# Empirical Bayes Estimation of Panel Data Models

Researcher
Khan Bahadar
Registration Number: 20. SE/Ph.D. Et /S12

Supervisors

Prof. Dr. Asad Zaman
Vice Chancellor
Pakistan Institute of Development Economics (PIDE)
Islamabad

Prof. Dr. Eatzaz Ahmad
Quaid-i-Azam University Islamabad

School of Economics
International Institute of Islamic Economics (IIIE),
International Islamic University, Islamabad (IIUI)
Pakistan.

# Empirical Bayes Estimation of Panel Data Models


Researcher

Khan Bahadar

Registration Number: 20. SE/Ph.D. Et /S12


Accepted by the

International Institute of Islamic Economics (IIIE)

International Islamic University, Islamabad (IIUI),

as partial fulfilment of the requirements for the award of degree of

DOCTOR OF PHILOSOPHY IN ECONOMETRICS


Supervisors

Dr. Asad Zaman

Dr. Eatzaz Ahmad

# ABOUT THE AUTHOR

The name of the author of this dissertation is Khan Bahadar. He belongs to village Islampur, Saidu Sharif, Swat, Khyber Pakhtunkhwa, Pakistan. He received his early education up to $5^{th}$ class and then up to $8^{th}$ class from Government Primary School Islampur Swat and Government Middle School Islampur Swat, respectively. He has done Matric from Government High School Chitor Swat. He has completed Inter and Bachelor level studies from Government Post Graduate Jahanzeb College Saidu Sharif Swat and M.Sc. in Statistics from Government Post Graduate College Nowshera in Nowshera. He stood first in M.Sc. Statistics. Further, he has earned M. Phil degree from International Islamic University, Islamabad. Recently, he has completed Ph.D. in Econometrics from International Islamic University, Islamabad under the supervision of Ex-Vice Chancellor (VC) of Pakistan Institute of Development Economics (PIDE) Prof. Dr. Asad Zaman, a well reputed scholar around the world and Prof. Dr. Eatzaz Ahmad, Professor at Quaid-e-Azam University Islamabad.

Mr. Khan Bahadar has attained several positions through Khyber Pakhtunkhwa Public Service Commission (KPPSC). In 2010, he had been recommended by KPPSC for the post of Subject Specialist in Statistics. In 2012, he had been recommended by KPPSC for the post of Lecturer in Statistics. In 2019, he had been recommended by KPPSC for the post of Assistant Professor of Statistics. Mr. Khan Bahadar stood first in the written examination for the post of Assistant Professor of Statistics. Moreover, in 2019 he also stood first in the written examination for the post of Associate Professor of Statistics but due lack of experience he has not been called for the final interview of the post and hence could not availed the post of Associate Professor that time.

Mr. Khan Bahadar is currently working as Assistant Professor of Statistics.

# DECLARATION

I, Khan Bahadar, hereby declare that this thesis, neither as a whole nor as a part thereof, has been copied out from any source. It is further declared that I have carried out this research by myself and have completed this thesis on the basis of my personal efforts under the guidance and help of my supervisors. If any part of this thesis is proven to be copied out or earlier submitted, I shall stand by the consequences. No portion of work presented in this thesis has been submitted in support of any application for any other degree or qualification in International Islamic University Islamabad or any other university or institute of learning.

Khan Bahadar

# In the Memory of My Late Parents

## Mother

Jazakallah, for your unconditional love during your life. I am honoured to have you as my mother. You had cared me a lot, but I could not do anything for you. I am extremely happy on achieving my goals but feel really sorry that you are no more here in this world.

## Father

Today with the blessing of Allah (SWT) I have achieved whatever you were expecting of me but have no idea that how this could be conveyed to you as you are no longer in this world.

May Allah (SWT) grant eternal rest to the souls of my Parents.

Ameen.

# Dedication

This Thesis is fully dedicated to my wife **Dr. Seema Ali Khan** and Sons **Ifhaam Khan** and **Baasim Khan**.

APPROVAL SHEET

**"Empirical Bayes Estimation of Panel Data Model"**

By

**Mr. Khan Bahadar**

**Reg. No:** 20-SE/Ph.D (Et)/S12

Accepted by the International Institute of Islamic Economics (IIIE), International Islamic University Islamabad (IIUI), as partial fulfillment of the requirements for the award of degree of PhD in Econometrics

Supervisor I: ____________________
(Pro. Dr. Eatzaz Ahmed)
Professor, QAU, Islamabad

Supervisor II: ____________________
(Prof. Dr. Asad Zaman)
Professor, Ex-VC, PIDE, Islamabad

Internal Examiner: ____________________
(Prof. Dr. Abdul Rashid)
Professor, IIIE, IIUI

External Examiner I : ____________________
Dr. Tanweer-ul-Islam
Associate Professor, S3H, NUST, Islamabad

External Examiner II : ____________________
Dr. Saud Ahmed Khan
Assistant Professor, PIDE, Islamabad

Head,
School of Economics
International Institute of Islamic Economics
International Islamic University, Islamabad

Director General,
International Institute of Islamic Economics
International Islamic University, Islamabad

Date of Viva Voce Examination: August 19, 2022

# Acknowledgement

It is Allah (SWT) who helps me in all respects of life and gets solved my problems with unbelievable sources. So first of all I am thankful to Almighty ALLAH.

Thereafter, first and foremost I would like to thank one of my supervisors and a worldwide well known scholar in the field of Statistics and Econometrics Prof. Dr. Asad Zaman, Vice Chancellor Pakistan Institute of Development Economics (PIDE) for supporting and guiding me during the entire Ph.D. studies.

I also want to thank my supervisor Prof. Dr. Eatzaz Ahmed, Professor at Quaid-i-Azam University Islamabad for his extraordinary help during thesis.

I am very much thankful to all my family members, sister and brothers, teachers, relatives, friends, colleagues, students for their support and courage and all those who helped me one way or the other.

There are many more who deserve to be thanked, but in short, I am thankful to all those who had direct or indirect positive effect upon my studies.

# ABSTRACT

This thesis mainly concentrates on the empirical Bayes estimation of all standard classical panel data models. The regression model that is used in this thesis consists of $K$ orthogonal regressors. The estimation technique that is used for the coefficients of the regressors is empirical Bayes. The strategy that is followed for the estimation of regression coefficients is single coefficient estimation, i.e., one coefficient from among all $K$ coefficients is estimated at a time. By the virtue of single orthogonal regression coefficient estimation with empirical Bayes several ideas have been contributed with the help of this dissertation.

The first contribution and main purpose of this thesis is to illustrate that all the standard frequentist panel data linear regression models and their corresponding estimators emerge from a single origin. That is to say, all the frequentist panel data linear regression models are the special cases of the Bayesian linear regression model and similarly all the frequentist panel data estimates are the special cases of the empirical Bayes estimate.

The second contribution is the origination of; (i) all the standard frequentist panel data linear regression models from a Bayesian linear regression model and (ii) the estimates corresponding to standard frequentist panel data linear regression models from the empirical Bayes estimate, analytically as well as numerically.

In fact, the frequentist panel data linear regression models and the corresponding estimators can be attained from the empirical Bayes by imposing certain restrictions upon the prior precision parameter. At this time, the structure of the empirical Bayes estimate is the precision weighted arithmetic mean of the prior mean and data mean, and where the precision weighted arithmetic mean of the ordinary least squares coefficients estimates as the estimate of the prior mean and the precision based on g-prior as the estimate of the prior precision. Therefore, the prior precision parameter in empirical Bayes has significant role in producing different frequentist estimators of panel data models. For every value of the prior precision parameter, different frequentist estimators are acquired. Tending the value of prior precision parameter from zero to infinity yields infinite frequentist estimators. In this approach no pre-testing for model selection and estimation technique is required rather the value of the prior precision parameter itself will led to the appropriate model and the corresponding estimate.

The third and one of the major contributions of this thesis is the development of new empirical Bayes estimator. As mentioned, the prior precision parameter has an important role in producing panel data models and their corresponding estimators, but it has to be estimated, so in this thesis a modified version of the empirical Bayes formula has been developed which is independent of the prior precision parameter and hence is free of estimating the prior precision parameter.

Finally, a real-world data of five European countries, namely, Austria, France, Italy, Sweden, and Britain, on two variables the "Gross Domestic Product" and "Consumption" for the period 1970 to 2016 have been used for depicting the estimation procedure for the new empirical Bayes estimator. A balanced panel has been employed. Further, the data have been taken from the world statistics. The empirical Bayes estimates that have been computed from this data set have the tendency towards the random coefficients model estimates.

**Keywords:** Empirical Bayes, Frequentist Estimates, Prior Precision Parameter

**JEL Code:** C11

# Table of Contents

# List of Tables

# List of Abbreviations

| | |
|---|---|
| OLS | Ordinary Least Squares |
| MLE | Maximum Likelihood |
| MLE | Maximum Likelihood Estimator |
| CB | Classical Bayes |
| EB | Empirical Bayes |
| MOM | Method of Moments |
| GLS | Generalized Least Squares |
| GMM | Generalized Method of Moments |
| DV | Data Variance |
| $(DV)^i$ | Data Variance of the $i^{th}$ unit of the panel |
| $(DV_k)^i$ | Data Variance of the $k^{th}$ regressor in the $i^{th}$ unit of the panel |
| DP | Data Precision |
| $(DP)^i$ | Data Precision of the $i^{th}$ unit of the panel |
| $(DP_k)^i$ | Data Precision of the $k^{th}$ regressor in the $i^{th}$ unit of the panel |
| PV | Prior Variance |
| $(PV)^i$ | Prior Variance of the $i^{th}$ unit of the panel |
| $(PV_k)^i$ | Prior Variance of the $k^{th}$ regressor in $i^{th}$ unit of the panel |
| PP | Prior Precision |
| $(PP)^i$ | Prior Precision of the $i^{th}$ unit of the panel |
| $(PP_k)^i$ | Prior Precision of the $k^{th}$ regressor in the $i^{th}$ unit of the panel |
| RC | Random Coefficient |

# List of Synonyms

| | |
|---|---|
| Frequentist | Classical |
| Panel Data | Longitudinal Data |
| Subjects | Units |

# Definition of Technical Terms

Density: In probability theory, a probability density function (PDF), or density of a continuous random variable, is a function that describes the relative likelihood for this random variable to take on a given value.

Data Density:

Prior Density:

Posterior Density:

Conjugate Prior:

g-Prior

d-prior

Likelihood Function:

Panel Data:

# Chapter 1

# Introduction

In Statistics and Econometrics, the more valid the data, the better the result, or to put it another way, the better the estimates of the parameters. Sufficient large sample size facilitates the researchers in the determination of the average values of the observations and evade the errors in testing that could have been caused by an inadequate small sample size. For this reason, a researcher should examine as many data points as possible in addition to different aspects of heterogeneity in the data set in order to obtain valid and better results

In Econometrics, basically there are three main types of data: the cross-sectional data, the time series data and the panel data. The panel data, in fact, is the combination of cross-sectional data and timeseries data. Panel[1] data econometrics has evolved rapidly over the last few decades and the researchers often deal in econometrics using panel data in the analysis of relationship between variables. Hsiao (2007) says, the panel data is preferred and considered better in characteristics than either pure time series data or purely cross-sectional data, because it addresses both sources of variation in the data, i.e., the time dimension as well as the cross-sectional dimension. The key difference between time series and panel data is that time series focuses on a single individual at multiple time intervals while panel data (or longitudinal data) concentrates on numerous individuals at multiple time intervals.

There are a number of advantages of panel data:

- Panel data can be used to model both the common and individual behaviours of groups.
- Panel data contains more information, more variability, and more efficiency than pure time series data or cross-sectional data.
- Panel data usually contain more degrees of freedom and more sample variability than cross-sectional or timeseries data.

[1] Panel data is also called longitudinal data, there is a bit difference between these two, however, in the literature they are being used interchangeably or synonymously.

- Panel data can capture the complexities of human behaviour better than a single cross-section or time series data.
- Panel data can detect and measure statistical effects that pure time series or cross-sectional data cannot.
- Panel data can minimize estimation biases that may arise from aggregating groups into a single time series.
- Panel data make more accurate inference of model parameters.

There are a number of other advantages of panel data as well.

Panel data involve at least two dimensions, a cross-sectional dimension, and a time dimension. In panel data if both temporal[2] and spatial sources of variations are modelled meticulously, then better results are expected. More reliable inference regarding the coefficients of the model can be anticipated in panel data (e.g., Hsiao et al., 1995).

Like other econometric models, the panel data econometrics has also huge variety of models. Chang, Swamy, Hallahan and Tavlas (1998) prescribed the facts of the econometric model which are as follows: (i) in economic relationships, the true functional forms of the models are almost unknown; (ii) from every econometric model, at least one unrevealed explanatory variable is excluded; (iii) it is very baseless or futile to assume that the excluded explanatory variables are uncorrelated with the candidate explanatory variables of the model; (iv) economic data are subject to measurement errors in many cases, so that is why they are only the approximate values to the underlying exact values. According to them, if an econometric model is consistent with all of the above realities of building of economic model, then it would be causal and meaningful. In other words, the elucidations attached with the model parameters are reliable with these realisms.

Some most commonly used standard models related to panel data in the standard econometrics or statistics books are: the subject specific panel data linear regression model, the subjects common panel data linear regression model, the fixed effects panel data linear regression model, the random effects panel data linear regression model and the random coefficient panel data linear regression model, etc. (see. Wooldridge

[2] Time based.

(2010), Greene (2003) etc). These panel data models are from frequentist[3] theory, which is to be explained shortly.

Further, the most commonly used frequentist techniques for the estimation of the unknown regression coefficients or parameters of the aforementioned models, is the ordinary least squares, abbreviated as, (OLS). Once the model is specified, the method of ordinary least squares is applied to estimate the parameters of the model.

The specification of the model also needs some pretesting. If after the model specification test, the subject specific model is deemed. Then the frequentist techniques of estimation for the model parameters yield the vectors of coefficient estimates one specific to each subject or unit of the panel.

Similarly, if subsequently to the model specification test, the subjects common model is supposed. Then the frequentist techniques of estimation for the model parameters yield a vector of coefficient estimates which is common to all subject of the panel.

Further, if next to the model specification test, the fixed effects model is believed. Then the frequentist techniques of estimation for the model parameters yield the vectors of coefficient estimates one for each unit of the panel, in which all the corresponding estimates of the coefficients across the subjects are constant and common to all subject of the panel except the intercepts, the intercepts are fixed to each subject of the panel separately.

Additionally, if succeeding to the model specification test, the random effects model is thought. Then the frequentist techniques yield the vectors of coefficient estimates, one for each unit of the panel, in which all the corresponding estimates of the coefficients across the subjects are constant and common to all subject of the panel except the intercepts. Here, the intercept is the mean of all the intercepts, and an intercept for a particular subject is considered to be randomly scattered from this mean of all subjects of the panel, consequently the mean of the intercepts is taken as the intercept for all units.

Finally, if following to the model specification test, the random coefficients model is considered. Then the frequentist techniques of estimation for the model parameters yield a vector consisting of the mean of the

---

[3] The word frequentist refers to the classical school of thought. The frequentist School of thought rely only on the data information. This is also to note that, the frequentist school of thought is also called the classical school of thought.

coefficient estimates, which is common to all subjects of the panel, and the vector of coefficient estimate specific to each subject of the panel is considered to be randomly scattered from this common mean vector.

Apart from panel data models and their corresponding estimates, there are other model and estimation technique for panel data.

Basically, there are mainly two school of thoughts, namely, the classical or the frequentist school of thought and the Bayesian school of thought. The frequentist school of thought relay only upon the available information or data set and having nothing to do with any additional information (if available) obtained from other sources. Fundamentally, a frequentist method makes decisions and predictions on the underlying truths of the experiment using only data from the current experiment.

Contrary to the frequentist school of thought, the Bayesian school of thought takes into consideration the additional information, technically called the prior[4] information, along with the data information. Thus, the Bayesian school of thought is consistent with the theory of more valid information for the better estimates and analysis, as mentioned above.

The Bayesian school of thought combines the data information with the prior information and produces estimates, which are far better than the frequentist estimates in many cases and are evident in the literature (See. Zaman (1996) or Carrington and Zaman (1994)). Bayesian statistics take a more bottom-up approach to data analysis. This means that past or relevant knowledge of similar experiments is encoded into a statistical device known as a prior, and this prior is combined with current experiment data to make a conclusion on the test at hand. The biggest distinction is that Bayesian probability specifies that there is some prior probability while the frequentist does not.

The Bayesian approach:

- defines the prior distribution that incorporates the subjective beliefs about a parameter. The prior can be uninformative or informative.
- gather data.
- update the prior distribution with the data using Bayes' theorem to obtain a posterior distribution.
- analyse the posterior distribution and summarize it, i.e., mean, median, standard deviation, quantiles, etc.

[4] In Bayesian school of thought, the information other than the data information are technically called prior information.

A fundamental aspect of Bayesian inference is updating the beliefs in light of new evidence. Essentially, one starts out with a prior belief and then updates it in light of new evidence. An important aspect of this prior belief is the degree of confidence in it.

The controversy between "frequentism" and "Bayesianism" is at least 200 years old (although these terms are much younger) and still unresolved. A pragmatic view is that to accept both viewpoints depending on the context. Samaniego and Reneau (2012) "A Note on the Comparison of the Bayesian and Frequentist Approaches to Estimation", presented a landmark study on the comparison of Bayesian and frequentist point estimators. Their findings indicate that Bayesian point estimators work well in more situations than were previously suspected. In particular, their comparison reveals how a Bayesian point estimator can improve upon a frequentist point estimator even in situations where sharp prior knowledge is not necessarily available.

## 1.1. Statement of the Problem

As stated above, in the frequentist set up there are many panel data linear regression models and accordingly many estimation techniques. Therefore, for a researcher there exists choices or uncertainties among the model selection as well as in the estimation techniques. Specifically, first to model the data with the most appropriate model and then estimate its parameters with a reasonable estimation technique. Thus, choosing the panel data model for any given scenario can be difficult and time consuming. Sometimes for a practitioner or a researcher, it becomes unfeasible to make an appropriate choice among the combinations of the available panel data models and the corresponding estimation techniques. Different panel data model and the estimation technique have their own specific requirements and assumptions of application. Some of the assumptions among various models and/or the estimation techniques are conflicting to each other. In such case of conflicting assumptions either among the panel data models or the estimation techniques, the selection choice becomes even more tedious. For example, the simple OLS theoretically assumes that all the cross-sections in the panel have different sets of parameters, i.e., one set of parameters for each cross section of the panel. Similarly, the pooled OLS assumes that all the cross-sections in the panel have identical set of parameters, i.e., one set of parameters for all cross sections of the panel. Further, the fixed effects assumes that all the cross-sections in the panel have different intercept parameters while the slope parameters are identical for all cross sections of the panel. Furthermore, the random effects assumes that all the cross-sections in the panel have identical intercept parameters but the variability among them is due to random causes and all the slope parameters are identical for all cross-sections of the panel. Finally, the random coefficient assumes that all the cross-sections in the panel have identical set of parameters but the variability among them is due to random causes. "Most statistical analyses are largely concerned with the appropriateness of the regression model being used. However, for the data at hand, an important and often neglected problem concerns the issue of misclassification error, which leads to bias and loss of efficiency in estimation. It is possible to correct this bias if the probabilities of the misspecification can be unbiasedly estimated from, for instance, the validation data." (See. Samuel Mwalili (2006)).

This should also be kept in mind that the complexity of model selection and estimation technique do not exist only in panel data. Besides panel data models, the existence of multiple models either in time series or cross-sectional data is also a common phenomenon. In empirical analysis of either time series or cross-sectional data, the existence of multiple models also leads to difficulty in the choice of selection of an appropriate model. That is to say, which of the models would be an appropriate model and to be used for the problem at hand. One way to handle this sort of problem is to nest all the candidate models into a general model, then the rival candidate model could be obtained as a special case of this general model by imposing certain restrictions upon the parameters of this general model to get the most convenient and parsimonious model. The researcher can then test the validity of the assumptions by certain hypothesis testing methods and it could be insured that the imposed restriction does not result in loss of predicting power (see, Hausman (1978)).

After a parsimonious model, the next step, either in time series or cross-sectional analysis, is the estimation of model parameters. Different techniques of estimation are used in different scenarios. Unfortunately, no crystal-clear law for any technique to be applied is defined. The multiplicity of available estimation techniques also causes problems in the estimation process. Hence the time series and cross-sectional data have also its own related issues.

The current frequentist literature of panel data mainly captures countable major sources of heterogeneity among the units of the panel and the corresponding models that will be discussed shortly. That is why the current researches have also been restricted only to these countable models. In fact, there are many sources of variation or heterogeneity in the resulted models. The frequentist literature of panel data linear regression models summarised from Hsiao (1986) or Wooldridge (1995) assumes one of the possibilities for the parameters vectors of models, that are given below;

i) totally different parameters vector for each unit,

ii) a unique parameters vector for all the units,

iii) a common parameters vector, but random fluctuations have also been considered for each unit as well,

iv) different intercept terms while all the slopes parameters of all the units are identical, but the difference in the intercepts is due to subject specific heterogeneity,

v) different intercept terms but all the slopes parameters are identical for all the units, and the difference in the intercepts is due to random fluctuations not due to subject specific heterogeneity.

In these five possibilities expressed above, the first possibility refers to the subject specific model; the second reveals the subjects common coefficients model; the third directs to the random coefficients model; the fourth describes the fixed effects model and the fifth shows the random Effects model. For further reference, (see, Arellano (2003), Baltagi (2001), Hsiao (2003), Mátyás and Sevestre (2008), and Nerlove (2002), etc).

However, it is injustice unfair with the widespread literature of panel data to deal only with these limited possibilities of heterogeneity presented above, and this is the gap in the literature of panel data models. There is a room for the discovery of some more new models in the existing frequentist literature of panel data models, other than those discussed above. Therefore, now this is the need of the hour to explore some new models in the frequentist literature of panel data models, and to fill up this gap.

Secondly, what important to note is that the frequentist considers all of the above panel data linear regression models are of different origin. As for as we think, all the panel data models are not of different origins. A uniform framework for all of the above models as well as for the estimation techniques is possible and exists. The uniform framework would be the best solution for the model and the corresponding estimation technique selection problems discussed above.

Similar to the case of either time series or cross-sectional data, in case of panel data models, it also seems possible to nest all the existing panel data linear regression models into a single general framework and then obtain the most suitable and parsimonious model from it. We argue that all of these possibilities (all of the above panel data models) can be derived with the help of a Bayesian linear regression model by using specialized Bayes estimation techniques.

In linear regression models the key entities for the classification of the models are the regression coefficients or the regression parameters. The whole model almost depends upon the characteristics of the regression parameters i.e., in panel data case, if the regression parameters are specific to each unit of the panel, then the corresponding regression model is termed as the unit specific panel data linear regression model. On the other hand, if the regression parameters are common for all the linear regression models corresponding to different units of the panel, then a single model with these common coefficients is applied to all the units of the panel and is termed as the units common linear regression model. Further, if the regression parameters of various linear regression models are different from each other, and if these differences can be acknowledged as random, then a single model like in the case of units common model with random fluctuations is sufficient and this linear regression model is called the random coefficient linear regression model. There are some other possible structures of the regression parameters too, which are discussed in this thesis in the coming sections. At the moment the above models are sufficient to clear the picture and the scenario.

The problem with the regression analysis is that the regression parameters do not exist in reality, they do not exist numerically. They only lie in the mind of the researchers. Therefore, to identify the regression parameters the corresponding parameter estimates are used for this purpose. If the parameter estimates match to any of the above hypothetical structure of the regression parameters, then the whole model is narrated in that way, and the model is considered of that type and is being named. This methodology, for the classification of different panel data models will be followed here.

The derivation of all the frequentist estimators of panel data models from a single empirical Bayesian estimator witness that the corresponding frequentist panel data models are also the derived versions of the Bayesian linear regression model. As models are classified on the basis of the vectors of regression coefficients and estimators produce the estimates of the vectors of coefficients, Hausman (1978). This will also confirm that in fact all the frequentist models are not of different origins, rather they have been emerged from a single Bayesian linear regression model with some restrictions imposed upon its prior parameters. Therefore, this thesis is intended to derive all the limiting cases of the Bayesian regression coefficients estimator and the resultant frequentist regression coefficients estimator of panel data models. The derivation of the frequentist regression coefficients estimators from the Bayesian estimator will be presented first analytically and then numerically and will be of main concern. This will confirm the claim that in fact all the frequentist panel data linear regression models can be derived from a single Bayesian linear regression model. This will also be shown that all the frequentist estimators have a common origin rather than a separate one for each one of them. In other words, all the frequentist panel data linear regression models are the specialized form of the Bayesian linear regression model and the frequentist estimators are the limiting cases of the Bayesian estimator as prior precision goes from zero to infinity all the frequentist panel data estimator will be produced.

## 1.2. Aims of the Study

This thesis is aimed to nest all of the frequentist panel data linear regression models in a general Bayesian linear regression model. This issue can be addressed indirectly, by deriving all the frequentist panel data linear regression models estimators from a single Bayesian linear regression model estimator. Further, the frequentist panel data linear regression models are not of different origins rather all of the models have a unique common origin. It is also aimed to derive all the frequentist[5] panel data linear regression model estimators from Bayesian linear regression model estimator, first analytically, and then numerically.

## 1.3. Objectives of the Study

The main objectives of this study are;

i) to encompass all the standard frequentist panel data linear regression models into a single general Bayesian linear regression model.

[5] Frequentist and classical are the two names of the same school of thoughts, and they shall be used interchangeably throughout. This school of thought is considering only data information and hence relying only on it.

ii) to derive and confirm analytically as well as numerically that all the frequentist panel data linear regression models are the specialized cases of the Bayesian linear regression models.

iii) to discover some new linear regression models and estimators in the frequentist literature of panel data models.

iv) to develop a modified empirical Bayes estimator for all standard panel data models.

## 1.4. Motivation for the Study

As there are so many frequentist panel data linear regression models in the literature of panel data, the application of which depend upon different conditions and assumptions. Due to the complexity in the assumptions among different models, it becomes more difficult for a researcher to choose any particular model for a particular problem at hand. In the frequentist literature of panel data models some ad-hoc methods are used for model selection. The efficacy of these ad-hoc methods for model selection have also been severely questioned in the literature, (See. Clark and Linzer(2012 )), and thus the risk of misspecification of models always exists. It is therefore very much essential to have a uniform and general specification of models, which could be appropriate for many research problems, in order to control the misspecification risk and also to avoid pretesting for model selection. For this purpose, the Bayesian linear regression model seems to be the most convenient choice of models, which has the potential to encompass all the existing frequentist panel data linear regression models and also to avoid the pretesting procedure for model selection as such.

The Bayesian linear regression model, (See. Zaman (1996)), has also the ability of reproducing the typical or conventional frequentist panel data linear regression model after imposing some restrictions upon its prior precision parameters. This model has also shown very fruitful results in the empirical analyses of data in the past. This intuition leads us to explore the possibility of a common framework for specification of panel data models with the help of Bayesian linear regression model. As in frequentist set up, the model specification test is prerequisite, and the efficiency of such tests have been criticised severely in the literature (see. Hausman (1978)). Therefore, it was aimed to get rid of such ad-hoc based pretesting procedures for model selection and to minimize the misspecification risk.

