School of Economics, Finance and Property

Applications of Information Theory to Economics and Finance

Ranjodh Singh

This thesis is presented for the Degree of

Doctor of Philosophy

of

Curtin University

April 2019

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made.

This thesis contains no material which has been accepted for the award for any other degree or diploma in any university.

Signature :

Date :

Acknowledgement

I am grateful for the privilege of being able to do a doctoral degree. Despite the numerous challenges, it is a blessing to be able to spend a number of years focusing on one's interests.

I want to thank my parents - Baldev Singh and Paramjit Kaur for encouraging me to pursue this degree. They have supported me immensely during this phase of my life. I hope to remind them that their once young wise crow has now grown into an older doubtful crow. To my wife - Rupinder Kaur, I would like to thank her for her patience and for providing *silent* times for me to focus at home. To my children - Akaljot, Sargun and Prabhroop, thank you for asking me what I did at the office all day. You have no idea how much clarity this question gave me. To my brother - Randhir Singh, thanks for asking how the PhD was going on a regular basis. I would like to thank all the relatives who encouraged me to finish this degree. Last but definitely not least to Sohna Singh for providing good company over long walks as well as lots of wet kisses.

I will always be thankful to my supervisor, Associate Professor Felix Chan for him accepting me as his student. I have benefited from our conversations which have spanned many topics over a number of years. I am deeply indebted to him for introducing me to many well known scholars whose work I have always admired. Lastly, I am thankful for all the opportunities/experiences that he was able to provide for me. These definitely helped me to stay in the PhD program. Without these, I would not have been able to finish the degree. Lastly, I am thankful to him for his understanding and patience with regard to my circumstances.

I wish to express my gratitude to Professor Mark Harris for introducing me to Health Economics. I have very much enjoyed working in this area and hope to continue to do so. Professor Harris's suggestions and support have enabled me to finish this work.

I wish to thank Dr Hiroaki Suenaga for all his help with the administrative aspects of the PhD program. This included the candidacy process, annual reports and finally the submission stage. In addition to that, he also provided very valuable feedback/suggestions on some sections of the thesis.

I wish to thank Dr John Gould for encouraging me to begin this path and more importantly believing in me even when I was doubtful. To my friends who have been on a similar journey - J-Han and Joyce. Thank you for providing me with much needed breaks.

I would like to acknowledge the financial support (UPA) from Curtin University. I also received funding for research assistance activities from two ARC projects. I would like to acknowledge that the research undertaken for chapter six was supported by a scholarship through the Population Mobility and Labour Markets project under the Cooperative Research Centre for Remote Economic Participation (CRC-REP). The CRC-REP was managed by Ninti One Limited and funded by the Australian Government's Cooperative Research Centres program.

Finally, I would like to thank the School of Economics and Finance, in particular Professor Helen Cabalu and Ms Sian Flynne for their support and understanding.

Related Thesis Publications

Listed below are the journal articles, conference papers and technical report which form part of this thesis. Specifically, chapter 3 consists of the referred journal article as well as the third conference paper. Chapter 4 contains the second conference paper. Chapter 5 consists of the first conference paper and lastly, chapter 6 consists of the technical report (CRC-REP working paper). I would like to acknowledge that the co-authors contributed immensely by providing valuable feedback and suggestions. However, I was responsible for conducting the data analysis as well as writing and compiling the final publications. Also listed below are conferences and workshops where I presented this research.

- Referred Journal Articles

  1. Singh R.B., Gould J., Chan F., Yang J.W., *Liquidation discount - A novel application of ARFIMA-GARCH, Journal of Empirical Finance*, Volume 36, 2016, Pages 151-161, ISSN 0927-5398,
     https://doi.org/10.1016/j.jempfin.2016.01.012.

- Referred Conference Papers

  1. Chan F., Harris, M. and Singh R. (2015). *Modelling Body Mass Index Distribution using Maximum Entropy Density*. In Weber, T., McPhee, M.J. and Anderssen, R.S. (eds) MODSIM2015, 21st International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2015, pp. 1036-1042. ISBN: 978-0-9872143-5-5.
     https://www.mssanz.org.au/modsim2015/E5/chan2.pdf

  2. Chan F. and Singh R. (2013). *Testing intra-daily seasonality using Maximum Entropy Density*. In Piantadosi, J., Anderssen, R.S. and Boland J. (eds) MODSIM2013, 20th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December

2013, pp.1173–1179. ISBN: 978-0-9872143-3-1.
https://www.mssanz.org.au/modsim2013/F1/chan2.pdf

3. Chan F., Gould J., Singh, R. and Yang, J.W. (2013). *Time series properties of liquidation discount*. In Piantadosi, J., Anderssen, R.S. and Boland J. (eds) MODSIM2013, 20th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2013, pp.1215–1221. ISBN:978-0-9872143-3-1.
https://www.mssanz.org.au/modsim2013/F3/chan2.pdf

- Technical Reports

  1. Dockery AM, Singh R, Harris M and Holyoak N. 2017. *Projecting Aboriginal and Torres Strait Islander populations for remote communities: a small number approach*. CRC-REP Working Paper CW030. Ninti One Limited.

- Presentations

  – MODSIM 2013, Adelaide, Australia.

  – Financial Markets and Corporate Governance 2014, Brisbane, Australia.

  – Australian Health Economics Workshop 2014, Perth, Australia.

  – 28th PhD Conference in Economics and Business 2015, Brisbane, Australia.

  – Curtin Business School PhD Conference 2015, Perth, Australia.

  – MODSIM 2015, Gold Coast, Australia.

  – Econometrics Group Meeting 2016, Hamilton, New Zealand.

  – CRC Workshop 2016, Alice Springs, Australia.

  – School of Economics and Finance PhD Seminar 2017, Perth, Australia.

Abstract

This thesis applies methods from the discipline of Information Theory to selected problems in both Economics and Finance. One concept that is central to Information Theory is Entropy. Entropy aims to measure the degree of uncertainty or randomness in a probability distribution. Rather than arbitrarily choosing a distribution as a suitable model, one could choose a distribution that possesses the maximum uncertainty while satisfying certain characteristics such as integrating to one and having specific moments (supplied by the modeller). A set of these characteristics form the constraints (or conditions) in the optimisation problem which involves maximising Shannon's entropy. The solution to this optimisation problem is known as the Maximum Entropy Density (MED) and it is the most *non-committal* distribution function. As such, it is an excellent choice for modelling random phenomenon. This thesis applies this framework to four different applications across Economics and Finance.

The global financial crisis led to a renewed focus on risk management along with an increased need to develop more robust methods to measure financial risks. Among these risks is liquidity i.e. an investor may find it difficult to convert some of their assets to their true market values in cash. As such, the investor may have to sell at a discount in order to maintain a certain level of liquidity. This thesis develops a model for liquidation discount rate, the discount an investor has to bear in the event of a liquidation sale of a portfolio of Australian stocks. The results indicate that the log liquidation discount rate possesses a long memory property. Hence, an ARFIMA-GARCH model is used to capture its dynamics. The resulting model is used for forecasting future discount rates as well as creating a new measure of liquidation discount-at-risk. This at-risk measure provides fund managers with a likelihood of an expected discount rate. A MED is estimated from the liquidation data and subsequently used to calculate the liquidation discount-at-risk. This process allows fund managers to budget for the future cost of liquidity for a given liquidation horizon and confidence level.

For the second application, this thesis examines the seasonality in financial returns. This topic has been of major interest to financial investors ever since the creation of financial markets. In the existing literature, the majority of the studies focus on the seasonal behaviour of returns in terms of the mean and variance. This behaviour/periodic change for returns is not considered in terms of higher moments. Using a MED, this thesis to generalises the method for detecting seasonal behaviour in return data. The proposed method will enable the analyst to capture seasonal/periodic changes in higher moments. This method is applied to study seasonality in weekdays (day of the week effect) and also for trading hours within a given day (time of the day effect). By comparing MED densities (parameters) across each weekday or time of the day, one can test for differences between different days or for different times of the day. This methodology is applied to examine the seasonal behaviour of returns for foreign exchange rates. The results indicate that Wednesday is significantly different from the rest of the weekdays. Secondly, the 12 p.m. to 2 p.m. time slot is significantly different from the rest of the trading hours in the day. Both these results indicate changes in higher moments and these results have implications for funds management.

For the third application, this thesis proposes a novel multivariate framework for MED. Unlike existing approaches, this framework is able to handle both a high number of variables as well as a high number of moment constraints with relative ease. In addition to this, the framework allows the each MED parameter to be function of one or more exogenous variables. The statistical properties (consistency and asymptotic normality) are provided for this multivariate framework. Next, this framework is used to model the distribution of Body Mass Index (BMI) for an individual given their socioeconomic attributes. From the results, it is evident that these attributes affect different moments of the estimated BMI distribution. These findings can be used to develop policies that reduce obesity levels.

Finally, the thesis proposes a new method of modelling population changes. It is based on the cohort component model whereby the individuals in a given age group category transition to the next age group category over time. Based on this, the change in population for each age group is calculated and is modelled using a varying limit censoring regression model. Prior to the modelling stage, the Kullback-Leibler (KL) divergence measure is used to assess if there is difference in the population across two different time periods. These methods are applied to model indigenous populations in regional and remote Australia. The results indicate that the distributions differ significantly over time.

Additionally, the model predicts an increase in the overall indigenous population by approximately 15% from 2011 to 2016. In the absence of population forecasts for regional and remote communities in Australia, these results bring direct benefits to researchers and planning agencies. The results also have important ramifications for services providers such as housing, health, education and infrastructure.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The seminal paper by Shannon (1948) introduced an information criterion known as Entropy which measures the degree of uncertainty or randomness in a probability distribution function. Jaynes (1957) used this concept to develop the Principle of Maximum Entropy. This method consists of maximizing an objective function (entropy) subject to a number of constraints. These constraints attempt to capture some of the desired properties of the probability distribution itself. Such as the fact that the distribution function has to integrate to one. This also includes constraining the distribution to have a particular mean. More generally, these constraints can specify all the raw moments (up to certain order) that the distribution should contain. These moments are supplied by the modeller and as such are inputs into the optimisation problem. Solving this problem yields a maximum entropy density (MED). This is a method of estimating a probability distribution using a set of moment conditions which represent the only relevant information the modeller knows. According to Jaynes (1957), the resulting distribution is the most non-committal among all other distributions i.e. it is the optimal distribution with contains the desired properties as well as possesses the most amount of uncertainty. Given this property, the principle of Maximum Entropy can be applied to many problems across various disciplines including Physics, Engineering, Chemistry, Biology, Ecology, Computer Science and Economics.

This thesis attempts to apply the maximum entropy framework to problems in Economics and Finance. In particular, the work presented here is related to a field known as Information and Entropy Econometrics (IEE). Golan and Perloff (2002) states that this field builds on the foundations of Information Theory and the Principle of Maximum Entropy.

Examples of papers in this field include Maasoumi (1993) which provides a detailed introduction to the principle of maximum entropy as well as the applicability of the method to problems in Economics and Econometrics. The main focus is on highlighting the applicability of information theoretic measures in modelling economic inequality and market concentration. Maasoumi and Racine (2002) studied stock returns by using a entropy measure which captures dependence. Moreover, this entropy measure is able to capture non-linear dependence unlike traditional measures. Rockinger and Jondeau (2002) used the entropy principle to develop a method to model time varying conditional moments. Using this idea, Chan (2009) proposed a more computationally efficient method and modelled the structure of the moments in terms of the MED parameters. Wu (2003) uses the maximum entropy density to approximate the distribution of income. Here the moment constraints are introduced sequentially (rather than simultaneously) in order to assess the changes the resulting distribution. This allows one to assess the impact of each additional moment on the resulting distribution. Golan and Maasoumi (2008) provide an overview of information theoretic and entropy methods and their applicability to problems in Portfolio Analysis and Finance, Empirical Likelihood estimation and the non-parametric methods. For more recent developments, refer to Armstrong et al. (2019), Gao and Han (2019) and Handley and Millea (2019).

This thesis attempts to continue this line of research given the challenging nature of problems in Economics and Finance as well as the flexibility of the maximum entropy framework. A brief summary and contribution for each chapter in this thesis is provided in the paragraphs below.

Chapter 2 provides the necessary background to the main concepts/terms discussed in the thesis. This includes the formulation of the principle of maximum entropy and the derivation of the solution i.e. MED. The existence and uniqueness conditions for the solution are discussed. Next, the techniques used to estimate the MED in subsequent chapters are outlined. For the sake of completeness, the alternative formulations of the maximum entropy problem are briefly discussed. The necessary background on Kullback-Leibler Divergence is provided. This quantity is used to assess the divergence in two distributions in chapter 6 of the thesis.

Chapter 3 introduces a novel approach to measure liquidation discount for a given portfolio. In other words, what discount rate can one expect from a fire sale of a portfolio of stocks? Does the discount rate depend on the time of the day? Does the size of the portfolio to be liquidated affect the discount rate? These are some of the questions this study attempts to answer. In addition to answering these questions, prediction tools are developed to assist prudent fund managers in the liquidation process. It is expected that these will enable them to achieve optimal outcomes as well as comply with regulatory standards. This approach takes advantage of the order book information to compute the discount factor. This is done for a different sized (small,medium and large) portfolios over a time period. The resulting series is known as a liquidation discount rate time series. An important finding is that the log liquidation discount rate series has a long memory property. This evidence implies that the current liquidation discount rates are affected by discount rates from very early periods. Using this new finding, a forecasting framework was constructed in order to predict future liquidation discount rates. Furthermore, a liquidation discount-at-risk measure was formulated to measure the level of uncertainty of a predicted discount rate. A MED was used to provide benchmark values for this risk measure. This is of direct benefit to fund managers who are able to budget for future discount rates rather than simply accepting the rate provided by the market on the day of the sale.

Chapter 4 develops a novel method for detecting seasonality in financial returns. Specifically, this involves comparing estimated MEDs across different time segments in order to detect periodic changes. The proposed method provides a more richer/thorough method of examining seasonality given that the comparison is done at a distribution/density level. Existing studies have focused on singular measures such as mean or variance in order to explain seasonal behaviour. Furthermore, this method becomes even more relevant where the seasonal behaviour is also present in higher moments i.e. cannot be detected using mean and/or variance. The results in this chapter indicate that this scenario occurs in high frequency data. The findings from this chapter have implications for financial trading and funds management.

