a function with respect with age. A varying coefficient PCR model is built based on this function. The coefficients are functions of age instead of constant. We are able to include age as another parameter into the model, and reduce the number of model from 61 to 8 (we have 8 aircraft types). The errors of the varying coefficient PCR is bigger than PCR because of the error in the polynomial fitting.

Another way to reduce the number of model is to model the average value of MV growth of aircraft in a similar group. We have made such groups in Chapter 2. It turns out the by modeling the average value of MV growth of aircraft for a group, the error is even smaller than the varying coefficient PCR, though still bigger than the PCR. We know that we are ignoring the small difference among the aircraft in the same group; for example, there is a tendency to underestimate the MV growth of younger aircraft and overestimate the MV growth of older aircraft in bad years. But when we want to reduce the calculation, this small difference can still be accepted.
Chapter 4 Model the market value growth with copula

In this chapter, one can see a comprehensive analysis of the whole dependence structure among involved variables (MV growth and five explanatory variables in Table 10). We construct a model of multivariate distribution based on copula functions. Through the joint distribution function, one can understand the relationship of several variables interacting simultaneously, not in isolation of one another.

Prediction is made based the conditional distribution of MV growth, conditionalized on the economy variables. The mean and the median of the samples simulated from the conditional distribution are taken as predicted values of MV growth. What’s more, with the whole dependence structure described by copula, we are able to see the interactive influences between aircraft of different ages, sizes or manufacturers.

Section 4.1 Copula model

The high correlations among all involved variables suggest that copula can be applied to solve this problem. Copula is a great tool for modeling dependence of several variables. It can split the whole relationships among variables into the marginals and dependence structure. Applying copula, we are able to describe the whole joint distribution between all the variables interested in the model. [20]

A $d$-dimensional copula $C: [0,1]^d \to [0,1]$ is a cumulative distribution function with uniform marginal [21],[22].

Usefulness of copula function is motivated by the following observation if the real valued random variable $Y$ has a distribution function $F$ and $F$ is continuous, then $F(Y) \sim U[0,1]$. It indicates that random variables from any given continuous distribution can be converted into random variables having a uniform distribution, and this method of converting random variables is called probability integral transform. It holds exactly provided that the distribution being used is
the true distribution of the random variable; if the distribution is the one fitted to the data, then
the result will hold approximately in large samples. Moreover by Sklar’s Theorem\(^{22}\)\(^{[23]}\) every joint
distribution function on \(\mathbb{R}^d\) can be described by a copula function and marginal distributions. In
other words, as long as we chose a copula and some marginal distributions we will get a
multivariate distribution function with the given marginal.

The inverse of the probability integral transform is to convert random variables from a
uniform distribution to a known distribution. It is also called inverse transform. With an arbitrary
distribution, we can simulate samples of random variables from uniformly distributed ones.

Based on these conclusions, one is able to construct any multivariate distribution function
with a copula and known marginal distribution. For current problem, we need to model the joint
distribution of the six variables. Because the marginals are not known, the work is thus split into
two parts: modeling the marginal (the marginal distribution function) and then modeling the
dependence structure (the copula). For prediction, conditional copula need to be used to get the
conditional distribution of interested variable. One can see the procedure of modeling in the
following chart.

We will still take new Boeing 737-800 as the example, and use the same notation for
variables in Chapter 3\(^{23}\); the new notation for the estimated marginal cdfs are \(F_Y, F_{X_1}, F_{X_2}, \ldots, F_{X_5}\),
respectively.

\(^{22}\) See Appendix A for Sklar’s theorem.

\(^{23}\) \(Y\) -- MV growth of new Boeing 737-800; \(X_1\) -- GDP growth; \(X_2\) -- Oil price change; \(X_3\) -- Revenue growth;
\(X_4\) -- RPK growth; \(X_5\) -- Passenger growth.
1. Modeling the marginal distribution functions

There are only 10 historical data available for each variable (year 2001 to year 2010). The information contained in the empirical distribution is thus quite limited. From the histogram present in Figure 24, one can hardly be confident about the type of marginal functions.

From the distributions we know, normal distribution can still be a good option as a parametric method to fit the data. Another way to handle this is to apply a non-parametric method: kernel density estimation. We will apply both methods to model the marginal distribution functions. A comparisons of performance of these two methods is present in section 4.2.

![Histograms for six variables](image)

**Figure 24 histograms for six variables**

a) Fitted normal distribution

It is assumed that each variable in Figure 23 follows a normal distribution. A one-dimensional normal density distribution holds a formula as following:

\[
f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2}
\]  

(4.1)
As we can see, there are two parameters—mean $\mu$ and std.deviation $\sigma$ for a normal distribution. The historical data can present the theoretical distribution. We need to estimate these two parameters from the known data.

We take the mean of the historical data as the estimated mean $\hat{\mu}$, and the std.deviation of training data as estimated std.deviation $\hat{\sigma}$.

The estimated parameters of the normal distribution fitting for GDP growth is $\hat{\mu} = 3.9396$ and $\hat{\sigma} = 2.5627$. We can see the comparison between fitted normal distribution and the empirical distribution in Figure 25.

**Figure 25 Empirical distribution vs. normal distribution fitting for GDP growth**

In Figure 25, the left plot is the comparison between empirical histogram and the fitted normal pdf, and the right plot is the comparison between empirical cdf and fitted normal cdf. We can see that fitting is acceptable. However, as the distribution of MV growth seems to have a left tail, the fitting is not very good around both tails.

We use Kolmogorov-Smirnov test for testing the hypothesis of normality. At 5% significance level, the hypothesis that data comes from a normal distribution with parameter $\hat{\mu}$ and $\hat{\sigma}$ is accepted. For all the variables we modelled, the hypothesis cannot be rejected at 5% significance level.

b) Kernel density estimation
First, we will see how it works with kernel density estimation. Kernel density estimation is a fundamental data smoothing problem where inferences about the population are made, based on a finite data sample\textsuperscript{24}. It is a non-parametric way for density estimation. We can see the results of kernel density estimation with gaussian kernel in Figure 26.

\textbf{Figure 26 Empirical distribution vs. kernel estimation for GDP growth}

In Figure 26, the left plot is the comparison between empirical histogram and the pdf estimated by kernel density estimation, and the right plot is the comparison between empirical cdf and cdf obtained by kernel density estimation.

In the upper plot, we can see that the kernel density estimation is close to the histogram, but is smooth and continuous. In the lower plot, the cdf of kernel density estimation follow the empirical cdf quite well except in the two tails, which is quite understandable because of the smoothness of the kernel density estimation. In conclusion, the kernel estimation fits the data well in general.

We still use Kolmogorov-Smirnov test for testing the fitting-goodness of the margins. The null hypothesis is that data and simulated samples from kernel density estimation come from the same parent continuous distribution. The alternative hypothesis is that they are from different distribution. At 5% significance level, the null hypothesis cannot be rejected separately for each

\textsuperscript{24} Introduction to kernel density estimation can be seen in Appendix A
margin. In other words, the marginal distribution functions modeled by kernel density estimation fit the data well enough.

This implies that both options (gaussian and kernal density) can be applied for the available data. We will apply both methods to model the marginal functions $F_{X_1}, F_{X_2}, \ldots, F_{X_j}, F_Y$. In the following parts, it is considered that these marginal functions have been known.

2. Model of the dependence structure

After obtaining the marginal functions, we are able to transform values of each variable into the $[0,1]$ interval applying probability integral transform. In this part, we will present the way of modeling the dependence structure.

The procedure of modeling dependence is independent of modeling marginal functions. However, one needs the results of modeling marginal. In order to explain how to model the dependence structure after knowing the marginal functions, we use the results of margins modeled by kernel density estimation in this part as an example.

In Figure 26, we can see the scatter plots for the probability integral transform of GDP growth versus the probability integral transform of new Boeing 737-800. It is difficult to tell the type of the copula from the available 10 data. Hence based on the available data we would be reluctant to choose some ‘exotic’ copula e.g. with tail dependence.
Thus we will use normal copula to model the dependence. A very important reason is that it is convenient to obtain conditional distribution from joint normal distribution. Based on the available data, it would be the most simple and convenient way to choose normal copula.

A normal copula has the following form:

\[
C_\Sigma(u_1, \ldots, u_6) = \Phi_\Sigma(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_6))
\]

where \(\Phi\) is one-dimension standard normal distribution; \(\Phi_\Sigma\) is a six-dimension standard joint normal distribution with correlation matrix \(\Sigma\) of gaussian marginal variables; \(u_i \in [0,1], i = 1, \ldots, 6\).

With the marginal function modeled in previous part, we are always able to transform the original data into interval \([0,1]\). So as long as the correlation matrix \(\Sigma\) can be estimated, the dependence structure among the multivariable is determined.

The value of correlation matrix \(\Sigma\) is obtained from the training data. By applying probability integral transform and marginal cdfs, we can get the standard normal variable as follows:

\[
Z_j = \Phi^{-1}\left(F_{X_j}(X_j)\right), j = 1, 2, \ldots, 5, \quad Z_6 = \Phi^{-1}\left(F_Y(Y)\right)
\]

where \(\Phi^{-1}\) is the inverse cdf of the standard normal variable. The copula (4.2) can be represented by \(Z_j\) as

\[
C_\Sigma(z_1, \ldots, z_6) = \Phi_\Sigma(z_1, \ldots, z_6)
\]

The parameter \(\Sigma\) of normal copula is the correlation matrix for \(Z_j\), which also represent the dependence structure among \(X_j\) and \(Y\). Theoretically, the parameter \(\Sigma\) should be the correlation matrix of variable \(Z_j, j = 1, \ldots, 6\).

\[
\Sigma = \text{corr}(Z_1, \ldots, Z_6)
\]

We can transform the vector of the training data \(\{x_{ij}\}_{i=1}^{10}, j = 1, \ldots, 5\) and \(\{y_i\}_{i=1}^{10}\) into \(\{z_{ij}\}_{i=1}^{10}, j = 1, \ldots, 6\) with the form (4.3)\(^{25}\).

The correlation matrix can thus be estimated from the training data.

\[
\hat{\Sigma} = \text{corr}(z_{11}, \ldots, z_{21}, z_{61})
\]

For a joint normal copula, each pair of bivariate margins is also a normal copula. The parameter (correlation coefficient) of the normal copula for the bivariate data shown in scatter

\[z_{1i} = \Phi^{-1}\left(F_{X_1}(x_{1i})\right), \ldots, z_{5i} = \Phi^{-1}\left(F_{X_5}(x_{5i})\right), z_{6i} = \Phi^{-1}\left(F_Y(y_i)\right), i = 1, \ldots, 10\]
plot of Figure 27 is 0.8055. Now simulating 100 samples from a bivariate copula with parameter 0.8055, and then applying inverse of the probability integral transform to the simulated uniform samples with the known marginals, we are able to see the density of the bivariate distribution in Figure 27.

![Figure 27](image)

**Figure 27** Scatter plots of 100 samples, generated from the bivariate distribution of variable for GDP growth and variable for MV growth of new Boeing 737-800

The goodness of fit of this density can be tested with a statistical test. We will use the statistical energy test introduced in [24]. The null hypothesis is that the two data are from the same parent distribution. One can test if the simulated large size of two-dimensional samples and the original data come from the same parent distribution. For the bivariate distribution in Figure 28, the p-value is 0.7132; it is accept that the simulated 100 samples and the data come from the same parent distribution.

For the multivariate distribution for the six variables, the steps for modeling are the same: we simulate large number of samples from the normal copula with the calculated correlation matrix $\Sigma$, and then applying inverse of the probability integral transform with the known margins. In this way, we get samples which can represent the joint distribution and dependence structure of the original data.
By combining the marginal model and the selected copula, we are able to get the final form of the whole model as following:

\[
C_2(x_1, \ldots, x_5, y) = \Phi_2 \left( \Phi^{-1} \left( F_{x_1}(x_1) \right), \ldots, \Phi^{-1} \left( F_{x_5}(x_5) \right), \Phi^{-1} \left( F_y(y) \right) \right) \tag{4.7}
\]

where \( \Phi^{-1} \) is the inverse cdf of the standard normal variable, \( \Sigma \) is the correlation matrix obtained from the data; \( \Phi_2 \) is a multivariate normal distribution with parameter \( \Sigma; F_{x_1}, \ldots, F_{x_5} \) and \( F_y \) is the marginal function modelled from the data.