The aforementioned were the main points which have compelled to work upon this direction with the help of this dissertation.

## 1.5. The Contributions Made by this Study

This work has been intended to contribute in the existing literature of panel data Econometrics and Statistics in the following ways.

1. With the help of this thesis, the major contribution in the literature of panel data is that all the frequentist panel data linear regression models and/or estimators have been derived from a single Bayesian linear regression model and/or estimator. In other words, the Bayesian linear regression model has encompassed all the existing classical[6] panel data linear regression models or conversely the frequentist panel data linear regression models have been derived from the Bayesian linear regression model.
2. Some new panel data linear regression models which are ignored (or were seem impossible to exist) in the current frequentist literature of panel data, are discovered. As with the help of this research new models are added to the literature and this is also a significant contribution in the literature of panel data models.
3. The techniques of computing the prior precision parameters from data give rid the researchers from pretesting for model specification or selection in panel data.
4. The work in this "thesis" lets the data free to tell by itself; that what is the convenient model for the data and no specification of a model in advance are mandatory. Ad-hoc bases of panel data model selection are avoided and as a result the problem of misspecification of models (in terms of candidateship) have been minimized.

## 1.6. Significance of the Study

This study mainly focuses on the estimation techniques of panel data linear regression models. Panel data models are actually the extracts of a single Bayesian linear regression model. In this study not only the existing panel data linear regression models are derived from the Bayesian linear regression model, but also some new models got birth in the literature of panel data. With the help of this study pretesting, like Hausman Specification Test etc., for model selection are no more needed. For any sort of panel data, if the Bayesian

[6] The term "Classical" over here means the School of thought as opposite to the Bayesian.

linear regression model has been used, the data itself will design a model which will be the most appropriate one. With the help of this study the researchers will be able:

a. to get rid of pretesting procedures for model selection.
b. to let the data free to speak what structure of parameter vector is needed to model the data.
c. as more models would be discovered so the researches shall not be restricted only to these countable existing frequentist panel data models.

# Chapter 2

## Literature Review

The statistical problems that can be confronted in either pure time series data or pure cross-sectional data can be controlled to a great extent by using panel data. In panel data analysis different sources of variations can be studied, like temporal change and spatial change. In other words, the time-varying analysis and unit-varying analysis are being done in panel data analysis.

The conventional methodology of econometrics considers the parameters of the linear regression model as invariant or fixed over time and over space. This implies that a single vector of parameters will suffice the need regarding the relationship between the explained and explanatory variables over the time and/or across the individual units of the panel. That is why the traditional linear regression model explicitly adopts that the sample data generating process of economic structure, remains fixed across the units of the panel and over different time periods. Therefore, the traditionalists believe that the pooled or the fixed parameters linear regression model describes the true functional structure between the regressand and the regressors.

In reality, when dealing with the time series data, the response coefficient might change until and unless the time variant factors are being modelled properly. Likewise, when dealing with the cross-sectional data, the response to an explanatory variable will vary for different units of the panel. As long as the data generating process is uncontrolled and unobservable experiments, the assumption of the fixed or constant parameters is very poor and highly restrictive.

As opposed to the fixed parameter models, there are other types of models called the varying coefficients models. These models assume the parameters to be variant over time or across the units of the panel. By allowing the parametric variation from one observation to another over time or across the units of the panel and challenge the assumption of the constant parameters in the conventional methodology of econometrics.

There are so many valid reasons for the justification of variations in the econometric model parameters. The key cause of parameter dissimilarity which is addressed in the literature can be categorised as;

specification errors primarily caused by the characteristics of the missing variables and/or misspecification of the real functional relationship (e.g., nonlinearities in true relationships); measurement error in the variables and cumulation bias.

As the true model or the data generating process is unknown, and does not exist numerically in reality, it is only the imagination. Therefore, some writers have stated some facts about the econometric models which have been revealed in the introductory chapter of this thesis. The study of Chang, Swammy, Hallan and Tavlas (1998) in this regard, discourses the question that how to verify whether a particular econometric model agrees with an underlying model or stochastic law or not, as defined by Pratt and Sclaifer (1988). In the real world problems, the fixed or constant coefficients models are very limited representations and fall short of being consistent with the econometric model building realities, therefore, they cannot be considered as causal. In one of the very important articles by Pratt and Schlaifer (1984), they have shown that the OLS estimates will almost and always be inconsistent by the nature of the stochastic laws. Also, in many basic theoretical econometric books, after analysing many of the estimation techniques for the linear regression models it has been concluded that the OLS estimates are always in consistent and inefficient and even biased by the nature due to the said law.

Considering the simple representation of the simple linear regression model, $y = x\beta + u$, the writers emphasise the statement that the model parameters can be consistently estimated with the help of OLS, if and only if the regressor of the model is uncorrelated with disturbance term of the model, that is, if $x$ is uncorrelated with $u$. Now from practical point of view or from real world issues it is very much difficult for anyone to judge whether the regressor $x$ and the disturbance term $u$ are correlated or not. The importance of the correctly interpreting the error process $u$ was interpreted by them. The interpretation of $u$ made by the econometricians takes the following form: $u$ is the effect of those omitted variables, say $w$, which along with the included variables $x$, suffice to determine the average value of the dependent variable $y$ but are not explicitly included in the model. If these excluded variables or factors were identified, the scientist would able to decide whether the joint effect of these excluded variables is likely to be correlated with $x$ or not; the authors emphasize that the econometricians never suggest identifying all of the excluded variables from the model.

Some authors, like Malinvaud, states that it is even impossible to identify all of the excluded variables. Although, if all the excluded variables, say, have been identified the inclusion of them in the model also leads to some incurable problem too, like the loss of degree of freedom, the high R-squared etc. Therefore, the assumption that $x$ being independent of the error term is either untrue or worthless and if this assumption is made so it is purely on adhoc bases.

Swamy and Tavlas (1995) gives explanations for using the Random Coefficients models, as opposed to the fixed or constant parameters model discussed above, in the context of a "class of functional forms" approach to model evaluation. The logic behind the functional form is that many estimation techniques in econometrics are concerned with the data generating process. They make particular assumptions about the functional form of the data generating process. Although the exact or true functional form of the data generating process is not known. Sometimes the economic theory explains the candidate variables that are very probable to be involved in the stochastic economic law, but it does not have much to say about the functional forms between the dependent and independent variables. The assumption about the functional form in the traditional econometric methods is very much essential, adding simply the disturbance term to the mathematical form, straight away, is not meaningful. The specific functional form plays a pivotal role in the results of the estimation and the estimates heavily depend upon the functional form. Further the effects of the omitted variables on the estimates of the coefficients of the included variables cannot be known in advance. It is very worthless to assume that every included regressors of the model are uncorrelated with every excluded variable from the model that affects the dependent variable. These issues have been addressed by Swamy and Tavlas (1995) through an article, namely "class of functional forms". This approach basically illustrates that one can begin with a wide class of functional form and check whether the answer to a specific question being addressed is fundamentally the same for any particular functional form in the class. If the functional forms in the class are found to give markedly answer, then the modification of the class will be desirable.in making the modification of the class, careful interactions with the data are necessary. They pointed out that the RC models are very significant as they signify the key intermediary steps in the problem of deriving major classes of functional forms.

A random coefficient model encompasses a variety of models including the fixed coefficients model. The fixed coefficients models are assumed as the special cases of the RC model. The specification errors, to a significant extent, are less serious in case of following the RC as compared to any of its special cases. Klein (1989) reasons the random parameters and systematic changes in the parameters may be indication of nonlinearities that have not been modelled effectively in a model's specification. According to Granger (1993) the time-varying parameters model provide satisfactory approximation to nonlinear associations.

The RC assumes each coefficient of an econometric equation of a model as a stochastic or random. Thus, each coefficient of the model comprises of two components; namely the deterministic and the stochastic. The first component is deterministic, that changes directly in line with the regressor, and the second part is stochastic, that may be simply a white noise process or may follow any process that may even be complex. In case of specification error such as the omitted variables, the incorrect functional form and measurement error in variables, it very poor assumption to assume that the simple error term added to the intercept will be sufficient to the effects of these natural sources of errors. If the parameters of the model are variable, then it is quite clear that the OLS estimators are unbiased hut seriously inefficient. Cooley and Prescott (1973) and Rosenberg (1973b) have shown a remarkable improvement in efficiency in simulation by using the RC specification. Additionally, when the parameters of the model are stochastic then it has been shown that the OLS sampling theory severely underestimate the parameter estimation error variance. Thus, the random coefficient specification is removing the downward bias in the estimated variance of error. Rosenberg (1973b) described in a simulation study that the OLS error variance rises to five times as compared to the efficient variance and the OLS sampling underestimates OLS error variance by a factor of twenty or more. The area of the time-varying parameter regression models have received significant attention in the literature.

The literature on the wide class of time-varying parameter models can be classified into the three main categories, which are described as under:

(i) Systematic but non-stochastic variation models (See. e.g. Quandt (1960), Belsley (1973));

(ii) Random Coefficient Models (See e.g. Hildreth and Houck (1968), Froehlich (1973), Hsiao (1974), Swamy and Tinsley (1980), Başçı, & Zaman (1998));

(iii) Random but not necessarily stationary coefficient models (See. Kalman filter models, Cooley and Prescott (1976)).

Regarding the first class of the models i.e., systematic but non-stochastic variation models, the vector of the coefficients can be expressed as a deterministic function of some observables, which might possibly be nonlinear and may contain the regressors themselves with the variation of the systematic parameter the theory of OLS is applicable.

Next, the second as well as the third class of the models emerges when the variation of the parameter contains an element which is a realisation of any stochastic process in addition to a deterministic portion that may be a function of observable. Further, for the stochastic coefficient models some more information must be put on the structure of how the coefficients changes across the data points, if possible, the estimation procedures are to be developed. There are numerous models and estimation techniques in the literature towards this end.

The third and final class is the time-varying regression models, these are also known as the sequential or Markov parameter models. In this type of models, the stochastic parameter process includes random drifts. For the estimation purpose of these models the Kalman filter technique has been vastly used, after the introduction of it with Kalman (1960) and Kalman and Bucy (1961). For deep understanding about the application of the Kalman filter technique in case of time-varying regression models, see Raj and Ullah (1982), also Chow (1984) and Nicholls and Pagan (1985).

For a variety of other estimators developed for the regression of nonstationary time varying parameter, see, Rosenberg (1973), Cooly and Prescot (1976). Also, for its Bayesian counterpart, see, Sarris (1973) and Liu and Hanssens (1981).

The study in this thesis does not concentrate on the time-varying regression models, this mainly concentrates on the unit-varying regression coefficients parameter in case of panel data, which are very much similar to time-varying regression above.

In "stochastic frontier analysis" the use of panel data was introduced by Pitt and Lee (1981) and Schmidt and Sickles (1984). The maximum likelihood analysis presented by Pitt and Lee (1981), the assumption of the constant coefficients over time allows the researchers to use a "within" estimator in case of fixed effects model, this is stated by Schmidt and Sickles (1984). Frequentist analysis of this model can be seen in Meeusen and Van den Broeck (1977) with an Exponential inefficiency distribution, Aigner, Lovell and Schmidt (1977) (half-Normal), Stevenson (1980) (truncated Normal) and Greene (1990) (Gamma). The application of this model in case of cross-sectional data abound. They are usually conducted in the setup of one of the references cited earlier. Further, the Bayesian analysis of the above model was introduced for cross-sectional data in Van den Broeck, Koop, Osiewalski, and Steel (1994) under different inefficiency distributions. The literature incorporating a Bayesian approach to panel data models with applications in stochastic frontier analysis has been growing in the last two decades. The approach was first suggested by Van den Broeck (1994), which considers the Bayesian method under the composed error model. Koop et al. (1997) has established the Bayesian setup where the random and fixed effects models are defined; they also applied Gibbs Sampling to analyse their model.

Liu, J., et, al (2013) consider two longitudinal data models with unobservable heterogeneous time-varying effects. The first one with unit effects treated as a random function of time, the second one with the common factors which are unknown in number and their effects are unit specific. This paper has two very important features and can be considered as the generalization of the conventional panel data models. Firstly, the unit specific effects that are considered to be heterogeneous across subjects as well as along time varying are treated non parametrically, following the essence of the model from Bai (2009) and Kneip et al. (2012), and Ahn et al. (2013). For an extended discussion of these and other models used in panel work in the productivity field (See. Sickles, Hsiao, and Shang (2013)).

Bayesian numerical integration methods are described in Osiewalski and Steel (1998) and used to fully perform the Bayesian analysis of the stochastic frontier model using both cross-sectional and panel data. However, the subject effects are assumed to be time-invariant by Liu, et, al (2013), which is inappropriate in many settings; for example, in the stochastic frontier analysis, the technical inefficiency levels typically adjust over time. In order, to address the temporal behaviour of individual technical efficiency effects, Tsionas (2006) considers a dynamic stochastic frontier model using Bayesian inference, where the inefficiency levels are assumed to evolve log-linearly.

Bayesian methods have many advantages. First, following the Bayesian perspective of random coefficient models, (Swamy (1970); Swamy and Tavlas (1995) the panel data model will not subjectively assume a common functional form for all the individuals as the subjective processes may vary among individuals and fixed parametric values of the parameters that describe this functional relationship may not be well-defined. Moreover, a Bayesian approach may circumvent the theoretically complex as well as the computationally intense nature of nonparametric or semiparametric regression techniques (Yatchew, 1998) and the need to rely on asymptotic theory for inference (Koop and Poirier, 2004).

## 2.1. Deficiencies in Frequentist Panel Data Models

There are several deficiencies related to panel data, some of which have been outlined below. From frequentist perspective, model selection is one of the major problems in panel data models. When conducting the research in the area of social sciences, it is very common to deal with the data that are bunched or grouped into higher-level units. The most important challenge while modelling such data appears when the regressand shows group level variability beyond what can be described by the regressors alone. Dealing with such cases, simply the fit of standard linear regression model or the generalized linear regression model without properly accounting for the grouped nature of the observations can lead to very poor fit of the models and also leads to misleading estimates of both the effect of the regressand and precision of the estimates. (Beck and Katz 1995; Greene 2012). The two well-known approaches for the remedy of this problem are the fixed effects and the random effects models approaches. There is a lot of written work in the literature over the theoretical properties of these two approaches. (for example, Kreft and DeLeeuw (1998); Robinson (1998)), application of these methods in applied work are often very confusing— occasionally contradictory (Gelman and Hill (2006)). Inadequate guidance can be understood by a researcher to decide between the fixed effects and random effects model to be used to model the data at hand. Researchers often use the Hausman test to resolve this problem. This is intended to apply to tell the researcher how significantly the parameter estimates differ between the two approaches. As demonstrated by Clark and Linzer (2012), after an extensive study of analysing Hausman technique from various perspective the finally reached to the conclusion that Hausman test is neither necessary nor sufficient condition for deciding between fixed effects model and random effects models. According to them what matters is the;

i) size of the data set (both number of units and number of observations)

ii) level of correlation between the covariates,

iii) subject effects,

iv) extent of within unit variation in the regressors relative to the regressand.

To summarise the discussion of this chapter, we prefer to use the panel data as compared to either pure time series data or pure cross-sectional data. Because panel data has more advantages over either of the two types of data. Secondly, we prefer the varying parameter models as opposed to the fixed parameter models because the former are more flexible and can bitterly model the data. For example, the fixed parameter model in case of panel data, assuming no temporal effects, assume the parameter are identical for all the units of the panel, which in reality are very poor assumptions. There must exist some unit effects or heterogeneity. Therefore, the fixed parameter cannot be highly recommended for the analysis of panel data.

On the other hand, assuming a strong heterogeneity among the units of the panel and employing different parameters model to each unit of the panel to bitterly accounts for the subject specific heterogeneity among the units of the panel. Then the common characteristic among the units of the panel is lost and hence the purpose for which the panel data has been preferred over either time series data or corss-sectional data is no longer exist. Even though if using the panel data and account for the heterogeneity among the units of the panel, there is not credible techniques in the classical literature that could help us in deciding whether the heterogeneity among the units of the panel is random or fixed. The commonly used tool for deciding this is Hausman test and as it has already been mentioned that this tool is not very trustworthy. The only option that is left with us the random coefficient setup, as it has great potential and can encompass variety of models. If this has been chosen, then the problem has been controlled to a certain extent. There lies another issue in the estimation of the model parameters, that is, either to follow the frequentist approach, which solely depends upon the sample data available to the researcher or the Bayesian approach, which along with sample data rely on additional information as well.

The Bayesian approach has shown very fruitful results in the literature as it utilises more information as compared to the frequentist approach. Therefore, the Bayesian techniques have been preferred, and this is the reason that panel data, random coefficient model and Bayesian techniques of estimation are the main ideas of this thesis.

# Chapter 3

# Methodological Framework

In this chapter we describe the methodological framework that is carried out in the chapters to be come in this thesis.

In chapter 4, the theoretical development of the $K$ regressors panel data linear regression model for the $i^{th}$ unit of the panel has been described, which includes, the variables in the model, the dimensions of the variables, the assumptions of the variables, the associated parameters of the variables in the model, the interpretations of the parameters of the model, the dimensions and assumptions of the parameters, the variances and covariances of the parameters, the precisions of the parameters, etc. will be described. Thereafter, the Bayesian structure of the parameters of the model, its assumptions, distributions of the parameters and the errors etc. will be presented. Next, various densities which are needed for the Bayesain analysis/computation, namely the data density, the prior density, the posterior density, the marginal density, their parameters, the corresponding structure and related assumptions will be expressed. After that, in the next section of chapter four, the Bayes estimators, the classical Bayes estimator, its theory, structure, assumption, drawbacks, etc. are discussed. Then the empirical Bayes estimate, its theory, its structure, assumption, superiority over classical Bayes estimator, the advantages, and the estimates of the prior parameters, etc. are discussed. The estimator for the prior mean used here is derived from the marginal density of $\hat{\beta}^i$. This estimator is the precision weighted arithmetic mean of the ordinary least squares coefficients estimates $\beta^i$. The estimator of the prior variance is, the Zellner's (1971) g-prior, which is proportional to the data precision. Lastly, the empirical Bayes estimate for a single $k^{th}$ orthogonal regressor of the $i^{th}$ unit of the panel data linear regression model is developed which is the main objective of this chapter. This completes chapter 4.

In chapter 5, the analytical derivations of the frequentist estimators from the empirical Bayes estimate will be derived. The frequentist estimates will be derived from the empirical Bayes estimate by imposing certain restrictions upon the prior precision parameters denoted by "$\rho_k$". This is shown easily that by imposing different restrictions on $\rho_k$ yield different frequentist estimators of the standard panel data linear regression models. Moving the prior precision parameter $\rho_k$ in different directions "i.e., when $\rho_k \to 0$, when $\rho_k \to \infty$ and finally when $0 < \rho_k < \infty$," yield almost all of the frequentist standard panel data linear regression models and the corresponding estimators for these models.

In the frequentist framework there are very limited panel data linear regression models as compared to the need of the researchers. The subject specific coefficients model, which assumes the coefficients to be constant with-in the subject but variable among different subjects of the panel. In other words, each subject of the panel will have its own distinct coefficients vector. On the other hand, the subjects common coefficients model assumes a single coefficients vector for all the units of the panel. It further assumes homogeneity of the coefficient vector with-in and between the units. Further, the fixed effects model assumes the intercept to be unit specific and the slope coefficients constant across all the units of the panel. Similarly, the random effects model assumes the intercept to be randomly fluctuated across the units of the panel and the slope coefficients are constant across all the units of the panel like in the case of fixed effects model. Finally, the random coefficients model which assumes some common characteristics and some random characteristics for all the regressors including the intercept as well as the slope coefficients.

There are number of questions regarding the frequentist standard panel data models. The first question that comes in mind is that in most of the models described above i.e., subjects common coefficients model, fixed effects coefficients model and the random effects coefficients, the slope coefficients have been restricted to be common and the intercept may or may not be variable. This restriction seems very impractical and very unlikely that all the regressors across all the units of the panel are constant. Further, in neither of the above cases it is assumed that the intercept may be common or constant across the units and some or all of the slopes coefficients may vary. This situation is very likely to occur in practice where the intercept may be common for all of the units but the slope coefficients may vary. This possibility has been completely ignored in their literature of panel data model.

The frequentists have categorized the regression coefficients in two major categories, in the first category, they put the intercepts coefficients and in the second category, they keep the rest of the regression coefficients. In many cases, they assume the second category as constant for all the units of the panel especially in the three cases mentioned above. Though, there are many regression coefficients in the second category as compared to the first category due to multiple regressors and the frequentist still assume them constant across all the units of the panel.

The next question is that the frequentist panel data models take into consideration the data information only, they do not consider any additional information to which Bayesian set up referred as the prior information. The problem in taking only the data information is that if for instance the data collected for a specific time period (or time periods) or for a particular unit (or units) due to any reason is very odd, extreme or outlying then the estimates from the frequentist techniques with the help of this data will be very irrelevant, meaningless and useless, because of outlier. The frequentist does not use other sources of data that could help and show that this special or particular case is exceptional and out of pattern. On the other hand, in the same assumed situation the empirical Bayes outperforms. The empirical Bayes estimate helps a lot in this regard, as it uses additional information. The empirical Bayes estimate has also the potential to produce all the frequentist estimates of panel data linear regression models. Therefore, the frequentist estimates of all the standard panel data linear regression models will be derived here from the empirical Bayes estimate and thus corresponding to each vector of the coefficient estimates the panel data linear regression model will be presented. The Bayesian linear regression model has the capability to produce a number of new models as well, which have not been discovered so for in the literature.

## 3.1. The Details of the Data Used in the Example

In order to show numerically that all the frequentist estimates of the panel data linear regression models can be attained from the empirical Bayes estimate of the panel data linear regression model. We take an example and the data from the book "Basic Econometrics" by Gujarati, Fourth Edition. This example is based on the data of Table 16.1 in it. The data in Table 16.1, is taken from one of the most popular studies on investment theory conducted by Y. Grunfeld. The rationale behind selecting the above example is to show that the estimates produced by the Bayesian technique, under certain restrictions, exactly match to the estimates produced by the frequentist approach in the book. The description of the study of Y. Grunfeld is as under. Grunfeld was concerned in discovering out that how real gross investment ($Y$) depends on the real value of the firm ($X_2$) and real capital stock ($X_3$).

It is to mention here that the original study covered number of companies. For demonstrative determinations we have picked data over four companies, the General Electric (GE), the General Motor (GM), U.S. Steel (US), and Westinghouse. For each company the data is available for the afore mentioned three variables from the period 1935-1954. Hence, there are four units in the panel and 20 time periods, so it results us 80 observations in all. The relationship between $Y$ and $X_2$ and $X_3$ is expected to be positive, a priori. Now, this implies that we can either run four time series regression one corresponding to each of the companies or can run 20 cross-sectional regression one corresponding to each time period. If running the latter case, the 20 cross-sectional regression one corresponding to each time period, we should care about the degree of freedom problem. Moreover, as mentioned earlier repeatedly, in the conventional econometrics, each of the frequentist estimates is considered to be of different origin. But now here the empirical Bayes estimates of the panel data linear regression model will nullify the theory of the separate origin of each frequentist estimate. This derivation will certify that all the frequentist estimates of panel data regression models are of the same origin and can be derived from that unique or common origin. The frequentist estimates to be derived from the empirical Bayes estimate includes the subject specific coefficients estimates, the subjects common coefficients estimates, the random coefficients estimates, the fixed effects coefficients estimates and the random effects coefficients estimates. For this purpose, for each of the above frequentist estimates, the theoretical description of derivation, the particular prior precision parameters with the help of which those particular coefficients estimates are being derived, then the corresponding derived frequentist coefficients estimates are presented in the next chapter.

# Chapter 4

# Deriving Empirical Bayes Estimate for Panel Data Linear Regression Model

This is the main chapter of this dissertation which consists of the empirical findings. In this chapter, the analytical derivation of the empirical Bayes estimate for the coefficient of the $k^{th}$ orthogonal regressor in the $i^{th}$ unit of the panel data linear regression model has been carried out. The empirical Bayes estimate that will be developed here is for a $K$ orthogonal regressors linear regression panel data model. But each regression coefficient in the model will be estimated individually and independently. This is a single or individual regression coefficient estimation approach. In other words, a single regression coefficient out of all $K$ regression coefficients will be estimated at a time one by one. In this chapter, we first develop theoretically the basic concepts regarding the $K$ orthogonal regressors panel data linear regression model for the $i^{th}$ unit of the panel and then the corresponding Bayes estimate. But before this, as for the analytical workout of the Bayesian linear regression model and the corresponding empirical Bayes regression estimate, we need prerequisite the frequentist panel data linear regression model, their estimates and some other quantities. As the Bayesian estimates utilizes these frequentist estimates for computation etc.

Therefore, below we discuss, the frequentist $K$ orthogonal regressors linear regression model for the "$i^{th}$" unit of the panel.

## 4.1. The General Panel Data Linear Regression Model

The general $K$ *orthogonal* regressors panel data linear regression model, for the $i^{th}$ unit of the panel, in the frequentist set up may be considered as under,

$$Y^i = X^i\beta^i + \varepsilon^i \qquad \text{for} \qquad i = 1, 2, \dots, N. \tag{4.1a}$$

where,

$$Y^i = \left(Y_1^i,\ Y_2^i,\ \dots, Y_t^i, \dots\ Y_T^i\right)',\ \beta^i = \left(\beta_1^i,\ \beta_2^i,\ \dots, \beta_k^i, \dots, \beta_K^i\right)',\ \varepsilon^i = \left(\varepsilon_1^i,\ \varepsilon_2^i,\ \dots, \varepsilon_t^i, \dots, \varepsilon_T^i\right)' \tag{4.1b}$$

and

$$X^i = \begin{bmatrix} X_{11}^i & X_{12}^i & \cdots & X_{1k}^i & \cdots & X_{1K}^i \\ X_{21}^i & X_{22}^i & \cdots & X_{2k}^i & \cdots & X_{2K}^i \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{t1}^i & X_{t2}^i & \cdots & X_{tk}^i & \cdots & X_{tK}^i \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{T1}^i & X_{T2}^i & \cdots & X_{Tk}^i & \cdots & X_{TK}^i \end{bmatrix} \tag{4.1c}$$

Note: We assume all the regressors in $X$ to be orthogonal.

Further, $Y^i$ is the dependent variable of order $T \times 1$, $X^i$ is the matrix of the orthogonal regressors of order $T \times K$, $\beta^i$ is the corresponding $K \times 1$, coefficients vector for the matrix of the independent variables "$X^i$" and $\varepsilon^i$ is the vector of random residuals of order $T \times 1$. Here, throughout in this dissertation, the superscript "$i$" represents the unit number in the panel, the first subscript "$t$" represents the time period, the second subscript "$k$" represents the regressor's number. The two subscripts "$tk$" denote, the time period and the regressor number respectively.