Chapter 5 introduces the multivariate form of the maximum entropy problem. In addition to that, the framework also allows for the MED parameters to be functions of exogenous variables. As a result, the exogenous/independent variables affect the shape and scale of the estimated MED. Both of these are relatively novel additions. To date, the existing literature does not contain this exact formulation. Notable mentions are attempts to model the joint distribution of the variables using Copulas. To ensure that the estimator is unbiased and well behaved, consistency and asymptotic normality proofs are provided for the multivariate framework. An empirical application of the proposed framework is used to model the distribution of Body Mass Index (BMI) of an individual given their socioeconomic attributes or risk factors. This is achieved by estimating a conditional maximum entropy density. The conditioning variables are the socioeconomic attributes pertaining to the individual. Previous studies have attempted to model the average/mean level of BMI for an individual given their risk factors. The advantage of the proposed methodology is that the impact of risk factors can be studied on different aspects (moments) of the distribution in addition to the mean. This is an extension to the *standard* models which estimate impact on BMI at the mean level. From this study, it is evident that some risk factors affect higher moments (in addition to the mean) of an individual's BMI distribution. For example, under the proposed framework, the researcher can measure the impact a risk factor has with regard to the variance of an individual's BMI distribution. In fact, this can be extended to higher moments such as the skewness and kurtosis. Changes in risk factors are examined closely to assess the impact on the entire BMI distribution, in particular the tail regions which corresponds to the probability of obesity. Given the national as well as global movement to combat obesity levels, policy makers need to not only identify the risk factors, but also measure their impact. The analysis illustrated in this chapter measures the impact of risk factors with regard to obesity in a clear and concise manner. It is expected this will lead to policies that tackle the burden of disease caused by obesity.

Chapter 6 develops a novel method of modelling the change in population over time. This consists of three stages. In the first stage, the Kullback-Leilber divergence measure is used assess if the population distribution has changed over time. If this measure detects a change, then the second stage involves modelling this change in for each age group in the population. The approach used here is based on the cohort-component model where individuals belonging to one age group transition to another next consecutive age group over time. The change in population observed for each age group is modelled using a varying limit censoring regression model. In the third stage, this model is used to produce forecasts for the next time period. This methodology was applied to examine trends in indigenous populations living in regional and remote Australia. The projections produced by the model are of direct benefit to researchers as well as government agencies since currently no projections are available below the state level. Given these forecasts, policy makers and planners will now be able to properly gauge the level of services (housing, health, education and infrastructure) required by each community. This will led to a sustainable economic and social environment for all regional and remote communities.

Chapter 7 summarises the results of each chapter in the thesis. This chapter also lists some potential future work to be carried out in this area.

Chapter 2

Background

This chapter contains necessary background material required for the subsequent chapters. This includes an introduction to the Principle of Maximum Entropy as presented by Jaynes (1957). Solving this optimisation problem yields the Maximum Entropy Density. A step by step solution of this problem is presented. Next, the conditions of existence and uniqueness are outlined. A brief section on the estimation is also presented. Section 2.0.4 outlines the flexibility of the Principle of Maximum Entropy. This is achieved by modifying the moments constraints as well the objective function. This allows a variety of different classes of MED distributions. Section 2.0.5 introduces a well known entropy based measure - the Kullback-Leibler Divergence. The definition as well as some properties of this measure are outlined. This divergence measure is applied to problems discussed in Chapter 6. Lastly, the multivariate form of the principle of maximum entropy is introduced.

2.0.1 Principle of Maximum Entropy

Shannon (1948) proposed the idea of entropy as a measure of the amount of uncertainty or randomness. The continuous random variable version of the entropy presented in Shannon's seminal paper is defined as

$$E = -\int_{\mathbf{A}} f(y) \log f(y) \, dy \tag{2.1}$$

where $f(y)$ denotes a probability density function and $\mathbf{A}$ is the region of integration. Jaynes (1957) proposed the principle of Maximum Entropy. This consists of maximising the entropy functional subject to certain moment constraints. These constraints consist of the desired properties of the density function. This problem can be presented in a non-linear optimisation framework:

$$\text{max. E} = -\int_{\mathbf{A}} f(y) \log f(y) dy. \tag{2.2}$$

subject to

$$\int_{\mathbf{A}} \quad f(y)\, dy = 1. \tag{2.3}$$

$$\int_{\mathbf{A}} \quad y^{\ell} f(y)\, dy = \mu_{\ell} \text{ where } \ell = 1, 2, ..., k. \tag{2.4}$$

Here, μ_{ℓ} represents the ℓ^{th} raw moment of the density. The first constraint ensures that the resulting density function integrates to 1 over a region **A**. The remaining k constraints specify the raw moments of the density. The solution to this problem yields a density that is known as the Maximum Entropy Density (MED). The resulting MED can be interpreted as the density that possesses the most amount of uncertainty[1] compared to all other densities that satisfy the given constraints.

Solving this problem using Lagrange's method yields the Hamiltonian function:

$$M(f) = -\int_{\mathbf{A}} f(y) \log f(y)\, dx + \lambda_0' \left[\int_{\mathbf{A}} f(y)\, dy - 1 \right] + \sum_{\ell=1}^{k} \lambda_{\ell} \left[\int_{\mathbf{A}} y^{\ell} f(y)\, dy - \mu_{\ell} \right].$$

Maximising the function $M(f)$ yields

$$f(y) = \exp(\lambda_0) \exp\left(\sum_{\ell=1}^{k} \lambda_{\ell}\, y^{\ell} \right)$$

where $\lambda_0 = \lambda_0' - 1$. Here $\exp(\lambda_0)$ can be expressed as

$$\exp(\lambda_0) = \left[\int_{\mathbf{A}} \exp\left(\sum_{\ell=1}^{k} \lambda_{\ell}\, y^{\ell} \right) dy \right]^{-1} = Q.$$

The quantity Q denotes the normalising constant which ensures that the first constraint is satisfied. Hence the MED can be expressed as

$$f(y) = Q^{-1} \exp\left(\sum_{\ell=1}^{k} \lambda_{\ell}\, y^{\ell} \right). \tag{2.5}$$

[1] In Jaynes (1957) paper, the phrase *most non-committal* is used.

From the above derivation, one can see that the MED is a Generalised Exponential distribution (equation 2.5). The λ_ℓ values represent parameters of the MED. These parameters are responsible for controlling different aspects of the distribution such as shape, scale and location of the distribution. Special cases of this family include the exponential and normal distribution (See Proposition 1 in Chan (2009)). With the higher number of parameters, this distribution allows a great deal of flexibility with regard to shape of the distribution. This property is especially desirable from a modelling perspective.

Different values of μ_i will produce different λ_ℓ values and this will create distributions which vary with respect to their mean, variance, skewness, kurtosis etc. Hence, the information from raw moments is transformed into MED parameters. The MED parameters (λ) are essentially non-linear functions of the raw moments that form the constraints (See Proposition 2 of Chan (2009)). This following result captures the relationship between the MED parameters and the raw moments:

$$\frac{\partial \lambda_\ell}{\partial \mu_i} = (\mu_{\ell+i} - \mu_\ell \mu_i)^{-1} \ \forall\ \ell,\ i = 1, \ldots, k. \tag{2.6}$$

From the above expression, it is evident that higher moments must exist for the result to hold. Even though the constraints consist of k moments, based on the result above the μ_{2k} must exist.

2.0.2 Existence and Uniqueness

Given the non-linear nature of the optimisation process, one is required to verify the existence of solution i.e. the resulting MED. The MED problem is connected to the classical moment problem in probability theory. For a given sequence of moments, does there exist a probability measure which satisfies those moments and is it unique. Note, that classical moment problem is divided into three problems - the Hamburger moment problem, the Hausdorff moment problem and the Stieltjes moment problem. Each of these problems are characterised by different support for the measure.

Frontini and Tagliani (1997) provides the existence conditions required for the MED. This existence condition is related to the positivity of the Hankel determinants. A Hankel matrix (H_k) is matrix that consists of moments as elements. A Hankel matrix of order ℓ is written as

$$H_\ell = \begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_\ell \\ \mu_1 & \mu_2 & \cdots & \mu_{\ell+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_\ell & \mu_{\ell+1} & \cdots & \mu_{\ell+\ell} \end{pmatrix}. \tag{2.7}$$

The necessary and sufficient condition for the existence of a maximum entropy solution is that the determinants of the Hankel matrix is positive i.e. $|H_\ell| > 0$ for $\ell = 0, 1, ..\frac{k}{2}$. Note the k is the total number of moments considered in the optimisation This result is true for the Hamburger case. For the Stieltjes case, the existence conditions are the same albeit using a modified definition of the Hankel determinant. For further details, please refer to Frontini and Tagliani (1997). There is an alternative proof for existence by Mead and Papanicolaou (1984). It involves investigating the properties of a *potential*, Γ. This is a Legendre transformation which expresses a function in terms of μ and λ.

$$\Gamma = \ln Z + \sum_{n=1}^{N} \mu_n \lambda_n$$

Here, Z denotes the normalising constant of the MED. Firstly, this potential is shown to be convex everywhere and furthermore it has a unique absolute minimum if the finite moment sequence is monotonic.

Given that a solution exists, there is a possibility that it may not be unique. In other words, solving a non-linear optimisation problem could yield multiple solutions. The existence condition stated above do not guarantee uniqueness of the solution. The uniqueness relies on the asymptotic behaviour of the ratio between a Hankel determinant and a *adjusted* Hankel determinant [2]. Frontini and Tagliani (1997) use the concept of entropy-convergence to prove uniqueness of solution. Zellner and Highfield (1988) also showed the uniqueness of the resulting MED (equation 2.5). This was done by relating the first order conditions of both the optimisation problem (equation 2.2) and the maximum likelihood problem. The results relating existence and uniqueness of the MED are quite important, especially with regard to empirical work.

Another aspect with regard to empirical work which needs to be considered is the existence of the moments itself. The constraints (equations 2.3 and 2.4) in the optimisation problem require the existence of population moments (up to order k). In empirical studies, one only has the luxury of a sample. A natural question to ask is that can the sample

[2] This determinant is computed from a Hankel matrix where the first column and row are deleted

provide any information about the existence of the population moments. Although a large sample may help alleviate this issue, it still cannot guarantee the existence of population moments. The result by Hill (1975) can be used to verify the existence of the highest sample moment available for a given *iid* sample. This method estimates the tail index and can be extended to include dependence within the sample. Subsequently, if the sample moments do exist (up to a certain order) then one can argue that they are consistent estimators of the population moments i.e. $\hat{\mu}_\ell \rightarrow \mu_\ell$. Hence, under this condition sample moments can be substituted in place of population moments.

2.0.3 Estimation

One possible method of estimating the MED parameters is via Maximum Likelihood Estimation (MLE). The MLE procedure takes advantage of the fact that the structure of the MED is known and as such one can write the log-likelihood of the MED and estimate its parameters. The log-likelihood for the MED derived above (eq. 2.5) is given by

$$L(\lambda_\ell; y_i) = \sum_{i=1}^{n} \log f(y_i) = -n \log Q + \sum_{i=1}^{n} \sum_{\ell=1}^{k} \lambda_\ell \, y_i^\ell$$

Here the aim is to find values of λ_ℓ that maximise $L(\lambda_\ell; y_i)$. In order for such a solution to exist, L must satisfy a set of regularity conditions. Given that the resulting MED is part of the exponential family, such results have already been proven in the wider literature (see Barndorff-Nielsen (2014)).

For computational convenience, the first order derivatives are computed. Given the estimation framework above, one can apply this data driven technique to estimate a MED. In doing so, the estimated MED will closely match the characteristics of the sample that was observed. This allows the researcher to estimate a density that can adapt/accommodate a wide variety distributional aspects such as skewness, fat-tails and even multi-modality.

Another computational method which can be used to estimate the MED parameters (λ_ℓ) involves solving a set of non-linear equations:

$$\begin{aligned}
\int_A f(y) &= 1 \\
\int_A y\, f(y) &= \hat{\mu}_1 \\
\int_A y^2\, f(y) &= \hat{\mu}_2 \\
&\vdots \\
\int_A y^\ell\, f(y) &= \hat{\mu}_\ell \\
&\vdots \\
\int_A y^k\, f(y) &= \hat{\mu}_k
\end{aligned}$$

Here $f(y)$ denotes the MED (eq. 2.5) and $\hat{\mu}_\ell$ denotes the estimated ℓ^{th} population moment. Note that the integration bounds are specified by the user. This flexibility allows the estimation to be carried out in the appropriate range as per the context. Solving the set of above equations will yield λ_ℓ for $\ell = 1 \ldots k$. For $k \leq 2$, it is possible to derive the solution of the parameter analytically. However, no closed form expressions exist for $k > 2$ (Rockinger and Jondeau (2002)). Hence, a numerical routine is required.

2.0.4 Alternative Distributions

The objective function i.e. Shannon Entropy can be replaced by an alternative entropy such as the Renyi or Tsallis Entropy. As such the resulting MED would differ depending on which objective function was used. The number of moment constraints (equations 2.3 and 2.4) used also affects the final form of the the resulting MED. For example, only using the first constraint (density must integrate to 1) produces a uniform distribution as a MED. This is expected since, in the absence of any other information (higher order moments), the MED allocates an equal probability across all possible values of the random variable. As additional information is added in the form of more moment constraint/s, the resulting MED shifts away from the uniform distribution. Other modifications include changing each of the moment constraints themselves. Instead of using the raw moment constraints, one could use log moment constraints or absolute value moment constraints. In its most general form, a constraint can be expressed as

$$\int_{\mathbf{A}} g(y)\, f(y)\, dy = h(\mu_\ell) \text{ where } \ell = 1, 2, \ldots k.$$

Different modifications of the constraints yield different MEDs. Park and Bera (2009) provide a table containing the resulting MED for different moment conditions. Examples of these MEDs include the Exponential, Normal, Log-Normal, Beta, Gamma, Wishart and various other distributions. Given the variety of distributions that can be accommodated in this framework, it provides the flexibility required for empirical work. However, the existence and uniqueness of solution have only been proven for the raw moment constraints with Shannon's Entropy as the objective function.