To test goodness of fit of the whole model, we need a multivariate test for multi-dimensional sample. We will still use the statistical energy introduced in [24]. The null hypothesis is that two samples are from the same parent distribution, and the alternative one is that the two samples are from different distributions. A six-dimensional samples with size 1000 is generated from the model, and the test gives us the p-value of 0.7770 which indicates that the null hypothesis cannot be rejected. The copula model fits the data well and can be applied for further analysis.

3. Predictions (conditional distribution of \( Y | X_1, \ldots, X_5 \))

The model obtained in step 1 and step2 describes the marginal distribution and dependence structure of the variables at the same time. For each year, we have observed data for variable \( X_1, X_2, \ldots, X_5 \) and need to predict the corresponding value for variable \( Y \).

With the form (4.3), we can always make transformation between \( X_j \) (or \( Y \)) and \( Z_i \). So the problem can be converted into how to predict \( Z_6 \) when we have observed values for \( Z_1, Z_2, \ldots, Z_5 \).

According to the setup in (4.5), we can partition the correlation matrix \( \Sigma \) as follows:

\[
\Sigma = \begin{pmatrix}
\Sigma_5 & r_6 \\
\text{T} r_6 & 1
\end{pmatrix}
\]

where \( \Sigma_5 \) is the correlation matrix of the transformed variable \( Z_1, \ldots, Z_5 \), and \( r_6 = (r_{1,6}, \ldots, r_{5,6}) \) is the vector of correlations between the observed variables \( Z_1, \ldots, Z_5 \) and \( Z_6 \).

When we have observed value \( \bar{Z} = (z_1, z_2, \ldots, z_5) \) for variable \( (Z_1, Z_2, \ldots, Z_5) \), we can obtain the conditional cdf for \( Z_6 \) simply using properties of joint normal distribution:

\[
F_{Z_6|z_1, \ldots, z_5}(z_6|z_1, \ldots, z_5) = \Phi \left( \frac{z_6 - \bar{\mu}}{\bar{\sigma}_6} \right) \tag{4.8}
\]

where \( \bar{\sigma}_6^2 = 1 - r_6^T \Sigma_5^{-1} r_6, \bar{\mu} = r_6^T \Sigma_5^{-1} \bar{Z} \).
As we can see, in fact we have the conditional normal distribution function for variable $Z_6$ with expectation $\mu = r_6^T \Sigma_{65}^{-1} \bar{Z}$ and variance $\sigma^2 = 1 - r_6^T \Sigma_{65}^{-1} r_6$.

We can generate large number of samples from the conditional distribution (4.8), and transform the simulated samples to the initial scale applying the inverse transformation of (4.3) as $Y = F_6^{-1}(\Phi(Z_6))$.

In this way, we are able to simulate samples for conditional distribution of $Y|X_\eta, \ldots, X_5$. With a large size (e.g., 1000) they approximate the theoretical conditional distribution well. For each year, we have a set of values for $X_1, \ldots, X_5$, and thus a conditional distribution for $Y|X_1, \ldots, X_5$. In Figure 29, one can see the histogram and empirical cdf of the simulated samples of the conditional distribution of MV growth for new Boeing 737-800 in year 2006.

![Figure 29 Conditional distribution of MV growth in year 2006, for new Boeing 737-800, marginal functions are modelled by kernel density estimation](image)

We can see the shape of pdf and cdf of the conditional distribution of MV growth in year 2006 for new Boeing 737-800. The mean of the distribution is 4.9202 while median 4.7561. $[-0.3995, 10.5978]$ are the $[0.025, 0.975]$ quantiles for this distribution.

For predicted value of MV growth, one can pick the expectation or the median of the distribution as the predictive value for Y. As we use simulated samples to represent the conditional distribution, the mean and median of the simulated samples are compared in terms of predictive power.
Section 4.2 Copula model-results

In Figure 30, one can see the results of Copula model (modeling the marginal function applying kernel density estimation) for new Boeing 737-800. The green line is the median of the conditional distribution, the lower dashed red line are the values for the 0.025 quantile of the distribution and the upper dashed red line are the values for 0.975 quantile of the distribution. We are confident that the observed values will fall between the two red lines with 95%.

![Figure 30 model versus truth, Copula model of new Boeing 737-800, modeling marginal function applying kernel density estimation](image)

As we can see that the model fits the data quite well in general. In year 2002, the truth value is near to the values of 0.025 quantile of the distribution; it indicates that the truth in year 2002 is much lower than the expectation of the model. We have mentioned before that, because of the 9-11 attack in year 2001, the airline industry was faced with great difficulties, and thus the aircraft market suffered strong impact. People’s confidence about the aircraft market was even lower than the true situation. According to our model, the market values in year 2002 should not be as low as it was according to the economy factors at that time. However, the factors influencing the aircraft become very complicated, and the economy factors cannot represent the total force on the aircraft market any more in year 2002.
Keep in mind that we have two methods to model the marginal functions and one can pick either median or mean as the final prediction. We apply both methods to see which results would be better. Once can check the comparison of the error in Table 23.

<table>
<thead>
<tr>
<th>Margin model</th>
<th>Statistic for prediction</th>
<th>MAE</th>
<th>MAEX</th>
<th>Absolute error in prediction 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>median</td>
<td>2.1596</td>
<td>4.7377</td>
<td>0.3470</td>
</tr>
<tr>
<td>Normal</td>
<td>mean</td>
<td>1.9988</td>
<td>4.6983</td>
<td>0.3480</td>
</tr>
<tr>
<td>Kernel</td>
<td>median</td>
<td>1.6395</td>
<td>4.8674</td>
<td>0.1040</td>
</tr>
<tr>
<td>Kernel</td>
<td>mean</td>
<td>1.7839</td>
<td>4.9177</td>
<td>0.2370</td>
</tr>
</tbody>
</table>

MAE measures the average error over year 2001 to year 2010, while MAEX measure the maximum error the model can make over the same period. We have two conclusions from comparison of MAE in Table 23: the kernel density estimation fits the data better than the normal fitting, the median of the distribution predict better than the mean of the distribution.

The result is quite understandable. The result of the kernel density estimation is more close to the real distribution function.

The median of the conditional distribution is more appropriate to be considered as the predicted value in general; the median of conditional distribution modeled by kernel density estimation makes the smallest MAE.

However, as we can see in Table 23, the mean value has a smaller MAEX, which indicates that the median will make a bigger error for some special case.

We need to provide some background to give the reason for this result. For Boeing 737-800, the year 2002 is the special time, and the practical situation is different from other years. If we deal with symmetric distribution then the expected value and the median are the same otherwise they will be different. As we can see in Figure 31, the conditional distribution in year 2002 has a big right tail. In this case, the mean lies to the left side of the median, and thus closer to the left tail; as the observation is quite extreme lying close to the left tail of the distribution, the mean will predict it better than the median.
Figure 31 the histogram of simulated samples from the conditional distribution of MV growth in year 2002, copula model of new Boeing 737-800

However, in principle we do not know where the true value will fall. Year 2002 is a special year because of the 9-11 attack. For the usual situation the median is the better option.

Table 24 show the average MAE and MAEX for all 61 aircraft between the four prediction options. It supports the conclusion we have in Table 23 that: kernel density estimation fits the data better, and median of the conditional distribution makes a better prediction in general.

Table 24 comparison of average errors for all 61 aircraft

<table>
<thead>
<tr>
<th>Margin model</th>
<th>Statistic for prediction</th>
<th>Average MAE</th>
<th>Average MAEX</th>
<th>Average Absolute Error in prediction 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>median</td>
<td>3.5227</td>
<td>8.8448</td>
<td>2.4368</td>
</tr>
<tr>
<td>Normal</td>
<td>mean</td>
<td>3.5225</td>
<td>8.8338</td>
<td>2.4478</td>
</tr>
<tr>
<td>Kernel</td>
<td>median</td>
<td><strong>2.8338</strong></td>
<td>9.0921</td>
<td>2.3293</td>
</tr>
<tr>
<td>Kernel</td>
<td>mean</td>
<td>2.9284</td>
<td>8.6175</td>
<td>2.2909</td>
</tr>
</tbody>
</table>

We will make some further discussion in next section about the internal properties of the aircraft. Except for these discussions of the results, when we refer to the results of copula model, it is the median values of the conditional distribution with the margins fitted by kernel density estimation.
Section 4.3 Incorporating the age in the model

Similarly to what we have done for the PCR model, we still hope that we can reduce the number of models we need to make. We expect that there would be some relationship between the parameters used in copula model and the age of the aircraft. We will still take Boeing 767-300ER as the example.

The most important parameter in copula model is the correlation matrix $\Sigma$. We will first check how the correlation between MV growth and the economic variables vary with the age of the aircraft. The scatter plots of age and correlation is present in Figure 32.

As we can see in Figure 32, the correlation between MV growth and airline industry variable decrease steadily with age. It indicates that the value of younger aircraft are more sensitive to the condition of airline industry. From the annual report of airline industry, we can figure out more information concerning the value of younger aircraft.

Younger aircraft also has a stronger correlation with GDP growth, especially the brand new ones. However, the influence of GDP growth towards MV growth of aircraft becomes quite stable when aircraft are ‘old enough’. For Boeing 767-300ER, GDP growth has a quite similar influence for aircraft older than 3 years.

Oil price change has stronger influence towards the MV growth of older aircraft as we can see in the second plot in Figure 32. The correlation between MV growth of younger aircraft and oil...
price change is also quite big, but it is difficult to tell if this is because of the strong correlation between MV growth and GDP growth, as we can see that the correlation between oil price change and MV growth is always positive, which may be different from our expectation.

Next we will check the relationship between the parameters of normal distribution fitted to margins and age. When we model the marginal function, we apply two methods: normal distribution fitting and kernel density estimation. There is no parameter for kernel density estimation; but we can see that for normal distribution fitting, mean and std. deviation need to be estimated. We can check the relationship between the estimated mean and std.deviation with the age of corresponding aircraft.

As we can see in Figure 33, the mean decreases with age, and the std.deviation increases with age. We have seen similar results in Section 3.5. There are more uncertainties in the MV growth of older aircraft. We can expect a much bigger fluctuation in the change of MV growth for older aircrafts.

We can apply polynomial fitting to describe the relationship between the parameters and age. One can see the polynomial fitting results in Appendix B.

Next, we will check the results of the basic model and the complete model (with parameter calculated from the fitted polynomial) separately.

Figure 33 Relationship between parameters of normal distribution fitting and age

As we can see in Figure 33, the mean decreases with age, and the std.deviation increases with age. We have seen similar results in Section 3.5. There are more uncertainties in the MV growth of older aircraft. We can expect a much bigger fluctuation in the change of MV growth for older aircrafts.

We can apply polynomial fitting to describe the relationship between the parameters and age. One can see the polynomial fitting results in Appendix B.

Next, we will check the results of the basic model and the complete model (with parameter calculated from the fitted polynomial) separately.
And with the fitted mean and std. deviation, we can see the comparison of the error between this fitted mean and std. deviation with the original copula model with marginal fitted by normal distribution.