Moreover, in frequentist set up, the model parameters vector $\beta^i$ is a vector of fixed constant quantities. It does not have any further structure. While this is not the case in the Bayesian set up. It is further to be noted here that, the balanced panel is assumed in this study i.e., all the units of the panel have equal number of observations for each of its variables or regressors and there is no missing data, also $T > K$ is assumed too.

Furthermore, the distribution of the residuals is assumed as given below,

$$\left[\varepsilon^i | (\sigma^2)^i\right] \underset{\sim}{ID} N\left(0,\ (\sigma^2)^i I_T\right) \tag{4.1d}$$

This implies that, $\left[\varepsilon^i | (\sigma^2)^i\right]$ is normally distributed with mean zero and covariance $\left[(\sigma^2)^i I_T\right]$.

Further,

$$\left([\varepsilon^1 | (\sigma^2)^1], [\varepsilon^2 | (\sigma^2)^2], \ldots, \left[\varepsilon^i | (\sigma^2)^i\right], \ldots, [\varepsilon^N | (\sigma^2)^N]\right) \text{ are all independent.} \tag{4.1e}$$

## 4.2. The Regression Coefficients Estimates

The *OLS or ML estimate* of $\beta^i$, with the help of (4.1d) above, is given as under,

$$\hat{\beta}^i = \left(X^{i\prime} X^i\right)^{-1} X^{i\prime} Y^i \tag{4.2a}$$

Note: Here in the above, the dash or prime " ′ " above "$X^i$", i.e., $X^{i'}$, depicts the transpose of the beneath variable $X^i$, and where transpose of a matrix is, the interchanging of the rows into columns of the matrix and vice versa.

*For analogous expression of (4.2a) above, (See. Carrington and Zaman ((1999),pp. 252)).*

## 4.3. The Covariances of the Regression Coefficients Estimates

Similarly, the variance covariance structure of $\hat{\beta}^i$ given in (4.2a) is as under,

$$Cov(\hat{\beta}^i) = \left[(\sigma^2)^i\left(X^{i'}X^i\right)^{-1}\right] \tag{4.3a}$$

The dimension of $Cov(\hat{\beta}^i)$ in (4.3a) above is $K \times K$, due to the fact of $K$ regressors in "$i^{th}$" regression model corresponding to "$i^{th}$" unit of the panel.

Let the data covariance of the regression coefficients estimate for the $i^{th}$ unit of the panel may be abbreviated and denoted by $(DV)^i$ then,

$$(DV)^i = \left[(\sigma^2)^i\left(X^{i'}X^i\right)^{-1}\right] \tag{4.3b}$$

## 4.4. The Precision of the Regression Coefficients Estimates

By definition, the precision of an estimate is equal to the inverse of its variance covariance structure, therefore, the precision of $\hat{\beta}^i$ which is denoted by $Prec(\hat{\beta}^i)$, equals as under,

$$Prec(\hat{\beta}^i) = \left[(\sigma^2)^i\left(X^{i'}X^i\right)^{-1}\right]^{-1} \tag{4.4a}$$

or equivalently,

$$Prec(\hat{\beta}^i) = \left[\left(X^{i'}X^i\right)\left\{(\sigma^2)^i\right\}^{-1}\right]. \tag{4.4b}$$

Let the data precision of the regression coefficients estimate for the $i^{th}$ unit of the panel may be abbreviated and denoted by $(DP)^i$ then,

$$(DP)^i = \left[\left\{(\sigma^2)^i\right\}^{-1}\left(X^{i'}X^i\right)\right]. \tag{4.4c}$$

Below, we discuss the additional workout that is needed for the Bayesian linear regression model.

## 4.5. The Bayesian View

From the Bayesian point of view, the information regarding the unknown parameters may be expressed in the form of a distribution. Prior to observe the data whatever information we do have is summarised with the help of a distribution called the prior distribution. When the sample data is observed, the Bayes estimator is used to update the prior density and obtained the posterior density. The posterior density describes the sum of the additional or prior information and the sample or data information.

## 4.6. The Bayesian Central Tendency Model

In Bayesian set up, the model parameters vector $\beta^i$ is considered to be a vector of variable quantities rather than a vector of fixed constant quantities, it has further a structure (fixed quantity plus random fluctuations), unlike in dealing with the classical linear regression model. The Bayesian setup demands us to put further a model on the structure of the parameters vector $\beta^i$. The *central tendency model*[7] seems suitable here; which is given as under,



$$\beta^i = \mathrm{B} + \upsilon^i \tag{4.6a}$$

where, B is called a hyperparameter and is the prior mean of $\beta^i$ in Bayesian set up, while $\upsilon^i$ is the disturbance term, and

$$\upsilon^i \sim N\left(0, \Lambda^i\right) \tag{4.6b}$$

The model parameter $\beta^i$ in Bayesian framework, is presumed to have a probability density with some mean B and variance $\Lambda^i$, and the coefficient vector for a particular case is considered as a random outcome from that density. Thus, this implies that $\beta^i$ is normally distributed with prior mean B and prior varianc-covariance $\Lambda^i$, this can be expressed as under,

$$\beta^i \sim N\left(\mathrm{B}, \Lambda^i\right) \tag{4.6c}$$

In the next section we describe various densities.

[7] For the central tendency model see Zaman (1996), Statistical Foundation for Econometric Techniques.

## 4.7. Various Densities Required for the Bayesian Analysis

The general procedure for the derivation of the Bayes estimate is as follows. First we describe some useful densities which have been using for the derivation purposes. These densities are already available in the Bayesian literature, but the purpose to present here is that to express them in our own symbols and notations which are easily understood, according to the needs and the goals of this dissertation. Therefore, the first density to be discussed below is called the data density.

## 4.8. The Data Density

The *data density* is the one that contains the data information or the sample information about the unknown population parameter in the form of a distribution or density. The data distribution can be described as; if $\hat{\beta}^i$ is the vector of the regression coefficients estimates and $\beta^i$ is the vector of the corresponding unknown parameters of the regression model, then the data density $f(\hat{\beta}^i)$ is considered to be the conditional density and is denoted by $f(\hat{\beta}^i | \beta^i)$ and pronounced as, the data density of $\hat{\beta}^i$ given $\beta^i$. The data density of $\hat{\beta}^i$ given $\beta^i$ is given as under,

$$[\hat{\beta}^i | \beta^i, (\sigma^2)^i,\ X^i] \sim N\ \left[\beta^i, (\sigma^2)^i \left(X^{i\prime} X^i\right)^{-1}\right] \tag{4.8a}$$

or, in the light of (4.3b),

$$[\hat{\beta}^i | \beta^i, (\sigma^2)^i,\ X^i] \sim N\ [\beta^i, (DV)^i] \tag{4.8b}$$

for reference, (See. *Corollary 3.3*, pp.(45), Statistical Foundation for Econometric Techniques, Zaman (1996)). It is further to explain that the data density is sometimes known as the conditional density.

From Bayesian perspective the existence of the prior or additional information regarding the unknown population parameter $\beta^i$ is assumed through a density, called the prior density. Therefore, the next important density that is to be described is the prior density.

## 4.9. The Prior Density

The density that contains useful information, other than the data, regarding the unknown parameter of interest $\beta^i$, is called the prior density. There are many types of the prior densities in the literature, one type of the prior densities is known as the conjugate prior.

In the Bayesian theory of probability, if the prior and posterior densities belong to the same family of the distributions, then the prior they are called the conjugate distributions. Further, the prior density is then called the conjugate prior density and the posterior is called the conjugate posterior density. Such priors are called the conjugate prior for the likelihood function. The conjugate prior is also known as the natural conjugate prior.

Now the prior we use here in the Bayes estimate are the normal *conjugate priors*. The kind of prior information required by the Bayesian techniques built on the natural conjugate priors are very demanding. Why we make this assumption about the prior because they are mathematically convenient and easy to compute. Representing the prior information in the form of a normal distribution takes to the estimates that have certain very unwanted properties.

Analytically suitable the natural conjugate priors are naturally not adequately flexible to appropriately express the prior information. However, they form the basis for *empirical*[8] and *hierarchical*[9] Bayes estimates.

Below we describe the normal conjugate prior density of $\beta^i$ and some other related useful information. The *prior density* conditional on hyperparameters $f(\beta^i|\mathrm{B},\Lambda^i)$ is as under,

$$[\beta^i|\mathrm{B},\Lambda^i] \sim N[\mathrm{B},\Lambda^i], \qquad \text{for all } \beta^i \qquad (4.9a)$$

This states that $\beta^i$ is distributed normally with prior mean vector B, and prior covariance matrix $\Lambda^i$. Where, B and $\Lambda^i$ are the parameters of the prior density and are called the *hyperparameters*, in order, to

[8] The Bayes estimator, where, the estimates of the hyperparameters, estimated from the density of the parameters, are used is called the hierarchical.

[9] The Bayes estimator, where, the hyperparameters further have prior structure or prior density, with the help of which first stage prior parameters are estimated from the density of the second stage prior parameters density, is called the empirical.

differentiate them from the parameters of the data density. *For analogous expression of (4.9a) above, (See. Carrington and Zaman (1994),pp. 256).*

Let the prior variance $\Lambda^i$ (hyperparameter) of the parameters of the $i^{th}$ unit of the panel may be abbreviated and denoted by $(PV)^i$ then,

$$(PV)^i = [\Lambda^i] \tag{4.9b}$$

Now with the help of (4.9b), the above (4.9a) can also be written as,

$$[\beta^i | \mathrm{B}, \Lambda^i\ ] \sim N\,[\mathrm{B}, (PV)^i], \qquad \text{for all } \beta^i. \tag{4.9c}$$

Similarly, let the inverse of the prior variance of the $i^{th}$ unit of the panel $[\Lambda^i]^{-1}$ i.e., the prior precision of the $i^{th}$ unit of the panel, may be abbreviated and denoted by $(PP)^i$ then,

$$(PP)^i = [\Lambda^i]^{-1} \tag{4.9d}$$

Now with the help of the above densities we obtain another valuable density, in fact, the most important density called the posterior density, this density is presented in the following subsection.

## 4.10. The Posterior Density

The posterior density is the desired density in the Bayesian setup as it contains all the valuable information about the unknown parameter of interest after observing the sample data. Thus, the *posterior density* has the virtue of containing all the useful information regarding the parameter $\beta^i$ which were either before in the data density or in the prior density. Therefore, in the light of (4.4c) and (4.9d) above, the *posterior density* $f(\beta^i | \hat{\beta}^i)$ is given as under,

$$\beta^i | \hat{\beta}^i \sim N\left\{[(DP)^i + (PP)^i]^{-1}[(DP)^i\hat{\beta}^i + (PP)^i\mathrm{B}], [(DP)^i + (PP)^i]^{-1}\right\}. \tag{4.10a}$$

The arithmetic mean or simply the mean of the posterior density contains a good single-point summary of this information. This mean is the optimal Bayes estimator for the class of loss functions it also includes the quadratic loss function.

Consequently, the prior to posterior transformation formula resultantly yields formula for the Bayesian estimates of the regression coefficients.

In other words, the posterior density has very useful information, the expected value or the mean of the posterior density is the Bayes estimate. Therefore, the expected value of the posterior density is shown below,

$$E[\beta^i|\hat{\beta}^i] = \left[(DP)^i + (PP)^i\right]^{-1}\left[(DP)^i\hat{\beta}^i + (PP)^i\mathrm{B}\right]. \tag{4.10b}$$

Let the mean of the posterior density or the Bayes estimate may be denoted by $M^i$, then,

$$M^i = \left[(DP)^i + (PP)^i\right]^{-1}\left[(DP)^i\hat{\beta}^i + (PP)^i\mathrm{B}\right]. \tag{4.10c}$$

Further, the *covariance matrix* of the posterior density is given as under,

$$Cov[\beta^i|\hat{\beta}^i] = \left[(DP)^i + (PP)^i\right]^{-1}. \tag{4.10d}$$

Let the *posterior covariance* may be denoted by $V^i$, then,

$$V^i = \left[(DP)^i + (PP)^i\right]^{-1} \tag{4.10e}$$

*(For analogous expression of (4.10a) - (4.10e) above, See. Statistical Foundation for Econometric Techniques, by Zaman (1996), pp. (45) equation (3.2), (3.3) and pp.(334) equation (17.15)).*

The posterior mean is the weighted arithmetic mean of the prior mean $B$ and the data mean $\hat{\beta}^i$. Furthermore, the weights that are entering in the formula of the posterior mean are just the corresponding precisions of the two measures or means, namely the precision of the data mean and the precision of the prior mean: both the means are multiplied by their precisions and the sum is then pre-multiplied by the inverse of the total of the precisions.

As it has been mentioned earlier that the mean of the posterior density is the Bayes estimate, so, in this particular case of linear regression, the mean of this posterior density is the Bayes linear regression estimate of the linear regression coefficients.

## 4.11. The Bayes Estimates of the Panel Data Linear Regression Model

Now, we describe two various approaches of the Bayesian estimate of the linear regression panel data models. The first kind we will label *classical* Bayes, simply because they were introduced first. Further, they are very easy to compute, and also based on the *natural conjugate priors,* where the conjugate prior is also known as the natural conjugate prior, as mentioned previously. This label is attached only for the convenience of reference to such types of estimates.

## 4.12. The Classical Bayes

The Bayes estimates based on normal priors refer to the "classical Bayes" since they were developed first. If the hyperparameters, i.e., $\{B, (PV)^i\}$, are known in advance, then $M^i$, which is the mean of the posterior density, is the classical Bayesian estimator for $\beta^i$, and denoted by $\widehat{M}^i_{CB}$. Thus, the classical Bayes estimate of the coefficients of the linear regression model from (4.10c) above is as under,

$$\widehat{M}^i_{CB} = \left[(DP)^i + (PP)^i\right]^{-1} \left[(DP)^i\hat{\beta}^i + (PP)^i\mathrm{B}\right], \tag{4.12a}$$

and similarly, the *covariance matrix* of the classical Bayes estimate in this case is given as under,

$$\hat{V}^i_{CB} = \left[(DP)^i + (PP)^i\right]^{-1}. \tag{4.12b}$$

The prior used here, in the classical Bayes, are assumed as known as the *conjugate priors*. The natural conjugate prior information, that gives the base, are very attractive for the classical Bayes estimation techniques.

If the prior information or in other words, the hyper parameters are not known then the classical Bayesian perception for the selection the values for the hyper parameters on subjective grounds does not work adequately well practically. This is the reason that the classical Bayes estimation techniques are not being seriously recommended for the use in the practical settings. There are some problems attached with classical Bayes, the main three complications related to the classical Bayes are being described below.

The first complication is the *"unbounded risk"*. It is worth noting that the Bayes estimate is biased towards the prior mean (See. Zaman (1996)). Thus, by choosing the prior variance sufficiently large, the set of values where the Bayes rule dominates ML or OLS arbitrarily large can be made larger, but at the same time, for the large value of $\hat{\beta}^i$, the risk of the Bayes rule goes to infinity.

Further, irrespective of the fact that how the prior parameters have been chosen, the risk of the ML dominates the Bayes rule arbitrarily for sufficiently large $\hat{\beta}^i$. This emerges the problems by using the Bayes rules. The cases where the prior information is certain and definite are very rare in practice. Thus, if the priors are not reasonable, then there is always a chance for tremendously poor results by the Bayes rule or classical Bayes.

The second is the "*choice of hyperparameters*", unluckily, there is no rule available in the literature that could help us in making good choice for the hyperparameters in real world applications. For the choice of the hyperparameters the classical Bayes technique needs the assessment of the prior density on subjective grounds.

There are occasions where rational choices can be made, in decision making environment, by assuming as if one has the known priors. The performance of Bayesian procedure mainly depends upon the choices of the hyperparameters, for different choices of the hyperparameters the Bayesian procedure performs very differently from one another.

The third is the "*conflicting prior and data*". Sometimes the prior information and the data information are conflicting to each other and hence such a situation always becomes a source of difficulties in the Bayesian analysis. However, when the conflict between the data information and prior information appears, in such situations it does makes sense to make use of both sources of data and aggregate the two data sets. In the situation like the one discussed here, if the researcher believes the prior to be more convenient and meaningful then he/she should discard the data and vice versa (See. Zaman (1996)).

## 4.13. Remedies of the Difficulties in the Classical Bayes Estimate

Now, related to the first difficulty of the classical Bayes estimate, that of the *unbounded risk*, if the prior variance is selected small and $\hat{\beta}^i$ is chosen near to the prior mean then the risk of the Bayes rule is good against the ML and very poor otherwise.

On the other hand, if the prior parameters are estimated from the sample data, then the data is automatically in conformity with the values of the hyperparameters and the case where the data confronts the prior is avoided.

Related to the second difficulty, uncertainty about the hyperparameters. As mentioned above to estimate the hyperparameters directly from data, this assumes the hyperparameters to be correct with certainty, but this is not the case, in reality. Attempts are made to account for the uncertainty related to the hyperparameter one way or the other.

To overcome this difficulty, another way is that the hyperparameters are not directly estimated rather estimate the decision rule to be used. This is the alternate and indirect way of addressing the problem.

Related to the third difficulty of the conflicting prior and data, this can also be minimized if the prior have been estimated from the data, conflict may still exist, however.

The fact behind the avoidance of the Bayesian technique in econometrics is that in many instances the reliable prior information is not available, and the conventional Bayesian estimates perform very badly if the prior information is somehow mis-specified.

The estimate where the prior parameters $\{B, (PV)^i\}$ are being estimated from the data set is another type of estimates and to be discussed after the classical Bayes estimate.

The Bayes estimate, where the data are used to estimate not only the parameters of the data density, but also the prior parameter, this is known as the "Empirical" Bayes approach. The empirical Bayes procedure involves proceeding with a classical Bayes analysis using these estimates as if they were the prior parameters.

There are three different ways to implement the empirical Bayes approach. In all cases, the marginal density of the observations is used to provide estimates for the hyperparameters. The simplest type of empirical Bayes procedure is based on directly estimating the hyperparameter.

In the coming subsection a very important density is presented that is very much meaningful for linking the data and the hyperparameters, this density is known as the marginal density. Now we are going to write the marginal density of $\hat{\beta}^i$.

## 4.14. The Marginal Density

The *marginal density* of $\hat{\beta}^i$ relative to data and prior in a compact form can be described as under,

$$m(\hat{\beta}^i) \sim N\left[\mathrm{B},\ (DV)^i\ +\ (PV)^i\right]. \tag{4.14a}$$

Why do we need to write the ***marginal density***[10] of $\hat{\beta}^i$ given prior parameters B and $\Lambda^i$? This is a crucial element of the empirical Bayes method. The marginal density links the data with the hyperparameters via the parameters of the data density.

Note that the data density depends on the parameters, and the prior density of the parameters depends on the hyperparameters so that the marginal density of the observation, after integrating out the parameters, will depend directly on the hyperparameters. This is the crucial step in empirical Bayes method. The marginal density thus allows estimation of the hyperparameters from the data.

Below we estimate the prior variance, that is needed for the empirical Bayes, from the data with the help of the marginal density of $\hat{\beta}^i$.

## 4.15. The Estimation of the Prior Variance

There are several different priors for the prior variance in the literature, i.e., d-prior or g-prior, flat prior, noninformative prior, etc., where d in d-prior stands for the diagonal and similarly g in g-prior stands for gamma. Below we describe the d and g-priors.

### 4.15.1. d-prior

The d-prior involves assuming that $\beta^i \sim N\left(\mathrm{B}, \Lambda^i\right)$, where $\Lambda^i$ is a diagonal matrix. As a general rule, putting in too many hyperparameters leads to instability in empirical Bayes procedures, and it would be a bit risky to try to estimate a general covariance matrix unless the number of hyperparameters is small relative to the dataset size. Thus, it is essential to keep the $\Lambda^i$ diagonal.

### 4.15.2. g-prior

An alternative formulation for the prior leads to very tractable formulae. Suppose we assume that $\beta^i \sim N\left(\mathrm{B}, \varphi(DV)^i\right)$, where '$\varphi$' is a scalar. Zellner (1971) and Ghosh et al. (1989), label this the g-prior and present some stories justifying the use of this prior. In fact, the prior is plausible, but its main justification is

[10] In joint density, the factor $m(\hat{\beta}^i)$, other than the posterior density, is of course, the marginal density of $\hat{\beta}^i$.

computational convenience and the fact that it frequently leads to reasonable/good results in empirical applications.

Note that it is typically not useful in small samples to estimate all features of the prior, and some parts of the prior must still be specified a priori. We will use the g-prior here, because only with the help of this we can reach to our destination and "g" stands for the gamma priors.

To estimate the prior[11] variance of $\beta^i$, with the help of g-priors, we need first to define the g-priors. By the definition of g-priors, the prior variance is proportional to the data variance. i.e.,

$$(PV)^i \propto \{(DV)^i\} \tag{4.15.2a}$$

or in the light of $\beta^i \sim N\left(\mathrm{B}, \varphi(DV)^i\right)$,

$$(PV)^i = \varphi\ (DV)^i. \tag{4.15.2b}$$

## 4.16. The Prior Precision Structure

The estimate of the prior precision of $\beta^i$, with the help of g-priors, will take the form as given below,

$$(PP)^i = \rho\ (DP)^i, \tag{4.16a}$$

where, $\rho$ is a $(K \times K)$ diagonal matrix here, and is called the prior precision parameters' matrix, also,

$$\rho = (\varphi)^{-1}. \tag{4.16b}$$

This prior precision parameter $\rho$ plays a very important role, in panel data linear regression models, the variation in this $\rho$ leads to different standard classical panel data linear regression estimates and the corresponding models.

## 4.17. The Estimate of the Prior Mean

The estimate of the prior mean is obtained from the marginal density of $\hat{\beta}^i$ relative to the prior B and $(PV)^i$. We have $m(\hat{\beta}^i) \sim N\left[\mathrm{B},\ (DV)^i\ +\ (PV)^i\right]$, thus the precision weighted arithmetic mean of the well-

[11] There are nonetheless certain difficulties introduced by estimating the prior in the usual empirical Bayes way.

known ordinary least squares estimates, the estimate of the prior mean with the help of g-priors is described as under,

$$\widehat{\mathrm{B}} = \left(\sum_{i=1}^{N}[(DP)^i + (PP)^i]\right)^{-1}\left(\sum_{i=1}^{N}\{[(DP)^i + (PP)^i]\hat{\beta}^i\}\right). \quad (4.17a)$$

The above in (4.17a) is the precision weighted arithmetic mean of the ordinary least squares coefficients estimates.

*For analogous expression of (4.16a), (See. Carrington and Zaman (1994), pp258), equation (7) and "Statistical Foundation for Econometric Techniques" by Zaman (1996), pp.334, equation (17.16)).* Now, having all the estimates needed for the empirical Bayes, below we describe the empirical Bayes estimates.

## 4.18. The Empirical Bayes Estimate

The empirical Bayes estimate of the regression coefficients of the panel data linear regression model is as under,

$$\widehat{M}_{EB}^{i} = \left[(\widehat{DP})^{i} + (\widehat{PP})^{i}\right]^{-1}\left[(\widehat{DP})^{i}\hat{\beta}^{i} + (\widehat{PP})^{i}\widehat{\mathrm{B}}\right], \quad (4.18a)$$

and *covariance matrix* of the empirical Bayes estimate in this case is given as under,

$$\widehat{V}_{EB}^{i} = \left[(\widehat{DP})^{i} + (\widehat{PP})^{i}\right]^{-1}. \quad (4.18b)$$

The above (4.18a) is the empirical Bayes estimate.

In the empirical Bayes technique described above, we estimate the hyperparameters, but then pretend that our estimates equal the hyperparameters with certainty and proceed to use classical Bayes formulas. Note: The model that is introduced by Hildreth and Houck (1968) and is called the Hildreth-Houck random coefficients model, is closely related to the empirical Bayes models, although estimation techniques and motivation for the models are different.

The estimate which alleviates all three of the difficulties associated with classical Bayesian estimates discussed above is called the empirical[12] Bayes.

[12] Empirical Bayes estimates can be viewed as attempts to remove or reduce some of these undesirable features of classical Bayes estimators.

Morris (1983), in an excellent though dated review of empirical Bayes methods, writes that the benefits of empirical Bayes depend on the validity of (4.6a) and less critically on the procedure used to estimate the hyperparameters. One of the main characteristics of the empirical Bayes estimates is the ability to compromise between the priors and the data information.

An important reason why the empirical Bayesian techniques have been avoided in econometrics, this is because the model and estimation techniques are new and relatively unfamiliar. Therefore, the empirical Bayes technique is vastly underutilized in econometrics. The same situation prevailed in statistics, where these techniques originated, until a series of articles by Efron and Morris (1972), (1973) and (1975), were written to address the four main obstacles to the use of the empirical Bayesian techniques.

A number of authors, such as, Efron and Morris, Carter and Rolph, and some others, have adapted the empirical Bayes techniques and successfully applied them to regression models. However, each has done so in his own way, and the unity behind the diverse techniques is hard to recognize.

Now, let us discuss, in following subsection, the empirical Bayes estimate single orthogonal regressor approach in this dissertation.

## 4.19. Deriving the Empirical Bayes Estimate in Single Orthogonal Regressor Case

In this section we develop the empirical Bayes estimate for a single $k^{th}$ orthogonal regressor, as this is the route to the goal of this dissertation. As here all the derivations are made about a particular $k^{th}$ regressor, therefore, we modify all of our previous symbols and notations, and specify them for a $k^{th}$ regressor.

For this purpose, let,

$$(DV_k)^i = (\hat{\sigma}_k^2)^i(A_{kk}^i) \tag{4.19A}$$

be the $k^{th}$ diagonal entity of $(DV)^i$ "the data variance matrix" given in (4.3b) and let,

$$(DP_k)^i = \left[(\hat{\sigma}_k^2)^i(A_{kk}^i)\right]^{-1} \tag{4.19B}$$

be the corresponding $k^{th}$ diagonal entity of $(DP)^i$ "the prior variance matrix" given in (4.4c).

Then the marginal density of $\hat{\beta}_k^i$ is as,

$$\hat{\beta}_k^i \sim N(B_k, (DV_k)^i + (PV_k)^i) \tag{4.19a}$$

This implies that $\hat{\beta}_k^i$ is normally distributed with mean $B_k$ and variance $\left[(DV_k)^i + (PV_k)^i\right]$.

Note: The variance of the marginal density of $\hat{\beta}_k^i$ is the sum of the data variance and prior variance.