2.0.5 Kullback-Leibler Divergence

Suppose that one is attempting measure how to *far apart* or *different* two distributions are from each other. Given two discrete distributions $p(i)$ and $q(i)$, this concept can be expressed as the following hypothesis:

$$H_0 : q_i = p_i$$
$$H_1 : q_i \neq p_i$$

The Kullback-Leibler (KL) Divergence (Kullback and Leibler (1951)) measures the divergence between two probability distributions. In other words, it measures how much a given distribution diverges from a *benchmark* distribution. The KL divergence can be expressed as

$$KL(p|q) = -\sum_i p_i \log \frac{p_i}{q_i} \tag{2.8}$$

If the two distributions are similar, then the KL divergence would be close to zero. As they grow apart the divergence measure would significantly differ from zero. This critical values of KL divergence can be obtained from the double F-distribution under normality. The KL divergence measure is also closely linked to the chi-squared goodness of fit test statistic.

If the benchmark distribution was a discrete uniform distribution, then the above expression is equivalent to the Shannon Entropy. The KL divergence presents a more general case. This measure is always non-negative i.e. $KL(p|q) \leq 0$. However, it does not satisfy all the properties of a distance metric. For example, swapping the given distribution with the benchmark distribution will not yield the same result as the original metric i.e. $KL(p|q) \neq KL(q|p)$.

Chapter 3

Time Series Properties of Liquidation Discount

3.1 Introduction

Investors sell assets for two primary reasons. The first reason is purely utilitarian i.e. cash flow from proceeds of the sales are required by the investor for personal/business reasons. The second is speculation. Here, the investor does not require the cash, however may feel that the price of asset may fall as such selling it may be pre-emptive. In the first case, the investor may have to bear a small loss due to the urgent nature of the sale. As such, accessing stock market liquidity entails a real cost for investors. This chapter presents a practically-oriented ”liquidation discount” measure for the cost of stock market liquidity. Our perspective is that of a stock market investor with a long spot position in a ”large” portfolio who has concern for the possible need to liquidate the portfolio in the future. Our liquidity cost measure, liquidation discount, is the market impact discount in value yielded by the *instantaneous* sale of the portfolio in one parcel, relative to its in-hand market value calculated from the prevailing pre-liquidation market prices of the portfolio components. For a notional market order to sell a large portfolio in one parcel, this study empirically identifies the instantaneous liquidation discount (due to market impact) *would have been* each day of the sample period. By modelling the time series behaviour of liquidation discount, this study is able to forecast the level and uncertainty of future liquidation discount.

Executing a large market order sale in one parcel might entail considerable market impact cost (representing an average price discount for immediate liquidity), but will benefit from a high degree of average price certainty and cash flow immediacy. To mitigate market impact cost, but with some loss of price certainty and cash flow immediacy, a large trade may be split up into smaller parcels and submitted to the market gradually over some interval of time (Chan and Lakonishok (1993) and Chan and Lakonishok (1995)). Bertsimas and Lo (1998) derive portfolio transaction strategies that optimise the trade-off between market impact cost and the risk associated with price volatility over the execution horizon. Engel et al. (2012) introduce ”liquidation value at risk” to assess the price risk versus market impact cost trade-off for different order execution approaches. To ensure the empirical credibility of the liquidation discount time series, this study disregards the flexibility to split the notional sell order into multiple parcels and hence avoid the need to assume any debatable theoretical extrapolation of the price and liquidity ripple effects that would be transmitted from any one trade parcel to subsequent trade parcels.

A portfolio's notional single-parcel liquidation discount is explicitly measurable in an order driven market with an open limit order book, and is negatively indicative of the instantaneously available stock (i.e., quantity or depth) of order book liquidity in the form of bid-side limit orders. It is to be emphasised that this study is not presenting or suggesting an optimal liquidation strategy. The objective to this study is to obtain a meaningful and quantifiable measure of liquidity cost. Our liquidity cost measure and its time series dynamics serve as an empirically verifiable benchmark for other liquidation strategies that might have more opaque liquidity costs such as multi-parcel or private (e.g., dark pool) transactions.

The Australian Securities Exchange (ASX) is an order driven market with an open limit order book. From 2006 to 2011, this study measures the daily morning and afternoon liquidation discount for variously-sized value-weighted portfolios of the leading 10 stocks listed on the ASX. The properties of the log transformations of these time series are modelled with the Autoregressive Fractionally Integrated Moving Average-Generalized Autoregressive Conditional Heteroskedasticity (ARFIMA-GARCH) model first developed by Baillie et al. (1996). That is, the mean component of each series is modelled using the long-memory ARFIMA model (Granger and Joyeux (1980) and Hosking (1981)), including ARMA and ARIMA as special cases (with d=0 and d=1 respectively); and the variance component of each series is modelled with the GARCH model (Bollerslev (1986)). Model performance is assessed in terms of out-of-sample forecasting mean squared error (MSE).

For the sake of model parsimony and generalisability across the various log liquidation discount time series, with only modest compromise in terms of forecasting MSE performance, this study proposes a single parsimonious ARFIMA(1,d,0)-GARCH(1,1) model. Using this model, a liquidation discount-at-risk measure is formulated using MED procedure mentioned in Chapter 2. The estimated MED provides critical values required in the calculation of liquidation discount at risk. This metric could allow portfolio managers to budget for the future cost of portfolio liquidity for a chosen liquidation horizon and confidence level.

This is a step towards improved recognition of liquidity risk within analytical models of general portfolio risk. The Bank for International Settlements (2001) (promotes such development of ”risk assessments that take account of market liquidity” and consideration of ”how such measures could be used in the disclosure of market risk”. For example, from the sample, when the cost of liquidity was at its highest during the onset of the Global Financial Crisis in late 2008, the manager of a value-weighted top-10 Australian stock portfolio worth about \$3 million would have been able to use the proposed model to budget for a cost of liquidity less than or equal to 0.5% (i.e., about \$15,000 or less) for a five trading day liquidation horizon, with 99% confidence.

3.1.1 Practical measurement of stock market liquidity

In an order driven market such the way that is considered in this study, bid [1] and ask [2] limit orders (submitted to the limit order book [3]) supply liquidity by offering the option to transact at prices somewhat less favourable than the prevailing market price, and market orders consume liquidity by exercising limit order options (thereby creating trading volume). The cost for liquidity suffered by market orders, in terms of an unfavourable transaction price, is compensation to limit orders for the risk of transacting with informed market orders. See O’Hara (1997) for an overview of information-based models of market microstructure.

Liquidity metrics determined from limit order book characteristics are variously reflective of the interconnected ”depth” and ”width” dimensions of liquidity as per the nomenclature of Harris (1990): depth being indicative of the trade size that can be accommodated at some degree of disadvantage to the market price; and width being indicative of the price disadvantage against the market price (i.e., market impact) that must

[1] potential buyers *bid* price

[2] potential sellers *asking* price

[3] list of bid and ask prices listed in descending order

be suffered for a given trade size. Sarr and Lybek (2002), Aitken and Comerton-Forde (2003) and Goyenko et al. (2009), for example, provide summary and comparison of a wide variety of approaches to measuring stock market liquidity. Aitken and Comerton-Forde (2003) find that order-based measures of liquidity, which, to some extent, take into account limit order book conditions, provide a better proxy for liquidity than trade-based measures reflective of trading volume. Couched in this terminology, conditional on a notional portfolio sell order size, the proposed liquidation discount measure is an average width measure weighted with respect to both the portfolio composition and the limit order depth stratification down through the price steps of the bid-side order books of each portfolio component. More simply and practically stated, liquidation discount is the discount in value yielded by the immediate sale of the portfolio relative to its pre-liquidation market value.

The proposed liquidation discount measure and its time series are purposed to be an unambiguous empirical reflection of the practical bid-side liquidity cost faced by investors with long equity portfolios (such as mutual funds). In contrast, measures that aim to summarise two-sided (bid and ask-side) market liquidity, while potentially useful as indicators of market conditions, tend not to be directly representative of practical trading concerns. For example, Aitken and Comerton-Forde (2003) review three order-based measures of liquidity. The first, time-weighted relative bid-ask spread, only reliably reflects the relative liquidity cost that can be expected within some period of time for an instantaneous round-trip trade in a small parcel of stock. The second, relative depth, is the maximum proportion of a stock's shares on issue that can be simultaneously both bought and sold for an unknown net position and at an unknown disadvantage to the market price. The third, ”new liquidity measure” (p56), uses historical limit order execution rates as weightings in the averaging of the values of all standing bids and all standing asks separately, and then combines the weighted bid and weighted ask values for a mid-point measure. None of these measures are directly representative of the practical liquidity requirements of market participants.

Cao et al. (2009) formulate a two-sided width measure built from bid and ask-side measures that are directly comparable with the proposed liquidation discount concept. Cao et al. (2009)'s demand side price impact measure, $LD(q)$ (see their equation (14)), formulates the average absolute (i.e., dollar) discount per share of a single-stock market sell order for q shares; whereas the proposed liquidation discount specification is a relative measure applied in a portfolio context. In similar fashion Cao et al. (2009) also specify a supply side price impact measure, which they combine with $LD(q)$ to obtain ”the scaled imbalance in price impact” (p32). Cao et al. (2009) do not concern themselves with forecasting liquidity or liquidity risk as this study does, but instead they demonstrate the role of the limit order book in price discovery.

Lillo and Farmer (2004) investigate the relationship between the flow of liquidity-consuming market orders and the stock of liquidity at the best bid or best ask. They specify ”relative liquidity” to be the ratio of market order size to the best-price limit order volume available to fill the market order. That is, Lillo and Farmer (2004)’s relative liquidity measure reflects demand for liquidity scaled by the best-price available supply of liquidity; whereas the proposed liquidation discount measure reflects the cost of accessing the bid-side supply of liquidity, inclusive of and beyond the best-bid price. Lillo and Farmer (2004) find that both the buy versus sell direction of market orders and relative liquidity follow long-memory processes in an anti-correlated way that makes the identified predictability of market order direction difficult to exploit for profit. This study finds that liquidation discount also demonstrates long-memory.

Demand for and supply of liquidity is often dependent on portfolio or market return performance Hameed et al. (2010). The manager of a poorly performing portfolio may be forced into a fire sale of the portfolio as a consequence of margin calls or capital withdrawals (Brunnermeier and Pedersen (2009); Shleifer and Vishny (2011)), leading to excess demand for liquidity feeding back into further downward price pressure and further demand for liquidity. As per the price pressure hypothesis of Scholes (1972), market prices can deviate from their information-efficient values as a consequence of imbalances in the demand for and supply of liquidity. The large negative cumulative average abnormal returns associated with mutual fund fire sales measured by Coval and Stafford (2007), for example, superficially seem not to concur with this study’s comparatively modest observations of liquidity cost: across the various portfolio scenarios this study measures a maximum liquidation discount of only about 0.5%, whereas Coval and Stafford (2007) report a fire sale cost of liquidity of the order of 14%. However, Coval and Stafford (2007) measurement indicates the cumulative cost of liquidity over a period of time conditional on (poor) portfolio performance, whereas the proposed liquidation discount time series distill the instantaneous cost of liquidity through time regardless of portfolio or market return performance.

For a given portfolio (which may in fact be a sub-portfolio/parcel of a larger portfolio), the portfolio manager need only check the current and lagged liquidity conditions for that portfolio to then use the proposed method to forecast the cost of liquidity for a chosen liquidation horizon. Despite the sample including the perturbation of the Global Financial Crisis, the log liquidation discount time series are all found to be stationary.If liquidation is actually exercised and a second portfolio is to then be considered for liquidation, the effect of the first portfolio's liquidation will be reflected in an updated observation of current liquidity conditions. A series of portfolio sales concentrated in time during a period of high liquidity cost (as would typically be the case for a fire sale) would entail a series of positively correlated liquidation discounts that could easily amount to a cumulative cost of liquidity of order to match Coval and Stafford (2007) measurement.

3.2 Liquidation discount

For a portfolio of N stocks, at time t for stock $x \in \{1, 2, ..., N\}$, s_t^x is the number of shares of stock x held in the portfolio. The bid-side limit order book for stock x is represented by length-m vectors $\mathbf{q}_t^x$ and $\mathbf{b}_t^x$:

$$\mathbf{q}_t^x = \begin{bmatrix} q_{t,1}^x \\ q_{t,2}^x \\ \vdots \\ q_{t,m}^x \end{bmatrix} \text{ and } \mathbf{b}_t^x = \begin{bmatrix} b_{t,1}^x \\ b_{t,2}^x \\ \vdots \\ b_{t,m}^x \end{bmatrix}$$

where $q_{t,1}^x$ is the number of shares for limit order purchase at the highest (best) bid price $b_{t,1}^x$; $q_{t,2}^x$ is the number of shares for limit order purchase at the second highest bid price $b_{t,2}^x$; ...; and $q_{t,m}^x$ is the number of shares for limit order purchase at the lowest (mth-best) bid price $b_{t,m}^x$. Additionally,

$$q_{t,0}^x = 0 \text{ and } b_{t,0}^x = 0.$$

Liquidation of s_t^x shares will "consume" the bid-side order book to depth $d_t^x \in \{0, 1, ..., m-1\}$ such that:

$$\sum_{i=0}^{d_t^x} q_{t,i}^x \leq s_t^x < \sum_{i=1}^{d_t^x+1} q_{t,i}^x.$$

Therefore the liquidated value of s_t^x shares, lv_t^x, is given by:

$$lv_t^x = \sum_{i=0}^{d_t^x} q_{t,i}^x b_{t,i}^x + \left(s_t^x - \sum_{i=0}^{d_t^x} q_{t,i}^x \right) b_{t,d_t^x+1}^x \,.$$

The market value of s_t^x shares, mv_t^x, is to be determined from the highest bid price plus a half-tick [4]:

$$mv_t^x = s_t^x \left(b_{t,1}^x + \frac{tick_t^x}{2} \right).$$

The portfolio liquidation discount, ld_t, is therefore:

$$ld_t = 1 - \left(\frac{\sum_{x=1}^{N} lv_t^x}{\sum_{x=1}^{N} mv_t^x} \right). \tag{3.1}$$

The equation (3.1) formulation for portfolio liquidation discount can be more intuitively summarised as the following ratio measure:

$$\text{portfolio liquidation discount} = \frac{\text{market value} - \text{instantaneous liquidation value}}{\text{market value}}.$$

Previously this has been described as a liquidity cost measure, however another interpretation of liquidation discount is that it indicates the relative amount of a portfolio's value that is (instantaneously) illiquid. Compare this to a debt-to-assets corporate leverage ratio. Just as a leverage ratio indicates the relative amount of a firm's asset-value that is financially inflexible, liquidation discount indicates the relative amount of a portfolio's value that has redemption inflexibility. Hence, similar in context to the way a leverage ratio is a measure of a firm's financial risk, liquidation discount can be considered a measure of a portfolio's liquidity risk. For the purpose of analysis, a log transformation is applied to portfolio liquidation discount ld_t (equation (3.1)) i.e. let $\ell_t = \log\ ld_t$.