**Table 25 Boeing 767-300ER, comparison between Basic copula model and Complete copula model, applying normal distribution modeling the marginal functions.**

<table>
<thead>
<tr>
<th>Method</th>
<th>Average MAE</th>
<th>Average MAEX</th>
<th>Average Error in 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic Copula model(normal)</td>
<td>5.3958</td>
<td>18.0745</td>
<td>2.0183</td>
</tr>
<tr>
<td>Complete Copula model(normal)</td>
<td>5.5861</td>
<td>19.5257</td>
<td>2.2417</td>
</tr>
</tbody>
</table>

**Table 26 for Boeing 767-300ER, comparison between Basic copula model and Complete copula model, applying kernel density estimation modeling the marginal functions.**

<table>
<thead>
<tr>
<th>Method</th>
<th>Average MAE</th>
<th>Average MAEX</th>
<th>Average Error in year 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic Copula model(kernel)</td>
<td>4.7953</td>
<td>16.7186</td>
<td>1.1272</td>
</tr>
<tr>
<td>Complete Copula model(kernel)</td>
<td>5.4478</td>
<td>17.2072</td>
<td>1.6490</td>
</tr>
</tbody>
</table>

As we can expect, that the fitted parameters will cause bigger errors. For a goal to get a better predicted result, it is not very meaningful that we include age into the model. But it is still interesting to see the relationship between parameters and age.

Another way to reduce the number of model is to model the average MV growth of the aircraft group we have made in Chapter 2. We can ignore the small difference among the MV growths of the aircraft in the same group, and use the average value to represent the MV growth of all aircraft in this group. This time we will use the kernel density estimation to model the marginal functions.

**Table 27 for Boeing 767-300ER, comparison between Basic copula model and clustered aircraft group, applying kernel density estimation modeling the marginal functions.**

<table>
<thead>
<tr>
<th>Method</th>
<th>Average MAE</th>
<th>Average MAEX</th>
<th>Average Error in year 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic Copula model(kernel)</td>
<td>4.7953</td>
<td>16.7186</td>
<td>1.1272</td>
</tr>
<tr>
<td>clustered aircraft groups(kernel)</td>
<td>5.0279</td>
<td>16.9982</td>
<td>1.5690</td>
</tr>
</tbody>
</table>
Still, grouping aircraft will cause more error. However, if we compare the results in Table 27 and Table 26, we can see that the error of clustered aircraft groups is smaller than error of the complete model. It still can be a good option for reducing the number of models.

However, as what we want is to get a better estimation of market value of aircraft, in the following parts, we will just use Basic Copula model (kernel density estimation) for each aircraft.

**Section 4.4 Influences between aircraft types**

We are interested in the interactive influence of aircraft of different types and ages. However, MV growths of all aircrafts are naturally correlated with each other as economy factors have strong impact on MV growth of all aircraft; one should remove the influence of economy factors when checking the correlations between MV growths of two different aircraft.

This can be realized with the help of copula. Let \( Y_1, Y_2 \) denote MV growths of two different aircraft, and \( X_1, X_2, ..., X_5 \) for the economy variables. We can model the joint distribution of \( X_1, X_2, ..., X_5, Y_1, Y_2 \) by following the same procedures in Section 4.1. By conditioning on the economy variables, we can get the conditional distribution \( F_{Y_1Y_2|X_1,...,X_5} \), which can represent the correlation of \( Y_1, Y_2 \) under a known economic environment.

Recall the form of copula (4.2) and the normal transform of variables (4.3), we can see that the conditional dependence between variable \( Y_1 \) and \( Y_2 \) can be described by the conditional correlation. We will compare the original Pearson’s correlation and the conditional correlation after knowing the economy factors to check the interactive influence of two different aircraft. Keep in mind that the size of the samples are only 10, and even under a significance level 0.1 (one tail), the correlation higher than 0.45 is considered as statistically significant.

For an aircraft, we are interested in three characteristics of the aircraft: its manufacturer, its size and age. We will make three groups of aircrafts to check the influence of these factors.

First we will check the interactive influence between competitors from two manufacturers. Boeing and Airbus are the two dominant manufacturers of commercial aircrafts. In global aircraft market including narrow-body and wide-body aircraft, Airbus and Boeing possess a duopoly since the end of the 1990s.

There is intense competition between Airbus and Boeing, and they have developed similar products to occupy the market. The competitors we want to investigate are Airbus A320-200 versus Boeing 737-800, Airbus A330-200 versus Boeing 767-300ER, Airbus A330-300 versus Boeing 777-200.
Table 28 direct competitors from two manufacturers

<table>
<thead>
<tr>
<th>MV growth of aircraft (new)</th>
<th>Pearson’s correlation</th>
<th>Conditional correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airbus A320-200 versus Boeing 737-800</td>
<td>0.9370</td>
<td>0.2877</td>
</tr>
<tr>
<td>Airbus A330-200 versus Boeing 767-300ER</td>
<td>0.8112</td>
<td>0.3508</td>
</tr>
<tr>
<td>Airbus A330-300 versus Boeing 777-200</td>
<td>0.8094</td>
<td>-0.0709</td>
</tr>
</tbody>
</table>

As we can see in Table 28 that after removing the influence of economical factors, we can see immediately that the correlation of MV growth between competitive aircraft from these companies is reduced greatly, especially for bigger aircrafts. There is now only small correlation between the market value growth of Airbus A330-300 and Boeing 777-200.

It indicates that interactive influence of MV growth of competitive aircraft from two manufacturers is quite small, which is a bit surprising. It is maybe because that the two companies have their own strategy towards the market. Sometimes the manufacturers lower the price of the aircraft for a big order, or even just preparing for the releasing a new model. One cannot simply judge the future trend of the market value of an aircraft according to the move of the market value growth of its direct competitor from another company.

Figure 34 Conditional distribution of MV growths for Airbus A320-200 and Boeing 737-800 in year 2002 and year 2005, size of the simulated samples =100
In Figure 34, we can see the comparison between the original bivariate distribution and the conditional distribution in year 2005 and year 2002 for Airbus A320-200 and Boeing 737-800. Year 2005 was a good year for aircraft market while in year 2002 the whole market suffered the regress because of the 9-11 attack. We can see that in year 2005, both MV growths of Airbus A320-200 and Boeing 737-800 are quite high, but looks independent of each other (because of the small conditional correlation value); while in year 2002, MV growths of both aircraft are scattered in a lower value.

Figure 34 support that the original strong correlation between MV growth of competitors is because of the influence of economy factors to the whole aircraft market; competitors from two different manufacturers have little influence towards each other.

Next we will check the influence between similar aircraft from the same manufactures. Generally, a manufacture have several families of aircraft. The difference for aircraft in the same families is quite small, and aircraft in the same families can be even considered as alternative for each other sometimes. In the eight types of aircraft we have, A319-100, A320-200 and A321-200 all belong to A320 family, Airbus A330-200 and Airbus A330-300 are both from A330 family and Boeing 767-300ER and Boeing 777-200 are two big wide-body aircrafts of Boeing.

**Table 29 members in one family in the same manufacturer**

<table>
<thead>
<tr>
<th>MV growth of aircraft (new)</th>
<th>Pearson’s correlation</th>
<th>Conditional correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airbus A320-200 versus Airbus A321-200</td>
<td>0.9493</td>
<td>0.7428</td>
</tr>
<tr>
<td>Airbus A330-200 versus Airbus A330-300</td>
<td>0.9450</td>
<td>0.9370</td>
</tr>
<tr>
<td>Boeing 767-300ER versus Boeing 777-200</td>
<td>0.9126</td>
<td>0.5718</td>
</tr>
</tbody>
</table>

Table 29 shows that the aircraft from the same/similar families produced by the same manufacturers have a strong influence on each other even after conditioning on the economical variables. This indicates that market values of members of the same family will behave in a quite similar manner even after the influences of economy variables have been removed.
Figure 35: The conditional distribution of MV growths for Airbus A330-200 and Airbus A330-300 in year 2002 and year 2005, size of the simulated samples = 100

As we can see in Figure 35, even after removing the influence of economy factors, there is still a strong correlation between the MV growth of new Airbus A330-200 and new Airbus A330-300. When the manufacturer makes the price of their product, they will adjust the price of aircraft of the same family at the same time. It indicates that for the manufacturers, they will make their plan more according to their own strategy to the market instead of the plan of their competitor.

One can also see how the size of the aircraft influences the pricing plan of the manufacturers in Table 30.

Table 30: Aircraft produced by the same manufacturer, but with quite different size

<table>
<thead>
<tr>
<th>MV growth of aircraft (new)</th>
<th>Pearson’s correlation</th>
<th>Conditional correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airbus A319-100 versus Airbus A330-300</td>
<td>0.8204</td>
<td>0.3413</td>
</tr>
<tr>
<td>Boeing 737-800 versus Boeing 777-200</td>
<td>0.8635</td>
<td>0.3269</td>
</tr>
</tbody>
</table>

In Table 30 we also see the conditional correlation of MV growth of aircrafts with quite different sizes produced by the same manufacture. Airbus A319-100 is a small model in A320 family, which is a series of medium-size narrow body aircrafts; Airbus A330-300 is a much bigger wide-body aircraft. Similarly, Boeing 737-800 is a much smaller aircraft compared to Boeing 777-
200, which is a large wide-body aircraft. We see that after removing the economical factors, the correlation is rather small, even smaller than for competitors from different manufacturers.

After these brief investigations we can conclude that MV growth is sensitive more to the aircraft characteristics rather than to the brand.

We can also look at relationships between aircrafts of the same type at different ages. In Table 31, we can see the samples generated from the conditional distributions of A320-200 at different ages.

**Table 31 the same aircraft type (Airbus A320-200) at different age**

<table>
<thead>
<tr>
<th>MV growth of Airbus A320-200</th>
<th>Pearson’s correlation</th>
<th>Conditional correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aircraft at age 0 versus aircraft at age 2</td>
<td>0.9896</td>
<td>0.9271</td>
</tr>
<tr>
<td>Aircraft at age 0 versus aircraft at age 10</td>
<td>0.7259</td>
<td>0.0796</td>
</tr>
<tr>
<td>Aircraft at age 9 versus aircraft at age 10</td>
<td>0.9875</td>
<td>0.9862</td>
</tr>
</tbody>
</table>

Table 31 shows the distribution of Airbus A320-200 at age 0 and at age 2 conditional on economic variables. After removing the influence of economy there is still a strong correlation (0.9271) between the market value growth of these two aircrafts. It indicates that the market value growth of the new aircraft has a strong influence towards the market value growth of the older ones.

However, we can also see that the conditional correlation between Airbus A320-200 at age 0 and at age 10 is almost zero. It tells that the influence of the new aircraft to old aircraft (such as aircraft at age 10) is reduced greatly.

It is easy to see that there is a strong conditional correlation between market value growths of aircraft A320-200 at age 9 and at age 10. It implies that, aircraft at similar ages have a strong impact on each other.

**Section 4.4 Conclusion**

Applying copula, we are able to model the whole dependency structure. The whole model can be splitted into two parts: model the marginal function and model the dependence structure by copula.
We apply two methods to apply the marginal functions. Fitting the data with a normal distribution is a parametric method, and kernel density estimation is a non-parametric way. Kolmogorov-Smirnov test shows that both methods can be used to describe the data.

For prediction, conditional distribution of MV growth is obtained from the conditional copula. We take the mean or median as the predicted value for MV growth. Comparison between the results shows that it is the median of the conditional distribution obtained from copula model with marginal function modeled by kernel density estimation that has a smaller error.

We also try to reduce the number of models by including age into the model. We can use polynomial fitting to describe the relationship of correlation value and age of aircraft, and the relationship of parameters of normal distribution and age of aircraft. From the scatter plot, we can see that the correlation between factors of airline industry and MV growth decreases with age, and std.deviation of MV growth increases with age of aircraft. It indicates that the factors influencing the MV growth of older aircraft become more complicated, which are not included in the five explanatory variables selected. However, the complete model including age has a bigger error for prediction.

Another way of reducing the number of models is to make models of the clustered groups. Again, this will result in bigger error in prediction.

The interactive influence of different aircraft are also analyzed with help of copula. We compare competitors from Airbus and Boeing, families of the same manufacturer, and aircraft of different size of the same manufacturers. The results shows that, after removing the economy factors, only families of the same manufacturers has strong impact towards each other. It implies that when manufacturer makes the price of aircraft, they mainly make their plans according to current market condition and strategy. The price of aircraft in another company does not influence their plan very much.