Further, for each $i = 1, 2, \dots, N$, the OLS estimate of the $k^{th}$ coefficient is an independent observation on $B_k$, but each of these estimates has different variance. We have $N$ independent observations on $B_k$ which have different variances then the best estimate of $B_k$ is the precision weighted average of the observations.

$$\hat{B}_k = \frac{\left[\sum_{i=1}^{N}\left\{\left[(DV_k)^i\right]^{-1} + \left[(PV_k)^i\right]^{-1}\right\}\hat{\beta}_k^i\right]}{\left[\sum_{i=1}^{N}\left\{\left[(DV_k)^i\right]^{-1} + \left[(PV_k)^i\right]^{-1}\right\}\right]} \tag{4.19b}$$

Depending on how we specify the prior variance, we can get different formulae for the empirical Bayes estimate of the prior mean $B_k$. A convenient assumption, which makes the formulae simplify is the Zellner's (1971) g-prior, where,

$$(PV_k)^i = \varphi_k (DV_k)^i. \tag{4.19c}$$

or equivalently,

$$\left[(PV_k)^i\right]^{-1} = \rho_k\left[(DV_k)^i\right]^{-1} \tag{4.19d}$$

and where,

$$\rho_k = (\varphi_k)^{-1} \tag{4.19e}$$

Now the formula for $\hat{B}_k$ in the light of (4.19c) simplifies,

$$\hat{B}_k = \frac{\left[\sum_{i=1}^{N}\left\{\left[(DV_k)^i\right]^{-1} + \left[\varphi_k (DV_k)^i\right]^{-1}\right\}\hat{\beta}_k^i\right]}{\left[\sum_{i=1}^{N}\left[(DV_k)^i\right]^{-1} + \left[\varphi_k (DV_k)^i\right]^{-1}\right]} \tag{4.19f}$$

Further, the formula for $\hat{B}_k$ in the light of (4.19d) simplifies to,

$$\hat{B}_k = \frac{\left[\sum_{i=1}^{N}\left\{\left[(DV_k)^i\right]^{-1} + \rho_k\left[(DV_k)^i\right]^{-1}\right\}\hat{\beta}_k^i\right]}{\left[\sum_{i=1}^{N}\left[(DV_k)^i\right]^{-1} + \rho_k\left[(DV_k)^i\right]^{-1}\right]} \tag{4.19g}$$

or equivalently,

$$\hat{B}_k = \frac{\left[\sum_{i=1}^{N}\left(\left[(DV_k)^i\right]^{-1}\hat{\beta}_k^i\right)\right]}{\left[\sum_{i=1}^{N}\left(\left[(DV_k)^i\right]^{-1}\right)\right]} \tag{4.19h}$$

Furthermore, in the light of (4.19c), the marginal density of $\hat{\beta}_k^i$ given in (4.19a) is as,

$$\hat{\beta}_k^i \sim N\left[B_k, (DV_k)^i + \varphi_k (DV_k)^i\right] \tag{4.19i}$$

or equivalently,

$$\hat{\beta}_k^i \sim N\left[B_k, (1+\varphi_k)(DV_k)^i\right] \tag{4.19j}$$

For empirical Bayes estimate, there are two steps in the derivation. The first step is the estimation of prior parameters from marginal density and the second step is to take these estimates and plug them into the classical Bayes estimate.

Since the classical Bayes estimate of $\hat{\beta}_k^i$, denoted by $CB(\hat{\beta}_k^i)$, is the precision weighted arithmetic mean of $\hat{\beta}_k^i$ and $B_k$, the weights are being the corresponding precisions, is given as,

$$CB(\hat{\beta}_k^i) = \frac{\left[(DV_k)^i\right]^{-1}\hat{\beta}_k^i + \rho_k\left[(DV_k)^i\right]^{-1}\hat{B}_k}{\left[(DV_k)^i\right]^{-1} + \rho_k\left[(DV_k)^i\right]^{-1}}. \tag{4.19k}$$

$$CB(\hat{\beta}_k^i) = \frac{1}{1+\rho_k}\hat{\beta}_k^i + \frac{\rho_k}{1+\rho_k}\hat{B}_k \tag{4.19l}$$

Now from the marginal density of $\hat{\beta}_k^i$ in (4.19j),

$$Mean(\hat{\beta}_k^i) = B_k \tag{4.19m}$$

and

$$Var(\hat{\beta}_k^i) = (1+\varphi_k)(DV_k)^i \tag{4.19n}$$

let,

$$Z = \left\{\frac{\hat{\beta}_k^i - \hat{B}_k}{\sqrt{(1+\varphi_k)(DV_k)^i}}\right\} \tag{4.19o}$$

let,

$$S_z = \sum_{i=1}^{N}\left(\left\{\frac{\hat{\beta}_k^i - \hat{B}_k}{\sqrt{(1+\varphi_k)(DV_k)^i}}\right\}^2\right) \tag{4.19p}$$

then

$$S_z = \sum_{i=1}^{N}\left(\left\{\frac{\hat{\beta}_k^i - \hat{B}_k}{\sqrt{(1+\varphi_k)(DV_k)^i}}\right\}^2\right) \sim \chi_N^2 \tag{4.19q}$$

since,

$$\frac{1}{S_z} \sim inv(\chi_N^2) \tag{4.19r}$$

$$E\left(\frac{1}{S_z}\right) = \frac{1}{N-2} \tag{4.19s}$$

or

$$E\left[\sum_{i=1}^{N}\left\{\frac{(1+\varphi_k)(DV_k)^i}{\left(\hat{\beta}_k^i - \hat{B}_k\right)^2}\right\}\right] = \frac{1}{N-2} \tag{4.19t}$$

$$E\left[\sum_{i=1}^{N}\left\{\frac{(N-2)(DV_k)^i}{\left(\hat{\beta}_k^i - \hat{B}_k\right)^2}\right\}\right] = \frac{1}{(1+\varphi_k)} \tag{4.19u}$$

$$E\left[\sum_{i=1}^{N}\left\{\frac{(N-2)(DV_k)^i}{\left(\hat{\beta}_k^i - \hat{B}_k\right)^2}\right\}\right] = \frac{1}{\left(1+\frac{1}{\rho_k}\right)} \tag{4.19v}$$

$$E\left[\sum_{i=1}^{N}\left\{\frac{(N-2)(DV_k)^i}{\left(\hat{\beta}_k^i - \hat{B}_k\right)^2}\right\}\right] = \frac{\rho_k}{(1+\rho_k)} \tag{4.19w}$$

thus,

$$\sum_{i=1}^{N}\left\{\frac{(N-2)(DV_k)^i}{\left(\widehat{\beta}_k^i-\widehat{B}_k\right)^2}\right\}\approx\frac{\rho_k}{(1+\rho_k)} \tag{4.19x}$$

as, the classical Bayes estimate for $\beta_k^i$ in (4.19l) is,

$$CB(\hat{\beta}_k^i)=\frac{1}{1+\rho_k}\hat{\beta}_k^i+\frac{\rho_k}{1+\rho_k}\hat{B}_k \tag{4.19y}$$

Now, using the empirical Bayes estimation methodology to derive an estimate for the $\frac{\rho_k}{(1+\rho_k)}$ that occurs in the classical Bayes estimate, when we do not know that value of the prior parameter $\rho_k$. Using (4.19x) the classical Bayes estimate in (4.19y) simplifies and known as empirical Bayes estimate, denoted by $EB(\hat{\beta}_k^i)$,

$$EB(\hat{\beta}_k^i)=\left(\left\{1-\left[\sum_{i=1}^{N}\left\{\frac{(N-2)(DV_k)^i}{\left(\hat{\beta}_k^i-\hat{B}_k\right)^2}\right\}\right]\right\}\hat{\beta}_k^i+\left\{\left[\sum_{i=1}^{N}\left\{\frac{(N-2)(DV_k)^i}{\left(\hat{\beta}_k^i-\hat{B}_k\right)^2}\right\}\right]\right\}\hat{B}_k\right) \tag{4.19z}$$

Here whichever part becomes positive must be taken in the first term and the second term must have Max (1, Sum). Hence, this was the desired form of the empirical Bayes estimate for a single $k^{th}$ orthogonal regressor.

## 4.20. Summary of the Chapter

In this chapter, first the panel data models and the associated stuff have been discussed, then the Bayesian procedure has been outlined. In the Bayesian procedure all the densities which are being used in the analysis have been presented and then the corresponding Bayesian estimates have been described. In the Bayesian estimates, first, the classical Bayes estimate for the linear regression model the corresponding pros and cons were discussed.

The classical Bayes estimates suffering from three main difficulties, these difficulties have been described here, also the remedies of the difficulties have been suggested. These difficulties can be overcome by the empirical Bayes estimate. After this, the empirical Bayes estimate for the panel data linear regression model, the corresponding pros and cons were discussed.

The empirical Bayes estimate discussed here in this chapter is of much importance, because from this empirical Bayes we derive all the frequentist panel data estimates and the associated models. All the frequentist panel data linear regression models will be derived from this empirical Bayes analytically and then numerically in chapter 5 and 6 respectively.

Finally, a new empirical Bayes estimate with the precision weighted arithmetic mean of the OLS estimates as the estimate of the prior mean and the g-prior precision as the prior precision, for single $k^{th}$ orthogonal regressor case, has been developed in 4.19z and this is the main objective and finding of this chapter as well.

# Chapter 5

## Analytical Derivations of the Frequentist Estimates

This chapter is the main chapter of this thesis, as one of the main objectives of this thesis is to show that all the standard frequentist panel data linear regression models and the corresponding estimators of these model, are of single[13] origin and all of them can be derived from this particular origin. The common origin we talk about is the Bayesian linear regression model and the empirical Bayes estimate.

This chapter discusses the derivations of the frequentist panel data linear regression models from the Bayesian panel data linear regression model and also the derivations of the frequentist estimators from the empirical Bayes estimate. Therefore, the whole analytical work is described here in this chapter.

Note: Here, the derivations will be carried out for $K$ orthogonal regressors model, but a single (for $k^{th}$ orthogonal) regressor estimation approach at a time.

The goals of derivations of the frequentist panel data estimates from the empirical Bayes estimate can be achieved, when;

i) the precision weighted arithmetic mean, of the ordinary least squares coefficients estimates, is used as the estimate of the prior mean,

ii) the Zellner's *g-prior* is used as the estimate of the prior variance (and the resulted prior precision), for the empirical Bayes estimate,

and then

iii) the prior precision parameter "$\rho_k$" takes different values i.e., zero, infinity or any value in-between zero and infinity. (Note: "$\rho_k$" is the prior precision parameter as defined above).

[13] Note: In the classical viewpoint, all the standard classical panel data model estimators have their own distinct origin, they are totally different from each other, while the Bayesian consider that all of the standard classical panel data model estimators have a single, unique and identical origin they are not of distinct origin.

Only with the help of the above conditions, regarding the empirical Bayes estimate, the goals will be achieved.

The five different possibilities or limiting cases of variation, in the prior precision parameter $\rho_k$, take to different frequentist estimates. These limiting cases are being discussed below in section (5.3) to (5.7) and hence the resulted frequentist estimates are also being derived.

Further, with each of the derived frequentist estimate of panel data models, the corresponding panel data linear regression model is also presented.

Below, we modify the $K$ orthogonal regressors' empirical Bayes estimate and the related quantities for a single $k^{th}$ regressor of $i^{th}$ unit of the panel, as this is the methodology of the derivation in this thesis.

## 5.1. Single Regressor Empirical Bayes Estimate

The empirical Bayes estimate of $\beta^i$ of $K$ regressors panel data linear regression model for the $i^{th}$ unit of the panel, with the precision weighted arithmetic mean of the ordinary least squares coefficients estimates as the estimate of the prior mean and the Zellner's g-prior as the estimate of the prior variance (and the resulted precision) from (4.18a), is given as under,

$$\widehat{M}_{EB}^i = \left[\left(\widehat{DP}\right)^i + \left(\widehat{PP}\right)^i\right]^{-1}\left[\left(\widehat{DP}\right)^i\hat{\beta}^i + \left(\widehat{PP}\right)^i\hat{\mathrm{B}}\right] \tag{5.1a}$$

or equivalently from (4.16),

$$\widehat{M}_{EB}^i = \left[\left(\widehat{DP}\right)^i + \rho\left(\widehat{DP}\right)^i\right]^{-1}\left[\left(\widehat{DP}\right)^i\hat{\beta}^i + \rho\left(\widehat{DP}\right)^i\hat{\mathrm{B}}\right]. \tag{5.1b}$$

In the light of (4.3b), the estimate of the data variance of the $k^{th}$ regression coefficient estimate of $i^{th}$ unit of the panel may be denoted as under,

$$\left(\widehat{DV}_k\right)^i = (\hat{\sigma}_k^2)^i \left(X_k^{i\prime} X_k^i\right)^{-1}, \tag{5.1c}$$

or

$$\left(\widehat{DV}_k\right)^i = (\hat{\sigma}_k^2)^i \ A_{kk}^i, \tag{5.1d}$$

because $X_k^i$, is a single orthogonal regressor and where,

$$A_{kk}^{i} = \left(X_k^{i\prime} X_k^{i}\right)^{-1}. \quad (5.1e)$$

Similarly, the corresponding data precision may be denoted and expressed as under,

$$\left(\widehat{DP}_k\right)^{i} = \left[(\hat{\sigma}_k^2)^i\right]^{-1}\left(A_{kk}^{i}\right)^{-1}, \quad (5.1f)$$

or

$$\left(\widehat{DP}_k\right)^{i} = \frac{1}{\left(\hat{\sigma}_k^2\right)^{i} A_{kk}^{i}}. \quad (5.1g)$$

Let us, the prior variance, may be denoted and defined as under,

$$\left(\widehat{PV}_k\right)^{i} = \varphi_k \left\{(\hat{\sigma}_k^2)^i A_{kk}^{i}\right\} \quad (5.1h)$$

or

$$\left(\widehat{PV}_k\right)^{i} = \varphi_k \left(\widehat{DV}_k\right)^{i}. \quad (5.1i)$$

The corresponding prior precision may be denoted and defined as under,

$$\left(\widehat{PP}_k\right)^{i} = \rho_k \left\{\frac{1}{\left(\hat{\sigma}_k^2\right)^{i} A_{kk}^{i}}\right\}, \quad (5.1j)$$

where

$$(\varphi_k)^{-1} = \rho_k \quad (5.1k)$$

then,

$$\left(\widehat{PP}_k\right)^{i} = \rho_k \left(\widehat{DP}_k\right)^{i}. \quad (5.1l)$$

Also, in the light of (4.2a), the ordinary least squares coefficient estimate for the $k^{th}$ orthogonal regressor of $i^{th}$ unit of the panel is given as under,

$$\hat{\beta}_k^{i} = \left(X_k^{i\prime} X_k^{i}\right)^{-1}\left(X_k^{i\prime} Y_k^{i}\right), \quad (5.1m)$$

because, $X_k^i$, is a single orthogonal regressor.

Analogously, the modified prior mean of the $k^{th}$ regression coefficient will be as under,

$$\text{Estimate of the Modified Prior Mean for } k^{th} \text{ Regression Coefficient} = \hat{B}_k. \quad (5.1n)$$

Thus, having all of the above sample counter parts, the empirical Bayes estimate for the $k^{th}$ regression coefficient of $i^{th}$ unit of the panel gets the form as under. Let this may be denoted by $EB(\hat{\beta}_k^i)$, then,

$$EB(\hat{\beta}_k^i) = \left(\left[\frac{1}{(\hat{\sigma}_k^2)^i}(X_k^{i\prime} X_k^i)\right] + \rho_k \left[\frac{1}{(\hat{\sigma}_k^2)^i}(X_k^{i\prime} X_k^i)\right]\right)^{-1}$$
$$\left(\left[\frac{1}{(\hat{\sigma}_k^2)^i}(X_k^{i\prime} X_k^i)\right]\left((X_k^{i\prime} X_k^i)^{-1}(X_k^{i\prime} Y_k^i)\right) + \rho_k \left[\frac{1}{(\hat{\sigma}_k^2)^i}(X_k^{i\prime} X_k^i)\right]\hat{B}_k\right), \tag{5.1o}$$

or

$$EB(\hat{\beta}_k^i) = \left[(\widehat{DP}_k)^i + \rho_k (\widehat{DP}_k)^i\right]^{-1} \left[(\widehat{DP}_k)^i \hat{\beta}_k^i + \rho_k (\widehat{DP}_k)^i \hat{B}_k\right]. \tag{5.1p}$$

Now on the basis of the above information, we can show how the empirical Bayes estimate produces all the frequentist coefficients estimates of the panel data linear regression models and then the corresponding panel data linear regression models.

Below we show all of the five cases mentioned above one by one analytically.

## 5.2. The Derivations of the Frequentist Estimates

In the following sections, we derive all the standard frequentist coefficients estimates of the $K$ orthogonal regressors panel data linear regression models from the empirical Bayes estimate. Further, a single regressor coefficient estimation technique will be adopted. Let see in the following subsection the individual or the subjects specific coefficients estimates and the corresponding subjects specific panel data linear regression model.

## 5.3 The Subjects Specific Coefficients Estimates

The case where the prior precision parameter $\rho_k$ tends to zero, corresponding to all the coefficients estimates of the model, the empirical Bayes estimate reduces to the subject specific coefficients estimates.

**The prior precision parameter $\rho_k$ tends to Zero:** The first limiting case is; when the prior precision parameter $\rho_k$ tends to zero corresponding to all the coefficients estimates, i.e. $\rho_k \to 0$, $\forall\ k$, in the empirical Bayes estimate. As a result, the prior precision becomes zero, which means a high level of impreciseness of the prior information and hence the common prior mean or the common characteristics of the units from the empirical Bayes estimate vanishes and resultantly only the subject specific individual estimate remains in the formula of the empirical Bayes and similarly the empirical Bayes estimate solely depends upon the data information. This case indicates complete heterogeneity among the units of the panel. Thus, in this case the empirical Bayes estimate, becomes the frequentist subject specific coefficient estimate. Now to show this, we proceed as follow.

From (5.2p) above, we have

$$EB\left(\hat{\beta}_k^i\right) = \left[\left\{\left(\widehat{DP}_k\right)^i + \rho_k \left(\widehat{DP}_k\right)^i\right\}^{-1} \left\{\left(\widehat{DP}_k\right)^i \hat{\beta}_k^i + \rho_k \left(\widehat{DP}_k\right)^i \widehat{\mathrm{B}}_k\right\}\right] \tag{5.3a}$$

$$EB\left(\hat{\beta}_k^i\right) = \left[\left\{(1 + \rho_k)\left(\widehat{DP}_k\right)^i\right\}^{-1} \left\{\left(\hat{\beta}_k^i + \rho_k \widehat{\mathrm{B}}_k\right)\left(\widehat{DP}_k\right)^i\right\}\right] \tag{5.3b}$$

$$EB\left(\hat{\beta}_k^i\right) = \left[(1 + \rho_k)^{-1} \left(\hat{\beta}_k^i + \rho_k \widehat{\mathrm{B}}_k\right)\right] \tag{5.3c}$$

applying the limits as $\rho_k \rightarrow 0$, on (5.3c) above,

$$\lim_{\rho_k \to 0} EB\left(\hat{\beta}_k^i\right) = \lim_{\rho_k \to 0} \left[(1 + \rho_k)^{-1} \left(\hat{\beta}_k^i + \rho_k \widehat{\mathrm{B}}_k\right)\right] \tag{5.3d}$$

$$\lim_{\rho_k \to 0} EB\left(\hat{\beta}_k^i\right) = \left[(1 + 0)^{-1} \left(\hat{\beta}_k^i + 0\widehat{\mathrm{B}}_k\right)\right] \tag{5.3e}$$

$$\lim_{\rho_k \to 0} EB\left(\hat{\beta}_k^i\right) = \hat{\beta}_k^i \tag{5.3f}$$

The estimate at the righthand side of (5.3f) above, is the frequentist subjects specific coefficients estimates for a single $k^{th}$ orthogonal regressor of the $i^{th}$ unit of the panel data linear regression model. Now the vector of all the $K$ regression coefficients estimates, in this case, will become as under,

$$\widehat{M}_{EB}^i = EB\left(\hat{\beta}^i\right) = \begin{pmatrix} \hat{\beta}_1^i \\ \hat{\beta}_2^i \\ \vdots \\ \hat{\beta}_k^i \\ \vdots \\ \hat{\beta}_K^i \end{pmatrix}, \qquad \text{for } i = 1, 2, \dots, N. \tag{5.3g}$$

Hence, it has been seen that as the precision parameter $\rho_k$ tends to 0, against all the coefficient estimates in the empirical Bayes estimate, then the empirical Bayes estimate tends to reduce to the frequentist subject specific coefficients estimate, and hence it witnesses that the subject specific coefficients estimate is the special case of the empirical Bayes estimate.

Now below we show the corresponding subjects specific coefficients panel data linear regression model for the $i^{th}$ unit of the panel.

### 5.3.1. The Subjects Specific Coefficients Panel Data Linear Regression Model

Now on the basis of the vector of the regression coefficients estimates above (5.3g), (vector of the subject specific regression coefficients estimates), the corresponding model can be expressed as,

$$Y^i = X_1^i\beta_1^i + X_2^i\beta_2^i + \dots + X_k^i\beta_k^i + \dots + X_K^i\beta_K^i + \varepsilon^i, \tag{5.3.1}$$

or in matrix notation the above model becomes as,

$$Y^i = X^i\,\beta^i + \varepsilon^i. \tag{5.3.2}$$

Hence, the subjects specific coefficients panel data linear regression model is the special case of the Bayesian linear regression model.

Below we show the next case as the prior precision parameter tends to infinity and the resulted subjects common coefficients estimates.

## 5.4. The Subjects Common Coefficients Estimates

The case where the prior precision parameter $\rho_k$ tends to infinity, corresponding to all the coefficients estimates of the model, the empirical Bayes estimate reduces to the subject common coefficients estimates and the resulted estimates so formed exactly constant and common across all the units of the panel.

**The prior precision parameter $\rho_k$ tends to Infinity:** The second limiting case is, when the prior precision parameter $\rho_k$ tends to infinity, corresponding to all the coefficients estimates, i.e. $\rho_k \to \infty\ \forall\ k$, in the empirical Bayes estimate. In this case the prior precision becomes very precise and hence the prior mean or the subjects common characteristics of the panel remains only and the subject specific characteristics vanish from the empirical Bayes estimate. Thus, the estimates so remains are the frequentist subjects common coefficients estimates of the panel data linear regression models. This case indicates complete homogeneity among the units of the panel. This time all the coefficients vector will be identical and will be equal to the common prior mean.

In order, to address the case where the prior precision parameter $\rho_k$ tends to infinity. First, simplifying the formula of the empirical Bayes and then taking the limits, as $\rho_k$ approaches infinity, on both sides of the simplified version of equation (5.2p) above. Thus, before applying the limits, equation (5.2p) can also be expressed as follow,

$$EB(\hat{\beta}_k^i) = \left[\left\{(\widehat{DP}_k)^i + \hat{\rho}_k\,(\widehat{DP}_k)^i\right\}^{-1}\ \left\{(\widehat{DP}_k)^i\hat{\beta}_k^i + \hat{\rho}_k\,(\widehat{DP}_k)^i\widehat{\mathrm{B}}_k\right\}\right] \tag{5.4a}$$

$$EB(\hat{\beta}_k^i) = \left[\frac{1}{\hat{\rho}_k}\left\{\left(\frac{1}{\hat{\rho}_k}\right)(\widehat{DP}_k)^i + (\widehat{DP}_k)^i\right\}^{-1}\hat{\rho}_k\left\{\left(\frac{1}{\hat{\rho}_k}\right)(\widehat{DP}_k)^i\hat{\beta}_k^i + (\widehat{DP}_k)^i\widehat{\mathrm{B}}_k\right\}\right] \tag{5.4b}$$

$$EB(\hat{\beta}_k^i) = \left[\left\{\left(\frac{1}{\hat{\rho}_k}\right)(\widehat{DP}_k)^i + (\widehat{DP}_k)^i\right\}^{-1}\left\{\left(\frac{1}{\hat{\rho}_k}\right)(\widehat{DP}_k)^i\hat{\beta}_k^i + (\widehat{DP}_k)^i\widehat{\mathrm{B}}_k\right\}\right] \tag{5.4c}$$

applying the limits as $\hat{\rho}_k \ \rightarrow \infty$, on the above,

$$\lim_{\hat{\rho}_k \rightarrow \infty} EB(\hat{\beta}_k^i) = \lim_{\rho_k \rightarrow \infty}\left[\left\{\left(\frac{1}{\hat{\rho}_k}\right)(\widehat{DP}_k)^i + (\widehat{DP}_k)^i\right\}^{-1}\left\{\left(\frac{1}{\hat{\rho}_k}\right)(\widehat{DP}_k)^i\hat{\beta}_k^i + (\widehat{DP}_k)^i\widehat{\mathrm{B}}_k\right\}\right] \tag{5.4d}$$

$$\lim_{\hat{\rho}_k \rightarrow \infty} EB(\hat{\beta}_k^i) = \left[\left\{(\widehat{DP}_k)^i\right\}^{-1}\left\{(\widehat{DP}_k)^i\widehat{\mathrm{B}}_k\right\}\right] \tag{5.4e}$$

$$\lim_{\hat{\rho}_k \rightarrow \infty} EB(\hat{\beta}_k^i) = \ \widehat{\mathrm{B}}_k. \tag{5.4f}$$

Thus, the estimate at the righthand side of the above (5.4f), is the subjects common coefficient estimate for a single $k^{th}$ orthogonal regressor of the $i^{th}$ unit of the panel data linear regression model.

Now the vector of all the $K$ regression coefficients estimates, in this case, will then become as under,

$$\widehat{M}_{EB}^i = EB(\hat{\beta}^i) = \begin{pmatrix} \hat{B}_1 \\ \hat{B}_2 \\ \vdots \\ \hat{B}_k \\ \vdots \\ \hat{B}_K \end{pmatrix}, \qquad \text{for } i = 1, 2, \dots, N. \tag{5.4g}$$

Hence, it has been seen that as the precision parameter $\rho_k$ tends $\infty$ for $k$ = 1, 2, 3, … , $K$, then the empirical Bayes estimate tends to reduce to the frequentist subjects' common coefficients estimate and hence it witnesses that the subjects common coefficients estimate is the special case of the empirical Bayes estimate too.

Further, since, in above (5.4f) $\lim_{\rho_k \to \infty} EB(\hat{\beta}_k^i) = \widehat{\mathrm{B}}_k$ and from (4.15d) we can write the modified version for $k^{th}$ regressor case as under,

$$\widehat{\mathrm{B}}_k = \left(\sum_{i=1}^{N}(\widehat{DP}_k)^i\right)^{-1}\left[\sum_{i=1}^{N}(\widehat{DP}_k)^i\,\hat{\beta}_k^i\right]. \tag{5.4h}$$

Here, in the light of (5.4f), the above (5.4h) can also be written as under,

$$\widehat{M}_{EB}^i = EB(\hat{\beta}^i) = \left[\sum_{i=1}^{N}(\widehat{DP}_k)^i\right]^{-1}\left[\sum_{i=1}^{N}(\widehat{DP}_k)^i\,\hat{\beta}_k^i\right]. \tag{5.4i}$$

The above (5.4i) is the frequentist subjects common coefficients estimate of all panel data linear regression models. Here in (5.4i) we have seen that as $\rho_k \to \infty$, the empirical Bayes estimate reduces to an estimate which is the precision weighted arithmetic of the ordinary least squares coefficients estimates. Now below we show the corresponding subjects common coefficients panel data linear regression model for the $i^{th}$ unit of the panel.