Data for this study pertain to stocks listed on the Australian Securities Exchange (ASX). The ASX is a purely order driven market with an open limit order book. In January of each year from 2006 to 2011 this study identifies the top 10 (highest weighted) stocks from the ASX/S&P200 index for inclusion in the notional portfolio. This 10-stock portfolio is, over time, variously constituted by 14 unique stocks (see Table 3.1). For these 14 different stocks daily time series of shares on issue are obtained from the Morningstar DatAnalysis database, and twice-daily (at 10:15, shortly after market open, and 15:45, shortly before market close) "snap-shots" of the bid-side limit order book from the Securities Industry Research Centre of Asia-Pacific (SIRCA) AusEquity database, for the period October 2006 to October 2011. This is achieved by identifying the starting

[4] Tick size is the smallest increment/movement by which a stock price can move. For the ASX market, this amount is \$0.01.

Table 3.1: Annual portfolio composition by company, in millions of shares, where $\gamma \in \{0.0001\%, 0.0002\%, \ldots, 0.0008\%\}$ is the portfolio size factor.

Company ticker	Jan 2006	Jan 2007	Jan 2008	Jan 2009	Jan 2010	Jan 2011
BHP	$3{,}590\gamma$	$3{,}496\gamma$	$3{,}356\gamma$	$3{,}356\gamma$	$3{,}356\gamma$	$3{,}356\gamma$
CBA	$1{,}289\gamma$	$1{,}290\gamma$	$1{,}316\gamma$	$1{,}471\gamma$	$1{,}534\gamma$	$1{,}549\gamma$
WBC	$1{,}831\gamma$	$1{,}851\gamma$	$1{,}878\gamma$	$2{,}880\gamma$	$2{,}974\gamma$	$3{,}005\gamma$
ANZ	$1{,}831\gamma$	$1{,}839\gamma$	$1{,}920\gamma$	$2{,}158\gamma$	$2{,}533\gamma$	$2{,}596\gamma$
NAB	$1{,}597\gamma$	$1{,}630\gamma$	$1{,}634\gamma$	$1{,}870\gamma$	$2{,}118\gamma$	$2{,}170\gamma$
WOW	$1{,}164\gamma$	$1{,}204\gamma$	$1{,}215\gamma$	$1{,}225\gamma$	$1{,}240\gamma$	$1{,}213\gamma$
TLS	$5{,}997\gamma$	$10{,}245\gamma$	$10{,}319\gamma$	$12{,}443\gamma$	$12{,}443\gamma$	$12{,}443\gamma$
WDC	$1{,}749\gamma$	$1{,}773\gamma$	$1{,}942\gamma$	$1{,}965\gamma$	$2{,}308\gamma$	
RIO	456.8γ	456.8γ	456.8γ		606.8γ	435.8γ
QBE		818.6γ	886.1γ	986.6γ		
WES					$1{,}005\gamma$	$1{,}005\gamma$
WPL	666.7γ					
CSL				603.0γ		
NCM						765.2γ

portfolio of top-10 stocks in January 2006 update the portfolio each subsequent January. In other words, in January of each year the portfolio is formed to hold proportion γ of the total shares on issue of each of the top 10 (by weighting) ASX/S&P200 companies. Although the companies in the portfolio are only updated annually, the quantity of shares of each company in the portfolio is updated with any change in total shares on issue. The proposed liquidation discount time series does not begin until October 2006 due to the SIRCA data being especially ”patchy” prior to October 2006. With respect to the proposed liquidation discount formulation: the number of stocks in the portfolio is $N = 10$; the SIRCA database only records bid-side depth to 20 steps, therefore $m = 20$; and for all the portfolio stocks a half-tick ($tick_t^x/2$) is \$0.005.

A top-10 ASX portfolio is constructed by specifying a portfolio size factor, $0 < \gamma \leq 1$, to be the fraction of the total shares on issue of each stock actually held in the portfolio. That is, defining S_t^x to be the total shares on issue for stock x at time t, the number of shares of stock x in the portfolio is given by $s_t^x = \gamma S_t^x$ (see Table 3.1). By this method the portfolio will be value-weighted. The maximum choice for γ is limited to ensure the notionally liquidated volume of each portfolio component never exceeds the 20-step depth limitation of the SIRCA database. For this investigation, consider values of γ between 0.0001% and 0.0008% inclusive. Figure 3.1 and the summary data in Table 3.2 show that, over the sample period, the smallest portfolio specification ($\gamma = 0.0001\%$) has an average value of about \$0.5 million, and the largest portfolio specification ($\gamma = 0.0008\%$) has an

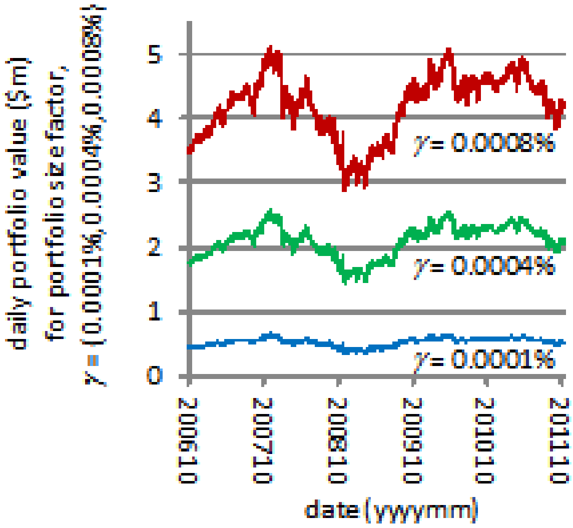


Figure 3.1: Daily portfolio value for portfolio size factor $\gamma \in \{0.0001\%, 0.0004\%, 0.0008\%\}$.

average value of about $4 million (all values are in Australian dollars).

Graphs and summary data for morning and afternoon daily liquidation discount and log liquidation discount time series for different portfolio size specifications of the value-weighted top-10 ASX portfolio are given by Figure 3.2 and Table 3.2. Figure 3.2 shows the increased cost of liquidity (i.e., liquidation discount) associated with increased portfolio size (represented by increased portfolio size factor, γ), and the comparatively high cost of liquidity associated with the heights of the Global Financial Crisis around October 2008. Figure 3.2 also shows that liquidity cost is, on average, higher in the morning at 10:15 than in the afternoon at 15:45. Table 3.3 presents Augmented Dickey-Fuller, KPSS and Phillips-Perron unit root test statistics for the log transformation of the liquidation discount time series: It is concluded that the time series are stationary and contain no structural breaks (this conclusion is further supported by Elliot-Rothenberg-Stock, Schmidt-Phillips and Zivot-Andrews unit root test statistics).

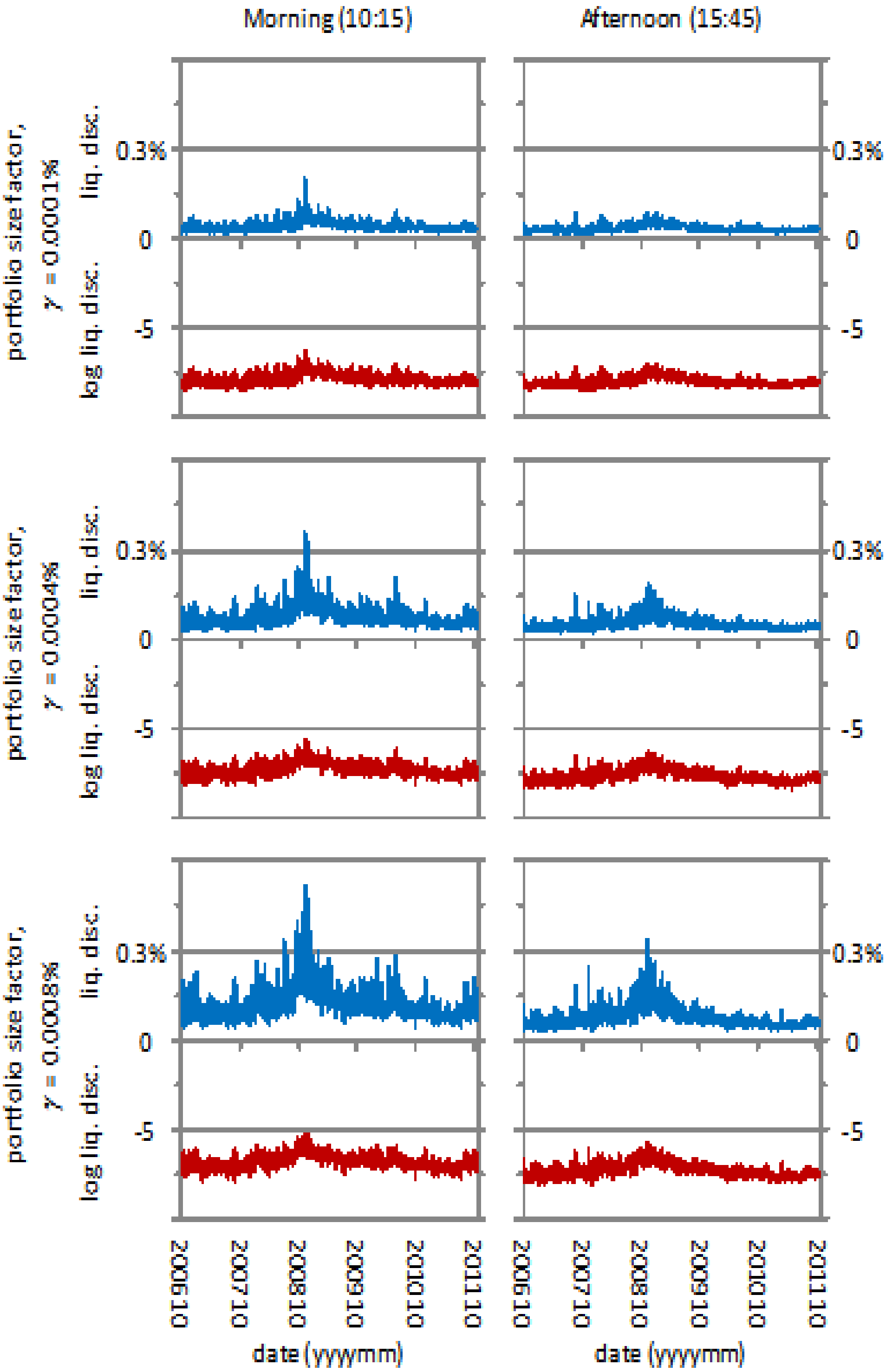


Figure 3.2: Daily portfolio liquidation discount and log liquidation discount for portfolio size factor γ ∈{0.0001%, 0.0004%, 0.0008%}.

Table 3.2: Summary statistics for daily morning and afternoon portfolio value and liquidation discount for portfolio size factor γ ∈{0.0001%, 0.0004%, 0.0008%}.

	Portfolio size and time		Obs	Min	Max	Mean	Median	Std. Dev.
Daily value	γ= 0.0001%	morning	1,229	\$0.34m	\$0.64m	\$0.53m	\$0.54m	\$0.061m
		afternoon	1,218	\$0.36m	\$0.64m	\$0.53m	\$0.54m	\$0.061m
	γ= 0.0004%	morning	1,229	\$1.38m	\$2.55m	\$2.10m	\$2.16m	\$0.244m
		afternoon	1,218	\$1.43m	\$2.55m	\$2.10m	\$2.16m	\$0.244m
	γ= 0.0008%	morning	1,229	\$2.75m	\$5.10m	\$4.21m	\$4.32m	\$0.488m
		afternoon	1,218	\$2.87m	\$5.11m	\$4.21m	\$4.33m	\$0.487m
Liquidation Discount	γ= 0.0001%	morning	1,229	0.0206%	0.196%	0.0414%	0.0377%	0.0152%
		afternoon	1,218	0.0202%	0.0883%	0.0338%	0.0311%	0.0098%
	γ= 0.0004%	morning	1,229	0.0282%	0.360%	0.0766%	0.0688%	0.0340%
		afternoon	1,218	0.0227%	0.183%	0.0521%	0.0460%	0.0213%
	γ= 0.0008%	morning	1,229	0.0402%	0.523%	0.123%	0.108%	0.0573%
		afternoon	1,218	0.0290%	0.332%	0.0776%	0.0668%	0.0369%

Table 3.3: Morning and afternoon log liquidation discount time series unit root test statistics and autocorrelation for portfolio size factor $\gamma \in \{0.0001\%, 0.0004\%, 0.0008\%\}$. For each of the tests, the null hypothesis states that a unit root exists.

			Unit root test statistic (* denotes rejection of null at 1% significance level)		
Portfolio detail		Obs	Augmented Dickey-Fuller	KPSS	Phillips-Peron
γ= 0.0001%	morn	1,229	-12.36*	2.459	-974.7*
	aftn	1,218	-11.73*	2.390	-656.5*
γ= 0.0004%	morn	1,229	-11.89*	2.632	-849.4*
	aftn	1,218	-11.09*	3.035	-616.9*
γ= 0.0008%	morn	1,229	-11.48*	2.607	-712.0*
	aftn	1,218	-10.31*	3.835	-507.9*

3.3 Modelling framework

Preliminary analysis of the log liquidation discount time series indicates the presence of slow decaying autocorrelation. This may suggest that the series may be fractionally integrated i.e., the series may have long-memory. Given this, the mean of the series is estimated with an Autoregressive Fractionally Integrated Moving Average (ARFIMA) model (Granger and Joyeux (1980); Hosking (1981)). This model seems to capture the mean value reasonably well, however the residuals from this model appear to be heteroskesdatic. In order to address this, the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model (Bollerslev (1986)) was used to model the variance of the residuals. Hence, as an overall model, consider the ARFIMA(r,d,s)-GARCH(p,q) model first developed by Baillie (1996):[5]

$$\phi_r(L)(1-L)^d(\ell_t^\gamma - \mu) = \theta_s(L)\varepsilon_t \,, \tag{3.2}$$

$$\varepsilon_t = \eta_t \sqrt{h_t},\ \eta_t \sim iid(0,1) \,, \tag{3.3}$$

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i h_{t-i}. \tag{3.4}$$

From the ARFIMA specification (equation (3.2)), r and s respectively denote the orders of the autoregressive and moving average parts of the model, and d represents the fractional difference term. Note that ARMA and ARIMA models are ARFIMA special cases with d

[5] Also see Ling (2003) for another application of this model.

equal to either 0 or 1 respectively. L is the lag operator where $Ly_t = y_{t-1}$ or more generally,

$$L^i y_t = y_{t-i}. \tag{3.5}$$

The auto-regressive operator is given by $\phi_r(L)$ where

$$\phi_r(L) = 1 - \phi_1 L - \phi_2 L^2 - ... - \phi_r L^r. \tag{3.6}$$

The moving average operator is given by $\theta_s(L)$ where

$$\theta_s(L) = 1 + \theta_1 L + \theta_2 L^2 + ... + \theta_s L^s. \tag{3.7}$$

The fractional differencing operator is $(1 - L)^d$ where

$$(1 - L)^d = 1 - dL - \frac{d(1-d)}{2!}L^2 - \frac{d(1-d)(2-d)}{3!}L^3 - ...$$

or as more frequently specified,

$$(1 - L)^{-d} = 1 + dL + \frac{d(1+d)}{2!}L^2 + \frac{d(1+d)(2+d)}{3!}L^3 + ...$$

From the GARCH specification (equations (3.3) and (3.4)), p and q respectively denote the orders of the autoregressive and moving average parts of the variance of the overall process. Baillie et al. (2002) further extend this model to an ARFIMA-FIGARCH approach which additionally allows for long-memory in the variance of the process. The squared residuals of the time series ARFIMA estimation do not exhibit long-memory (as evidenced by Ljung-Box test statistics for different autocorrelation lags), hence this study does not incorporate fractional integration into the variance model.