What’s more, for aircraft of the same type, market value growths of aircraft at similar ages have a strong impact towards each other. The influence of new aircraft to older aircraft becomes smaller for older aircraft. When one wants to check the market value growth of an aircraft at certain age, he should check the market value growth of the one-year-younger aircraft as the reference.
Chapter 5 Model performance

In this chapter, we will have a close look at the results of the two models presented in Chapters 3 and 4. We will compare the results of PCR model and copula model in this chapter.

What’s more, we would like to check the model performance for aircraft of different types at each age. As we are modeling aircraft for 8 types at different ages, it is quite possible that the models have different performance for different aircraft. We will check this difference and try to figure out why the difference exists.

Further, we will get the final prediction of market values based on the prediction of market value growth. We will also compare result in [8] and results of our models.

Section 5.1 Comparison between PCR and copula model

In this section, we will compare the results of PCR model and Copula model to see the difference between them. Here we will just use the original PCR model and copula model (kernel density estimation for marginal functions) without including age into the model.

Again, we will take new Boeing 737-800 as the example.
Figure 36 comparison between PCR and Copula, prediction versus time

We can see the comparison in Figure 36. The red solid line is for results of copula model while the dashed black line is for results of PCR. We can see that in fact, results of two methods are quite similar, and Copula model does a little bit better job than PCR model in average. We can also see that in year 2002, PCR has a better prediction than the Copula model.

What’s more, the 95% confidence interval of Copula is much narrower than the confidence interval of PCR. The point in year 2002 is out of the 95% confidence interval of copula model.

We can see the comparison of errors in a more clear way in Table 32. The result in Table 33 shows that Copula model has smaller MAE and MAEX, and bigger Max absolute error. From the results in Table 32, we can get the conclusion that in average, copula model performs better; but for some ‘unexpected’, PCR makes less mistakes.

<table>
<thead>
<tr>
<th>Model/ Error measure</th>
<th>MAE</th>
<th>MAEX</th>
<th>Error in year 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td>2.1928</td>
<td>3.7118</td>
<td>0.4586</td>
</tr>
<tr>
<td>Copula</td>
<td>1.7714</td>
<td>4.8674</td>
<td>0.3196</td>
</tr>
</tbody>
</table>
Errors for all aircraft are checked in Table 33. We take the average of MAE, MAEX and Average Max absolute error over the 61 aircraft modelled to see the general performance of PCR model and Copula model.

**Table 33 Comparison of errors between the model of PCR and Copula for all aircraft being modeled over year 2001 to year 2010**

<table>
<thead>
<tr>
<th>Model/ Error measure</th>
<th>Average MAE</th>
<th>Average MAEX</th>
<th>Average Error in year 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td>3.4943</td>
<td>8.3292</td>
<td>2.4485</td>
</tr>
<tr>
<td>Copula</td>
<td>2.8338</td>
<td>9.0921</td>
<td>2.3560</td>
</tr>
</tbody>
</table>

Again, it support the conclusion that in average, copula fits the data better.

For most aircraft, the ‘max error’ happened in year 2002. We have explained in Chapter 2 that because the updating time of aircraft market data is in July, the data in year 2002 represents the information from the second half year of 2001 and the first half year of 2002, which means that the influence of 9-11 attack is included in values of year 2002. The 9-11 attack makes a direct shock airline industry and any business related to airlines, and there are many complicated reasons for the recession of aircraft market. Both our models overestimate the market value in this year.

In year 2009 we observe another crisis for aircraft market. The difference between year 2002 and year 2009 is that crisis in year 2009 is not so instantaneous. The economy crisis began by the end of 2008, and affected the whole world through year 2009. As we can see that both models can fit data in year 2009 well. The market condition can be explained by the model with the low demand level because of the economy recession.

Our results show that copula model is not good at dealing with unexpected event even from the known data. Linear regression is not good at this as well, however, looks a little bit better than copula.

However, apparently regression is a much faster method compared to Copula as we can see in Table 25. Copula method needs many steps such as computing correlation matrix, generating samples, which takes lots of time. If speed of calculations is of importance the PCR is a better option, however 2.5 seconds of waiting time is not a problem.
Section 5.2 Error analysis with age, year and aircraft type

In this section, we will check if model performance for different aircraft is the same. The error analysis will be made with age, year and aircraft type separately.

First, we will check the error of model for aircraft at different ages. Boeing 767-300ER is an aircraft model produced in year 1986, and data for aircraft at age 0 to age 14 is available, we can see the relationship between error of model and age of aircraft in Figure 37.

Figure 37 MAE of model versus age

MAE is a measure of average error. In Figure 37, we can see the MAE of the models for Boeing 767-300 increases with ages for both PCR model and Copula model. In other words, the model performs worse for older aircraft when the aircraft type is the same.

The error of copula model for new Boeing 767-300ER is only 2.7184, and for aircraft of Boeing 767-300ER at age 14, the error goes to 6.7717. The PCR model has a similar pattern. It
indicates that the explanatory variables selected cannot predict the market value growth of older aircraft very well.

If we look into the MAE for both models of all the 8 aircraft, we can see that the error goes up steadily (Table 35).

<table>
<thead>
<tr>
<th>Aircraft / Model</th>
<th>PCR</th>
<th>Copula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average MAE for aircraft at Age 0</td>
<td>2.5570</td>
<td>1.9950</td>
</tr>
<tr>
<td>Average MAE for aircraft at Age 1</td>
<td>2.6454</td>
<td>2.1183</td>
</tr>
<tr>
<td>Average MAE for aircraft at Age 2</td>
<td>2.7039</td>
<td>2.1258</td>
</tr>
<tr>
<td>Average MAE for aircraft at Age 3</td>
<td>2.7523</td>
<td>2.2660</td>
</tr>
</tbody>
</table>

We can understand this, the factors influence the market value of older aircraft get more complicated. The five variables we have selected mainly stand for the demanding from the market. However, for older aircraft, factors such as a plan of releasing new models, a new environmental policy, may all have an impact on market value of the used aircraft. These factors are difficult to measure in quantitative way. We can see that market demand have determined the trend of market value largely; however, these unknown factors have their influence on the used aircraft as well, and the influence getting bigger for older aircraft.

Another thing we are worried about is that if the model treats each year equally as our data is a time series. For each aircraft, there is only ten data available. In order to have a general observation, we will put the standardized residuals for each aircraft together to see the performance of PCR model and Copula model in Figure 38.
It is easy to see that for year 2002 and year 2004 both model have a tendency to overestimate the market value growth. For these two years, errors for most of (all) the aircrafts are below zero for both PCR and Copula model.

For year 2002, it is an understandable behaviour. As the crisis in year 2002 struck the airline industry directly, and threaten mainly the industry related the airline, thus including aircraft market. People are not confident about the future of aircraft market and this dampened the crisis in year 2002. Compared to the year 2009, as the general economy spread relatively slower (compared to an attack), people can predict more from the economy trend about the condition of the aircraft market, and thus the model still does a good job.

Year 2004 is the turning point of the aircraft market. The market was starting to recover from the recession of year 2001 to year 2003. According to our model, the market should have recovered faster with the economy demand. People are still not confident about the aircraft market in practice. This may explain why in general, both models will overestimate the market value growth of aircraft.

Next, we will check the performance of two methods for the 8 aircraft types. As we have seen that age of the aircraft is another factors for model performance, we will only compare the
new aircraft of these 8 aircraft types. Again, we use MAE and max absolute error to test the fitness of the model. One can see the comparison in Figure 31.

![Figure 39 comparison between errors of new aircraft model, left: MAE, right: MAEX](image)

Figure 39 shows the MAE (left plot) and MAEX (right plot) of two models for all new aircraft of the 8 types we have modeled. As we can see, the smallest error of the model is for airbus A321-200 and the largest one is for Boeing 767-300ER.

In fact, it seems that the behavior of Boeing 767-300ER is different from other aircraft. The determining coefficient for model of new Boeing 767-300ER is only equal to 0.7582, which is not as high as other aircraft types. We can conclude that the demand index (the principal component) cannot explain the data of Boeing 767-300ER very well.

In Figure 40, we can see time trend of the prediction of two models and truth over year 2001 and year 2011. The data of year 2001 to year 2010 is for training and data of year 2001 is for prediction.
As we can see, year 2002 and year 2006 deviate from the model line greatly. We have presented that for most aircraft, there is a systematical prediction error in year 2002. There is something unexpected in year 2006 for Boeing 767-300ER.

It may can be explained by some strategy of Boeing company in year 2006. The market value growth in year 2006 of new Boeing 767-300ER is 12.7288%, which is quite high compared to other aircraft. As for its competitor Airbus A330-200, the market value growth in year 2006 is 3.2596%, which is a quite normal number; for its similar Boeing type, Boeing 777-200, the market value growth in year 2006 is 9.9526%, also higher than average growth; however, the prediction of year 2006 for Boeing 777-200 fits data well.

We can also check the error of prediction in year 2011 for new aircraft of 8 types.

**Table 36 error in year 2011, model of new aircraft**

<table>
<thead>
<tr>
<th>New aircraft</th>
<th>Error</th>
<th>New aircraft</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PCR</td>
<td>Copula</td>
<td>PCR</td>
</tr>
<tr>
<td>Airbus A319-100</td>
<td>-2.9086</td>
<td>-1.7534</td>
<td>Airbus A330-200</td>
</tr>
<tr>
<td>Airbus A320-200</td>
<td>-1.5078</td>
<td>-2.6357</td>
<td>Airbus A330-300</td>
</tr>
<tr>
<td>Airbus A321-200</td>
<td>-0.8231</td>
<td>-1.9286</td>
<td>Boeing 767-300ER</td>
</tr>
<tr>
<td>Boeing 737-800</td>
<td>-0.4586</td>
<td>-0.3196</td>
<td>Boeing 777-200</td>
</tr>
</tbody>
</table>
In year 2011, Copula and PCR model has quite similar results. The maximum mistake is made by PCR model for Airbus A319-100: the market value growth is overestimated by 2.9086%.

**Section 5.3 market value**

After modeling the market value growth, we are able to calculate the market value with the form:

\[
MV(\text{year}_n) = \left( 1 + \frac{MV \text{growth}(\text{year}_n)}{100} \right) \times MV(\text{year}_{n-1})
\]

We always have known the true market values in previous year, so by estimating the market value growth, we are able to estimate the market value in current year. In Figure 41, we can see the comparison of the true and modeled market values in the past 11 years. Year 2001 to year 2010 are used as training data for the model, and year 2011 are used for prediction.

![new Boeing 737-800](image)

**Figure 41 true and modeled market values versus time**

Figure 41 shows that the model does a good job. The estimation in year 2002 is not very good as we have expected. Table 37 shows the money loss of two models for Boeing 737-800 from age 0 to age 3.
Table 37 Money loss for Boeing 737-800, over year 2001 to year 2010

<table>
<thead>
<tr>
<th>Aircraft age</th>
<th>Money loss of PCR model (million)</th>
<th>Money loss of Copula model (million)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>MAX</td>
</tr>
<tr>
<td>0-year-old</td>
<td>0.7762</td>
<td>1.4678</td>
</tr>
<tr>
<td>1-year-old</td>
<td>0.8414</td>
<td>1.5591</td>
</tr>
<tr>
<td>2-year-old</td>
<td>0.8601</td>
<td>1.8661</td>
</tr>
<tr>
<td>3-year-old</td>
<td>0.8018</td>
<td>2.2435</td>
</tr>
</tbody>
</table>

As we can see that, over year 2001 to year 2010, for new Boeing 737-800, the average money loss made by PCR model is 0.7762 (million USD), and 0.6301 by Copula model; The maximum money loss in the past ten years is 1.4678 (million USD) for PCR model, and 1.9242 (million USD) for Copula model. These errors are still tolerable for an asset worth tens of millions.