## 5.4.1. The Subjects Common Coefficients Panel Data Linear Regression Model

Now on the basis of the vector of the regression coefficients estimates above (5.4g), (vector of the subjects common regression coefficients estimates), the corresponding panel data linear regression model can be expressed as,

$$Y^i = X_1^i B_1 + X_2^i B_2 + \ldots + X_k^i B_k + \ldots + X_K^i B_K + \varepsilon^i. \tag{5.4.1}$$

or in matrix notation the above model becomes as,

$$Y^i = X^i\,\beta + \varepsilon^i. \tag{5.4.2}$$

Hence, the subjects common coefficients panel data linear regression model is the special case of the Bayesian linear regression model.

### 5.4.2. A Special Case of Subjects Common Coefficients Estimate

As it has been shown in (5.4i) above, that as the prior precision parameter $\rho_k \to \infty$ against all the regression coefficients, the empirical Bayes estimate gives the precision weighted arithmetic mean of the ordinary least squares estimates.

One of the special cases of the estimate given in (5.4i) above, is that when the homoscedasticity or the constant variance across all the units of the panel is assumed, i.e. $(\hat{\sigma}_k^2)^i = (\hat{\sigma}_k^2)$ for all $i = 1, 2, \ldots, N$, then (5.4i) above takes the form as under,

putting $(\hat{\sigma}_k^2)^i = (\hat{\sigma}_k^2)$ in above,

$$\lim_{\rho_k \to \infty} EB(\hat{\beta}_k^i) = \left[\sum_{i=1}^{N} \frac{1}{(\hat{\sigma}_k^2)} (X_k^{i\prime} X_k^i)\right]^{-1} \left[\sum_{i=1}^{N} \frac{1}{(\hat{\sigma}_k^2)} (X_k^{i\prime} X_k^i)\, \hat{\beta}_k^i\right] \tag{5.4.2a}$$

$$\lim_{\rho_k \to \infty} EB(\hat{\beta}_k^i) = \left[\sum_{i=1}^{N} X_k^{i\prime} X_k^i\right]^{-1} (\hat{\sigma}_k^2)\, \frac{1}{(\hat{\sigma}_k^2)} \left[\sum_{i=1}^{N} (X_k^{i\prime} X_k^i)\, \hat{\beta}_k^i\right] \tag{5.4.2b}$$

$$\lim_{\rho_k \to \infty} EB(\hat{\beta}_k^i) = \left[\sum_{i=1}^{N} X_k^{i\prime} X_k^i\right]^{-1} \left[\sum_{i=1}^{N} (X_k^{i\prime} X_k^i)\, \hat{\beta}_k^i\right] \tag{5.4.2c}$$

The estimate given (5.4.2c) above is the aggregate estimate for all the units of the panel, this estimate exactly matches to the estimate given in, Zaman (1996), equation 10.2.

Below we show the third case when the prior precision parameter remains in-between zero and infinity and the corresponding random coefficients estimates.

## 5.5. The Random Coefficients Estimates

The case where the precision parameter $\rho_k$ is in-between zero and infinity, corresponding to all the coefficients estimates of the model, the empirical Bayes estimate reduces to the random coefficients estimates and the resulted estimates so formed differ from both of the above two cases.

**The prior precision parameter $\rho_k$ in-between Zero and Infinity:** The third case is, when the prior precision parameter $\rho_k$ remains in-between zero and infinity corresponding to all the coefficients estimates i.e. $0 < \rho_k < \infty$ for any $k$ out of all $K$ regressors, in the empirical Bayes estimate. This time the empirical

Bayes estimate neither solely depends upon the data information nor the prior information, rather, on the combination of these two. It means, that there is neither complete homogeneity nor complete heterogeneity, rather, some common and some random structure exists in the units of the panel.

The case where the prior precision parameter $\rho_k$ tends to $R$, where $R \in [0, \infty]$ any real number.

$$EB(\hat{\beta}_k^i) = \left[\left\{\left(\widehat{DP}_k\right)^i + \rho_k \left(\widehat{DP}_k\right)^i\right\}^{-1} \left\{\left(\widehat{DP}_k\right)^i \hat{\beta}_k^i + \rho_k \left(\widehat{DP}_k\right)^i \hat{\mathrm{B}}_k\right\}\right], \tag{5.5a}$$

taking the limits, as $\rho_k$ approaches to $R$, on both sides of (5.5a) above.

$$EB(\hat{\beta}_k^i) = \left[\left\{\left(\widehat{DP}_k\right)^i + R \left(\widehat{DP}_k\right)^i\right\}^{-1} \left\{\left(\widehat{DP}_k\right)^i \hat{\beta}_k^i + R \left(\widehat{DP}_k\right)^i \hat{\mathrm{B}}_k\right\}\right]. \tag{5.5b}$$

The coefficients estimate at the righthand side of (5.5b) above, is the random coefficient estimate for a single $k^{th}$ orthogonal regressor of the $i^{th}$ unit of the panel data linear regression model. Now the vector of all the $K$ regression coefficients estimates, in this case, will then become as under,

$$\widehat{M}_{EB}^i = EB(\hat{\beta}^i) = \begin{pmatrix} \hat{\tilde{B}}_1 \\ \hat{\tilde{B}}_2 \\ \vdots \\ \hat{\tilde{B}}_k \\ \vdots \\ \hat{\tilde{B}}_K \end{pmatrix}, \qquad \text{for} \qquad i = 1, 2, \ldots, N. \tag{5.5c}$$

Note: In the above, hat "∧" is placed upon tilde "~" in order, to represent the random fluctuations, and also to differentiate between the common coefficients and the random coefficients estimates.

Hence, it has been seen that as the prior precision parameter $\rho_k$ remains in-between 0 and ∞, for $k$ = 1, 2, 3, … , $K$, then the empirical Bayes estimate tends to reduce to the frequentist random coefficients estimate and hence it witnesses that the random coefficients estimate is the special case of the empirical Bayes estimate too.

Note: The empirical Bayes estimate itself is the random coefficients estimate.

This is the estimate of more importance; it has a huge potential to be fitted in many circumstances. This estimate encompasses so many possibilities existing in panel data models. In the above cases, all the prior precisions estimates were tending either to zero, infinity or was in-between zero and infinity, which yielded the corresponding estimates.

In the third case, the prior precision can take infinite values and has the potential to produce infinite frequentist random coefficients estimates of panel data linear regression models, and these models are what explored with the help of this thesis, for the very first time in the history of panel data literature.

If the prior precision parameter is close to zero, the estimated coefficient will resemble the unit specific coefficients estimate, on the other hand, if the prior precision parameter is close to infinity, then the estimated coefficient resembles the subjects common coefficients estimates of all the units of the panel.

Also, if the prior precision becomes moderately precise, i.e., neither too precise nor too imprecise, then the common characteristics as well as the specific characteristics of the units or the common prior mean and the subject specific individual units estimates, both become the parts of the empirical Bayes estimate and such estimates are known as the random coefficients estimates of the panel data models.

If starting the journey of the prior precision parameter from being zero and tending it towards infinity, one starts to have the subject specific coefficients estimates and then slowly and gradually modifying structure towards the subjects common coefficients estimates for all the units of the panel.

The estimates between these two extremes i.e. zero and infinity, are known as the random coefficients estimates of the panel data linear regression models as discussed above.

The speed of the convergence of estimates from subject specific coefficients estimates to subjects common coefficients estimates, actually depends upon the speed of the convergence of the prior precision parameters from zero to infinity.

Now below we show the corresponding random coefficients panel data linear regression model for the $i^{th}$ unit of the panel.

### 5.5.1. The Random Coefficients Panel Data Linear Regression Model

Now on the basis of the vector of the regression coefficients estimates above (5.5c), (vector of the random coefficients regression estimates), the corresponding model can be expressed as,

$$Y^i = X_1^i\tilde{B}_1 + X_2^i\tilde{B}_2 + \dots + X_k^i\tilde{B}_k + \dots + X_K^i\tilde{B}_K + \varepsilon^i. \tag{5.5.1}$$

or in matrix notation the above model becomes as,

$$Y^i = X^i\,\tilde{\beta} + \varepsilon^i. \tag{5.5.2}$$

Hence, the random coefficients panel data linear regression model is the special case of the Bayesian linear regression model.

Below we show the next case as the prior precision parameter $\rho_k$ tends to 0, for $k = 1$, and $\rho_k$ tends to $\infty$ for $k = 2, 3, \dots, K$.

## 5.6. The Fixed Effects Coefficients Estimates

The case where the prior precision parameter $\rho_k$ tends to zero, corresponding to the intercept estimate and tends to infinity, corresponding to the rest of the coefficients estimates of the model, the empirical Bayes estimate reduces to the fixed effects coefficients estimates and the resulted estimates so formed have exactly the fixed effects characteristics.

**The prior precision parameter $\rho_k$ tends to Zero for $k = 1$and tends to infinity for $k = 2, 3, \dots, K$:** The fourth limiting case is; when the prior precision parameter $\rho_k$ tends to zero corresponding to the first coefficients estimates i.e. $\rho_k \to 0$ for $k = 1$, and tends to infinity, corresponding to all the remaining $(K-1)$ coefficients estimates i.e. $\rho_k \to \infty$ for $k = 2, 3, \dots, K$, in the empirical Bayesian estimator, as a result, the empirical Bayes estimate reduces to the fixed effects coefficients estimates.

When the prior precision parameter $\rho_k$ for $k = 1$, tends to zero and $\rho_k$ tends to infinity, for $k = 2, 3, \dots, K$, then let us see what happens to the empirical Bayes estimate.

The case where the precision parameter $\rho_k$ tends to zero, the empirical Bayes estimate reduces to the subject specific coefficients estimates, as already shown in (5.3f) above, i.e.

$$\widehat{M}_{EB}^i = EB(\beta^i) = \hat{\beta}_k^i. \tag{5.6a}$$

Similarly, where the prior precision parameter $\rho_k$ tends to infinity, the empirical Bayes estimate reduces to the subjects common coefficients estimates, this is also been shown in (5.4) above, i.e.

$$\widehat{M}_{EB}^{i} = EB(\hat{\beta}^{i}) = \widehat{B}_{k}. \tag{5.6b}$$

Thus, if the first regressor, in a regression model, is a $(T \times 1)$ vector of ones and the remaining $(K-1)$ regressors fixed variables then the combination above yields the fixed effects coefficients estimates. So, the combination of the estimates at the righthand side of the (5.6a) and (5.6b) above, constitutes the fixed effects coefficient estimates for a single $k^{th}$ orthogonal regressor of the $i^{th}$ unit's panel data linear regression model. Now the vector of all the $K$ regression coefficients estimates, in this case, will then become as under,

$$\widehat{M}_{EB}^{i} = EB(\hat{\beta}^{i}) = \begin{pmatrix} \hat{\beta}_1^{i} \\ \widehat{B}_2 \\ \vdots \\ \widehat{B}_k \\ \vdots \\ \widehat{B}_K \end{pmatrix}, \qquad \text{for } k\ i = 1, 2, \dots, N. \tag{5.6c}$$

Hence, it has been seen that as the prior precision parameter $\rho_k$ tends to 0 for $k = 1$, and $\rho_k$ tends to $\infty$, for $k = 2, 3, \dots, K$, then the empirical Bayes estimate tends to reduce to the frequentist fixed effects coefficients estimate and hence it witnesses that the fixed effects coefficients estimate is the special case of the empirical Bayes estimate too. Now below we show the corresponding fixed effects coefficients model.

### 5.6.1. The Fixed Effects Coefficients Panel Data Linear Regression Model

Now on the basis of the vector of the regression coefficients estimates above (5.6c), (vector of the fixed effects regression coefficients estimates), the corresponding model can be expressed as,

$$Y^{i} = X_1^{i}\beta_1^{i} + X_2^{i}B_2 + \dots + X_k^{i}B_k + \dots + X_K^{i}B_K + \varepsilon^{i}. \tag{5.6.1}$$

or in matrix notation the above model becomes as,

$$Y^{i} = X^{i}\,\beta^{i} + \varepsilon^{i}. \tag{5.6.2}$$

Hence, the fixed effects coefficients panel data linear regression model is the special case of the Bayesian linear regression model.

Below we show the next case as the prior precision parameter $\rho_k$ remains in-between 0 and $\infty$ for $k$ = 1, and $\rho_k$ tends to $\infty$ for $k$ = 2, 3, … , $K$., and the resulted random effects coefficients estimates

## 5.7. The Random Effects Coefficients Estimates

The case where the prior precision parameter $\rho_k$ remains in-between zero and infinity for the intercept coefficients estimate and tends to infinity, corresponding to the rest of all the coefficients estimates of the model, the empirical Bayes estimate reduces to the random effects coefficients estimates and the resulted estimates so formed have exactly the random effects characteristics.

**The prior precision parameter $\rho_k$ Remains in-between Zero and Infinity for $k$ = 1and $\rho_k$ tends to infinity for $k$ = 2, 3, … , $K$:** The fifth limiting case is; when the prior precision parameter $\rho_k$ remains in-between zero and infinity i.e. $0 < \rho_k < \infty$ for $k = 1$, corresponding to the first coefficients estimates and tends to infinity, i.e. $\rho_k \to \infty$ for $k$ = 2, 3, … , $K$, corresponding to all the remaining $(K-1)$ coefficients estimates in the empirical Bayesian estimator. The empirical Bayes estimate reduces to the random effects coefficients estimates.

When the prior precision parameter $\rho_k$ remains in-between 0 and $\infty$ for $k = 1$, and $\rho_k$ tends to $\infty$ for $k$ = 2, 3, … , $K$, then let us see what happens to the empirical Bayes estimate.

The case where the prior precision parameter $\rho_k$ remains in-between Zero and infinity i.e. $R$, the empirical Bayes estimate becomes the combination of the random estimates as already shown in (5.5) above, i.e.

$$EB(\hat{\beta}_k^i) = \left[\left\{(\widehat{DP}_k)^i + R\,(\widehat{DP}_k)^i\right\}^{-1} \left\{(\widehat{DP}_k)^i \hat{\beta}_k^i + R\,(\widehat{DP}_k)^i \widehat{\mathrm{B}}_k\right\}\right]. \tag{5.7a}$$

Similarly, where the precision parameter $\rho_k$ tends to infinity, the Bayes estimate, as before, reduces to the subjects common coefficients estimates, this is also been shown in (5.4) above, i.e.

$$EB(\hat{\beta}_k^i) = \widehat{\mathrm{B}}_k. \tag{5.7b}$$

Thus, if the first regressor is a $(T \times 1)$ vector of ones and the remaining $(K-1)$ regressors are fixed or non-random variables, like in the case of the fixed effects, then the combination above yields the random effects coefficients estimates.

Thus, the combination of the estimates at the righthand side of the (5.7a) and (5.7b) above, constitutes the random effects coefficient estimates for the panel data linear regression model.

Now the vector of all the $K$ regression coefficients estimates, in this case, will then become as under,

$$\hat{M}^i_{EB} = EB(\hat{\beta}^i) = \begin{pmatrix} \hat{\tilde{\beta}}^i_1 \\ \hat{B}_2 \\ \vdots \\ \hat{B}_k \\ \vdots \\ \hat{B}_K \end{pmatrix}, \quad \text{for} \quad i = 1, 2, \dots, N. \tag{5.7c}$$

Hence, it has been seen that as the precision parameter $\rho_k$ tends to 0 for $k = 1$. and $\rho_k$ tends to $\infty$ for $k = 2, 3, \dots, K$, then the empirical Bayes estimate tends to reduce to the frequentist random effects coefficients estimate and hence it is evident that the random effects coefficients estimate is the special case of this Bayes estimate too.

Now below we show the corresponding random effects coefficients model.

### 5.7.1. The Random Effects Coefficients Panel Data Linear Regression Model

Now on the basis of the vector of the regression coefficients estimates above (5.7c), (vector of the random effects regression coefficients estimates), corresponding model can be expressed as,

$$Y^i = X^i_1 \tilde{\beta}^i_1 + X^i_2 B_2 + \dots + X^i_k B_k + \dots + X^i_K B_K + \varepsilon^i. \tag{5.7.1}$$

or in matrix notation the above model becomes as,

$$Y^i = X^i \beta + \varepsilon^i. \tag{5.7.2}$$

Hence, the random effects coefficients panel data linear regression model is the special case of the Bayesian linear regression model.

So, this way all the frequentist panel data model estimators and the corresponding panel data linear regression models have been derived from the empirical Bayes estimate.

## 5.8. Summary of the Chapter

In this chapter, we showed analytically that all the frequentist estimates of the panel data linear regression model and the corresponding panel data linear regression models can be derived from the Bayesian linear regression model and the empirical Bayes estimate of the panel data linear regression model, when the empirical Bayes estimate uses the precision weighted arithmetic mean of the ordinary least squares coefficients estimate as the estimate of the prior mean and the g-prior as the estimate of the prior variance.

We derived five different standard frequentist estimates, namely the subject specific estimate, the subjects common coefficients estimate the random coefficients estimate, the fixed effects coefficients estimate and the random effects coefficients estimate, and the corresponding panel data linear regression models by considering a $K$ orthogonal regressors panel data linear regression model, and a single $k^{th}$ regressor estimation technique at a time.

The frequentist estimates and the corresponding panel data linear regression models above were obtained from the empirical Bayes estimate by assigning five different values to the prior precision parameter. These values were either zero or infinity or a value in-between zero and infinity or the combination of these three values.

Now, the finding of the chapter are, that all the frequentist estimates and the corresponding standard panel data regression model are of a unanimous origin and they have been derived from this unanimous origin, hence they are the special cases of this unanimous origin. This unanimous origin is the empirical Bayes estimate. Hence this achieves our goal of deriving all the frequentist panel data linear regression models and the corresponding estimates from the Bayesian linear regression model and the empirical Bayes estimate, respectively.

Now in the next chapters we show numerically with help of numerical examples that all the frequentist estimates of the panel data linear regression model can be derived from the empirical Bayes estimate of the panel data linear regression model. This will more authenticate the claim of deriving all the frequentist panel data linear regression models and the corresponding estimates from the Bayesian linear regression model and the empirical Bayes estimate, respectively. So, let us see all the frequentist estimates of the panel data linear regression mode in the next chapters.

# Chapter 6

# The Numerical Derivation of the Frequentist Estimates

In the previous chapter, all the frequentist estimators and the corresponding panel data linear regression models have been derived from the empirical Bayes estimate and the corresponding Bayesian linear regression model. All the assumptions, conditions and technicalities have been made clear over there.

Now here, in this chapter, we numerically derive, all the frequentist estimates of the panel data linear regression models that have been derived analytically in the previous chapter, from the empirical Bayes estimate. In fact, as said before, the frequentist estimates of the panel data linear regression models are the special cases of the empirical Bayes estimate, for different values of the prior precision parameters and hence this chapter mainly notarize this fact.

The numerical derivation will be done here, with the help of a numerical example. We take the data and an example from, Gujrati (2003), Basic Econometrics, $4^{th}$ edition, the details of the book and data have already been discussed, in details, in chapter 4, we do not mention it again.

One more thing to note is that, for all of the derivations, the programming has been done with the help of "MATLAB 2017a". The programming is done according to single regressor estimation approach. Here, in this chapter only the theoretical descriptions and the results are given in different sections corresponding to different frequentist estimates. The results are shown with the help of different tables. For each type of the frequentist estimates, the prior precision parameters, with the help of which those particular frequentist estimates are derived, have been discussed.

## Derivation of the Frequentist Estimates Numerically

In this section we derive all the frequentist estimates of the panel data linear regression models from the corresponding empirical Bayes estimates. This chapter is the numeric counter part of the previous chapter therefore, all the assumptions are the same as to that of the previous chapter.

## 6.1. The Derivation of the Subject Specific Coefficients Estimates

Here, we describe and derive numerically the frequentist subject specific coefficients estimates. In the light of the panel data literature, if the intercepts terms, as well as, all the slopes coefficients, in $K$ regressors panel data linear regression models, are unit specific, across all the units of the panel, (i.e. $\beta_k^i \neq \beta_k^j$, for $i \neq j$ or at least all $\beta_k^i$ are not equal to $\beta_k$, for $k$ = 1, 2, … , $K$), then such coefficients correspond to the subject specific panel data linear regression model and the corresponding coefficients estimates are known as the subject specific coefficients estimates. The following is the subject specific coefficients panel data analytic model.

$$Y^i = X^i \beta^i + \varepsilon^i \qquad \text{for } i = 1, 2, \ldots, N. \qquad (6.1a)$$

$$\beta^i = \left(\beta_1^i,\ \beta_2^i,\ \ldots, \beta_k^i, \ldots, \beta_K^i\right)' \qquad (6.1b)$$

The distribution of $\beta_k^i$ is normal with mean, $\mu_{\beta_k}^i$ and variance, $\left[\sigma_{\beta_k}^2\right]^i$, both subject specific across the panel,

$$\beta_k^i \sim N\left(\mu_{\beta_k}^i, \left[\sigma_{\beta_k}^2\right]^i\right) \qquad \text{for } k = 1, 2, \ldots, K \qquad (6.1c)$$

The description of the above model and the associated terms is as given in chapter 3.

Further, in the previous chapter, as we have seen analytically that, the empirical Bayes estimate of panel data linear regression model, reduces to the subject specific coefficients estimates, when the prior precision parameters, corresponding to each regression coefficient (in the empirical Bayes estimate), is set to zero, i.e. $\rho_k \to 0$, for $k$ = 1, 2, … , $K$. So, here making this as the base for the numerical computation i.e. MATLAB programming, we take the prior precision parameters $\rho_k$ equal to zero, for all the regression coefficients estimates, in the empirical Bayes estimate. Thus, Table (6.1a), shows the estimates for the prior precision parameters of the empirical Bayes estimate corresponding to each regression coefficient estimate. Since there are three regressors in each model of a unit in the panel, therefore, the first, second and third prior precision parameters, are lying in the first, second and third row of Table (6.1a), and correspond to the intercept term, the second and the third regression coefficients estimates of the model, respectively.

Now, when we put the prior precision parameters as zero corresponding to each regression coefficient estimate in the MATALB Programming for the empirical Bayes, discussed above, it gave the subject specific coefficients estimates corresponding to each regressor, as given in Table (6.1b). Further, from Table (6.1b), it can be seen that the intercepts estimates are different across the units of the panel and are subject specific, similarly, the coefficients estimates of the second and third regressor of any one unit are different from the corresponding coefficients estimates of any other unit across the units of the panel, and these should be the way they are, for being the subject specific coefficients estimates. Thus, the coefficients estimates of Table (6.1b) are the subjects specific coefficients estimates.

Now, it is evident that when all the prior precision parameters corresponding to each regression coefficient estimate, are set to zero, the empirical Bayes estimate reduces to the subject specific coefficients estimates. Hence, it is concluded that as the frequentist subject specific coefficients estimates are derived from the empirical Bayes estimate, therefore, they are the special cases of the empirical Bayes estimate. The subject specific coefficients model is also known as the subject specific fixed coefficients model, the individual units model, sometimes also the ordinary least squares linear regression model etc.

Below, we show the tables of the prior precision parameters and the derived frequentist subject specific coefficients estimates. Table (6.1a) contains the prior precision parameters and Table (6.1b) contains the corresponding subject specific coefficients estimates.

Table 6.1a: The Prior Precision Parameters for the Subject Specific Coefficients Estimates

| Prior Precision Parameters for the Subject Specific Coefficients Estimates | |
|---|---|
| Prior Coefficients Estimates | Prior Precision Parameter |
| Intercept | 0 |
| Second Regressor | 0 |
| Third Regressor | 0 |

Table 6.1b: The Subject Specific Coefficients Estimates

| Variable | The Subject Specific Coefficients Estimates | | | |
|---|---|---|---|---|
| | Unit: 1 | Unit: 2 | Unit: 3 | Unit: 4 |
| Intercept | -9.9563 | -149.4667 | -50.078 | -0.5804 |
| Second Regressor | 0.0265 | 0.1192 | 0.1710 | 0.0531 |
| Third Regressor | 0.1517 | 0.3715 | 0.4080 | 0.0917 |

## 6.2. Derivation of the Subjects Common Coefficients Estimates

Here, we describe and derive numerically the frequentist subjects common coefficients estimates. In the light of the panel data literature, if the intercept term, as well as, all the slopes coefficients, of $K$ regressors panel data linear regression model, are common for all the units of the panel, (i.e. $\beta_k^i = \beta_k^j = \beta_k$, for $k$ = 1, 2, ... , $K$, common for all the units of the panel), then such coefficients correspond to the subjects common panel data linear regression model and the corresponding coefficients estimates are known as the subjects common coefficients estimates. The following is the subject common coefficients panel data analytic model.

$$Y^i = X^i\beta + \varepsilon^i \qquad \text{for } i = 1, 2, \dots, N. \tag{6.2a}$$

$$\beta = (\beta_1,\ \beta_2,\ \dots, \beta_k, \dots, \beta_K)'. \tag{6.2b}$$

The distribution of $\beta_k$ is normal with mean, $\mu_{\beta_k}$ and variance, $[\sigma^2_{\beta_k}]$, both subject common across the panel,

$$\beta_k \sim N\left(\mu_{\beta_k}, [\sigma^2_{\beta_k}] = 0\right) \qquad \text{for } k = 1, 2, \dots, K. \tag{6.2c}$$

Further, in the previous chapter as we have seen analytically that, the empirical Bayes estimate of panel data linear regression model, reduces to the subjects common coefficients estimates, when the prior precision parameters, corresponding to each regression coefficient estimate (in the empirical Bayes estimate) is set to infinity, i.e. $\rho_k \to \infty$, for, $k$ = 1, 2, ... , $K$. So here making this as the base for the numerical computation i.e. MATLAB programming, we take the prior precision parameters $\rho_k$ equal to infinity, ($\rho_k$ equal to $(10)^{10}$ = 10000000000, as the approximate value of infinity for, $k$ = 1, 2, ... , $K$), for all the regression coefficients estimates, in the empirical Bayes estimate. Thus, Table (6.2a), shows the prior precision parameters corresponding to each regression coefficients estimate. The first, second and third prior precision parameters Table in (6.2a), correspond to the intercept term, the second and third regression coefficients estimates, respectively.

Now, when we put the prior precision parameters as infinite in the formula of empirical Bayes [MATALB Programming], it gave the subjects common coefficients estimates, as given in Table (6.2b), for all the units of the panel. Further, from Table (6.2b), it can be seen that the coefficients estimates of the intercept term are identical for all the units across the panel, similarly, the coefficients estimates of the second regressor are also identical for all the four units across the panel, in the same way, the coefficients estimates of the third regressor are too identical for all the units across the panel, and these should be the way they are, for being the subjects common coefficients estimates. In other words, all the units of the panel have a common intercept estimate, similarly, all the units have a common second regression coefficient estimate and the third regression coefficient estimate is also common for all of the units of the panel. Thus, the coefficients estimates of Table (6.2b) are the subjects common coefficients estimates. Now, it is evident that when all the prior precision parameters corresponding to all regression coefficients estimates, set to infinity, the empirical Bayes estimate reduces to the subjects common coefficients estimates. Hence, it is concluded that the frequentist subjects common coefficients estimates are derived from the empirical Bayes estimate and they are the special cases of the empirical Bayes estimate. The subjects common coefficients model is also known as the common coefficients model, the constant coefficient model, the pooled coefficients model or the pooled ordinary least squares regression model.