For 16 log liquidation discount time series (separate morning and afternoon time series with portfolio size factor, γ, equal to 0.0001%, 0.0002%, . . . , 0.0008%), the ARFIMA(r,d,s)-GARCH(p,q) model is fitted for each lag structure combination of $r \in\{0, 1, 2, 3\}$, $s \in\{0, 1, 2, 3\}$, $p \in\{1, 2, 3\}$, and $q \in\{1, 2, 3\}$. Differencing parameter d is set either to 0 or 1, or is estimated ($d \in (-1, 0.5)$). In total, 432 (=4×4×3×3×3) models are fitted to the first 1000 observations of each of the 16 time series, and then used to forecast the subsequent/final 200+ observations of each time series. The forecasts are then compared to the actual observations. Let $d_{t,\alpha}$ denote the difference between the actual value and the forecast on tth day for a given α i.e.

$$d_{t,\alpha} = \ell_{t,\alpha} - \hat{\ell}_{t,\alpha}. \tag{3.8}$$

The optimal model for each time series is determined to be that which minimises the out-of-sample forecasting mean-squared-error (MSE). The MSE is defined as

$$\text{MSE}_{\alpha} = \frac{1}{f_j} \sum_{t=N_j-f_j+1}^{N_j} {d_{t,\alpha}}^2 \tag{3.9}$$

where f_j is the forecast horizon and N_j is the total number of observations for a given time period j. Similarly, the mean absolute deviation (MAD) is also computed for comparison purposes,

$$\text{MAD}_{\alpha} = \frac{1}{f_j} \sum_{t=N_j-f_j+1}^{N_j} |d_{t,\alpha} - \bar{d}_{t,\alpha}|. \tag{3.10}$$

where $\bar{d}_{t,\alpha}$ the mean value of the deviations for a given α. The models producing the minimum MSE_{α} across all α are selected. The results from both measures are consistent for most values of α.

The models are estimated using the Ghalanos (2012) package in *R*. This package takes into account user specification of the type and order of the mean and variance models, the optimization routine, initial values, residual distribution, etc. (please refer to package documentation for further details). The model estimation results are robust with respect to a range of initial values. The identified optimal models are presented in Table 3.4.

Table 3.4 results indicate the ARIMA specification ($d = 1$) to be optimal for three afternoon time series, and negative fractional differencing is identified for two morning time series [6]. The evidence indicates that the full-sample (1200+ observations) log liquidation discount time series are stationary (e.g., see Table 3.3) and fractionally integrated. Table 3.5 identifies out-of-sample forecast mean-squared-error minimising models optimised with respect to only the final 200+ time series observations. Due to the potential for small sample peculiarities, it is not surprising that a much reduced sample of only 200+ observations might be marginally better modelled with order 1 integration. However, long-memory fractional differencing ($d \in (0, 0.5)$) is predominantly the optimal differencing structure for the time series. Furthermore, the optimal autoregressive and moving average lag structures are predominantly $r = 1$ and $s = 0$ respectively. A single optimal lag structure for the variance of the process is not clearly evident, but lag orders greater

[6] Variance estimates for d parameter are available upon request.

Table 3.4: Optimal (i.e., out-of-sample forecast mean-squared-error minimizing) ARFIMA(r,d,s)-GARCH(p,q) lag and differencing structures for the morning and afternoon log liquidation discount time series with portfolio size factor γ ∈{0.0001%, 0.0002%, . . . , 0.0008%}.

Portfolio size (γ)	Morning (10:15)					Afternoon (15:45)				
	r	d	s	p	q	r	d	s	p	q
0.0001%	3	-0.101	0	2	1	0	0.479	1	2	1
0.0002%	1	0.388	0	2	3	1	1	2	2	2
0.0003%	3	-0.518	2	2	1	1	1	2	3	2
0.0004%	1	0.388	0	3	2	1	1	3	1	1
0.0005%	1	0.386	0	1	3	1	0.388	0	1	3
0.0006%	1	0.389	0	1	3	1	0.392	0	1	3
0.0007%	1	0.392	0	2	3	1	0.396	0	2	3
0.0008%	1	0.407	0	2	3	1	0.400	0	3	3

than 1 are generally favoured.

3.3.1 Proposed Model

For the sake of model parsimony and generalisability across all 16 log liquidation discount time series, a single parsimonious ARFIMA(1,d,0)-GARCH(1,1) model is proposed, which can be expanded as follows:

$$\begin{aligned} \ell_t &= E_t[\ell_t] + \varepsilon_t \qquad (3.11) \\ &= \mu + (\phi_1 + d)(\ell_{t-1} - \mu) \\ &\quad -\left(\phi_1 - \frac{(1-d)}{2}\right)\frac{d}{1!}(\ell_{t-2} - \mu) - \left(\phi_1 - \frac{(2-d)}{3}\right)\frac{d(1-d)}{2!}(\ell_{t-3} - \mu) \\ &\quad -\left(\phi_1 - \frac{(3-d)}{4}\right)\frac{d(1-d)(2-d)}{3!}(\ell_{t-4} - \mu) - ... + \varepsilon_t \ , \end{aligned}$$

where

$$\varepsilon_t = \eta_t \sqrt{h_t}, \ \eta_t \sim iid(0,1) \ ,$$

and

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \ .$$

For each time series, Table 3.5 presents the parameter estimates and out-of-sample forecast MSE associated with the proposed parsimonious model, and the MSE sacrifice in comparison to the individual optimal model structure given by Table 3.4. The out-of-sample forecast MSE for the proposed parsimonious model is generally about 5% for the morning log liquidation discount time series, and about 2% for the afternoon series. Furthermore, the MSE sacrifice associated with the parsimonious model is a very modest 8 basis points or less for every time series. Although adoption of the parsimonious model entails only modest MSE sacrifice, it yields reasonably stable or steadily changing parameter estimates across the time series.

Table 3.5 shows that the base level of log liquidation discount, μ, increases monotonically with the portfolio size factor, γ: in fact, the exponential of μ is an almost perfect linear function of γ such that $\mu = \ln(104.3\gamma{+}0.00026)$ at 10:15, and $\mu = \ln(44.41\gamma{+}0.00026)$ at 15:45. Equation (3.11) shows how the level of persistence of lagged log liquidation discount depends on both the autoregression parameter, ϕ_1, and the fractional difference parameter, d. The estimates for d are similar across all 16 time series, however the ϕ_1 estimates vary.

Unit-root tests confirm the absence of non-stationarity, and existence of variance is confirmed by $\alpha_1 + \beta_1 < 1$. The values of $\alpha_1 + \beta_1$ is close to 1, implying the rate of decay of variance to its unconditional level is slow.

There is a clear distinction between the 10:15 and 15:45 log liquidation discount processes. For a given portfolio size factor, the base level of log liquidation discount (μ) and its base level variance (ω) are (almost always) higher at 10:15 – from an unconditional perspective, liquidation at 15:45 is to be preferred. Further discernment of the intra-day log liquidation discount pattern would be a useful extension of this study.

Table 3.5: For each of the morning and afternoon log liquidation discount time series with portfolio size factor $\gamma \in \{0.0001\%, 0.0002\%, \ldots, 0.0008\%\}$: parameter estimates for the proposed parsimonious ARFIMA(1,d,0)-GARCH(1,1) model (see equations (**??**) to (3.4)); out-of-sample forecast mean-squared-error; and increase in out-of-sample forecast mean-squared-error compared to the optimal model structure presented in Table 3.4.

	Portfolio size factor, γ	μ (e^{μ})	ϕ_1	d	ω	α_1	β_1	$\alpha_1 + \beta_1$	Out-of-sample forecast MSE	MSE increase compared to optimal model
Morning (10:15)	0.0001%	-7.93 (0.0004)	-0.265	0.378	0.00059	0.0162	0.9743	0.991	0.0271	0.0004
	0.0002%	-7.66 (0.0005)	-0.277	0.401	0.04754	0.0000	0.3778	0.378	0.0379	0.0002
	0.0003%	-7.47 (0.0006)	-0.252	0.387	0.00056	0.0120	0.9810	0.993	0.0437	0.0006
	0.0004%	-7.30 (0.0007)	-0.247	0.385	0.00067	0.0148	0.9771	0.992	0.0476	0.0004
	0.0005%	-7.15 (0.0008)	-0.235	0.384	0.00065	0.0156	0.9767	0.992	0.0513	0.0005
	0.0006%	-7.02 (0.0009)	-0.226	0.387	0.00056	0.0144	0.9788	0.993	0.0546	0.0007
	0.0007%	-6.91 (0.0010)	-0.222	0.391	0.00058	0.0150	0.9781	0.993	0.0573	0.0008
	0.0008%	-6.83 (0.0011)	-0.237	0.406	0.01866	0.0672	0.7195	0.787	0.0598	0.0007
Afternoon (15:45)	0.0001%	-8.10 (0.0003)	-0.146	0.378	0.00021	0.0195	0.9746	0.994	0.0116	0.0004
	0.0002%	-7.97 (0.0003)	-0.159	0.373	0.00085	0.0337	0.9493	0.983	0.0178	0.0007
	0.0003%	-7.85 (0.0004)	-0.173	0.378	0.00078	0.0328	0.9533	0.986	0.0208	0.0006
	0.0004%	-7.74 (0.0004)	-0.176	0.383	0.00022	0.0218	0.9744	0.996	0.0227	0.0004
	0.0005%	-7.64 (0.0005)	-0.185	0.388	0.00011	0.0206	0.9773	0.998	0.0231	0.0003
	0.0006%	-7.55 (0.0005)	-0.189	0.392	0.00007	0.0207	0.9778	0.999	0.0231	0.0003
	0.0007%	-7.47 (0.0006)	-0.194	0.396	0.00003	0.0210	0.9780	0.999	0.0236	0.0004
	0.0008%	-7.40 (0.0006)	-0.199	0.400	0.00002	0.0206	0.9786	0.999	0.0244	0.0004

3.3.2 Future liquidation discount level and uncertainty

Portfolio managers face liquidation discount risk. The proposed parsimonious ARFIMA(1,*d*,0)-GARCH(1,1) log liquidation discount model can combine the expected future level and uncertainty of log liquidation discount into a single liquidation discount-at-risk measure, $ldaR_{t,n}^{1-c}$, where positive integer n specifies a liquidation horizon n trading days into the future, and c specifies the confidence that the liquidation discount-at-risk will not be breached.

Future log liquidation discount at any horizon is given by:

$$\ell_{t+n} = E_t[\ell_{t+n}] + \eta_{t+n}\sqrt{h_{t+n}} \qquad \eta_{t+n} \sim iid(0,1). \tag{3.12}$$

Therefore liquidation discount-at-risk is given by:

$$ldaR_{t,n}^{1-c} = \exp\left(E_t[\ell_{t+n}] + \Phi_\eta^{-1}(c)\sqrt{h_{t+n}}\right). \tag{3.13}$$

In equation (3.13), $\Phi_\eta^{-1}(\cdot)$ is the inverse cumulative distribution function for η_{t+n}. This function may be estimated by a Maximum Entropy Density (MED). For further information on this, please refer to Chapter 2.

The continuous version of the entropy presented in Shannon (1948) and Jaynes (1957) is defined as

$$\mathrm{H} = -\int_{\mathbf{A}} f(y)\log f(y)dy. \tag{3.14}$$

where $f(y)$ is the probability density function and A represents the region of integration. The principle of Maximum Entropy involves maximising H subject to some moment condition. For example,

$$\begin{aligned} &\int_{\mathbf{A}} \quad f(y)\,dy = 1. \\ &\int_{\mathbf{A}} \quad y^{\ell} f(y)\,dy = \mu_\ell \text{ where } \ell = 1, 2, ..., k. \end{aligned} \tag{3.15}$$

In the above equations, μ_ℓ represents the i^{th} moment of the distribution. Solving this non-linear optimisation problem yields the following solution:

$$f(y) = Q^{-1}\exp\left(\sum_{\ell=1}^{k} \lambda_\ell \, y^\ell\right). \tag{3.16}$$

The quantity (Q) denotes the normalising constant which ensures that the first constraint is satisfied. One can see that the resulting density is the generalised exponential distribution where λ_ℓ value are the MED parameters.

Assuming that the distribution of η_t is time-invariant, one can use the model outputs to generate this random variable for each portfolio size (γ) and time period (morning and afternoon). Note that this variable is essentially the standardised residual, $\eta_t = \varepsilon_t/\sqrt{h_t}$. Next, the first four moments of each of these series are estimated and subsequently to estimate the MED for each portfolio size across both time periods. The resulting MEDs are numerically integrated to obtain the one sided tail probabilities (1% sig. level). By comparison the resulting critical values are larger than those of the Normal distribution. This is expected since the empirical value of the kurtosis for each portfolio size and time period is greater than 3. For the purpose of facilitating convenient practical application, the Student's t distribution with 12 degrees of freedom is proposed as a familiar "off-the-shelf" distribution (for which critical values are readily available) that closely, but conservatively, approximates the heavy right tails of the MED estimations. In this regard, the 1% right tail critical value of the Student's t distribution with 12 degrees of freedom is 2.681, whereas the 1% right tail critical values obtained by numerical integration of the MED estimations range from about 1.7 to 2.6.