We have presented that the model has a worse performance for older aircraft, but the loss of money for older aircraft seems similar to the money loss of younger ones in Table. It is because the older aircraft are much cheaper than the young aircraft. The PCR and Copula model only estimate the market value growth, which is a percentage change of the market value. So when calculating the market value based on the market value growth, the smaller market value and bigger error mediate the money loss.

In Table 38, we are able to see the money loss for new aircraft of the 8 types. Except the PCR model of Airbus A330-300, the money loss made by the model will be smaller than 1 million USD.

Table 38 Money loss in 2011 prediction for new aircraft

<table>
<thead>
<tr>
<th>New aircraft</th>
<th>Money loss(million)</th>
<th>New aircraft</th>
<th>Money loss(million)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td>Copula</td>
<td>PCR</td>
<td>Copula</td>
</tr>
<tr>
<td>Airbus A319-100</td>
<td>-0.8915</td>
<td>Airbus A330-200</td>
<td>0.4121</td>
</tr>
<tr>
<td>Airbus A320-200</td>
<td>-0.5654</td>
<td>Airbus A330-300</td>
<td>1.8982</td>
</tr>
<tr>
<td>Airbus A321-200</td>
<td>-0.3473</td>
<td>Boeing 767-300ER</td>
<td>-0.2214</td>
</tr>
<tr>
<td>Boeing 737-800</td>
<td>-0.1747</td>
<td>Boeing 777-200</td>
<td>-0.7981</td>
</tr>
</tbody>
</table>

With the modeled market value, we are also able to compare the results of PCR and Copula model to the results in [8]. For a given type, we will take the average and max value of the
model error over year 2001 to year 2011 for aircraft at each age. For example, for Airbus A319-100, the average error is calculated by taking the average of errors of models for Airbus A319-100 at age 0 to Airbus A319-100 at age 5 over year 2001 to year 2011.

**Table 39 Comparison of money loss, model in [8], PCR model and Copula model, over year 2001 to year 2011**

<table>
<thead>
<tr>
<th></th>
<th>Model in [8]</th>
<th>PCR</th>
<th>Copula</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Max</td>
<td>Average</td>
</tr>
<tr>
<td>Airbus A319-100</td>
<td>1.1303</td>
<td>2.5329</td>
<td>0.8538</td>
</tr>
<tr>
<td>Airbus A320-200</td>
<td>1.2201</td>
<td>3.5782</td>
<td>0.8201</td>
</tr>
<tr>
<td>Airbus A321-200</td>
<td>1.1019</td>
<td>4.4593</td>
<td>0.6994</td>
</tr>
<tr>
<td>Boeing 737-800</td>
<td>1.2133</td>
<td>3.6815</td>
<td>0.8199</td>
</tr>
<tr>
<td>Airbus A330-300</td>
<td>3.0987</td>
<td>9.3491</td>
<td>1.5035</td>
</tr>
<tr>
<td>Boeing 767-300ER</td>
<td>3.4350</td>
<td>12.5657</td>
<td>2.0631</td>
</tr>
</tbody>
</table>

The result in Table 39 shows that, PCR model and Copula model makes a better prediction than the model in [8] both in the average and max measure of errors. We are sure that the two models present in this paper improved the results in [8] greatly.

**Section 5.3 Conclusion**

In this chapter, we check the performance of PCR and Copula models. Comparison between two models is made, and we also analyze the error of the model towards age, year, and aircraft type.

The comparison between PCR model and Copula model shows that Copula has a little better estimation of MV growth in average. Compared to PCR model, copula can take the maximum information in the explanatory variables; because for PCR method, there is a small part of information excluded from the principal component. It can explain why in average, Copula model has smaller error.

Copula also has a narrower 95% confidence interval, which may be helpful for the practical application of the model. For most aircraft, data in year 2002 is out of the 95% confidence interval, implying that the condition in year 2002 cannot be explained very well by the five explanatory variables. In fact, PCR model has a better prediction for year 2002 for most of the aircraft.
At the same time, as both PCR and Copula model are built based on the five explanatory variables selected in Chapter 2. We can see that there are some common problems with these two models.

First, for older aircraft, the estimation of MV growth is worse. The error increases steadily when the age of the aircraft modeled goes up. It is because the factors influencing the market value of older aircraft become much more complicated. Some policies of the government, new technology improvement all have some impact to the aircraft market value, and these influence on the older aircraft is stronger compared to the influence on young aircraft.

There is a system error of the model in year 2002. We can see that for copula method, models of all aircraft overestimate the market value growth in year 2002. The impact is from the 9-11 attack in year 2001. Despite of that the economy factor show that the recession should not go that far, the fear of terrorism beats the airline industry heavily and the aircraft market as well.

Another year that the market value growth is overestimated is year 2004. Though the economy factors had showed the sign of market recovery, people are still not confident about the aircraft market. The market value growth does not increase as fast as the model expect.

These two years imply that for year with unexpected event or for year of turning point, the model does not have a good performance. Again, it is because there is too much complicated reasons for the market except the market demand.

We also find that Boeing 767-300ER behave differently in year 2006 from other aircraft. The market value of Boeing 767-300ER increases sharply in year 2006; the truth is even out of the 95% confidence intervel of copula model. We can only explain this as some special strategy of Boeing company.

Based on the market value growth obtained from PCR or Copula model, we are able to calculate the market value. The results show that for Copula model, the average loss of money from is within 2 million for all aircraft. What’s more, both models improve the results in [8] greatly.
Chapter 6 Conclusions

This thesis presents a comprehensive analysis on the aircraft market. It is a complex process to make an aircraft appraisal in practice. However, based on the ‘average condition of an aircraft’ assumption, the problem is simplified greatly.

An exploration of the data available is made first to see the interrelationship between market value and other economy variables. It is found that it is the market value growth (the percentage change of market value) has a strong correlation with the growth of other economy factors. In general economy, it is the GDP growth and crude oil price change which have statistically significant correlation with the market value growth with all aircraft. In fact, the economy growth is the driven factor for traffic demand. Airline industry is the customer of the aircraft market. We pick revenue growth, passenger growth and RPK growth to indicate the health of the airline industry.

We model the market value growth with these five indicative factors selected. Two methods, Principal component regression and Copula, are applied separately. Though the predictions of market value growth of two models are similar, we can get some conclusions from two different aspects.

Principal component analysis is applied to handle the high correlation among the five explanatory variables. The principal component selected can be regarded as the average of the five variables; it represents the demand from the current economy environment. With the this demand indicator, we apply linear regression to model the market value growth. It is found that the market value growths of older aircraft are more sensitive to the change of the market. We also want to incorporate age into the model in two ways. One way is to use polynomial fitting to describe the relationship between age and model coefficient; another way is to group the aircraft that have similar behavior.

Copula is a natural way to model the whole dependency structure. By conditioning the economy factors, we are thus able to get the estimation of the market value growth for each
aircraft. Further, when modeling the dependence structure of economy factors and market value growths of two aircraft, one can check the conditional correlation between market value growth of two aircraft conditioning the economy factors. We analysis the interactive influence of these two aircraft in this way, and find that it is the characteristic of aircraft that determines its market value growth, as manufacturers of the aircraft do not have a strong influence. What’s more, for a given type, aircraft of similar age has a strong influence towards each other.

In average, Copula model has a better estimation, and PCR model has a smaller ‘maximum error’ over year 2001 to year 2010. The prediction for yer 2011 turns out to be quite good with money loss less than 1 million.

There are some common problems with the two models as they basically use the same predictive variables. The predictions for year 2002 and year 2004 are always higher than the truth. It transfer two pieces of information: the market in the year with ‘unexpected ’ event, and the year of ‘turning point’ cannot be explained simply by demanding level. People’s confidence towards the market may be lower than the true market condition after a sudden suffering from loss.

What’s more, one may find that the prediction for MV growth of older aircraft is worse than the prediction for MV growth of younger ones. In fact, one can see that the correlation coefficients between the indicative factors and market value growth for older aircraft are smaller; we can also see that a bigger variance for MV growth of older aircraft. It indicates that the influential factors for MV growth of older aircraft are more complicated than the five selected variables, and the MV growth of older aircraft are more sensitive to the change of market conditions.

Market values are calculated based on the prediction of market value growth. Results of two models are both better than results of model in [8], which support that it is a good way to pick the market value growth for modeling.
Bibliography


1. Pearson’s correlation coefficient

The product moment correlation of random variables $X, Y$ with finite expectation of $E(X), E(Y)$ and finite variances $\sigma_X^2, \sigma_Y^2$, is

$$\rho(X, Y) = \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y}$$

And if we can given $N$ pairs of samples $(x_i, y_i)$ from the random vector $(X, Y)$, the sample/ population product correlation can be calculated as following:

$$\rho(X, Y) = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2 \sum_{i=1}^{N} (y_i - \bar{y})^2}}$$

Where $\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$ and $\bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i$.

2. Spearman’s rank correlation coefficient

The rank correlation of random variables $X, Y$ with cumulative distribution functions $F_X$ and $F_Y$ is

$$\rho_r(X, Y) = \rho(F_X(X), F_Y(Y))$$

If we are given $N$ pairs of samples $(x_i, y_i)$ for the random vector $(X, Y)$, the raw data must be converted to ranks $(R_i, S_i)$ at first, where $R_i$ is the rank of $x_i$ among the other $x$’s, $S_i$ the rank of $y_i$ among the other $y$’s. Then the rank correlation is defined to be a product moment correlation of ranks can be calculated as following:

$$\rho_r(X, Y) = \frac{\sum_{i=1}^{N} (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^{N} (R_i - \bar{R})^2 \sum_{i=1}^{N} (S_i - \bar{S})^2}}$$

Where $\bar{R} = \frac{1}{N} \sum_{i=1}^{N} R_i$ and $\bar{S} = \frac{1}{N} \sum_{i=1}^{N} S_i$.

3. Significance test of Pearson’s /Spearman’s correlation coefficient

Suppose that we have obtained a Pearson’s /Spearman’s correlation coefficient $\rho$ for a sample with size $N$. We would like to verify the hypothesis as following:
\[ H_0: \rho = 0; \text{against} \]
\[ H_1: \rho \neq 0 \]

statistics \( t \) is defined as

\[
t = \frac{\rho \sqrt{N - 2}}{\sqrt{1 - \rho^2}}
\]

\( t \) follows a student distribution with \((N - 2)\)-degree freedom, i.e. \( t \sim t(N - 2) \)

One need to check the probability value (p-value) of the statistics \( t \) according to the student distribution \( t(N - 2) \). When the p-value is smaller than the significance level (for example, 0.05), we would reject the null hypothesis that the correlation coefficient is zero. In other words, only when p-value is smaller than the significance level, we would think that the correlation is statistically significant.

4. Classical multiple regression

Given a data set \( \{y_i, x_{i1}, \ldots, x_{ip}\}_{i=1}^n \) of \( n \) statistical units, a linear regression model assumes that the relationship between the dependent variables \( y_i \) and the \( p \)-vector of regressors \( X_i \) is linear. This relationship is modeled through a error variable \( \varepsilon_i \), which is an unobserved random variable that adds noise to the linear relationship between the dependent variable and regressors.

If we write the relation in matrix notation, we can get the model with the following form:

\[
Y = X\beta + \varepsilon
\]

Where

\[
Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}_{n \times (p+1)},
\]

\[
\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}
\]

\( X \) is called design matrix; \( X_i \) is called predictor variable/explanatory variable/independent variable/ regressor, and \( Y \) is called response variable/responding variable/dependent variable/ regressand under various contexts; \( \beta \) is the parameter needed to be estimated; \( \varepsilon \) is a random variable, which represents the difference of the model from the truth; it is assumed that \( \{\varepsilon_i\}_{i=1}^n \) is independent and identically distributed (i.i.d.) in a normal distribution with expectation zero.