Below we show the tables of the prior precision parameters and the derived frequentist subject common coefficients estimates. Table (6.2a) contains the prior precision parameters and Table (6.2b) contains the corresponding subject common coefficients estimates.

Table 6.2a: The Prior Precision Parameters for the Subjects Common Coefficients Estimates

| Prior Precision Parameters for the Subjects Common Coefficients Estimates | |
|---|---|
| Prior Coefficients Estimates | Prior Precision Parameter |
| Intercept | $(10)^{10} = 10000000000$ |
| Second Regressor | $(10)^{10} = 10000000000$ |
| Third Regressor | $(10)^{10} = 10000000000$ |

Table 6.2b: The Subjects Common Coefficients Estimates

| | The Subjects Common Coefficients Estimates | | | |
|---|---|---|---|---|
| Variables | Unit: 1 | Unit: 2 | Unit: 3 | Unit: 4 |
| Intercept | -2.0839 | -2.0839 | -2.0839 | -2.0839 |
| Second Regressor | 0.0542 | 0.0542 | 0.0542 | 0.0542 |
| Third Regressor | 0.2105 | 0.2105 | 0.2105 | 0.2105 |

## 6.3. Derivation of the Fixed Effects Coefficients Estimates

Here, we describe and derive numerically the frequentist fixed effects coefficients estimates. In the light of the panel data literature, if the intercept term, in $K$ regressors panel data linear regression model, is unit specific, (i.e. $\beta_k^i \neq \beta_k^j$, for $i \neq j$ and $k = 1$), but all the slopes coefficients, are common for all units of the panel, (i.e. $\beta_k^i = \beta_k^j = \beta_k$, for $i \neq j$ and $k = 2, 3, \ldots, K$), then such coefficients correspond to the fixed effects panel data linear regression model and the corresponding coefficients estimates are known as the fixed effects coefficients estimates. The following is the fixed effects coefficients panel data analytic model.

$$Y^i = X^i\beta^i + \varepsilon^i \qquad \text{for } i = 1, 2, \ldots, N. \tag{6.3a}$$

$$\beta^i = \left(\beta_1^i,\ \beta_2,\ \ldots, \beta_k, \ldots, \beta_K\right)'. \tag{6.3b}$$

The distribution of $\beta_k^i$ for $k = 1$, is normal with mean, $\mu_{\beta_k}^i$ and variance, $\left[\sigma_{\beta_k}^2\right]^i$, both subject specific,

$$\beta_k^i \sim N\left(\mu_{\beta_k}^i, \left[\sigma_{\beta_k}^2\right]^i\right) \qquad \text{for } k = 1 \tag{6.3c}$$

Further, the distribution of $\beta_k$ for $k = 2, 3, \ldots, K$, is normal with mean, $\mu_{\beta_k}$ and variance, $\left[\sigma_{\beta_k}^2\right]$, both subject common across the panel,

$$\beta_k \sim N\left(\mu_{\beta_k}, \left[\sigma_{\beta_k}^2\right] = 0\right) \qquad \text{for } k = 2, 3, \ldots, K. \tag{6.3d}$$

Furthermore, In the previous chapter, as we have seen analytically, the empirical Bayes estimate of panel data linear regression model, reduces to the fixed effects coefficients estimates, when the prior precision parameter, corresponding to the estimate of the intercept is set to zero, i.e. $\rho_k \to 0$, for $k = 1$, and corresponding to all the slopes coefficients is set to infinity, i.e. $\rho_k \to \infty$, for $k = 2, 3, \ldots, K$. So, here making this as the base for the numerical computation i.e. MATLAB programming, we take the prior precision parameter $\rho_k$, equal to zero, for $k = 1$ and equal to infinity, for $k = 2, 3$. Thus, Table (6.3a), shows the prior precision parameters of the first, second and third regression coefficients estimates.

Now, by putting the prior precision parameters, from Table (6.3.a), in the formula of empirical Bayes, [MATALB Programming], it gave the fixed effects coefficients estimates as given in Table (6.3b). Further, from Table (6.3b), it can be seen that the coefficients estimates of the intercept term are not identical or common for all the units across the panel, but the coefficients estimates of the second and third regressor are identical to their respective counterparts across all the units of the panel. Thus, all the cells' entities in the first row of the Table (6.3b) are subject specific and exactly match to their counterparts in the subject specific coefficients estimates of Table (6.1b).

Further, all the cells' entities in the second and third row of the Table (6.3b) are identical, and exactly match to their respective counterparts in the subject common coefficients estimates of Table (6.2b) and these should be the way they are for being the fixed effects coefficients estimates. Thus, the coefficients estimates of Table (6.3b) are the fixed effects coefficients estimates.

Now, it is evident that when the prior precision parameters corresponding to the intercept is set to zero and corresponding to all the slopes set to infinity, the empirical Bayes estimate reduces to the fixed effects coefficients estimates. Hence, it is concluded that the frequentist fixed effects coefficients estimates are the special cases of the empirical Bayes estimate.

Below we show the tables of the prior precision parameters and the derived frequentist fixed effects coefficients estimates. Table (6.3a) contains the prior precision parameters and Table (6.3b) contains the corresponding fixed effects coefficients estimates.

Table 6.3a: The Prior Precision Parameters for the Fixed Effects Coefficients Estimates

| Prior Precision Parameters for the Fixed Effects Coefficients Estimates | |
|---|---|
| Prior Coefficients Estimates | Prior Precision Parameter |
| Intercept | 0 |
| Second Regressor | $(10)^{10} = 10000000000$ |
| Third Regressor | $(10)^{10} = 10000000000$ |

Table 6.3b: The Fixed Effects Coefficients Estimates

| Variables | The Fixed Effects Coefficients Estimates | | | |
|---|---|---|---|---|
| | Unit: 1 | Unit: 2 | Unit: 3 | Unit: 4 |
| Intercept | -9.9563 | -149.4667 | -50.078 | -0.5804 |
| Second Regressor | 0.0542 | 0.0542 | 0.0542 | 0.0542 |
| Third Regressor | 0.2105 | 0.2105 | 0.2105 | 0.2105 |

## 6.4. Derivation of the Random Effects Coefficients Estimates

Here, we describe and derive numerically the random effects coefficients estimates. In the light of the panel data literature, if the intercept term is not unit specific, (i.e. $\beta_k^i = \beta_k + e_k^i$, for $k = 1$, and where $e^i$ are random fluctuations), but is randomly fluctuated from a common mean with constant variance, and all the slope coefficients, are common for all units of the panel, (i.e. $\beta_k^i = \beta_k^j = \beta_k$, for $i \neq j$ and $k = 2, 3, \dots, K$), then such coefficients correspond to the random effects panel data linear regression model and the corresponding estimates are known as the random effects coefficients estimates. The following is the random effects coefficients panel data analytic model.

$$Y^i = X^i \beta^i + \varepsilon^i, \qquad \text{for } i = 1, 2, \dots, N. \tag{6.4a}$$

$$\beta^i = \left(\beta_1^i,\ \beta_2,\ \dots, \beta_k, \dots, \beta_K\right)', \tag{6.4b}$$

$$\beta^i = \left(\left[\beta_1 + e_1^i\right],\ \beta_2,\ \dots, \beta_k, \dots, \beta_K\right)'. \tag{6.4c}$$

The distribution of $\beta_k^i$, for $k = 1$, is normal with mean, $\mu_{\beta_k}$ and variance, $\left[\sigma_{\beta_k}^2\right] = \left[\sigma_{e_k}^2\right]$,

$$\beta_k^i \sim N\left(\mu_{\beta_k}, \left[\sigma_{\beta_k}^2\right] \neq 0\right) \qquad \text{for } k = 1, \tag{6.4d}$$

and

$$e_k^i \sim N\left(0, \left[\sigma_{e_k}^2\right]\right) \qquad \text{for } k = 1, \tag{6.4e}$$

then,

$$\beta_k^i \sim N\left(\mu_{\beta_k}, \left[\sigma_{e_k}^2\right]\right) \qquad \text{for } k = 1. \tag{6.4f}$$

Further, the distribution of $\beta_k$, for $k = 2, 3, \dots, K$, is normal with mean, $\mu_{\beta_k}$ and variance, $\left[\sigma_{\beta_k}^2\right]$, both subject common across the panel,

$$\beta_k \sim N\left(\mu_{\beta_k}, \left[\sigma_{\beta_k}^2\right] = 0\right) \qquad \text{for } k = 2, 3, \dots, K. \tag{6.4g}$$

Furthermore, in the previous chapter, as we have seen analytically that, the empirical Bayes estimate, reduces to the random effects coefficients estimates, when the prior precision parameter, corresponding to the intercept, is set in-between zero and infinity, i.e. $0 < \rho_k < \infty$ for $k = 1$, and corresponding to the rest of all the slopes, is set to infinity, i.e. $\rho_k \to \infty$, for $k = 2, 3, \dots, K$, in the formula of empirical Bayes. So, here making this as the base for the numerical computation i.e. MATLAB programming, we take the prior precision parameter; $\rho_k$ as in-between zero and infinity for $k$ = 1, and $\rho_k$ equal to infinity for $k$ = 2, 3, (for $\rho_k$ we take $(10)^1$ = 10, as the approximate value in-between zero and infinity). Thus, Table (6.4a), shows the estimates for the prior precision parameters of the empirical Bayes estimate corresponding to each regression coefficients estimate.

Now, by putting the prior precision parameter accordingly as discussed, in the formula of empirical Bayes, [MATALB Programming], it gave the random effects coefficients estimates as given in the Table (6.4b). From Table (6.4b), it can be seen that the coefficients estimates of the intercept terms are randomly fluctuated, but the coefficients estimates of the second and third regressor exactly match to their fixed effects as well as the common coefficients counterparts. Thus, the coefficients estimates of Table (6.4b) are the random effects coefficients estimates. Now, it is evident that the frequentist random effects coefficients estimates are derived from the empirical Bayes estimate. Hence, it is concluded that the frequentist random effects estimates are the special cases of the empirical Bayes estimate. The random effects coefficients model is also known as the random intercept model.

Below, Table (6.4a) contains the prior precision parameters and Table (6.4b) contains the corresponding random effects coefficients estimates.

Table 6.4a: The Prior Precision Parameters for the Random Effects Coefficients Estimates

| The Prior Precision Parameters for the Random Effects Coefficients Estimates | |
|---|---|
| Prior Coefficients Estimates | Prior Precision Parameter |
| Intercept | $(10)^1$ = 10 |
| Second Regressor | $(10)^{10}$ = 10000000000 |
| Third Regressor | $(10)^{10}$ = 10000000000 |

Table 6.4b: The Random Effects Coefficients Estimates

| Variables | The Random Effects Coefficients Estimates | | | |
|---|---|---|---|---|
| | Unit: 1 | Unit: 2 | Unit: 3 | Unit: 4 |
| Intercept | -2.7996 | -15.4823 | -6.4470 | -1.9472 |
| Second Regressor | 0.0542 | 0.0542 | 0.0542 | 0.0542 |
| Third Regressor | 0.2105 | 0.2105 | 0.2105 | 0.2105 |

## 6.5. Derivation of the Random Coefficients Estimates

Here, we describe and derive numerically the random coefficients estimates. In the light of the panel data literature, if the intercept term as well as all the slopes coefficients, in the $K$ regressors panel data linear regression model, are neither unit specific nor common, (i.e., $\beta_k^i = \beta_k + e_k^i$, for $k$ = 1, 2, 3, ... , $K$, and where the $e^i$ are random fluctuations), but are randomly fluctuated around a common mean with constant variance, then such coefficients correspond to the random coefficients panel data linear regression model and the corresponding estimates are known as the random coefficients estimates. The following is the random coefficients panel data analytic model.

$$Y^i = X^i\beta^i + \varepsilon^i, \qquad \text{for } i = 1, 2, \ldots, N. \qquad (6.5a)$$

$$\beta^i = \left(\beta_1^i,\ \beta_2,\ \ldots, \beta_k, \ldots, \beta_K\right)', \qquad (6.5b)$$

$$\beta^i = \left(\left[\beta_1 + e_1^i\right], \left[\beta_2 + e_2^i\right], \ldots, \left[\beta_k + e_k^i\right], \ldots, \left[\beta_K + e_K^i\right]\right)'. \qquad (6.5c)$$

The distribution of $\beta_k^i$, for $k$ = 1, 2, 3, ... , $K$, is normal with mean, $\mu_{\beta_k}$ and variance, $\left[\sigma_{\beta_k}^2\right] = \left[\sigma_{e_k}^2\right]$, both subject common across all the units of the panel,

$$e_k^i \sim N\left(0, \left[\sigma_{e_k}^2\right]\right), \qquad \text{for } k = 1, 2, 3, \ldots, K, \qquad (6.5e)$$

then.

$$\beta_k^i \sim N\left(\mu_{\beta_k}, \left[\sigma_{e_k}^2\right]\right), \quad \text{for } k = 1, 2, 3, \ldots, K. \qquad (6.5d)$$

Further, in the previous chapter as we have seen analytically that, the random coefficients estimates have been deduced from the empirical Bayes estimate, when the prior precision parameter, corresponding to all the coefficients estimates, are set in-between zero and infinity, i.e., $0 < \rho_k < \infty$, for all $k$ = 1, 2, 3, ... , $K$, in the formula of empirical Bayes. So, here making this as the base, we take the prior precision parameter $\rho_k$ as in-between zero and infinity. Here again as before, we put the prior precision parameters $\rho_k$ equal to $(10)^1 = 10$, as the approximate value, in-between zero and infinity. Thus, Table (6.5a), shows the prior precision parameters of the empirical Bayes estimate corresponding to each regression coefficient estimate in the panel data linear regression model. In simple words, in this case, all the regressors share a common value as the prior precision parameters and this time this common value is the value in-between zero and infinity.

Now, when we put the prior precision parameters, discussed above, corresponding to all the coefficients estimates, in the formula of empirical Bayes [MATALB Programming], it gave the random coefficients estimates as given in Table (6.5b). Further, from Table (6.5b), it can be seen that all the coefficients estimates in the table are not identical, rather, seem randomly fluctuated. Thus, all the cells' entities in the first, second and third row of the table are different within the row from one another, but the difference is due to random fluctuations and this should be the way they are for being the random coefficients estimates. Thus, now we can say that, the coefficients estimates of Table (6.5b) are the random coefficients estimates. Now, it is evident that when the prior precision parameters corresponding to all the coefficients estimates are set in-between zero and infinity, the empirical Bayes estimate reduces to the random coefficients estimates. The random coefficients model is also called the varying parameters model.

Below we show the tables of the prior precision parameters and the derived frequentist random coefficients estimates. Table (6.5a) contains the prior precision parameters and Table (6.5b) contains the corresponding random coefficients estimates.

Table 6.5a: The Prior Precision Parameters for the Random Coefficients Estimates

| Prior Precision Parameters for the Random Coefficients Estimates | |
|---|---|
| Prior Coefficients Estimates | Prior Precision Parameter |
| Intercept | $(10)^1 = 10$ |
| Second Regressor | $(10)^1 = 10$ |
| Third Regressor | $(10)^1 = 10$ |

Table 6.5b: The Random Coefficients Estimates

| Variables | The Random Coefficients Estimates | | | |
|---|---|---|---|---|
| | Unit: 1 | Unit: 2 | Unit: 3 | Unit: 4 |
| Intercept | -2.7996 | -15.4823 | -6.4470 | -1.9472 |
| Second Regressor | 0.0517 | 0.0602 | 0.0649 | 0.0541 |
| Third Regressor | 0.2025 | 0.2251 | 0.2285 | 0.1997 |

This completes the goal of this thesis as the all the standard panel data models estimates have been derived numerically from the empirical Bayes estimate.

## Summary of the Chapter

In the above five sections, we have derived the five standard frequentist panel data linear regression models coefficients estimates from the empirical Bayes estimate and then the corresponding models. The main purpose in all of the above five sections was, to derive and justify that the frequentist estimates for panel data linear regression models can be derived from this single empirical Bayes estimate. For this derivational purpose, we took different numerical prior precision parameters corresponding to different regression coefficient estimates, in the formula of the empirical Bayes. We either took the prior precision parameters equal to zero or infinity or any other value in-between zero and infinity. Only with these special numerical prior precision parameters, we could derive all the desired frequentist coefficients estimates of panel data linear regression models from the empirical Bayes estimate.

The choice of zeros as the prior precision parameters was crystal clear, and also the choice of the approximate value for the infinity of the prior precision parameter was clear to somehow, but the choices of the values in-between zero and infinity of the prior precision parameters was infinite and hence it was not as clear as the previous two cases, in terms, that what value for the prior precision parameters to be assumed, as in-between zero and infinity.

We arbitrarily took $(10)^1 = 10$, as the in-between zero and infinity, numerical prior precision parameters and derived the required random estimates, namely the random effects coefficients estimates and the random coefficients estimates. In all of the above sections, our aim was to derive these particular frequentist estimates from the empirical Bayes estimate. We have seen that, in fact, all the standard frequentist panel data models estimates have been derived from the empirical Bayes estimate.

After a huge research and attempts (attempted many different prior means and variances estimates) finally we reached to the goal and found that, this was only possible with, when the empirical Bayes estimate use the precision weighted arithmetic mean of the ordinary least squares coefficients estimates as the estimate of the prior mean and the Zellner's g-prior as the prior variance and the corresponding precision and then the prior precision parameter in the formula take different values i.e. either zero or infinity or values in-between these two limits. The precision is actually the measure of the uncertainty of an estimate and hence the empirical

Bayes estimate incorporates this uncertainty. Therefore, in the above numerical derivations, we were restricted only to use these predetermined values for the prior precision parameters. Below we present all of the above different frequentist estimates derived from the empirical Bayes estimate altogether in a single table for observing the special characteristic of each of the frequentist estimates.

## 6.6. Comparison of All the Frequentist Estimates

The table below shows all the derived frequentist estimates from the empirical Bayes estimate, in order, to observe the characteristics of each of the frequentist estimates. In this table all the frequentist coefficients estimates have been stacked one upon the other to constitute a single table.

Table 6.6: Comparison of All the Frequentist Coefficients Estimates from the Empirical Bayes Estimates

| **Table of Coefficients Estimates of the Standard Classical Panel Data Model Derived from the Empirical Bayes Estimates** | | | | | |
|---|---|---|---|---|---|
| **Variables** | **Serial No** | **Units** | **Intercept** | **Second Regressor** | **Third Regressor** |
| **Subject Specific Coefficients Estimates** | 1 | Unit: 1 | -9.9563 | 0.0266 | 0.1517 |
| | 2 | Unit: 2 | -149.4667 | 0.1192 | 0.3715 |
| | 3 | Unit: 3 | -50.0780 | 0.1714 | 0.4087 |
| | 4 | Unit: 4 | -0.5804 | 0.0531 | 0.0917 |
| **Common Coefficients Estimates** | 5 | Unit: 1 | -2.0839 | 0.0542 | 0.2105 |
| | 6 | Unit: 2 | -2.0839 | 0.0542 | 0.2105 |
| | 7 | Unit: 3 | -2.0839 | 0.0542 | 0.2105 |
| | 8 | Unit: 4 | -2.0839 | 0.0542 | 0.2105 |
| **Fixed Effects Coefficients Estimates** | 9 | Unit: 1 | -9.9563 | 0.0542 | 0.2105 |
| | 10 | Unit: 2 | -149.4667 | 0.0542 | 0.2105 |
| | 11 | Unit: 3 | -50.0780 | 0.0542 | 0.2105 |
| | 12 | Unit: 4 | -0.5804 | 0.0542 | 0.2105 |
| **Random Effects Coefficients Estimates** | 13 | Unit: 1 | -2.7996 | 0.0542 | 0.2105 |
| | 14 | Unit: 2 | -15.4823 | 0.0542 | 0.2105 |
| | 15 | Unit: 3 | -6.4470 | 0.0542 | 0.2105 |
| | 16 | Unit: 4 | -1.9472 | 0.0542 | 0.2105 |
| **Random Coefficients Estimates** | 17 | Unit: 1 | -2.7996 | 0.0517 | 0.2052 |
| | 18 | Unit: 2 | -15.4823 | 0.0602 | 0.2251 |
| | 19 | Unit: 3 | -6.4470 | 0.0649 | 0.2285 |
| | 20 | Unit: 4 | -1.9472 | 0.0541 | 0.1997 |

Now in Table 6.6 above, the empirical Bayes estimate, for zero as the prior precisions parameters corresponding to all the three coefficients, gave the subject specific coefficients estimates. Thus, the first 4 ×

3 portion of the coefficients estimates in the table, i.e. row no.1 to row no. 4 and column named intercept to column named third regressor, contains the subject specific coefficients estimates, for all four units of the panel.

The first column and the corresponding rows, as mentioned above, in this specified portion of Table 6.6, contains the estimates of the intercepts of all the four units of the panel, here, it can easily be seen that all the estimates of the intercepts are different from each other.

Similarly, the second column, in this specified portion of the table, contains the estimates of the second regression coefficient of all the four units of the panel, here, it can also be seen easily, that all the estimates of the second regression coefficients are different from each other.

In the same way, the third column, in the said portion of the table, contains the estimates of the third regression coefficients of all the four units of the panel, here, it can be seen again easily, that all the estimates of the third regression coefficients, are different from each other, and this way all the coefficients estimates discussed here, exactly match in characteristics to the subject specific coefficients estimates.

Next, the empirical Bayes estimate, for infinity as the prior precisions parameters, corresponding to all the three coefficients, gave the subjects common coefficients estimates. Thus, the third, $4 \times 3$ portion of the coefficients estimates in the table, i.e. row no.9 to row no. 12 and column named intercept to column named third regressor, contains the subjects common coefficients estimates, for all the four units of the panel.

The first column and the corresponding rows, as mentioned above, in this specified portion of Table 6.6, contains the estimates of the intercepts of all the four units of the panel, here, it can easily be seen that all the estimates of the intercepts are common, constant and identical to each other.

Similarly, the second column, in this second specified portion of the table, contains the estimates of the second regression coefficients of all the four units of the panel, here too, it can be seen easily, that all the estimates of the second regression coefficients are identical to each other.

In the same way, the third column in this said portion of the table, contains the estimates of the third regression coefficients of all the four units of the panel, here again, it can easily be seen, that all the estimates of the third regression coefficient are identical to each other and this way all the coefficients estimates discussed here, exactly match in characteristics to the subjects common coefficients estimates.

Further, the empirical Bayes estimate, for zero as the prior precisions parameters corresponding to the intercepts coefficients, and infinity, corresponding to both the slopes coefficients, gave the fixed effects coefficients estimates. Thus, the third $4 \times 3$ portion of the coefficients estimates in the table, i.e., row no.9 to row no. 12 and column named intercept to column named third regressor, contains the fixed effects coefficients estimates, for all the four units of the panel.

The first column and the corresponding rows, as mentioned above, in this specified portion of the table, contains the estimates of the intercepts for all the four units of the panel, here, it can easily be seen that all the estimates of the intercepts are different from each other, as in this case they are subject specific like the intercepts in the case of the subject specific coefficients estimates above, and therefore, they exactly match to them as well.

Similarly, the second column, in this third specified portion of the table, contains the estimates of the second regression coefficients of all the four units of the panel, here too like in the case of the subjects common coefficients estimates, it can be seen easily, that all the estimates of the second regression coefficients are identical to each other, and this should be in the case of the fixed effects coefficients estimates, as, in the fixed effects coefficients estimates, the slopes coefficients must exactly match to their counterparts slopes coefficients in the subjects common coefficients estimates.

In the same way, the third column in this said portion of the table, contains the estimates of the third regression coefficients of all the four units of the panel, here again, it can easily be seen, that all the estimates of the third regression coefficients are identical to each other, here too like in the case of the subjects common coefficients estimates, it can be seen easily, that all the estimates of the third regression coefficients are identical to each other. Here too the third regression coefficients estimates must exactly match to the slopes coefficients of tis counterpart in the subjects common coefficients estimates, and this way all the coefficients estimates discussed here, exactly match in characteristics to the fixed effects coefficients estimates.

Furthermore, the empirical Bayes estimate, for a value in-between zero and infinity as the prior precisions parameters corresponding to the intercepts coefficients, and infinity, corresponding to both the slopes coefficients, gave the random effects coefficients estimates. Thus, the fourth $4 \times 3$ portion of the coefficients estimates in the table, i.e. row no.13 to row no. 16 and column named intercept to column named third regressor, contains the random effects coefficients estimates, for all the four units of the panel.

The first column and the corresponding rows, as mentioned above, in this specified portion of the table, contains the estimates of the intercepts of all the four units of the panel, here, it can easily be seen that all the estimates of the intercepts are fluctuating from each other, but differently from either the case of the subject specific coefficients estimates or the fixed effects estimates, above. Theoretically they should be the way they are in the case of the random effects coefficients estimates.

Further the nature and characteristics of the slopes coefficients in the random effects coefficients estimates should exactly equal to either of the subjects common coefficients estimates or the fixed effects coefficients estimates, and here they are like them, and they have already been explained. Therefore, here we do not feel any need of them again and do not discuss them again. Thus, all the coefficients estimates discussed here, exactly match in characteristics to the random effects coefficients estimates.

Finally, the empirical Bayes estimate, for values in-between zero and infinity as the prior precisions parameters corresponding to all the coefficients estimates, gave the random coefficients estimates. Thus, the fifth $4 \times 3$ portion of the coefficients estimates in the table, i.e. row no.17 to row no. 20 and column named intercept to column named third regressor, contains the fixed effects coefficients estimates, for all the four units of the panel.

The first column and the corresponding rows as mentioned above, in this specified portion of the table, contains the estimates of the intercepts of all the four units of the panel, here, it can easily be seen that all the estimates of the intercepts are fluctuating from each other, similar to the case of the random effects coefficients estimates, but differently from either the case of the subject specific coefficients estimates or the fixed effects estimates, above. Theoretically they should be the way they are in the case of the random coefficients estimates.

Similarly, the second column, in this fifth specified portion of the table, contains the estimates of the second regression coefficients of all the four units of the panel here too, it can easily be seen that all the estimates of the second regression coefficients are fluctuating from each other but differently from the case of the subject specific coefficients estimates, above. Theoretically they should be the way they are in the case of the random coefficients estimates.

In the same way, the third column in this said portion of the table, contains the estimates of the third regression coefficients of all the four units of the panel here again, it can easily be seen that all the estimates of the third regression coefficients are fluctuating from each, but here too, differently from the case of the subject specific coefficients estimates, above. Theoretically, it also should be the way they are in the case of the random coefficients estimates, and this way all the coefficients estimates discussed here, exactly match in characteristics to the random coefficients estimates.

Thus, from this detailed analysis of Table 6.6 above, it is concluded that all the frequentist estimates of the panel data model have been derived from the empirical Bayes estimate. Further, all of the frequentist estimates are the special cases of the empirical Bayes estimate. Finally, all the frequentist estimates have the common origin as the empirical Bayes estimate.

# Chapter 7

## Computation of the Empirical Bayes Estimate

In chapter 5, we have shown and seen analytically in details, that all the standard frequentist estimators of panel data linear regression models, are the limiting or special cases of the empirical Bayes estimate. The empirical Bayes estimate used there took the precision weighted arithmetic mean of the ordinary least squares coefficients estimates as the estimate of the prior mean, and the Zellner's g-prior as the estimate of the prior variance. Only with the help of this combination, of the estimate of the prior mean and prior variance, the frequentist estimates of panel data linear regression model could be derived from the empirical Bayes estimate.