The expectation of future log liquidation discount is an ARFIMA(1,d,0) process such that:

$$E_t[\ell_{t+n}] = \mu + \phi_1^n(\ell_t - \mu) + \sum_{i=0}^{\infty}\left[(\ell_{t-i} - E_{t-i}[\ell_{t-i}])\sum_{j=1}^{n}\left(\phi_1^{n-j}\frac{\prod_{k=0}^{(i+j)-1}(k+d)}{(i+j)!}\right)\right], \tag{3.17}$$

where the formulation of $E_{t-i}[\ell_{t-i}]$ can be determined from equation (3.11) as a function of historic realised log liquidation discount.

The variance term, h_{t+n}, is a GARCH(1,1) process determined according to Baillie and Bollerslev (1992) such that:

$$E_t[h_{t+n}] = \frac{\omega}{1-\alpha_1-\beta_1} + (\alpha_1+\beta_1)^{n-1}\left(\omega + \alpha_1\varepsilon_t^2 + \beta_1 h_t - \frac{\omega}{1-\alpha_1-\beta_1}\right). \tag{3.18}$$

For a liquidation horizon of 5 trading days and a 99% confidence level (with $\Phi_\eta^{-1}(0.99)$ conservatively approximated as 2.681 using the Student's t distribution with 12 degrees of freedom), Figure 3.3 displays the morning and afternoon time series of portfolio liquidation discount-at-risk, $ldaR_{t,5}^{1\%}$, and realised liquidation discount, ld_{t+5}, for different portfolio size factors and in accordance with the parsimonious model parameterisations presented in Table 3.5. Bear in mind that, in the context of the time series, the realised liquidation discount is the liquidity cost that *would have been* suffered if liquidation had actually occurred after five days. It is assumed that the particular notional portfolio manager never does liquidate, but the time series do reflect the impact of many "real" portfolio managers that did actually liquidate various portfolios at various times. If the notional portfolio manager did actually join the fray and liquidate at some point in time, then the time series would notionally deviate from the historic record; nevertheless the proposed liquidation discount-at-risk model could still validly be applied with the input of the notionally updated liquidity conditions.

Figure 3.3 shows that the proposed liquidation discount-at-risk model has successfully been able to forecast 99%-confidence upper-budgets for the future cost of liquidity through time for variously-sized top-10 Australian stock portfolios. Generally the cost of liquidity is low. However, during the Global Financial Crisis, for a portfolio with 0.0008% size factor (worth about \$3 million at the time – see Figure 3.1), the five-day 99%-confidence upper-budget morning cost of liquidity forecast is of the order of a not inconsequential 0.5%. Although this cost is not great on its own, multiple related portfolio sales concentrated in time can obviously entail a significant accumulation of liquidity cost.

3.4 Conclusion

As a measure for the cost of portfolio liquidity, the *liquidation discount* measure was introduced, this being the market impact discount in value yielded by the immediate sale of a portfolio relative to its in-hand market value calculated from prevailing market prices. For any portfolio, the day-to-day liquidation discount that would be suffered if liquidation were undertaken is variable, and thereby a source of risk for portfolio managers. For variously-sized top-10 Australian stock portfolios this study finds that the log-liquidation discount is best modelled with an Autoregressive Fractionally Integrated Moving Average-Generalised Autoregressive Conditional Heteroskedasticity (ARFIMA-

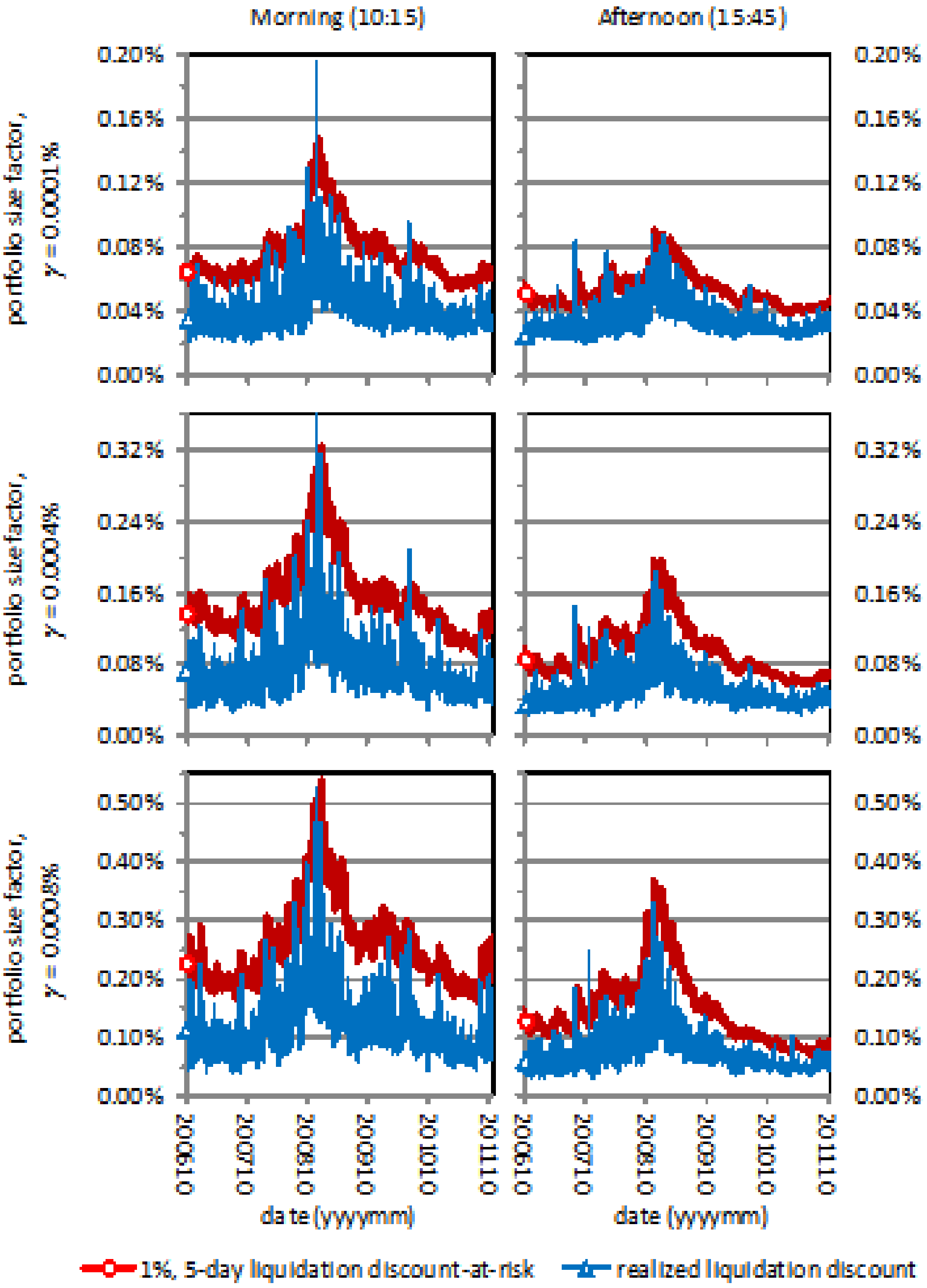


Figure 3.3: Daily portfolio 1%, 5-day liquidation discount-at-risk ($ldaR_{t,5}^{1\%}$) and 5-day forward realized liquidation discount as per the parsimonious model parameterizations presented in Table 3.5 for portfolio size factor $\gamma \in \{0.0001\%, 0.0004\%, 0.0008\%\}$.

GARCH) process – in particular a parsimonious ARFIMA(1,*d*,0)-GARCH(1,1) model was proposed. Using this model, the expected future level and uncertainty of log liquidation discount is combined into a single liquidation discount-at-risk measure, with which portfolio managers can budget for the future cost of portfolio liquidity for a chosen liquidation horizon and confidence level.

Chapter 4

Detecting Intra-Daily Seasonality in Returns Data

4.1 Introduction

There is some evidence that financial returns contain repeated patterns over specific time horizons. Traditionally, analysts have searched for these patterns over longer time horizons such as yearly, bi-annually or quarterly time periods. However, there are cases where these seasonal patterns exist over a much smaller time horizons such as daily or even hourly time periods. This has been made possible due to the availability of ultra-high frequency data (Engle (1996)). This type of seasonality is more commonly referred to as intra-daily seasonality. Examples of intra-daily seasonality include the weekday effect, time of the day effect and weekend effect. By identifying these patterns, investors attempt to position themselves in order to gain from such effects.

This study attempts to generalise the method for detecting seasonal patterns in return data. Previous studies have traditionally focused on detecting changes in the average/mean component of seasonal behaviour of returns. Here, seasonality is captured at a distributional level i.e. periodic changes in the distribution of returns across a time segment. For example, if the time segment is weekdays and one could model the return distribution for each one of the weekdays. If there are significant differences in the return distribution across the weekdays then by construction there will be periodic changes in return distribution over time. For the purposes of this study, this phenomenon is referred to as intra-daily seasonality.

Intra-daily seasonality has generated interest amongst researchers. Traditionally, researchers have explored the changes in the mean and/or variance of returns across a given time segment in order to detect seasonal patterns. French (1980) and Gibbons and Hess (1981) showed that the average daily return for Monday was negative compared to the positive returns for the rest of the weekdays. Rogalski (1984) investigated the behaviour of returns over trading and non-trading periods. Smirlock and Starks (1986) conducted a similar study replacing daily data with hourly data. Doyle and Chen (2009) introduced the wandering weekday effect which states that the weekday effect is not fixed, but changes over time. More recently, Hamid (2018) and Tse (2018) investigate seasonal behaviour of returns across large and small stocks as well as currencies.

However, periodic changes in returns could be found in higher moments (beyond mean and variance) for a given time segment. Given this hypothesis, the objective of this study is to test for intra-daily seasonality using Maximum Entropy Density (MED). The MED estimation is essentially a method which produces a density function. Comparing densities allows one to assess if there are periodic differences in higher moments. Specifically, this study attempts to detect seasonal patterns over weekdays and through the hours of a given trading day. A comparison of MEDs across these two time segments of the return data allows one to test for the existence of intra-daily seasonality.

Section 4.2 provides some brief details about the MED. This section also outlines the practical considerations of implementing this technique. Section 4.3 discusses the properties of the data used in this study. Section 4.4 provide evidence whether or not intra-daily seasonality exists over different time segments of the data. Finally, section 4.5 summarises the major findings of this study along with its limitations.

4.2 Methodology

Shannon (1948) proposed the idea of entropy as a measure of the amount of uncertainty or randomness. Following that Jaynes (1957) proposed the principle of maximum entropy. This allowed one to maximise Shannon's entropy subject to certain moment conditions.

The continuous version of the entropy presented in Shannon (1948) and Jaynes (1957) is defined as

$$E = -\int_{\mathbf{A}} p(x) \log p(x)\, dx \tag{4.1}$$

where $p(x)$ is a probability density function and $\mathbf{A}$ represents the set in which the integration occurs. The principle of Maximal Entropy (ME) involves maximizing E subject to various moment conditions. The moment conditions are as follows:

$$\int_{\mathbf{A}} p(x)\,dx = 1 \tag{4.2}$$

$$\int_{\mathbf{A}} x^i\,p(x)\,dx = m_i \text{ where } i = 1, 2, ..., k. \tag{4.3}$$

In the above equations, m_i represents the i^{th} raw moment of the distribution. Solving this non-linear optimisation problem yields the following solution:

$$p(x) = Q^{-1}\exp\left(\sum_{i=1}^{k} \lambda_i x^i\right) \tag{4.4}$$

where Q is the normalising constant. Further details on derivation of the results as well as other details on principle of ME can be found in chapter 2.

One can see that the resulting density is the generalised exponential distribution (equation (4.4)). The MED estimation produces estimates the parameters (λ_i) of this distribution function. Given that these estimates characterise the MED, comparing these estimates allows one to compare different aspects of the distribution. In particular, higher moments (beyond the mean and/or variance) of the distribution. Repeated patterns in one or more moments between time segments over a period implies the presence of intra-daily seasonality.

Straight forward computation of the MED for return data is problematic since the MED estimation procedure assumes that a random variable is independent and identically distributed (iid) [1]. This condition ensures that consistent estimators are produced. However, return data has a correlation structure and as such the observations are not iid. This correlation structure needs to filtered out prior to the MED estimation. The following model is considered for this purpose:

$$\phi_r(L)r_t = \mu + \theta_s(L)\epsilon_t \tag{4.5}$$

$$\epsilon_t = \eta_t\sqrt{h_t} \quad \eta_t \sim \text{iid}(0, 1) \tag{4.6}$$

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i h_{t-i} \tag{4.7}$$

Equation (4.5) represents the Autoregressive Moving Average (ARMA) model where r and s denote the order of autoregressive and moving average parts of the model respec-

[1] Under certain mixing conditions, the weak law of large numbers may apply.

tively. It attempts to capture the dynamics of the mean of the return data r_t. Note that this model is a special case of equation 3.2, where $d = 0$. In the above model, μ represents the drift term and L is the lag operator as defined in equation 3.5. The autoregressive operator, $\phi_r(L)$ and the moving average operator, $\theta_s(L)$ are defined in the same way as equations 3.6 and 3.7 respectively. Both these polynomials have their unit roots outside the unit circle and share no common roots. This model assumes that the conditional variance of residuals is constant over time. However, estimating the ARMA model for the return data shows that the residuals are not constant over time. Furthermore, the ARCH test results are significant for the first and second lags. This suggests evidence of GARCH effects. Therefore, the variance of the process is modelled using a GARCH(p, q) model (equation (4.7)) where p and q represent the order of the autoregressive and moving average parts of the model respectively. Bollerslev (1986) introduced the GARCH model whilst extending the work of Engle (1982) on ARCH models. This allows one to model a time varying conditional variance. An ARMA(1,1) - GARCH(1,1) model is used to filter out the time dynamics of the return data. Using higher orders did not result in a significantly improved fit. The residuals of this model are standardised and checked to ensure that that no autocorrelation is present in the first and second moment. Subsequently, the MED is estimated from these residuals.