We will use a linear model based on assumptions as following:
1) Constant variance: different response variables have the same variance in their errors, regardless of the values of the predictor variance

\[ \text{Var}(\varepsilon_i) = \sigma^2, \quad \text{for } i = 1, 2, \ldots, n \]

This assumption may be violated in the context of time series data.

2) Independence: The errors of the response variables are uncorrelated with each other

\[ \text{Cov}(\varepsilon_i, \varepsilon_j) = 0, \quad i \neq j, i, j = 1, 2, \ldots, n \]

In combination of assumption 3) and 4), we can get the conclusion that

\[ \text{var}(\varepsilon) = \sigma^2 I_n \]

3) normality is assumed, that the errors have normal distribution conditional on the regressors:

\[ \varepsilon | X \sim N(0, \sigma^2 I_n) \]

5. Ordinary least squares estimation (OLS)

Ordinary least squares estimation (OLS) is a common method for estimating the parameters in a linear regression model. It will find out the estimated parameters \( \hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p)^T \) which can minimizes the sum of squared vertical distances between the observed response in the data and the responses predicted by the linear approximation.

\[
\left( \hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p \right) = \min_{(\beta_0, \beta_1, \ldots, \beta_p)} \sum_{i=1}^{n} \left( y_i - \left( \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} \right) \right)^2
\]

we have a simple form for OLS estimation of parameters as following:

\[ \hat{\beta}_{OLS} = (X^T X)^{-1} X^T Y, \quad \text{if } \det(X^T X) \neq 0 \]

The variance-covariance matrix of the estimated regression coefficients is obtained as follows:

\[ C = \hat{\sigma}^2 (X^T X)^{-1} \]

The estimated standard error of \( \hat{\beta}_j \) is

\[ se(\hat{\beta}_j) = \sqrt{C_{jj}} \]

6. The determining coefficient R²

Following the notation in 4 and 5. We have observed values \( y_i \) and the modelled values \( \hat{y}_i \). The ‘variability’ of the data set is measured through different sums of squares:

\[ SS_{tot} = \sum_{i=1}^{N} (y_i - \bar{y})^2 \]

is the total sum of squares;
\[ SS_{reg} = \sum_{i=1}^{N} (\hat{y}_i - \bar{y})^2 \]
is the sum of regression squares;
\[ SS_{err} = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \]
is the sum of residuals squares.

Where \( \bar{y} \) is the mean of the observed value:
\[ \bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i \]

The determining coefficient \( R^2 \) is defined
\[ R^2 = 1 - \frac{SS_{err}}{SS_{tot}} \]

7. the significance test for determining coefficient \( R^2 \)
We apply F distribution for the significance test of determining coefficient \( R^2 \).
\[ H_0: \beta_0 = \beta_1 = \cdots = \beta_p = 0; \text{against} \]
\[ H_1: \text{at least one of the } \beta_i \neq 0 \]
the F-statistics \( t \) defined below follow a F-distribution
\[ t = \frac{SS_{reg}/p}{SS_{err}/N - (p + 1)} \sim F(p, N - (p + 1)) \]

When the p-value of statistics \( t \) is smaller than the significance level, we would think that the model as a whole is accepted.

8. significance test for individual coefficient
Hypothesis
\[ H_0: \beta_j = 0 \]
\[ H_1: \beta_j \neq 0 \]
The test statistic for the test is based on student distribution
\[ t = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \sim t(N - 2) \]
Where the standard error, \( se(\hat{\beta}_j) \) is defined in 5. We would reject the null hypothesis and accept that the contribution of this coefficient to the model is statistically significant only when p-value is smaller than the significance level.

9. Kaiser-Meyer-Olkin (KOM) test
The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is an index used to examine the appropriateness of factor analysis. High values (between 0.5 and 1.0) indicate factor analysis is appropriate. Values below 0.5 imply that factor analysis or principal component analysis may not be appropriate.

10 Bartlett’s test of Sphericity
The Bartlett’s test of sphericity is to test the null hypothesis that the variables are uncorrelated in the population. In other words, the population correlation matrix is an identity matrix. An identity matrix is the matrix in which all of the diagonal elements are 1 and all off diagonal elements are 0. Where this study needs to reject this null hypothesis. The study indicates strong relationship between all variables where in the correlation table all the diagonal elements are 1. In Bartlett’s test, this study needs to reject the null hypothesis of uncorrelated variable or non-identity matrix.

11. Durbin-Watson statistics
In statistics, the Durbin–Watson statistic is a test statistic used to detect the presence of autocorrelation (a relationship between values separated from each other by a given time lag) in the residuals (prediction errors) from a regression analysis.

If \( e_t \) is the residuals associate with the observation at time \( t \), then the test statistic is

\[
d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=2}^{T} e_t^2}
\]

To test for positive autocorrelation at significance \( \alpha \), the test statistic \( d \) is compared to lower and upper critical values (\( d_{L,\alpha} \) and \( d_{U,\alpha} \)):
- If \( d < d_{L,\alpha} \), there is a statistical evidence that the error terms are positively autocorrelated.
- If \( d > d_{U,\alpha} \), there is no statistical evidence that the error terms are positively autocorrelated.
- If \( d_{L,\alpha} < d < d_{U,\alpha} \), the test is inconclusive.

12. Principal component analysis
Given a data matrix \( X^T \) with zero empirical mean. This can be done by subtracting the empirical (sample) mean of the distribution from the original data set.

\[
X^T = (X_1, X_2, \ldots, X_p)^T
\]

Where \( X_t = (x_{i1}, x_{i2}, \ldots, x_{in})^T \) is one of the variables with sample number \( n \).

Applying singular value decomposition (SVD) on \( X \), we will get that

\[
X = W\Sigma V^T
\]
Where the $p \times p$ matrix $W$ is the matrix of eigenvectors of the covariance matrix $XX^T$; the $p \times n$ matrix $\Sigma$ is the rectangular diagonal matrix with nonnegative real numbers on the diagonal, and the $n \times n$ matrix $V$ is the matrix of eigenvectors of $X^TX$. Then the principal components transformation is given by

$$Z^T = X^TW = V\Sigma^TW = V \Sigma^T$$

When $p < n - 1$, $Z^T$ is uniquely defined. From the definition of SVD, we know that $W$ is an orthogonal matrix, so each row of $Z^T$ is simply a rotation of the corresponding row of $X^T$. The first column of $Z^T$ is made up of the ‘scores’ of the cases with respect to the ‘principal’ component, the next column has the scores with respect with the ‘second principal’ component, and so on. The singular values in $\Sigma$ are the square roots of the eigenvalues of the matrix $XX^T$. Each eigenvalue is proportional to the portion of the ‘variance’ that is correlated with each eigenvector.

13. Kolmogorov-Smirnov test

In statistics, the Kolmogorov–Smirnov test (K–S test) is a nonparametric test for the equality of continuous, one-dimensional probability distributions that can be used to compare a sample with a reference probability distribution, or to compare two samples.

The empirical distribution function $F_n$ for $n$ i.i.d. observation $X_i$ is defined as

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{X_i \leq x}$$

Where $I_{X_i \leq x}$ is the indicator function, equal to 1 if $X_i \leq x$ and equal to 0 otherwise.

The Kolmogorov-Smirnov statistic for a given cumulative distribution function $F(x)$ is

$$D_n = \sup_x |F_n(x) - F(x)|$$

Where $\sup_x$ is the supremum of the set of distances.

14. 95% confidence interval of the predicted values

The model of PCR (3.2) contains two parameters $\gamma_0, \gamma_1$. Applying OLS method, we can have the estimation $\hat{\gamma}_0, \hat{\gamma}_1$ for $\gamma_0, \gamma_1$. As long as there is an observed value of the component, one can have a predicted value for MV growth for the modelled aircraft as following:

$$\hat{Y} = \hat{\gamma}_0 + \hat{\gamma}_1 * x$$

The error of prediction is:

$$e_i = Y_i - \hat{Y}_i$$

The estimated variance of error $e$ is defined as
\[ \hat{\sigma}_e = \frac{\sum_{i=1}^{n} e_i^2}{n - 2} \]

and the statistic \( t \) will follow a t-distribution

\[ t = \frac{Y - \hat{Y}}{\hat{\sigma}_e} \sim t(\nu) \]

where \( \hat{\sigma}_e \) is the estimated standard variance of the error \( e \).\(^{26}\)

The 95% confidence interval of predicted values on the observation is obtained as follows:

\[ \hat{Y}_i \pm t(\nu)_{0.025} \cdot \hat{\sigma}_e \sqrt{1 + \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}} \]

15. Mean of absolute Error (MAE)

Given a set of errors \( \{e_i\}_{i=1}^{N} \), the mean of absolute error is defined as

\[ MAE = \frac{\sum_{i=1}^{N} |e_i|}{N} \]

16. Maximum of Absolute Error (MAXE)

Given a set of errors \( \{e_i\}_{i=1}^{N} \), the Maximum of absolute error is defined as

\[ MAXE = \max_{i=1,...,N} |e_i| \]

17. Kernel density estimation

In statistics, kernel density estimation (KDE) is a non-parametric way to estimate the probability density function of a random variable.

Let \( (x_1, x_2, ..., x_n) \) be an iid sample drawn from some distribution with an unknown density \( f \). We are interested in estimating the shape of the function \( f \). Its kernel density estimator is

\[ \hat{f}_h = \frac{1}{n} \sum_{i=1}^{n} K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^{n} K_h \left( \frac{x - x_i}{h} \right) \]

Where \( K(\cdot) \) is the kernel – a symmetric but not necessarily positive function that integrates to one, and \( h > 0 \) is a smoothing parameter called the bandwidth. A kernel with subscript \( h \) is called the scaled kernel and defined as \( K_h(x) = \frac{1}{h} K \left( \frac{x}{h} \right) \).

18. The joint normal distribution

\(^{26}\)One can check the std.variance of error \( e \) in appendix B
For a known multivariate normal distribution of a $k$-dimensional random vector $X = [X_1, X_2, ..., X_k]$ can be written in the following notation:

$$X \sim N(\mu, \Sigma)$$

With $k$-dimensional mean vector

$$\mu = [E(X_1), E(X_2), ..., E(X_k)]$$

And $k \times k$ covariance matrix

$$\Sigma = [\text{Cov}(X_i, X_j)], i,j = 1,2, ..., k$$

19. Sklar theorem (1959)

Consider a $d$-dimensional cdf $F$ with marginals $F_1, ..., F_d$. There exists a copula $C$, such that

$$F(x_1, x_2, ..., x_n) = C\left(F_{X_1}(x_1), F_{X_2}(x_2), ..., F_{X_d}(x_d)\right)$$

For all $x_i \in \mathbb{R}, i = 1, ..., d$. If $F_i$ is continuous for all $i = 1, ..., d$. Then $C$ is unique; otherwise $C$ is uniquely determined only on $\text{Ran} F_1 \times \cdots \times \text{Ran} F_d$, where $\text{Ran} F_i$ denotes the range of the cdf $F_i$. In the explanatory example with dice, the range was $\left\{\frac{1}{6},\frac{2}{6}, ..., \frac{6}{6}\right\}$, while for a continuous random variable this is always $[0, 1]$.