Then, next in chapter 6, we numerically derived, the frequentist estimates, of standard frequentist panel data linear regression models, from the same empirical Bayes estimate. In this case we assigned different values to the prior precision parameters, in order, to derive the desired frequentist estimates. This method or trick worked well, only to the extent to show that in fact, with these prior precision parameters the frequentist estimates of panel data models can be derived from the empirical Bayes estimate.

But in real world practice, this self-selection or subjective procedure, of the assignment of values to the prior precision parameters, is looking hard and confusing, as different users might use different values for the prior precision parameters and hence different estimates will be produced by different users, even if dealing with the same data set.

To solve this issue, we need a unique solution of the problem. Therefore, it is far better than assigning arbitrarily values to the prior precision parameters, instead, estimate them too, from the same data set. For this purpose, in the current chapter, we estimate the prior precision parameters from the same data set, that has been used in chapter 6, by empirical Bayes procedure. The empirical Bayes procedure to be adopted here will be as given in (4.18z). The benefit of this method is that, we do not need to estimate the prior precision parameters, corresponding to each regression coefficient estimate, separately, rather, these will be estimated from the data automatically along with the model parameters, by applying (4.18z). Let us see the empirical Bayes estimate.

## 7.1. The Empirical Bayes Estimate

As mentioned above, the empirical Bayes coefficients estimates shall be obtained here by applying (4.18z), the courtesy of this estimator is that we do not need to estimate the prior precision parameter "$\rho$" separately from the estimates of the model parameters, rather, both the prior precision parameters and the model parameters will be estimated simultaneously with the help of the said formula of the empirical Bayes. Thus, from (4.18z) we have,

$$EB(\hat{\beta}_k^i) = \left(\left\{1 - \left[\sum_{i=1}^{N}\left\{\frac{(N-2)(\widehat{DV_k})^i}{\left(\hat{\beta}_k^i - \hat{B}_k\right)^2}\right\}\right]\right\}\hat{\beta}_k^i + \left\{\left[\sum_{i=1}^{N}\left\{\frac{(N-2)(\widehat{DV_k})^i}{\left(\hat{\beta}_k^i - \hat{B}_k\right)^2}\right\}\right]\right\}\hat{B}_k\right). \tag{7.1a}$$

and the estimate of the prior mean $\hat{B}_k$, from (4.16h) is as under,

$$\hat{B}_k = \frac{\left[\sum_{i=1}^{N}\left(\left[(\widehat{DV_k})^i\right]^{-1}\hat{\beta}_k^i\right)\right]}{\left[\sum_{i=1}^{N}\left(\left[(\widehat{DV_k})^i\right]^{-1}\right)\right]}. \tag{7.1b}$$

Further, in the light of (4.2a), $\hat{\beta}_k^i$ can be expressed as follow,

$$\hat{\beta}_k^i = \left(X_k^{i\prime}X_k^i\right)^{-1}\left(X_k^{i\prime}Y^i\right). \tag{7.1c}$$

Similarly, in the light of (5.12b), the sample counterpart of $(DV_k)^i$ can be written as under,

$$\left(\widehat{DV}_k\right)^i = (\hat{\sigma}_k^2)^i\left(X_k^{i\prime}X_k^i\right)^{-1} \tag{7.1d}$$

Now having all of the above quantities or estimates, below, we show, the empirical Bayes coefficients estimates for the same data set of chapter 6, Table (7.1), shows these estimates. Table (7.1) contains the empirical Bayes estimates.

Table 7.1: The Empirical Bayes Coefficient Estimates

| Variables | The Empirical Bayes Coefficient Estimates | | | |
|---|---|---|---|---|
| | Unit: 1 | Unit: 2 | Unit: 3 | Unit: 4 |
| Intercept | -8.0823 | -114.3830 | -38.6532 | -0.9383 |
| Second Regressor | 0.0446 | 0.0768 | 0.0950 | 0.0538 |
| Third Regressor | 0.1919 | 0.2614 | 0.2732 | 0.1729 |

In the above, empirical Bayes procedure, the prior precision parameter has been estimated along with the parameters of the model.

## 7.2. Summary of the Chapter

We summaries the findings of chapter 7, and discuss and relate it to theoretical realities, in order, to reach some conclusions. Table (7.1), contains the important and valuable information of this chapter. Therefore, this is the main and useful table of this section, with the help of this we can reach to any valid conclusion. The estimates in Table (7.1) have been computed with the help of such version of the empirical Bayes estimate which have yielded all the frequentist panel data estimates analytically as well as numerically. These facts we have been seen in chapter 5 and 6 respectively. Now towards Table (7.1), theoretically, if the intercept terms as well as the slopes coefficients, of the panel data linear regression models corresponding to different units of the panel are different from each other due to random fluctuations. Then they are known as the random coefficients estimates. In the random coefficients estimates, the coefficients (intercepts and slopes) are deviating randomly from the common mean, and the deviation is due to random causes.

Now, related to our this specific problem, it is evident that all the coefficients estimates, including the intercepts and slopes, are different from one another across all the units of the panel. These are the most appropriate estimate for the data, as here the prior precision parameters have been estimated from the data and not by our choice. Hence the estimates coincide with theory of the random coefficients and justify the results to be random coefficients estimates. Since, the derived estimates, of Table (7.1), have all the characteristics of being the frequentist random coefficients estimates, therefore, they are considered as the random effects coefficients estimates, and they are the special cases of the empirical Bayes estimate.

# Chapter 8

# Empirical Bayes Estimates from the Real-World Data

In this chapter, we take the real-world problem, to show the procedure for the analysis without pretesting for model selection. Here we show a single regressor empirical Bayes estimation procedure in details as well as the tendency of the empirical Bayes estimates. Therefore, we take a simple case of linear regression model, for five European Union countries namely, Austria, France, Italy, Sweden and Britain. The data is on two variables, i.e. the "GDP" and "Consumption" for the period of 1970 to 2016. Each variable for each country consists of 47 observations. The data has been taken from the world statistics. The log transforms of the variables have been made in order, to condense the data. Further, the GDP as the regressor has been centralised, in order, to make the regressors orthogonal and bring independency in the intercepts and slope coefficients.

One of the most important things to be noted here is that we are not interested in the economic characteristics of the above countries, i.e., how the GDP of a country affects the expenditure of the country, or how much the GDP of a country affects the consumption of a country, rather we are interested in the econometric or the statistical characteristics of the estimates corresponding to these countries by applying this specialised empirical Bayes technique of estimation.

This was all about the data set and the corresponding variables. Below we show the model to be used here.

## 8.1. The Simple Linear Regression Model

The simple linear regression model of our study with two regressors, i.e. the intercept and the GDP, is given as under,

$$c_t^i = y_{1t}^i \beta_1^i + y_{2t}^i \beta_2^i + \epsilon_t^i \tag{8.1}$$

Here, the description of the above model is as under,

$c_t^i$ denotes, the consumption of the $i^{th}$ country at $t^{th}$ time period,

$y_{1t}^i$ denotes, the constant 1 for the intercept term of the $i^{th}$ country at $t^{th}$ time period,

$y_{2t}^i$ denotes, the GDP of the $i^{th}$ country at $t^{th}$ time period.

$\epsilon_t^i$ denotes, the random errors of the $i^{th}$ country at $t^{th}$ time period.

$\beta_1^i$ denotes, the intercept term of the $i^{th}$ country.

$\beta_2^i$ denotes, the slope coefficient of the $i^{th}$ country.

The above (8.1a) can also be described as given below,

$$c_t^i = 1\,\beta_1^i + y_{2t}^i\,\beta_2^i + \epsilon_t^i. \qquad (8.1a)$$

### 8.1.1. Matrix Form of the Variables

The matrix structures of the above variables are given as under,

$$C^i = \begin{bmatrix} C_1^i \\ C_2^i \\ \vdots \\ C_t^i \\ \vdots \\ C_T^i \end{bmatrix}_{T\times 1}, \quad Y_1^i = \begin{bmatrix} y_{11}^i = 1 \\ y_{12}^i = 1 \\ \vdots \\ y_{1t}^i = 1 \\ \vdots \\ y_{1T}^i = 1 \end{bmatrix}_{T\times 1}, \quad Y_2^i = \begin{bmatrix} y_{21}^i \\ y_{22}^i \\ \vdots \\ y_{2t}^i \\ \vdots \\ y_{2T}^i \end{bmatrix}_{T\times 1}, \qquad (8.1.1a)$$

Let, $$Y^i = [Y_1^i \quad \dot{Y}_2^i] \qquad (8.1.1b)$$

Then, the expended matrix form of $Y^i$ in (8.1.1b) above, will get the form given as under,

$$Y^i = \begin{bmatrix} 1 & y_{21}^i \\ 1 & y_{22}^i \\ \vdots & \vdots \\ 1 & y_{2t}^i \\ \vdots & \vdots \\ 1 & y_{2T}^i \end{bmatrix}_{T\times 2} \qquad (8.1.1c)$$

In order, to orthogonalize the intercept and the slopes coefficients, we centralize $Y_2^i$ in (8.1.1a) above and get the following,

$$\dot{Y}_2^i = \begin{bmatrix} \dot{y}_{21}^i \\ \dot{y}_{22}^i \\ \vdots \\ \dot{y}_{2t}^i \\ \vdots \\ \dot{y}_{2T}^i \end{bmatrix}_{T \times 1} \tag{8.1.1d}$$

Where, $\dot{y}_{2t}^i = (y_{2t}^i - \bar{y}_2^i)$, the top-head dot is to differentiate the centralized $\dot{y}_{2t}^i$ from the simple $y_{2t}^i$, and $\bar{y}_2^i$ is the mean.

After centralization of $Y_2^i$, (8.1.1c) can now be written as follow,

$$\dot{Y}^i = \begin{bmatrix} 1 & \dot{y}_{21}^i \\ 1 & \dot{y}_{22}^i \\ \vdots & \vdots \\ 1 & \dot{y}_{2t}^i \\ \vdots & \vdots \\ 1 & \dot{y}_{2T}^i \end{bmatrix}_{T \times 2} \tag{8.1.1e}$$

After the centralization, below we present the modified model as well.

### 8.1.2. The Modified Model After Centralization

The model below is the modified model after centralization of $Y_2^i$.

$$C_t^i = y_{1t}^i\, \beta_1^i + (y_{2t}^i - \bar{y}_2^i)\, \beta_2^i + \epsilon_t^i, \tag{8.1.2a}$$

$$C_t^i = y_{1t}^i\, \beta_1^i + \dot{y}_{2t}^i\, \beta_2^i + \epsilon_t^i, \tag{8.1.2b}$$

where,

$$\dot{y}_{2t}^i = (y_{2t}^i - \bar{y}_2^i) \tag{8.1.2c}$$

## 8.2. Some Basic Quantities

In the following subsections of (8.2), we present some of the important quantities that are needed for this estimation process.

### 8.2.1. The Model Parameters Estimates

Now to estimate the model (8.1.2b) parameters, we have the ordinary least squares coefficients estimate which are given as under. Let this may be denoted by $\hat{\beta}^i$

$$\hat{\beta}^i = \left(\dot{Y}^{i\prime}\dot{Y}^i\right)^{-1}\dot{Y}^{i\prime}C^i \tag{8.2.1a}$$

$$\left(\hat{\beta}^i\right) = \begin{bmatrix}\hat{\beta}_1^i\\ \hat{\beta}_2^i\end{bmatrix} \tag{8.2.1b}$$

These $\hat{\beta}_1^i$ and $\hat{\beta}_2^i$ are now the ordinary least squares coefficients estimates for the intercept term and the slope coefficients of model parameters (8.1.2b) respectively. We also need the variance of the ordinary least squares coefficients estimates. Below, we present the data variances of the ordinary least squares coefficients estimates.

### 8.2.2. The Data Variance of the Model Parameters Estimates

As mentioned above, we need the data variances of the model parameters i.e. the variances of the ordinary least squares coefficients estimate, therefore, by definition,

$$DV\left(\hat{\beta}^i\right) = (\sigma^2)^i\left(\dot{Y}^{i\prime}\dot{Y}^i\right)^{-1} \tag{8.2.2a}$$

further, by definition,

$$(\sigma^2)^i = \left(\frac{\epsilon_t^{i\prime}\epsilon_t^i}{T-K}\right) \tag{8.2.2b}$$

and in this orthogonal case, we have,

$$\left(\dot{Y}^{i\prime}\dot{Y}^i\right) = \begin{bmatrix}T & 0\\ 0 & \sum_{t=1}^{T}\left(\dot{y}_{2t}^i\right)^2\end{bmatrix} \tag{8.2.2c}$$

and

$$\left(\dot{Y}^{i\prime}\dot{Y}^i\right)^{-1} = \begin{bmatrix}(T)^{-1} & 0\\ 0 & \left(\sum_{t=1}^{T}\left(\dot{y}_{2t}^i\right)^2\right)^{-1}\end{bmatrix} \tag{8.2.2d}$$

thus finally,

$$(\sigma^2)^i\left(\dot{Y}^{i\prime}\dot{Y}^i\right)^{-1} = (\sigma^2)^i\begin{bmatrix}(T)^{-1} & 0\\ 0 & \left(\sum_{t=1}^{T}\left(\dot{y}_{2t}^i\right)^2\right)^{-1}\end{bmatrix} \tag{8.2.2e}$$

or

$$(\sigma^2)^i(\dot{Y}^{i\prime}\dot{Y}^i)^{-1} = \begin{bmatrix} (\sigma^2)^i(T)^{-1} & 0 \\ 0 & (\sigma^2)^i\left(\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right)^{-1} \end{bmatrix} \tag{8.2.2f}$$

and thus,

$$DV(\hat{\beta}^i) = \begin{bmatrix} (\sigma^2)^i(T)^{-1} & 0 \\ 0 & (\sigma^2)^i\left(\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right)^{-1} \end{bmatrix} \tag{8.2.2g}$$

Note: In (8.2.2g) the left-hand top diagonal or the first diagonal, quantity is the variance of the first regression coefficient estimate, i.e. the intercept while the right-hand bottom diagonal or the second diagonal, quantity is the variance of the second regression coefficient estimate, i.e. the slope.

Let, $DV(\hat{\beta}_1^i)$ denotes the data variance of the first regression coefficient and $DV(\hat{\beta}_2^i)$ denotes the data variance of the second regression coefficient.

Then,

$$DV(\hat{\beta}_1^i) = (\sigma^2)^i(T)^{-1} \tag{8.2.2h}$$

and

$$DV(\hat{\beta}_2^i) = (\sigma^2)^i\left(\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right)^{-1} \tag{8.2.2i}$$

On the basis of the above quantities, below in section (8.2.3), we show the corresponding data precisions.

### 8.2.3. The Data Precision of the Model Parameters Estimates

Here on the basis of the above data variances, the corresponding data precisions, which are denoted by $DP(\hat{\beta}_1^i)$ and $DP(\hat{\beta}_2^i)$ respectively for both of the regression coefficient estimates become as under,

$$DP(\hat{\beta}_1^i) = \left[(\sigma^2)^i\right]^{-1}T \tag{8.2.3a}$$

and

$$DP(\hat{\beta}_2^i) = \left[(\sigma^2)^i\right]^{-1}\left[\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right] \tag{8.2.3b}$$

Below, we describe the data densities for both coefficients estimates.

### 8.2.4. The Data Densities for Both of the Coefficients Estimates

The data density in terms of means and variances of $\hat{\beta}_1^i$ is given as under,

$$\hat{\beta}_1^i \underset{\sim}{IID} N\left(\beta_1^i, DV(\hat{\beta}_1^i)\right) \tag{8.2.4a}$$

Thus, $\hat{\beta}_1^i$ is normally distributed with mean $\beta_1^i$ and variance $DV(\hat{\beta}_1^i)$.

and similarly, the data density in terms of means and variances of $\hat{\beta}_2^i$ is also given as under,

$$\hat{\beta}_2^i \underset{\sim}{IID} N\left(\beta_2^i, DV(\hat{\beta}_2^i)\right) \tag{8.2.4b}$$

Thus, $\hat{\beta}_2^i$ is normally distributed with mean $\beta_2^i$ and variance $DV(\hat{\beta}_2^i)$. Now in the following section we describe the prior densities too.

### 8.2.5. The Prior Densities of Both Regression Coefficients

The prior density in terms of mean and variance of $\beta_1^i$ is given as under,

$$\beta_1^i \underset{\sim}{IID} N\left(\mathrm{B}_1, \Lambda_1^i\right) \tag{8.2.5a}$$

and similarly, the prior density in terms of mean and variance of $\beta_2^i$ is also given as under,

$$\beta_2^i \underset{\sim}{IID} N\left(\mathrm{B}_2, \Lambda_2^i\right) \tag{8.2.5b}$$

Here, $\mathrm{B}_1$ and $\mathrm{B}_2$ denotes the prior means of $\beta_1^i$ and $\beta_2^i$, i.e. the $1^{st}$ and $2^{nd}$ regression coefficient in the model respectively, and $\Lambda_1^i$ and $\Lambda_2^i$ denotes the prior variances of $\beta_1^i$ and $\beta_2^i$ respectively.

Let, $$(PV_1)^i = \Lambda_1^i, \tag{8.2.5c}$$

and $$(PV_2)^i = \Lambda_2^i, \tag{8.2.5d}$$

for the rest of the chapter. Now, in the following subsection we present the prior variance.

### 8.2.6. Computation of the Prior Variance

In order, to use the g-priors in general and the g-prior precision in specific, let us define the g-prior precision for both of the regression coefficients.

By the definition of the g-prior, the g-prior variance is proportional to the data variance. In our case,

$$(PV_1)^i \propto DV(\hat{\beta}_1^i) \tag{8.2.6a}$$

or

$$(PV_1)^i = \rho_1^{-1} DV(\hat{\beta}_1^i) \tag{8.2.6b}$$

and

$$(PV_2)^i \propto DV(\hat{\beta}_2^i) \tag{8.2.6c}$$

or

$$(PV_2)^i = \rho_2^{-1} DV(\hat{\beta}_2^i) \tag{8.2.6d}$$

Now plugging the value of $DV(\hat{\beta}_1^i)$ and $DV(\hat{\beta}_2^i)$ from (8.2.3a) and (8.2.3b) above in (8.2.6c) and (8.2.6d) respectively, we have,

$$(PV_1)^i = \rho_1^{-1}\left[(\sigma^2)^i (T)^{-1}\right] \tag{8.2.6e}$$

$$(PV_2)^i = \rho_2^{-1}\left[(\sigma^2)^i \left\{\sum_{t=1}^{T} \left(\dot{y}_{2t}^i\right)^2\right\}^{-1}\right] \tag{8.2.6f}$$

Now, in the following subsection we present the prior precision.

## 8.2.7. The Prior Precisions

Let, $(PP_1)^i$ and $(PP_2)^i$ denote the two prior precisions respectively, then, in terms of precision (8.2.6e) and (8.2.6f) above can also be expressed as,

$$(PP_1)^i = \rho_1 \left[\left\{(\sigma^2)^i\right\}^{-1} T\right] \tag{8.2.7a}$$

$$(PP_2)^i = \rho_2 \left[(\sigma^2)^{i^{-1}} \left\{\sum_{t=1}^{T} \left(\dot{y}_{2t}^i\right)^2\right\}\right] \tag{8.2.7b}$$

where,

$$(PP_1)^i = \left(\Lambda_1^i\right)^{-1} \tag{8.2.7c}$$

and

$$(PP)_2^i = \left(\Lambda_2^i\right)^{-1} \tag{8.2.7d}$$

Now, with the help of (8.2.6e) and (8.2.6f) above, the prior densities of both of the coefficients are as under,

$$\hat{\beta}_1^i \overset{IID}{\sim} N\left(\mathrm{B}_1, \{\rho_1^{-1}[(\sigma^2)^i(T)^{-1}]\}\right) \tag{8.2.7e}$$

and

$$\beta_2^i \overset{IID}{\sim} N\left(\mathrm{B}_2, \left\{(\rho_2)^{-1}(\sigma^2)^i\left(\sum_{t=1}^{T}\left(\dot{y}_{2t}^i\right)^2\right)^{-1}\right\}\right) \tag{8.2.7f}$$

Below we show the marginal densities of the estimates of the model parameters.

## 8.3. The Marginal Densities of the Estimates

Here, the marginal density for the $1^{st}$ regression coefficient estimate becomes as under,

$$m\left(\hat{\beta}_1^i\right) \overset{IID}{\sim} N\left(\mathrm{B}_1, \{[(\sigma^2)^i(T)^{-1}] + \rho_1^{-1}[(\sigma^2)^i(T)^{-1}]\}\right) \tag{8.3a}$$

Now, we describe the marginal density for the $2^{nd}$ regression coefficient estimate which becomes as under,

$$m\left(\hat{\beta}_2^i\right) \overset{IID}{\sim} N\left(\mathrm{B}_2, \left\{\left[(\sigma^2)^i\left(\sum_{t=1}^{T}\left(\dot{y}_{2t}^i\right)^2\right)^{-1}\right] + \rho_2^{-1}\left[(\sigma^2)^i\left(\sum_{t=1}^{T}\left(\dot{y}_{2t}^i\right)^2\right)^{-1}\right]\right\}\right) \tag{8.3b}$$

Below we show the estimates of the prior means.

## 8.4. The Estimates of the Prior Means

Below we show the estimates of the prior means separately, first, we show the estimates of the prior mean for the intercept and then the estimates of the prior mean for the slope. Let us see the estimates of the prior mean for the intercept in the following subsection.

### 8.4.1. The Estimates of the Prior Mean for the Intercept

Now, the estimate of the prior mean $\hat{\mathrm{B}}_1$ for the first regression coefficient may be given as follow,

$$\widehat{\mathrm{B}}_1 = \frac{\sum_{i=1}^{N}\left\{\left[\left\{(\hat{\sigma}_1^2)^i\right\}^{-1} T\right] + \hat{\rho}_1\left[\left\{(\hat{\sigma}_1^2)^i\right\}^{-1} T\right]\right\}\hat{\beta}_1^i}{\sum_{i=1}^{N}\left[\left\{(\hat{\sigma}_1^2)^i\right\}^{-1} T\right] + \hat{\rho}_1\left[\left\{(\hat{\sigma}_1^2)^i\right\}^{-1} T\right]} \tag{8.4.1a}$$

$$\widehat{\mathrm{B}}_1 = \frac{\sum_{i=1}^{N}\left\{(1+\hat{\rho}_1)\left[\left\{(\hat{\sigma}_1^2)^i\right\}^{-1} T\right]\right\}\hat{\beta}_1^i}{\sum_{i=1}^{N}\left\{(1+\hat{\rho}_1)\left[\left\{(\hat{\sigma}_1^2)^i\right\}^{-1} T\right]\right\}^{-1}} \tag{8.4.1b}$$

$$\widehat{\mathrm{B}}_1 = \frac{(1+\hat{\rho}_1)\sum_{i=1}^{N}\left[\left\{(\hat{\sigma}_1^2)^i\right\}^{-1} T\right]\hat{\beta}_1^i}{(1+\hat{\rho}_1)\sum_{i=1}^{N}\left[\left\{(\hat{\sigma}_1^2)^i\right\}^{-1} T\right]} \tag{8.4.1c}$$

$$\widehat{\mathrm{B}}_1 = \frac{\sum_{i=1}^{N}\left[\left\{(\hat{\sigma}_1^2)^i\right\}^{-1} T\right]\hat{\beta}_1^i}{\sum_{i=1}^{N}\left[\left\{(\hat{\sigma}_1^2)^i\right\}^{-1} T\right]} \tag{8.4.1d}$$

Here it is to be noted that, $\widehat{\mathrm{B}}_1$ becomes independent of $\hat{\rho}_1$.

Now, let us see the estimates of the prior mean for the slope in the following subsection.

## 8.4.2. The Estimates of the Prior Mean for the Slope

Also, the estimate of the prior mean $\mathrm{B}_2$ for the second regressor may be given as follow,

$$\widehat{\mathrm{B}}_2 = \frac{\sum_{i=1}^{N}\left\{\left[\left\{(\hat{\sigma}_2^2)^i\right\}^{-1}\left(\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right)\right] + \hat{\rho}_2\left[\left\{(\hat{\sigma}_2^2)^i\right\}^{-1}\left(\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right)\right]\right\}\hat{\beta}_2^i}{\sum_{i=1}^{N}\left[\left\{(\hat{\sigma}_2^2)^i\right\}^{-1}\left(\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right)\right] + \hat{\rho}_2\left[\left\{(\hat{\sigma}_2^2)^i\right\}^{-1}\left(\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right)\right]} \tag{8.4.2a}$$

$$\hat{B}_2 = \frac{\sum_{i=1}^{N}\left\{(1+\hat{\rho}_2)\left[\left\{(\hat{\sigma}_2^2)^i\right\}^{-1}\left(\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right)\right]\right\}\hat{\beta}_2^i}{\sum_{i=1}^{N}\left\{(1+\hat{\rho}_2)\left[\left\{(\hat{\sigma}_2^2)^i\right\}^{-1}\left(\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right)\right]\right\}} \tag{8.4.2b}$$

$$\widehat{\mathrm{B}}_2 = \frac{(1+\hat{\rho}_2)\sum_{i=1}^{N}\left[\left\{(\hat{\sigma}_2^2)^i\right\}^{-1}\left(\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right)\right]\hat{\beta}_2^i}{(1+\hat{\rho}_2)\sum_{i=1}^{N}\left[\left\{(\hat{\sigma}_2^2)^i\right\}^{-1}\left(\sum_{t=1}^{T}(\dot{y}_{2t}^i)^2\right)\right]} \tag{8.4.2c}$$

$$\widehat{\mathrm{B}}_2 = \frac{\sum_{i=1}^{N}\left[\left\{\left(\hat{\sigma}_2^2\right)^i\right\}^{-1}\left(\sum_{t=1}^{T}\left(\dot{y}_{2t}^i\right)^2\right)\right]\hat{\beta}_2^i}{\sum_{i=1}^{N}\left[\left\{\left(\hat{\sigma}_2^2\right)^i\right\}^{-1}\left(\sum_{t=1}^{T}\left(\dot{y}_{2t}^i\right)^2\right)\right]} \tag{8.4.2d}$$

Here again it is to be noted that, $\widehat{\mathrm{B}}_2$ becomes independent of $\hat{\rho}_2$.

## 8.5. The Empirical Bayes Estimates

Here, by applying the empirical Bayes estimate given (4.18z), we apply the empirical Bayes estimates without estimating the prior precision parameter, thus, the empirical Bayes estimates for the intercept term and the slope coefficients respectively, are described in the following sections.