Additionally, MED estimation assumes the existence of moments. In practise, sample moments are calculated and used in place of population moments and as such are always finite quantities. However, depending on the population distribution, these moments may not be finite. Using a large sample maybe helpful in alleviating some of this concern, but this does not guarantee the existence of moments. One possible way to determine the highest finite moment is to estimate the tail index. Hill (1975) introduced a method that achieves this. The Hill tail index estimator assumes that a distribution can be approximated by slowly varying function (See chapter 2). Although as per its original derivation the estimator requires *iid* random variable, this condition can be relaxed i.e. the sample data can be a sequence of dependant random variables. In this study the exchange rate return data does have a correlation structure and as such this condition will need to be relaxed. The Hill estimator is given by

$$\frac{1}{k}\sum_{i=1}^{k}\log\left[\frac{X_{(i)}}{X_{(k+1)}}\right] \tag{4.8}$$

where $X_{(j)}$ represents the j^{th} largest value out of n observations. In addition to the Hill estimator, this study also uses the method outlined in Berkes et al. (2003) to estimate the tail index. Specifically, this method produces a maximal moment exponent for a GARCH(1,1) process. This is considered suitable since the variance of the exchange rate returns can be be modelled using this process. In order to prove that the $2k^{th}$ moment is finite, the following inequality must be satisfied.

$$E[\alpha\eta_{t-1}^2 + \beta]^k < 1 \tag{4.9}$$

By substituting sample estimates and solving for k allows one to verify the highest finite moment available.

$$\frac{1}{T}\sum_{i=1}^{T}(\hat{\alpha}\hat{\eta}_t^2 + \hat{\beta})^k - 1 = 0. \tag{4.10}$$

The overall methodology of this study is as follows. A data series consisting of returns is segmented into weekdays. For example, all the Monday returns are extracted from the data. Each Monday segment is referred to as a *block*. Note that there is a time discontinuity between two consecutive Monday blocks. This has important implications with regard to filtering the correlation structure. An ARMA(1,1) - GARCH(1,1) model is only implemented at the block level since there are no time discontinuities within a block. Prior to this implementation, standard tests are carried out to ensure that the data did not contain unit roots.

The residuals for each block are standardised and checked to ensure that no autocorrelation exists. Using these residuals from each block, the highest moment available is estimated using the methods described above. This allows one to justify the use of higher moments in order to estimate a MED. Next, the first four moment constraints are used for the MED estimation. As such, there are four MED parameters i.e. λ_i where $i \in 1, 2, 3, 4$. Subsequently, the mean values for each parameter (λ_i) are computed from the all the blocks corresponding to a given weekday. These values represent the final MED parameters for a given weekday. For example, the λ_i values are averaged over all Monday blocks to get the overall MED parameters for the Monday segment. The final result is the distribution of exchange rate returns for all Mondays.

Testing for intra-daily seasonality is done in two parts. Firstly in order to verify the structure of the resulting MED, tests are conducted to assess if the computed mean values are significantly different from zero. Secondly, in order to check for intra-daily seasonality, t tests are conducted to assess if the mean values for a given parameter estimate are significantly different across the weekdays. Significant differences in the mean values of λ_i across weekdays indicates that the distribution of returns changes during the week. This pattern over a period of time corresponds to intra-daily seasonality. This entire process is then repeated to check for intra-daily seasonality across different time intervals over a trading day.

4.3 Data

The data used in this study consists of returns for nine foreign exchange (FX) rates. This data is sourced from the SIRCA database. The recording period is from the 30th of May 2008 to the 1st of February 2012. The bid price of each exchange rate is captured at a tick level (see chapter 2 for details). These prices are then translated to minute level data and subsequently the returns are calculated. Observations that occur in non-trading hours[2] are excluded from each of the return series. As mentioned in the previous section, this introduces time discontinuities in the dataset. Each continuous period is referred to as a block.

Figure 4.1 illustrates the returns for the Australian Dollar against the Japanese Yen (AUD/JPY) exchange rate across all blocks.

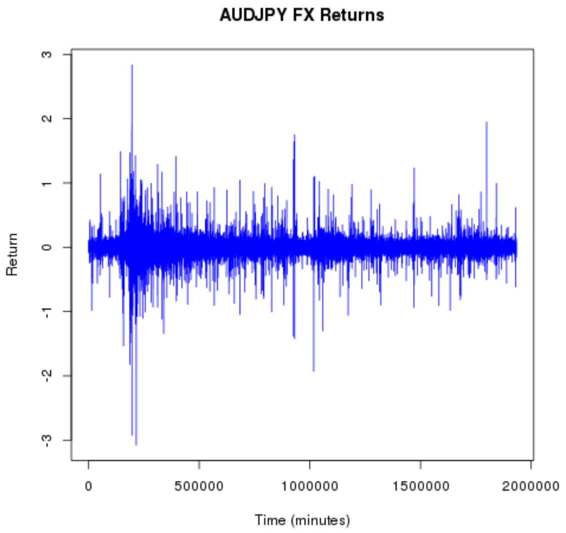


Figure 4.1: FX Returns: AUD/JPY

Figures A.1 in the appendix illustrate the returns for the remaining eight exchange rates. One common feature across all the plots is the impact of the global financial crisis

[2] Trading hours are assumed to be 10a.m. to 4p.m. weekdays

(GFC). This occurs between the 0 and 500,000 minutes where the returns spike in both directions before settling back into their normal behaviour.

In order to test for intra-daily seasonality across weekdays, each of the nine return series is segmented into weekdays. Each weekday consists of 192 blocks [3], each containing 360 observations (60 minutes × 6 hours per day). Table 4.1 shows the summary statistics for each weekday for the AUD/JPY returns over all the weekday blocks. Here, one can observe that the mean returns for Monday and Thursday are negative and that the median return across all the weekdays are all equal to 0. The summary statistics for the remaining eight return series are included in the appendix A.2.2.

Table 4.1: AUDJPY FX weekday summary statistics

Weekday	Minimum	Q1	Median	Mean	Q3	Maximum
Monday	-1.1440000	-0.0230100	0.0000000	-0.0000435	0.0230900	1.0280000
Tuesday	-0.8313000	-0.0241800	0.0000000	0.0003133	0.0244300	1.4760000
Wednesday	-1.4870000	-0.0243800	0.0000000	0.0001161	0.0244100	0.9999000
Thursday	-0.9309000	-0.0244900	0.0000000	-0.0000411	0.0246000	0.8752000
Friday	-1.3060000	-0.0243800	0.0000000	0.0003706	0.0244400	1.6460000

In order to test for intra-daily seasonality within a trading day, three time slots/intervals are introduced. These are 10:00 to 12:00 (noon), 12:00 to 14:00 and 14:00 to 16:00. These partitions were determined in order to acquire adequate number of observations within each time slot. Additionally, this also allows one to allocate a lunch hour in order to determine if there is evidence of a *lunch time effect* on returns. Using this setup, each time interval consists of 959 blocks, each containing 120 observations (two hour intervals). Table 4.2 shows the summary statistics for each time slot of the AUD/JPY returns over all the time slot blocks. Similar to the weekday results, the median return across all time slots is the same. However, none of the time slots have a negative mean value. The summary statistics for each time slot for the remaining return series are included in the appendix A.2.2.

Table 4.2: AUDJPY FX timeslot summary statistics

Timeslot	Minimum	Q1	Median	Mean	Q3	Maximum
10-12	-0.6339000	-0.0231100	0.0000000	0.0002926	0.0232900	0.8594000
12-14	-1.4870000	-0.0241500	0.0000000	0.0001307	0.0243200	1.4760000
14-16	-1.1440000	-0.0248100	0.0000000	0.0000062	0.0248400	1.6460000

[3] number of times a specific weekday occurs over the period of study

4.4 Results

4.4.1 Weekday Effect

As mentioned in previous sections, weekday effects have been defined as changes in the return distribution across the weekdays. The return distribution for each weekday across all blocks is modelled using a MED. In particular, given the existence of higher moments, a four parameter MED is estimated.

Figure 4.2 shows the box plots or each of MED parameters (λ_1 to λ_4) across all weekday blocks for AUD/JPY returns. These box plots illustrate the distribution of the MED parameter estimates. The mean value of each MED parameter for all time slots and FX returns has been added to the corresponding box plot. It is represented by a circle between the first and third quartile. In the case of AUD/JPY returns, the mean value of the second parameter λ_2, is always negative, whereas the remaining parameters have both negative and positive values across weekdays. There seems to be no discernible pattern with regard to this.

Section A.2.2 contains box plots for each weekday across all MED parameter estimates and the remaining FX returns. For a given FX return, these plots allow one to compare the distribution across weekdays for a given MED parameter. From the plots, there appears to be some difference across the weekdays. However, prior to assessing the differences across weekdays, one needs to firstly assess if the estimated MED parameters are indeed significant. This is achieved by testing the mean values of each of the estimated MED parameters across all weekdays. The parameters with significant mean values will form the MED for a given weekday i.e. this process allows one to estimate an MED for each weekday. The hypothesis of this test is given by equation (4.11). Here $\bar{\lambda}_{i,w}$ represents the mean value of the λ_i over all blocks for given weekday w.

$$\begin{aligned} H_0 &: \bar{\lambda}_{i,w} = 0 \\ H_1 &: \bar{\lambda}_{i,w} \neq 0 \end{aligned} \tag{4.11}$$

The test statistic for this hypothesis test is

$$t = \frac{\bar{\ell}_{i,w} - 0}{s_{\ell_{i,w}} / \sqrt{n_{i,w}}} \tag{4.12}$$

where $\bar{\ell}_{i,w}$ is the sample estimate of $\bar{\lambda}_{i,w}$. The standard deviation of ℓ_i over all blocks corre-

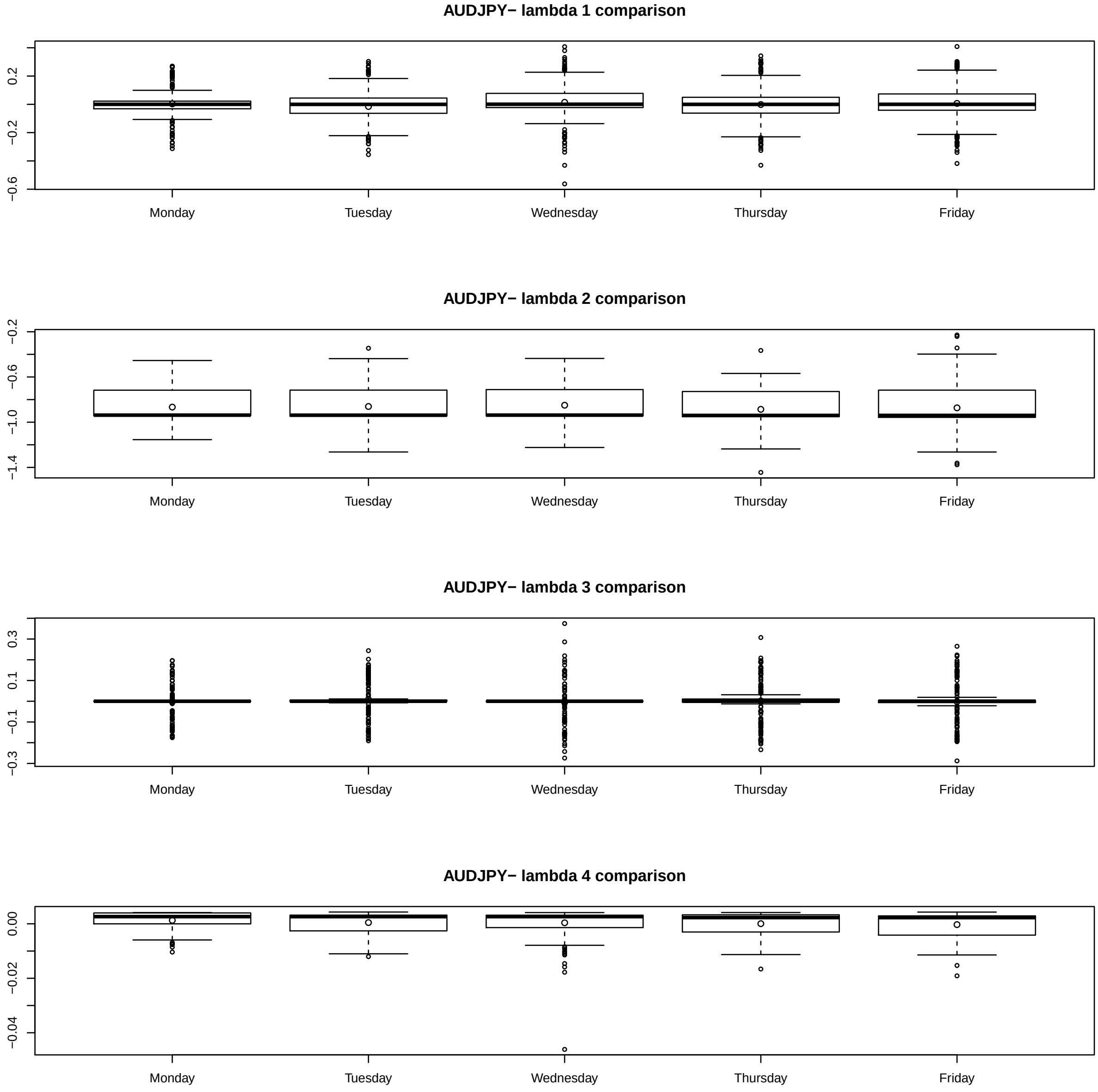


Figure 4.2: Weekday: Box plot of MED parameter estimates

sponding to a given weekday w is given by $s_{\ell_{i,w}}$ and $n_{i,w}$ denotes the number of blocks corresponding to ℓ_i for a given weekday w. This test statistic follows a Student's t-distribution if the null hypothesis is supported.

The results of these tests are shown in section A.2.2. The tables in this section show the estimated mean values of the MED parameters along with their corresponding t test statistic. This test is carried out for each MED parameter across all weekdays for each of FX returns. The results indicate that the value of $\bar{\lambda}_2$ is significant for all weekdays. There are mixed results for the remaining parameters. In the case of the AUD[4] and EURAUD returns, all values of $\bar{\lambda}_4$ are also significant. This is also almost the case for EURGBP

[4] This denotes the AUD/US FX return

(except Thursday). Other FX return series have at least one weekday where the value of $\bar{\lambda}_4$ is significant. Both $\bar{\lambda}_1$ and $\bar{\lambda}_3$ are significant for some cases albeit to a lesser extent.