20. The conditional normal distribution

If $\mu, \Sigma$ are partitioned as follows

\[
\mu = \begin{pmatrix}
\mu_1 \\
\mu_2
\end{pmatrix}_{k}
\text{with size \(q \choose N - q\)}
\]

\[
\Sigma = \begin{pmatrix}
\Sigma_{11} & \Sigma_{12} \\
\Sigma_{21} & \Sigma_{22}
\end{pmatrix}_{k \times k}
\text{with size \(q \times q\), \(q \times (N - q)\), \((N - q) \times q\), \((N - q) \times (N - q)\)}
\]

Then the distribution of $x_1 = [X_1, ..., X_q]$ conditional on $x_2 = [X_{q+1}, ..., X_k] = a$ is a multivariate normal $(x_1 | x_2 = a) \sim N(\bar{\mu}, \bar{\Sigma})$, where

$$\bar{\mu} = \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (a - \mu_2)$$

And covariance matrix

$$\bar{\Sigma} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$$
### Appendix B

**B.1 predicted market value based on PCR model**

**Airbus A319-100**

<table>
<thead>
<tr>
<th>Age</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>32.66</td>
<td>31.28</td>
<td>31.52</td>
<td>32.80</td>
<td>34.67</td>
<td>35.77</td>
<td>36.76</td>
<td>31.70</td>
<td>31.00</td>
<td>31.99</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>30.83</td>
<td>29.16</td>
<td>29.22</td>
<td>30.23</td>
<td>31.11</td>
<td>33.56</td>
<td>34.47</td>
<td>29.34</td>
<td>28.69</td>
<td>29.53</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>29.05</td>
<td>27.34</td>
<td>24.68</td>
<td>27.42</td>
<td>29.22</td>
<td>30.23</td>
<td>32.11</td>
<td>27.50</td>
<td>26.67</td>
<td>27.52</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>27.52</td>
<td>25.78</td>
<td>23.16</td>
<td>25.96</td>
<td>26.49</td>
<td>28.39</td>
<td>29.84</td>
<td>30.81</td>
<td>25.91</td>
<td>24.96</td>
<td>25.75</td>
</tr>
<tr>
<td>5</td>
<td>24.98</td>
<td>22.81</td>
<td>20.55</td>
<td>23.10</td>
<td>24.93</td>
<td>26.54</td>
<td>28.15</td>
<td>29.17</td>
<td>23.35</td>
<td>22.70</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>38.81</td>
<td>34.87</td>
<td>32.67</td>
<td>36.43</td>
<td>38.44</td>
<td>41.72</td>
<td>41.09</td>
<td>43.13</td>
<td>37.93</td>
<td>36.77</td>
<td>39.47</td>
</tr>
</tbody>
</table>

**Airbus A320-200**

<table>
<thead>
<tr>
<th>Age</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>38.81</td>
<td>34.87</td>
<td>32.67</td>
<td>36.43</td>
<td>38.44</td>
<td>41.72</td>
<td>41.09</td>
<td>43.13</td>
<td>37.93</td>
<td>36.77</td>
<td>39.47</td>
</tr>
<tr>
<td>1</td>
<td>36.71</td>
<td>32.79</td>
<td>30.60</td>
<td>34.09</td>
<td>35.77</td>
<td>38.55</td>
<td>38.47</td>
<td>40.49</td>
<td>35.17</td>
<td>33.77</td>
<td>36.30</td>
</tr>
<tr>
<td>2</td>
<td>34.86</td>
<td>30.93</td>
<td>28.74</td>
<td>32.27</td>
<td>33.68</td>
<td>36.21</td>
<td>36.13</td>
<td>38.16</td>
<td>32.83</td>
<td>31.02</td>
<td>33.59</td>
</tr>
<tr>
<td>3</td>
<td>33.28</td>
<td>29.24</td>
<td>26.98</td>
<td>30.71</td>
<td>31.85</td>
<td>34.02</td>
<td>34.05</td>
<td>36.18</td>
<td>30.77</td>
<td>28.50</td>
<td>31.18</td>
</tr>
<tr>
<td>4</td>
<td>31.71</td>
<td>27.65</td>
<td>25.34</td>
<td>29.17</td>
<td>30.04</td>
<td>32.07</td>
<td>32.10</td>
<td>34.28</td>
<td>28.80</td>
<td>26.19</td>
<td>29.01</td>
</tr>
<tr>
<td>6</td>
<td>28.87</td>
<td>24.77</td>
<td>22.27</td>
<td>26.06</td>
<td>26.92</td>
<td>28.13</td>
<td>28.44</td>
<td>30.61</td>
<td>25.07</td>
<td>22.03</td>
<td>24.96</td>
</tr>
<tr>
<td>7</td>
<td>27.62</td>
<td>23.55</td>
<td>21.00</td>
<td>24.70</td>
<td>25.28</td>
<td>26.05</td>
<td>26.66</td>
<td>28.80</td>
<td>23.41</td>
<td>20.19</td>
<td>23.04</td>
</tr>
<tr>
<td>9</td>
<td>25.14</td>
<td>21.43</td>
<td>18.78</td>
<td>22.21</td>
<td>22.49</td>
<td>22.56</td>
<td>23.25</td>
<td>25.33</td>
<td>20.36</td>
<td>17.00</td>
<td>19.49</td>
</tr>
<tr>
<td>10</td>
<td>23.92</td>
<td>20.52</td>
<td>17.88</td>
<td>21.05</td>
<td>21.05</td>
<td>20.71</td>
<td>21.61</td>
<td>23.64</td>
<td>19.01</td>
<td>15.64</td>
<td>17.86</td>
</tr>
<tr>
<td>11</td>
<td>22.74</td>
<td>19.69</td>
<td>17.10</td>
<td>19.96</td>
<td>19.83</td>
<td>18.64</td>
<td>20.02</td>
<td>21.99</td>
<td>17.81</td>
<td>14.41</td>
<td>16.31</td>
</tr>
<tr>
<td>12</td>
<td>21.74</td>
<td>18.91</td>
<td>16.38</td>
<td>18.99</td>
<td>18.68</td>
<td>17.06</td>
<td>17.91</td>
<td>20.57</td>
<td>16.70</td>
<td>13.27</td>
<td>14.84</td>
</tr>
</tbody>
</table>

**Airbus A321-200**

<table>
<thead>
<tr>
<th>Age</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boeing 767-300ER</td>
<td>age 1</td>
<td>age 2</td>
<td>age 3</td>
<td>age 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td>45.07</td>
<td>40.11</td>
<td>38.95</td>
<td>37.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td>42.29</td>
<td>37.57</td>
<td>35.98</td>
<td>34.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td>40.11</td>
<td>35.55</td>
<td>33.75</td>
<td>32.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>41.17</td>
<td>38.49</td>
<td>36.61</td>
<td>34.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>45.17</td>
<td>39.96</td>
<td>37.80</td>
<td>35.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>46.91</td>
<td>41.23</td>
<td>38.85</td>
<td>36.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>46.17</td>
<td>40.78</td>
<td>38.43</td>
<td>36.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>46.93</td>
<td>41.52</td>
<td>39.15</td>
<td>37.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>43.77</td>
<td>38.55</td>
<td>36.35</td>
<td>34.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>43.30</td>
<td>37.75</td>
<td>35.47</td>
<td>33.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>43.45</td>
<td>37.77</td>
<td>35.39</td>
<td>33.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Boeing 737-800</th>
<th>age 0</th>
<th>age 1</th>
<th>age 2</th>
<th>age 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>39.53</td>
<td>37.45</td>
<td>35.62</td>
<td>34.28</td>
</tr>
<tr>
<td>2002</td>
<td>35.47</td>
<td>33.47</td>
<td>31.60</td>
<td>30.01</td>
</tr>
<tr>
<td>2003</td>
<td>32.85</td>
<td>30.80</td>
<td>28.99</td>
<td>27.67</td>
</tr>
<tr>
<td>2004</td>
<td>36.89</td>
<td>34.71</td>
<td>32.91</td>
<td>31.34</td>
</tr>
<tr>
<td>2005</td>
<td>38.98</td>
<td>36.24</td>
<td>34.45</td>
<td>32.88</td>
</tr>
<tr>
<td>2006</td>
<td>42.35</td>
<td>39.14</td>
<td>37.08</td>
<td>35.26</td>
</tr>
<tr>
<td>2007</td>
<td>44.11</td>
<td>41.67</td>
<td>39.86</td>
<td>38.34</td>
</tr>
<tr>
<td>2008</td>
<td>45.37</td>
<td>42.69</td>
<td>40.64</td>
<td>38.81</td>
</tr>
<tr>
<td>2009</td>
<td>38.77</td>
<td>36.12</td>
<td>33.98</td>
<td>32.14</td>
</tr>
<tr>
<td>2010</td>
<td>38.20</td>
<td>35.41</td>
<td>33.10</td>
<td>31.05</td>
</tr>
<tr>
<td>2011</td>
<td>40.67</td>
<td>37.46</td>
<td>34.88</td>
<td>32.78</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Airbus A330-200</th>
<th>age 0</th>
<th>age 1</th>
<th>age 2</th>
<th>age 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>85.41</td>
<td>81.06</td>
<td>77.37</td>
<td>74.31</td>
</tr>
<tr>
<td>2002</td>
<td>75.59</td>
<td>71.47</td>
<td>68.44</td>
<td>66.59</td>
</tr>
<tr>
<td>2003</td>
<td>69.61</td>
<td>65.85</td>
<td>63.19</td>
<td>61.16</td>
</tr>
<tr>
<td>2004</td>
<td>79.31</td>
<td>73.65</td>
<td>69.70</td>
<td>66.80</td>
</tr>
<tr>
<td>2005</td>
<td>82.37</td>
<td>76.66</td>
<td>72.14</td>
<td>68.65</td>
</tr>
<tr>
<td>2006</td>
<td>90.27</td>
<td>82.58</td>
<td>77.13</td>
<td>72.73</td>
</tr>
<tr>
<td>2007</td>
<td>90.67</td>
<td>83.13</td>
<td>77.79</td>
<td>73.53</td>
</tr>
<tr>
<td>2008</td>
<td>99.13</td>
<td>91.45</td>
<td>84.97</td>
<td>79.56</td>
</tr>
<tr>
<td>2009</td>
<td>81.84</td>
<td>74.74</td>
<td>69.62</td>
<td>65.73</td>
</tr>
<tr>
<td>2010</td>
<td>78.82</td>
<td>71.19</td>
<td>65.65</td>
<td>61.71</td>
</tr>
<tr>
<td>2011</td>
<td>83.59</td>
<td>75.38</td>
<td>69.41</td>
<td>65.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Airbus A330-300</th>
<th>age 0</th>
<th>age 1</th>
<th>age 2</th>
<th>age 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>90.70</td>
<td>84.48</td>
<td>79.42</td>
<td>72.12</td>
</tr>
<tr>
<td>2002</td>
<td>80.29</td>
<td>75.21</td>
<td>71.09</td>
<td>64.48</td>
</tr>
<tr>
<td>2003</td>
<td>74.85</td>
<td>70.65</td>
<td>67.13</td>
<td>61.42</td>
</tr>
<tr>
<td>2004</td>
<td>83.88</td>
<td>77.92</td>
<td>73.31</td>
<td>65.75</td>
</tr>
<tr>
<td>2005</td>
<td>86.57</td>
<td>80.58</td>
<td>75.70</td>
<td>68.04</td>
</tr>
<tr>
<td>2006</td>
<td>91.50</td>
<td>85.07</td>
<td>79.29</td>
<td>69.82</td>
</tr>
<tr>
<td>2007</td>
<td>91.49</td>
<td>83.87</td>
<td>78.26</td>
<td>69.56</td>
</tr>
<tr>
<td>2008</td>
<td>99.56</td>
<td>91.59</td>
<td>84.54</td>
<td>73.06</td>
</tr>
<tr>
<td>2009</td>
<td>89.08</td>
<td>81.35</td>
<td>75.22</td>
<td>64.92</td>
</tr>
<tr>
<td>2010</td>
<td>85.81</td>
<td>77.23</td>
<td>70.80</td>
<td>60.30</td>
</tr>
<tr>
<td>2011</td>
<td>90.85</td>
<td>81.80</td>
<td>74.70</td>
<td>62.92</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Boeing 767-300ER</th>
<th>age 0</th>
<th>age 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>71.73</td>
<td>68.11</td>
</tr>
<tr>
<td>2002</td>
<td>60.01</td>
<td>54.83</td>
</tr>
<tr>
<td>2003</td>
<td>53.01</td>
<td>48.01</td>
</tr>
<tr>
<td>2004</td>
<td>59.26</td>
<td>53.72</td>
</tr>
<tr>
<td>2005</td>
<td>62.34</td>
<td>56.27</td>
</tr>
<tr>
<td>2006</td>
<td>62.25</td>
<td>56.22</td>
</tr>
<tr>
<td>2007</td>
<td>68.18</td>
<td>63.18</td>
</tr>
<tr>
<td>2008</td>
<td>71.48</td>
<td>66.52</td>
</tr>
<tr>
<td>2009</td>
<td>58.73</td>
<td>53.50</td>
</tr>
<tr>
<td>2010</td>
<td>57.33</td>
<td>52.30</td>
</tr>
<tr>
<td>2011</td>
<td>59.12</td>
<td>54.65</td>
</tr>
</tbody>
</table>
### Prediction of market value of used commercial aircraft