### 8.5.1. The Empirical Bayes Estimates for the Intercept

From (4.18z) we have,

$$EB\left(\hat{\beta}_k^i\right) = \left(\left\{1 - \left[\sum_{i=1}^{N}\left\{\frac{(N-2)\left[DV\left(\hat{\beta}_1^i\right)\right]}{\left(\hat{\beta}_1^i - \hat{B}_1\right)^2}\right\}\right]\right\}\hat{\beta}_k^i + \left\{\left[\sum_{i=1}^{N}\left\{\frac{(N-2)\left[DV\left(\hat{\beta}_1^i\right)\right]}{\left(\hat{\beta}_1^i - \hat{B}_1\right)^2}\right\}\right]\right\}\hat{B}_k\right) \tag{8.5.1a}$$

Modifying for the intercept we get,

$$EB\left(\hat{\beta}_1^i\right) = \left(\left\{1 - \left[\sum_{i=1}^{N}\left\{\frac{(N-2)\left[DV\left(\hat{\beta}_1^i\right)\right]}{\left(\hat{\beta}_1^i - \hat{B}_1\right)^2}\right\}\right]\right\}\hat{\beta}_1^i + \left\{\left[\sum_{i=1}^{N}\left\{\frac{(N-2)\left[DV\left(\hat{\beta}_1^i\right)\right]}{\left(\hat{\beta}_1^i - \hat{B}_1\right)^2}\right\}\right]\right\}\hat{B}_1\right) \tag{8.5.1b}$$

or

$$EB\left(\hat{\beta}_1^i\right) = \left(\left\{1 - \left[\sum_{i=1}^{N}\left\{\frac{(N-2)\left\{\left(\hat{\sigma}_1^2\right)^i (T)^{-1}\right\}}{\left(\hat{\beta}_1^i - \hat{B}_1\right)^2}\right\}\right]\right\}\hat{\beta}_1^i + \left\{\left[\sum_{i=1}^{N}\left\{\frac{(N-2)\left\{\left(\hat{\sigma}_1^2\right)^i (T)^{-1}\right\}}{\left(\hat{\beta}_1^i - \hat{B}_1\right)^2}\right\}\right]\right\}\hat{B}_1\right) \tag{8.5.1c}$$

The above in (8.5.1c), is the empirical Bayes estimate of the first regression coefficient i.e. the intercept term of the model independent of the prior precision parameter.

### 8.5.2. The Empirical Bayes Estimate for the Slope

Modifying (8.5.1a) for the slope we get,

$$EB\left(\hat{\beta}_2^i\right) = \left(\left\{1 - \left[\sum_{i=1}^{N}\left\{\frac{(N-2)\left[DV\left(\hat{\beta}_2^i\right)\right]}{\left(\hat{\beta}_2^i - \hat{B}_2\right)^2}\right\}\right]\right\}\hat{\beta}_2^i + \left\{\left[\sum_{i=1}^{N}\left\{\frac{(N-2)\left[DV\left(\hat{\beta}_2^i\right)\right]}{\left(\hat{\beta}_2^i - \hat{B}_2\right)^2}\right\}\right]\right\}\hat{B}_2\right) \tag{8.5.2a}$$

or

$$EB\left(\hat{\beta}_2^i\right) = \left(\left\{1 - \left[\sum_{i=1}^{N}\left\{\frac{(N-2)\left\{\left(\hat{\sigma}_2^2\right)^i\left[\sum_{t=1}^{T}\left(\dot{y}_{2t}^i\right)^2\right]^{-1}\right\}}{\left(\hat{\beta}_{22}^i - \hat{B}_2\right)^2}\right\}\right]\right\}\hat{\beta}_2^i + \left\{\left[\sum_{i=1}^{N}\left\{\frac{(N-2)\left\{\left(\hat{\sigma}_2^2\right)^i\left[\sum_{t=1}^{T}\left(\dot{y}_{2t}^i\right)^2\right]^{-1}\right\}}{\left(\hat{\beta}_2^i - \hat{B}_2\right)^2}\right\}\right]\right\}\hat{B}_2\right) \quad (8.5.2b)$$

The above in (8.5.2b), is the empirical Bayes estimate of the second regression coefficient i.e. the slope coefficient of the model. Now, it is very simple to find the empirical Bayes estimates from (8.5.1c) and (8.5.2b) above.

## 8.6. The Empirical Bayes Estimates

In this section we present, with the help of different tables, all the numerical findings.

Below we show the empirical Bayes estimates for the intercept terms.

Table 8.6a: The Empirical Bayes Estimates for the Intercepts

| Empirical Bayes Estimates for the Intercept | | | | | |
|---|---|---|---|---|---|
| | Austria | France | Italy | Sweden | Britain |
| Intercept Coefficient Estimates | 0.2694 | 0.3012 | 0.2861 | 0.2767 | 0.2921 |
| Standard Error of Intercepts | 0.0097 | 0.0183 | 0.0234 | 0.0065 | 0.0203 |

Table 8.6b: Empirical Bayes Estimates for the Slopes

| Empirical Bayes Estimates for the Slopes | | | | | |
|---|---|---|---|---|---|
| | Austria | France | Italy | Sweden | Britain |
| Slope Coefficient Estimates | 0.9804 | 0.9436 | 0.9302 | 0.9338 | 0.9989 |
| Standard Error of Slopes | 0.0084 | 0.0078 | 0.0064 | 0.0092 | 0.0074 |

Table 8.6c: The Empirical Bayes Coefficients Estimates for All Countries

| The Empirical Bayes Coefficients Estimates | | | | |
|---|---|---|---|---|
| Country | Intercept | | Slope | |
| | Coefficients Estimates | Standard Error of the Coefficients Estimates | Coefficients Estimates | Standard Error of the Coefficients Estimates |
| Austria | 0.2694 | 0.0097 | 0.9804 | 0.0084 |
| France | 0.3012 | 0.0183 | 0.9436 | 0.0078 |
| Italy | 0.2861 | 0.0234 | 0.9302 | 0.0064 |
| Sweden | 0.2767 | 0.0065 | 0.9338 | 0.0092 |
| Britain | 0.2921 | 0.0203 | 0.9989 | 0.0074 |

Here, Table (8.6a) above, contains the intercepts estimates with their corresponding standard errors for all the five countries of the analysis. Similarly, Table (8.6b) above, contains the slopes with their corresponding standard errors for all the five countries of the analysis. Further, Table (8.6c) above, contains the intercepts as well as the slopes with their corresponding standard errors simultaneously, for all the five countries of the analysis. Here as per the theoretical relationship between the consumption and GDP, the sign and magnitude of both the intercepts and the slopes seem very much consistent with the theory. That is, according to theory, the sign of the slope coefficient in the consumption function must be positive, and the magnitude of the slope coefficient must be smaller than unity, and so are the cases. Also, the coefficients estimates among different countries seem randomly fluctuating.

## Summary of the Chapter

Here, in this chapter we applied the empirical Bayes estimation techniques to the real-world data. The data was for five European Union countries namely, Austria, France, Italy, Sweden and Britain. The data was on two variables, i.e. the "GDP" and "Consumption" for the period of 1970 to 2016. Each variable for each county consists of 46 observations. The data had been taken from the world statistics. The log transforms of the variables have been made in order, to condense the data. Further, the GDP as the regressor has been centralised, in order, to make the regressors orthogonal and bring independency in the intercepts and slope coefficients. Thereafter, the whole procedure was described analytically.

Salient features of the estimates are that we have not done any pretesting procedure for model selection and neither we have decided in advance that either we fit the subjects specific coefficient model or the subjects common coefficients model or any other model of the frequentist set up, rather we estimated the coefficients by the empirical Bayes estimate for each unit. This is the quality of the empirical Bayes estimate that one have no need for pretesting for model selection, the technique itself will first design the vectors of the coefficients estimates and then resultantly the corresponding models, and this is what we have shown for the very first time in this thesis.

# Chapter 9

## Summary, Contributions, Conclusion, Recommendations and Direction for Future Studies

In this chapter we summarise the whole study conducted in this thesis, in order to reach some conclusions and then on the basis of these conclusions to assert some recommendations and directions for future studies. This dissertation concentrates on the derivation of the standard frequentist panel data linear regression models from the panel data Bayesian linear regression model.

Chapter one is the first and the introductory chapter, it contains the introduction to this dissertation. In the first part of the introduction chapter the statement of the problem is described, then we have briefly discussed the use and importance of the panel data. We have also discussed some important characteristics of the panel data and the related models for panel data, and similarly the corresponding frequentist estimation techniques necessary for the efficient estimation of these models. Next in turn, we have discussed the Bayesian techniques of estimation used for panel data models. Further, the aims of the study were described with the objectives of the study in chapter one. After discussing the objectives, the motivation for this study is also expressed in chapter one. Next, the contributions made with the help of this thesis have been presented. Finally, the significance of the study has been given.

Chapter two has mainly dealt with the literature review. The chapter has basically oscillated around the use or need of the Bayesian techniques. It also discussed the frequentist panel data models and their corresponding estimation techniques.

Chapter three has presented the methodological framework carried out in this dissertation.

Chapter four has developed the theoretical background of the models, the variables, the estimates, etc. It also contains the new empirical Bayes estimate.

Chapter five is spared to the analytical derivations of the frequentist estimates from the Bayesian estimates. In this chapter, from the empirical Bayes estimator with the precision weighted arithmetic mean of the ordinary least squares coefficient estimate as the estimate of the prior mean and the Zellner's g-prior as the estimate of the prior variance, the frequentist estimates of panel data linear regression models have been derived, by single regressor estimation method. In the derivation of the frequentist estimates, first, the subject specific coefficients estimate, second the subjects common coefficients estimate, third the fixed effects estimates, fourth, the random effects estimates and finally the frequentist random coefficients estimate have been derived, this has completed chapter five.

In chapter six, similar to chapter five the frequentist estimates have been derived numerically. Here the frequentist estimates have been derived from the empirical Bayes estimate, with the precision weighted arithmetic mean of the ordinary least squares coefficients estimates as the estimate of the prior mean and the Zellner's g-prior as the estimate of the prior variance, analogously to chapter five but numerically. We have shown for each coefficients estimates, in a table, the prior precision parameters with the help of which those particular coefficients estimates can be derived from the empirical Bayes estimate, then in another table, the corresponding derived frequentist coefficients estimates.

In chapter seven we estimated the prior precision parameter "$\rho$", because in chapter six, the prior precision parameter has been purposively and arbitrarily selected rather than estimated, as some special results had to derive, and these special results could only be derived with the help of these special selected values of the prior precision parameters. But one interesting think to note is that, in chapter four, we develop such an empirical Bayes estimate, which estimates the prior precision parameter along with model parameters simultaneously. So in chapter seven we used this estimator and did not estimate the prior precision parameters separately.

In chapter eight, the real-world problem has been taken into account, to show a single regressor empirical Bayes estimation procedure in details as well as the tendency of the empirical Bayesian estimates. Therefore, we took a simple linear regression model for five European Union countries namely Austria, France, Italy, Sweden and Britain. The data was on two variables i.e. the "GDP" and "Consumption" for the period of 1970 to 2016. Each variable for each country consisted of 46 observations. The data had been taken from the world statistics. The log transforms of the variables had been made in order, to condense the data.

Further, the GDP as the regressor had been centralised, in order to make the regressors orthogonal and bring independency in the intercepts and slope coefficients. And then the empirical Bayes estimate of chapter had been applied.

This was the whole story of the thesis. Now we describe the contributions made of the thesis.

## 9.1. The Contributions

In the contributions, of this thesis, the **first contribution** is that all the frequentist panel data models have been shown to be encompassed by a single model. The **second contribution** is the derivation of all basic standard frequentist panel data linear regression models from the Bayesian linear regression model. The **third contribution** is that it has been shown that all the frequentist panel data models have a single origin, which is the Bayesian model, and all the frequentist models are the special cases of this origin. The **fourth contribution** is that all the frequentist estimates corresponding to different frequentist panel data models can be derived from empirical Bayes estimate analytically as well as numerically. The **fifth main contribution** is the new version of the empirical Bayes estimate based on the weighted arithmetic mean of the ordinary least squares coefficients estimates as the estimate of the prior mean and the Zellner's g-prior as the estimate of the prior variance has been developed, where there is no need of estimation of the prior precision parameters first. The importance of this estimator is that this estimator has been derived from the same combination of the prior mean and prior variance of the empirical Bayes estimate which produces the estimators required for all the basic standard panel data linear regression models. The **sixth contribution** is that this technique give relief to the researchers from pretesting of model selection in panel data.

## 9.2. Advantage of the Research

The advantages of the research include; self-selection of the models in the light of data, no pretesting for model selection on ad-hoc basis, to estimates the models with more reliable techniques of estimation where the uncertainty is as an inherent part.

## 9.3. Conclusion of this Dissertation

After the detailed analysis, it has been found that the empirical Bayes estimate is the most flexible estimator for the panel data ease as compared to any of the frequentist estimators.

The first conclusion we have been reached at is that all the frequentist panel data linear regression models have a unique and common origin they are not of the different origin. The consideration of different origin by the frequentist School of thought is not very reasonable.

The second conclusion that we have been arrived at is that all the frequentist estimates are the derived forms or versions of the empirical Bayes estimate. The Bayesian estimator have encompassed all the frequentist estimates of the panel data models and therefore has the greatest potential.

The third conclusion that we derived, based on this dissertation, is that it is meaningful for many panel data sets to be modelled with the random coefficients model and the corresponding estimates. Except the two extreme cases, that is, either the subject specific coefficients models or the subjects common coefficients models, there are numerous random coefficients models and the resultant estimates.

The fourth conclusion is that when the frequentist statisticians or econometricians force the panel data to be fitted with only these basic five models is injustice.

The fifth and in fact the most important conclusion that has been derived is that by using empirical Bayes estimate the data itself will generate the most appropriate model without any pretesting etc.

## 9.4. Recommendations

Here in this dissertation, all the frequentist panel data models have been derived from the Bayesian model. As there are infinite values of the prior precision parameters and only two of them i.e., zero and infinity, take to two extreme cases namely, the subject specific coefficients estimates and the subjects common coefficients estimates and the rest of the values of the prior precision parameter take to the random coefficients estimates. So, therefore, there is very high likelihood of the random coefficients model that fits the data. In the real-world problem too, in most of the cases, the panel data is neither completely heterogeneous nor completely homogeneous but a combination of these two. Fortunately, the Bayes estimate spontaneously leads to the random coefficients estimates in many cases, and resultantly, the Bayesian estimation techniques are highly recommended. Therefore, the empirical Bayes estimation techniques are strongly recommended for practical uses.

## 9.5. Directions for Future Studies or Research

As all of the frequentist estimates of panel data models have been derived from the empirical Bayes estimate, it is therefore the most important, valuable and useful technique of estimation. Here, in this dissertation, we derived these frequentist estimates by adopting a single regressor estimation approach, so in directions for future studies, it is advised to derive the Bayes estimate of the panel data models with the same combination of the prior mean and prior variance that could estimate all regression parameters of a model simultaneously.

Further, the empirical Bayes estimate is a very precious tool in general and the estimator we developed in specific, but unfortunately it has not been used by the users or researchers to get facilitated from it, because it needs programming and many of the researchers are either unable to do programming or can do very little programming and hence this deprives them from the benefits of this beneficial technique. There is no such statistical package or software with the built-in functions for it, and friendly interface. Therefore, the future research may be conducted to design a friendly used software for the users of the empirical Bayes formula that is developed in this thesis.

# REFERENCES


Ahn, S. C., Lee, Y. H., and Schmidt, P. (2013). Panel Data Models with Multiple Time-Varying Individual Effects. *Journal of Econometrics*, *174*(1), 1-14.

Aigner, D., Lovell, C. K., and Schmidt, P. (1977). Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometrics*, *6*(1), 21-37.

Arceneaux, K., and Nickerson, D. W. (2009). Modeling Certainty with Clustered Data: A Comparison Of Methods. *Political Analysis*, *17*(2), 177-190.

Arellano, M. (2003). *Panel Data Econometrics*. Oxford University Press.

Bai, J. (2009). Panel Data Models with Interactive Fixed Effects. *Econometrica*, *77*(4), 1229-1279.

Baltagi, B. H., Song, S. H., and Jung, B. C. (2001). The Unbalanced Nested Error Component Regression Model. *Journal of Econometrics*, *101*(2), 357-381.

Başçı, S., and Zaman, A. (1998). Effects of Skewness and Kurtosis on Model Selection Criteria. *Economics Letters*, *59*(1), 17-22.

Beck, N., and Katz, J. N. (1995). What to Do (and Not to Do) With Time-Series Cross-Section Data. *American Political Science Review*, *89*(3), 634-647.

Belsley, D. A., and Kuti, E. (1973). Time-Varying Parameter Structures: An Overview. In *Annals of Economic And Social Measurement, Volume 2, Number 4* (pp. 375-379).

Carrington, W. J., and Zaman, A. (1994). Interindustry Variation in the Costs of Job Displacement. *Journal of Labor Economics*, *12*(2), 243-275.

Chang, I. L., Swamy, P. A. V. B., Hallahan, C., and Tavlas, G. S. (2000). A computational approach to finding causal economic laws. Computational Economics, 16(1), 105-136.

Chow, G. C. (1984). Random and Changing Coefficient Models. *Handbook of Econometrics*, *2*, 1213-1245.

Christou, C., Swamy, P. A. V. B., and Tavlas, G. S. (1998). A General Framework for Predicting Returns From Multiple Currency Investments. *Journal of Economic Dynamics and Control*, *22*(7), 977-1000.

Clark, T. S., and Linzer, D. A. (2015). Should I Use Fixed or Random Effects?. *Political Science Research and Methods*, *3*(2), 399-408.

Clark, T., and Linzer, D. A. (2012). Deciding Between Fixed and Random Effects. *APSA-Comparative Politics Newsletter*, *22*(1), 9-10.

Cooley, T. F., and Prescott, E. C. (1973). An Adaptive Regression Model. *International Economic Review*, 364-371.

Cooley, T. F., and Prescott, E. C. (1976). Estimation in The Presence of Stochastic Parameter Variation. *Econometrica: Journal of the Econometric Society*, 167-184.

Efron, B., and Morris, C. (1972). Limiting the risk of Bayes and empirical Bayes estimators—Part II: The empirical Bayes case. Journal of the American Statistical Association, 67(337), 130-139.

Efron, B., & Morris, C. (1973). Stein's estimation rule and its competitors—an empirical Bayes approach. Journal of the American Statistical Association, 68(341), 117-130.

Efron, B., & Morris, C. (1975). Data analysis using Stein's estimator and its generalizations. Journal of the American Statistical Association, 70(350), 311-319.

Frees, E. W. (2004). *Longitudinal and Panel Data: Analysis and Applications in the Social Sciences*. Cambridge University Press.

Froehlich, H. A. (1973). Natural and Man-Caused Slash in Headwater Streams. *Log Handb.*

Gelman, A. (2005). Analysis of Variance—Why It Is More Important Than Ever. *The Annals of Statistics*, *33*(1), 1-53.

Gelman, A., and Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models. Cambridge university press.

Ghosh, M., Saleh, A. M. E., and Sen, P. K. (1989). Empirical Bayes subset estimation in regression models. Statistics & Risk Modeling, 7(1-2), 15-36.

Granger, C. W., & Terasvirta, T. (1993). Modelling Non-Linear Economic Relationships. *OUP Catalogue*.

Greene, W. H. (1990). A Gamma-Distributed Stochastic Frontier Model. *Journal of Econometrics*, *46*(1-2), 141-163.

Greene, W. H. (2003). *Econometric Analysis*. Pearson Education India.

Gujarati, D. N. (2003). Basic Econometrics. Fourth edition McGraw-Hill. New York.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica: Journal of the Econometric Society*, 1251-1271.

Hildreth, C., and Houck, J. P. (1968). Some Estimators for A Linear Model with Random Coefficients. *Journal of the American Statistical Association*, *63*(322), 584-595.

Hsiao, C. (1974). Statistical Inference for A Model with Both Random Cross-Sectional And Time Effects. *International Economic Review*, 12-30.

Hsiao, C. (1986). Analysis of Panel Data, Econometric Society monographs No. 11, New York: Cambridge University Press.

Hsiao, C. (2003). Analysis of Panel Data (Vol. 34). *Econometric Society Monographs*.

Hsiao, C. (2007). Panel Data Analysis—Advantages and Challenges. *Test*, *16*(1), 1-22.

Hsiao, C., Mountain, D. C., and Ho-Illman, K. (1995). Bayesian Integration of End-Use Metering and Conditional Demand Analysis. Journal of Business and Economic Statistics, 13:315–326.

Hsiao, T. C., and Lauchli, A. (1986). Role of Potassium in Plant-Water Relations. *Advances in Plant Nutrition (USA)*.

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, *82*(1), 35-45.

Kalman, R. E., and Bucy, R. S. (1961). New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering*, *83*(1), 95-108.

Kennedy, P. (2003). *A Guide to Econometrics*. MIT press.

Klein, L. R. (1989). Econometric aspects of input-output analysis. RE Miller et al, 3-21.

Kneip, A., Sickles, R. C., and Song, W. (2012). A New Panel Data Treatment for Heterogeneity in Time Trends. *Econometric Theory, 28*(3), 590-628.

Koop, G. (2004). Dale Poirier. *State Space and Unobserved Component Models: Theory and Applications*, 152.

Koop, G., and Poirier, D. J. (2004). Bayesian Variants of Some Classical Semiparametric Regression Techniques. *Journal of Econometrics*, *123*(2), 259-282.

Koop, G., Ley, E., Osiewalski, J., and Steel, M. F. (1997). Bayesian Analysis of Long Memory and Persistence Using ARFIMA Models. *Journal of Econometrics*, *76*(1-2), 149-169.

Kose, M. A., Prasad, E. S., and Terrones, M. E. (2003). Financial integration and macroeconomic volatility. *IMF Staff papers, 50(1), 119-142.*

Kreft, I. G., and De Leeuw, J. (1998). *Introducing Multilevel Modeling*. Sage.

Liu, J. S., Lu, L. Y., Lu, W. M., and Lin, B. J. (2013). A survey of DEA Applications. *Omega*, *41*(5), 893-902.

Liu, J. S., Lu, L. Y., Lu, W. M., and Lin, B. J. (2013). Data Envelopment Analysis 1978–2010: A Citation-Based Literature Survey. *Omega*, *41*(1), 3-15.

Liu, J., Sickles, R., and Tsionas, E. G. (2013). Bayesian Treatments to Panel Data Models. *Unpublished manuscript, Rice University, Houston*, TX, 4.

Liu, L. M., and Hanssens, D. M. (1981). A Bayesian Approach to Time-Varying Cross-Sectional Regression Models. *Journal of Econometrics*, *15*(3), 341-356.

Mátyás, L., & Sevestre, P. (2008). The econometrics of panel data: fundamentals and recent developments in theory and practice (Vol. 46). Springer Science & Business Media.

Meeusen, W., and van Den Broeck, J. (1977). Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. *International economic review*, 435-444.

Morris, C. N. (1983). Parametric empirical Bayes inference: theory and applications. Journal of the American statistical Association, 78(381), 47-55.

Mwalili, S. M. (2006). Bayesian And Frequentist Approaches to Correct for Misclassification Error With Applications To Caries Research. Department of Mathematics, Katholieke Universiteit Leuven.

Nerlove, M. (2002). The History of Panel Data Econometrics, 1861-1997. *Essays in Panel Data Econometrics.*

Nicholls, D. F., and Pagan, A. R. (1985). 16 Varying Coefficient Regression. *Handbook of statistics, 5*, 413-449.

Osiewalski, J., and Steel, M. F. (1998). Numerical Tools for The Bayesian Analysis of Stochastic Frontier Models. *Journal of Productivity Analysis*, *10*(1), 103-117.

Pitt, M. M., and Lee, L. F. (1981). The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry. *Journal of Development Economics*, *9*(1), 43-64.

Pratt, J. W., and Schlaifer, R. (1984). On the Nature and Discovery of Structure: Rejoinder. *Journal of the American Statistical Association*, *79*(385), 29-33.

Pratt, J. W., and Schlaifer, R. (1988). On the Interpretation And Observation Of Laws. *Journal of Econometrics*, *39*(1-2), 23-52.

Quandt, R. E. (1960). Tests of The Hypothesis That A Linear Regression System Obeys Two Separate Regimes. *Journal of the American Statistical Association*, *55*(290), 324-330.

Raj, B., and Ullah, A. (1981). Econometrics: A Random Coefficient Approach. *Croom Helm.*

Robinson, D. G. (1998). A survey of probabilistic methods used in reliability, risk and uncertainty analysis: Analytical techniques 1 (No. SAND-98-1189). Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

Rosenberg, B. (1973). A Survey of Stochastic Parameter Regression. In *Annals of Economic and Social Measurement, Volume 2, number 4* (pp. 381-397).

Rosenberg, B. (1973). Random Coefficients Models: The Analysis of a Cross Section of Time Series by Stochastically Convergent Parameter Regression. In *Annals of Economic and Social Measurement, Volume 2, Number 4* (pp. 399-428).

Sarris, A. H. (1973). A Bayesian Approach to Estimation of Time-Varying Regression Coefficients. In *Annals of Economic and Social Measurement, Volume 2, Number 4* (pp. 501-523).

Schmidt, P., and Sickles, R. C. (1984). Production Frontiers and Panel Data. *Journal of Business & Economic Statistics, 2*(4), 367-374.

Stevenson, R. E. (1980). Likelihood Functions for Generalized Stochastic Frontier Estimation. *Journal of Econometrics, 13*(1), 57-66.

Swamy, P. A. (1970). Efficient Inference in A Random Coefficient Regression Model. *Econometrica: Journal of the Econometric Society*, 311-323.

Swamy, P. A., and Tavlas, G. S. (1995). Random Coefficient Models: Theory and Applications. *Journal of Economic Surveys, 9*(2), 165-196.

Swamy, P. A., and Tinsley, P. A. (1980). Linear Prediction and Estimation Methods for Regression Models with Stationary Stochastic Coefficients. *Journal of Econometrics, 12*(2), 103-142.

Tsionas, E. G. (2006). Inference in Dynamic Stochastic Frontier Models. *Journal of Applied Econometrics, 21*(5), 669-676.

Van den Broeck, J., Koop, G., Osiewalski, J., and Steel, M. F. (1994). Stochastic Frontier Models: A Bayesian Perspective. *Journal of Econometrics, 61*(2), 273-303.

Wilson, S. E., and Butler, D. M. (2007). A Lot More to Do: The Sensitivity of Time-Series Cross-Section Analyses to Simple Alternative Specifications. *Political Analysis, 15*(2), 101-123.

Wooldridge, J. M. (1995). Score Diagnostics for Linear Models Estimated by Two Stage Least Squares. *Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor CR Rao*, 66-87.

Wooldridge, J. M. (2005). Simple Solutions to the Initial Conditions Problem In Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity. *Journal of Applied Econometrics*, *20*(1), 39-54.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT press.

Yatchew, A. (1998). Nonparametric Regression Techniques in Economics. *Journal of Economic Literature*, *36*(2), 669-721.

Zaman, A. (1996). *Statistical Foundations for Econometric Techniques*. Academic Press.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*, 1st Edition, ISBN-13: 978-0471169376.

Zellner, A. (1986) On Assessing Prior Distributions and Bayesian Regression Analysis with g-Prior Distributions. In: Goel, P. and Zellner, A., Eds., Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti, Elsevier Science Publishers, Inc., New York, 233-243.