<table>
<thead>
<tr>
<th>Age</th>
<th>Boeing 777-200</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>109.36</td>
<td>103.61</td>
<td>97.57</td>
<td>102.93</td>
<td>108.52</td>
<td>110.85</td>
<td>119.17</td>
<td>125.52</td>
<td>114.63</td>
<td>114.19</td>
<td>118.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>100.28</td>
<td>95.49</td>
<td>90.52</td>
<td>94.28</td>
<td>99.82</td>
<td>102.36</td>
<td>109.78</td>
<td>115.75</td>
<td>104.34</td>
<td>103.68</td>
<td>106.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>94.94</td>
<td>89.28</td>
<td>84.27</td>
<td>87.49</td>
<td>93.27</td>
<td>95.90</td>
<td>102.33</td>
<td>107.98</td>
<td>95.47</td>
<td>94.94</td>
<td>97.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>90.61</td>
<td>84.61</td>
<td>79.01</td>
<td>82.21</td>
<td>87.99</td>
<td>90.62</td>
<td>96.53</td>
<td>101.39</td>
<td>88.27</td>
<td>87.70</td>
<td>90.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>86.43</td>
<td>80.54</td>
<td>74.89</td>
<td>77.87</td>
<td>83.36</td>
<td>85.81</td>
<td>91.25</td>
<td>95.54</td>
<td>82.46</td>
<td>81.61</td>
<td>84.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### B. 2 predicted market value based on copula model

#### Airbus A319-100

<table>
<thead>
<tr>
<th>Age</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>33.69</td>
<td>32.04</td>
<td>28.86</td>
<td>31.06</td>
<td>33.15</td>
<td>34.73</td>
<td>36.08</td>
<td>35.78</td>
<td>31.61</td>
<td>30.99</td>
<td>31.64</td>
</tr>
<tr>
<td>1</td>
<td>31.71</td>
<td>30.03</td>
<td>26.68</td>
<td>28.69</td>
<td>30.64</td>
<td>32.16</td>
<td>33.83</td>
<td>33.62</td>
<td>29.35</td>
<td>28.41</td>
<td>29.20</td>
</tr>
<tr>
<td>2</td>
<td>29.90</td>
<td>28.42</td>
<td>24.94</td>
<td>26.76</td>
<td>28.69</td>
<td>30.16</td>
<td>31.77</td>
<td>31.86</td>
<td>27.41</td>
<td>26.34</td>
<td>27.27</td>
</tr>
<tr>
<td>3</td>
<td>28.40</td>
<td>27.25</td>
<td>23.43</td>
<td>25.04</td>
<td>26.96</td>
<td>28.48</td>
<td>30.01</td>
<td>30.31</td>
<td>25.88</td>
<td>24.64</td>
<td>25.42</td>
</tr>
<tr>
<td>4</td>
<td>27.00</td>
<td>25.46</td>
<td>21.97</td>
<td>23.60</td>
<td>25.25</td>
<td>26.70</td>
<td>28.31</td>
<td>28.85</td>
<td>24.28</td>
<td>22.93</td>
<td>23.93</td>
</tr>
<tr>
<td>5</td>
<td>25.55</td>
<td>24.33</td>
<td>20.78</td>
<td>22.06</td>
<td>23.64</td>
<td>25.16</td>
<td>26.78</td>
<td>27.50</td>
<td>22.91</td>
<td>21.40</td>
<td>22.49</td>
</tr>
</tbody>
</table>

#### Airbus A320-200

<table>
<thead>
<tr>
<th>Age</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>38.70</td>
<td>34.74</td>
<td>32.67</td>
<td>35.38</td>
<td>39.23</td>
<td>41.00</td>
<td>41.08</td>
<td>43.25</td>
<td>39.14</td>
<td>37.77</td>
<td>39.89</td>
</tr>
<tr>
<td>1</td>
<td>36.86</td>
<td>32.89</td>
<td>30.84</td>
<td>32.89</td>
<td>36.28</td>
<td>38.00</td>
<td>38.59</td>
<td>40.56</td>
<td>36.01</td>
<td>34.65</td>
<td>36.46</td>
</tr>
<tr>
<td>2</td>
<td>35.15</td>
<td>31.34</td>
<td>29.20</td>
<td>30.91</td>
<td>34.01</td>
<td>35.69</td>
<td>36.23</td>
<td>38.33</td>
<td>33.32</td>
<td>31.97</td>
<td>33.69</td>
</tr>
<tr>
<td>Age 1</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005</td>
<td>2006</td>
<td>2007</td>
<td>2008</td>
<td>2009</td>
<td>2010</td>
<td>2011</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Age 0</td>
<td>44.48</td>
<td>42.05</td>
<td>39.93</td>
<td>42.89</td>
<td>45.42</td>
<td>46.91</td>
<td>46.48</td>
<td>47.64</td>
<td>44.14</td>
<td>43.27</td>
<td>43.91</td>
</tr>
<tr>
<td>Age 1</td>
<td>41.93</td>
<td>39.71</td>
<td>37.64</td>
<td>40.21</td>
<td>42.40</td>
<td>43.89</td>
<td>43.59</td>
<td>44.88</td>
<td>41.24</td>
<td>40.20</td>
<td>40.80</td>
</tr>
<tr>
<td>Age 2</td>
<td>39.99</td>
<td>37.82</td>
<td>35.65</td>
<td>37.89</td>
<td>39.83</td>
<td>41.24</td>
<td>41.05</td>
<td>42.40</td>
<td>38.90</td>
<td>37.72</td>
<td>38.17</td>
</tr>
<tr>
<td>Age 3</td>
<td>38.37</td>
<td>36.12</td>
<td>33.86</td>
<td>35.92</td>
<td>37.54</td>
<td>38.93</td>
<td>38.66</td>
<td>40.20</td>
<td>36.71</td>
<td>35.43</td>
<td>35.83</td>
</tr>
<tr>
<td>Age 4</td>
<td>36.72</td>
<td>34.87</td>
<td>32.52</td>
<td>34.22</td>
<td>35.52</td>
<td>36.83</td>
<td>36.53</td>
<td>38.17</td>
<td>34.76</td>
<td>33.37</td>
<td>33.61</td>
</tr>
</tbody>
</table>

**Airbus A330-200**

| Age 0 | 85.83| 76.38| 69.74| 76.47| 84.45| 89.02| 92.60| 95.97| 88.24| 77.90| 83.17|
| Age 1 | 81.82| 72.25| 66.07| 71.66| 78.21| 81.92| 85.09| 88.43| 79.57| 70.35| 74.92|
| Age 2 | 78.09| 69.76| 63.50| 67.76| 73.64| 76.67| 79.35| 82.84| 72.98| 64.83| 69.25|
| Age 3 | 75.43| 68.65| 61.72| 64.71| 69.81| 72.50| 74.79| 77.93| 67.95| 61.13| 64.81|

**Airbus A330-300**

| Age 0 | 88.52| 81.03| 75.27| 81.35| 87.25| 91.23| 93.18| 102.08| 93.01| 85.81| 92.83|
| Age 1 | 81.82| 72.25| 66.07| 71.66| 78.21| 81.92| 85.09| 88.43| 79.57| 70.35| 74.92|
B. 3 polynomial function of coefficients of PCR with respect to age

For principal component regression, we have form

```markdown
<table>
<thead>
<tr>
<th>Age</th>
<th>Boeing 767-300ER</th>
<th>Boeing 777-200</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

```
Applying polynomial fitting, we are able to model $Y_0, Y_1$ with respect to age $x$.

The following table shows the fitted line for Boeing 767-300ER

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>$x = 0,1,\ldots,5$</th>
<th>$x = 0,1,\ldots,12$</th>
<th>$x = 0,1,\ldots,4$</th>
<th>$x = 0,1,\ldots,3$</th>
<th>$x = 0,1,\ldots,8$</th>
<th>$x = 0,1,\ldots,14$</th>
<th>$x = 0,1,\ldots,3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airbus A319-100</td>
<td>$Y_0$</td>
<td>$y(x) = -0.1181x - 0.2204$</td>
<td>$y(x) = -0.00793x^2 - 0.173x + 0.1164$</td>
<td>$y(x) = 0.0168x^2 - 0.1363x - 0.4057$</td>
<td>$y(x) = 0.06456x^2 - 0.5668x - 0.13$</td>
<td>$y(x) = -0.009192x^2 - 0.3294x + 0.08019$</td>
<td>$y(x) = -0.01071x^2 - 0.05794x - 1.815$</td>
<td>$y(x) = -0.2773x + 0.8598$</td>
</tr>
<tr>
<td></td>
<td>$Y_1$</td>
<td>$y(x) = 0.2246x + 3.213$</td>
<td>$y(x) = -0.002615x^2 + 0.02974x^2 + 0.129x + 3.525$</td>
<td>$y(x) = -0.01513x^2 + 0.126x + 2.353$</td>
<td>$y(x) = 0.2316x^2 + 0.1129x + 3.702$</td>
<td>$y(x) = -0.007747x^2 + 0.018x + 3.508$</td>
<td>$y(x) = 0.001826x^3 - 0.03623x^2 + 0.3417x + 4.037$</td>
<td>$y(x) = 0.0106x^2 + 0.0848x + 3.131$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**B. 3 polynomial function of coefficients of copula with respect to age**

The following table shows the fitted line for parameter $s$ of model for Boeing 767-300ER

<table>
<thead>
<tr>
<th>parameter</th>
<th>Fitted trend line with respect with age</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_0$</td>
<td>$y(x) = -0.1181x - 0.2204$</td>
</tr>
<tr>
<td>$Y_1$</td>
<td>$y(x) = 0.2246x + 3.213$</td>
</tr>
<tr>
<td></td>
<td>Equation</td>
</tr>
<tr>
<td>------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>( y(x) = -0.01127x^2 - 0.0368x - 1.826 )</td>
</tr>
<tr>
<td><strong>Std.deviation</strong></td>
<td>( y(x) = 0.005264x^3 - 0.1094x^2 + 1.037x + 9.178 )</td>
</tr>
<tr>
<td><strong>Correlation 1</strong></td>
<td>(-0.0001605x^3 + 0.004093x - 0.0296x + 0.7191)</td>
</tr>
<tr>
<td><strong>Correlation 2</strong></td>
<td>(-4.986 \times 10^{-5}x^3 + 0.001173x^2 - 0.006577x + 0.6267)</td>
</tr>
<tr>
<td><strong>Correlation 3</strong></td>
<td>(-2.68 \times 10^{-5}x^3 + 0.0006395x^2 - 0.007128x + 0.8828)</td>
</tr>
<tr>
<td><strong>Correlation 4</strong></td>
<td>(0.0006204x^2 - 0.01631x + 0.8398)</td>
</tr>
<tr>
<td><strong>Correlation 5</strong></td>
<td>(-6.975 \times 10^{-5}x^3 + 0.00172x^2 - 0.01486x + 0.7915)</td>
</tr>
</tbody>
</table>

*Correlation 1 is the correlation coefficients between MV growth and GDP growth
*Correlation 2 is the correlation coefficients between MV growth and oil price change
*Correlation 3 is the correlation coefficients between MV growth and airline industry revenue growth
*Correlation 4 is the correlation coefficients between MV growth and RPK growth
*Correlation 5 is the correlation coefficients between MV growth and passenger growth