From the above results, one is able to identify the significant parameters of the MED for each weekday across all FX returns. These values form the final distribution for each weekday for a given FX return. Hence, a natural question to ask is that, are there significant differences in the MED parameters across the weekdays for a given FX return. A difference in parameters will suggest a difference in the structure of the underlying distribution. This allows one to test for weekday effects. Note that the test is being applied to the same parameter across different weekdays. For example, given a FX return, the first parameter for Monday is compared with the first parameter for Tuesday, Wednesday, Thursday and Friday. All combinations across weekdays are considered. More formally, test whether or not the $\bar{\lambda}_i$ values differ across the weekdays for each FX return. Equation (4.13) states the hypothesis of this test.

$$\begin{aligned} H_0 &: \bar{\lambda}_{i,w} = \bar{\lambda}_{i,w^*} \text{ where } w \neq w^* \\ H_1 &: \bar{\lambda}_{i,w} \neq \bar{\lambda}_{i,w^*} \end{aligned} \tag{4.13}$$

The t-test statistic for the above test is

$$t = \frac{\bar{\ell}_{i,w} - \bar{\ell}_{i,w^*}}{\sqrt{\frac{s_{\ell_{i,w}}}{n_{i,w}} + \frac{s_{\ell_{i,w^*}}}{n_{i,w^*}}}}. \tag{4.14}$$

This test assumes that the population variances are not equal. The results of these tests are shown in section A.2.2. The tables in this section contain test statistics for all possible combinations.

Overall, the results indicate that there are some weekday effects. These appear in five of the nine FX returns tested in this study. The results for AUDJPY, EURAUD and EURGBP (to a lesser extent) indicate that Monday returns are significantly different from the rest of the weekdays. Similarly, the results for EUR, EURAUD and EURGBP indicate that Friday returns are significantly different compared to the other weekdays. There are some sporadic differences amongst other weekdays, but Monday and Friday effects are most prominent. These results is very much aligned to the Monday and weekend effect in the intra daily seasonality literature. However, the contribution here is that the comparison is done at a distributional level (comparing MEDs) rather than just comparing means across weekdays. Because of this, the differences due to higher moments are also detected in the comparisons.

4.4.2 Time of the day Effect

The time of the day effect has been defined as a change in the return distribution across different time slots throughout the trading period over time. The return distribution for each time slot (for all blocks) is modelled using a MED. This is achieved by following the steps outlined in the methodology section.

Figure 4.3 shows the box plots or each of MED parameters (λ_1 to λ_4) across all time slot blocks for AUD/JPY returns. These box plots illustrate the distribution of MED parameter estimates. The mean value for each MED parameter for all time slots been included in the box plot. It is represented by a circle between the first and third quartile. In the case of AUD/JPY returns, the mean value of the first and second parameter λ_2, is always negative across all time slots. The mean value of the third parameter λ_3 is always positive across the three time slots, whereas the mean value of the fourth parameter is both negative and positive.

Section A.2.3 contains box plots for each weekday across all MED parameter estimates and the remaining FX returns. For a given FX return, these plots allow one to compare the distribution of a MED parameter across the three different time slots. From the plots, it is evident that there may be differences between these distributions. However, prior to assessing the differences across the time slots, one needs to firstly assess if the mean of each of the estimated MED parameters (from each block) is indeed significant. The hypothesis of this test is given by equation (4.15). Here $\bar{\lambda}_{i,t}$ represents the mean value of the λ_i over all blocks for given time slot t.

$$\begin{aligned} H_0 &: \bar{\lambda}_{i,t} = 0 \\ H_1 &: \bar{\lambda}_{i,t} \neq 0 \end{aligned} \tag{4.15}$$

The corresponding test statistic for this hypothesis test is of the same form as equation (4.12). Table 4.3 shows the mean MED parameter estimates and their corresponding test statistic values for the AUD/JPY returns. In this case, $\bar{\lambda}_2$ is significant across all time slots and more interestingly $\bar{\lambda}_4$ is significant for the 12-14 (lunch time) time slot. Additionally both, $\bar{\lambda}_1$ and $\bar{\lambda}_3$ are significant for the 14-16 time slot.

The significance results for the remaining FX returns are shown in section A.2.3. The

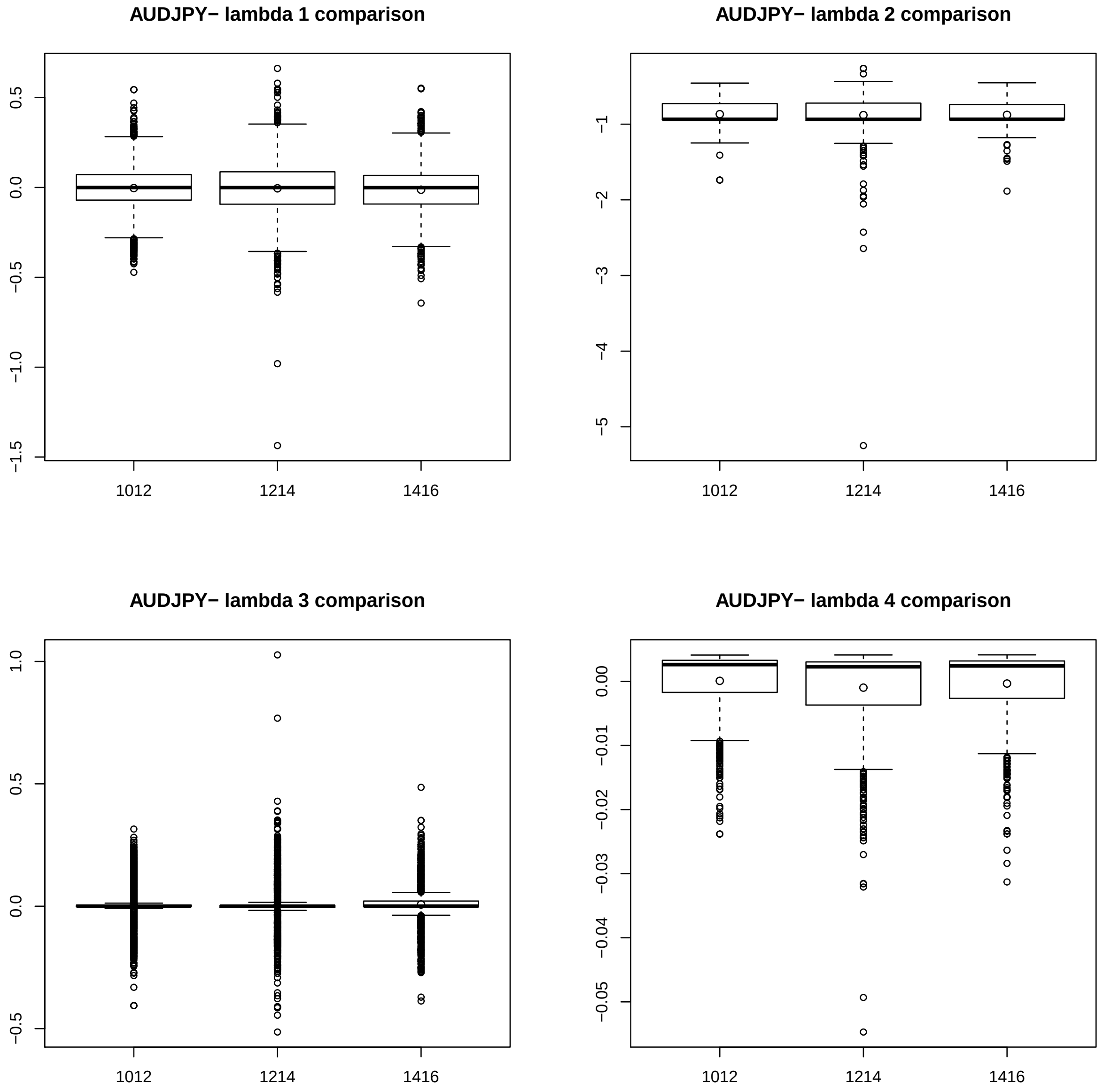


Figure 4.3: Time Slot: Box plot of MED parameter estimates

Table 4.3: Significance of MED parameters: AUD/JPY Returns

Time Slots	10-12		12-14		14-16	
AUDJPY	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	−0.0033	−0.6801	−0.0045	−0.7456	−0.0124	−2.3900
$\bar{\ell}_2$	−0.8675	−198.1785	−0.8799	−110.6193	−0.8766	−197.7718
$\bar{\ell}_3$	0.0048	1.6581	0.0020	0.5196	0.0069	2.2235
$\bar{\ell}_4$	0.0001	0.5027	−0.0010	−4.4280	−0.0003	−1.8458

tables in this section provide the estimated values of the mean MED parameters together with their corresponding t test statistic. Overall, the results indicate that the value of $\bar{\lambda}_2$ is significant for all time slots and FX returns. There are mixed results for the remaining parameters. In case of EUR[5], EURAUD, EURGBP, GBP[6] and JPY[7] returns, all values of $\bar{\lambda}_4$ are also significant. Both EURAUD and GBPAUD returns have significant $\bar{\lambda}_3$ values.

From the above results, one is able to identify the significant parameters of the MED for each time slot across all FX returns. These values form the final distribution for each weekday for a given FX return. Hence, a natural question to ask is that, are there significant differences in the MED parameters across the time slots for a given FX return. A difference in parameters will suggest a difference in the shape of the underlying distribution. Using this, one can test for time of day effects. Note that the test is being applied to the same parameter across different weekdays. For example, given a FX return, the first parameter for 10-12 time slot is compared with the first parameter for the 12-14 time slot and subsequently compared with the 14-16 time slot. All combinations across time slots are considered. More formally, test whether or not the $\bar{\lambda}_i$ values differ across the time slots within a trading day for each FX return. Equation (4.16) states the hypothesis of this test.

$$\begin{aligned} H_0 &: \bar{\lambda}_{i,t} = \bar{\lambda}_{i,t^*} \text{ where } t \neq t^*. \\ H_1 &: \bar{\lambda}_{i,t} \neq \bar{\lambda}_{i,t^*} \end{aligned} \tag{4.16}$$

The corresponding test statistic for this hypothesis test is of the same form as equation (4.14).

Table 4.4 shows the comparison results for MED parameters across time slots for AUD/JPY returns. This table contains the test statistic values resulting from the difference in the mean values of the MED parameters across time slots. For a given FX return, there are three time slots and four parameters resulting in 12 possible combinations/tests. Based on the results below, the λ_4 parameter corresponding to the lunch time slot (12-14) is significantly different to λ_4 values corresponding to the remaining time slots. This result is consistent with the existing evidence of lunch time effects in stock returns. However, this analysis allows one to detect these effects by observing differences in higher moments.

The time slot comparison results for the remaining FX returns are shown in section A.2.3. The tables in this section provide the t test statistics resulting from the comparison of mean values of the MED parameters across all time slot combinations. For a given MED parameter and FX return, a significant difference between two time slots is signalled

[5] This denotes the EUR/US FX return

[6] This denotes the GBP/US FX return

[7] This denotes the JPY/US FX return

Table 4.4: Comparing MED parameters across time slots: AUD/JPY

Comparisons	1012 vs 1214	1012 vs 1416	1214 vs 1416
AUDJPY	t statistic	t statistic	t statistic
$\Delta\ \bar{\ell}_1$	0.1543	1.2720	0.9845
$\Delta\ \bar{\ell}_2$	1.3656	1.4630	−0.3607
$\Delta\ \bar{\ell}_3$	0.5942	−0.4900	−1.0039
$\Delta\ \bar{\ell}_4$	3.8102	1.6888	−2.2484

by a *large* t test statistic. Overall, the results suggest that there is evidence of time of the day effects. These appear in seven of the nine returns tested. For instance, the results for AUD, AUDJPY, EUR and EURJPY indicate that 12p.m. to 2p.m. timeslot is significantly different from the rest of time slots. Similarly, the results for EURGBP and GBP indicate that the 2p.m. to 4p.m. time slot is significantly different from the rest of time slots. The results for JPY show that the 10a.m. to 12p.m. time slot is significantly different from the rest of the timeslots. Lastly, EURAUD and GBPAUD results showed no time of the day effects. These results are consistent with the existing literature on intra daily seasonality. In addition to lunch time effects, there may morning effects (market open) effects as well as afternoon effect (market close). The effects seem to be prevalent in different markets. However, unlike previous studies, this study finds evidence of such effects at a distributional level. In particular, such effects are present in higher moments.

4.5 Conclusion

This chapter has introduced a method for testing intra-daily seasonality using a MED. The resulting density allows for a more richer comparison across different time segments (weekdays or time intervals). Specifically, one is able to check for differences in higher moments. This is especially important when differences in lower moments are not significant. This is precisely one of the results found in this study. It shows that the $\bar{\lambda}_4$ value for Wednesday is significantly different from the rest of the weekdays. But the remaining $\bar{\lambda}$ values are not significantly different across the weekdays. Similarly, the $\bar{\lambda}_4$ value is significant for the 12p.m. to 2p.m. interval. Lastly, the results are limited to properties of the data used in this study.

Chapter 5

Modelling the Distribution of Body Mass Index

5.1 Introduction

The MED described in the previous sections is a univariate density. However, in many scenarios, one is interested in modelling the joint distribution of two or more variables i.e. in this case, estimating a multivariate distribution function. As motivated by Joe (1997), multivariate models for non-normal variables is an important area to consider. More specifically, understanding the dependence concepts across multiple variables is necessary in order to develop a multivariate model. As such, the principle of maximum entropy can be adapted to a multivariate setting. An entropy functional consisting of a multivariate function can be maximized subject to constraints which specify the dependence structure across the variables. The resulting density is known as a multivariate MED. Compared to the literature on univariate functions, there appears to be relatively little on estimating a multivariate MED. One possible reason for this may be the lack of theoretical development for the multivariate setting. Unlike the univariate MED case, where the uniqueness and existence conditions have been proven, currently no such conditions exist for the multivariate MED case. However, from a practical viewpoint, the need to model dependence across multiple variables has led a number of empirical developments. Based on a literature review, there appear to be two major approaches used.

The first approach is mentioned in Kapur (1989) (Chapters 4 and 5). Here the Shannon entropy is maximised subject to a number of constraints which specify the moments for each variable as well as the correlations across all pairs of variables. This approach is demonstrated only for the bivariate case with second order moments. As the number of variables increases and/or the order of the moments, the number of constraints in the optimisation problem increases rapidly. Thus, creating unnecessary complexity with regard to parameter estimation.

The second approach consists of estimating *entropy copula* models. Here, the Shannon entropy of the copula density function is maximised subject to constraints satisfied by the copula. These copula constraints include the integration to one constraint, conditions on the marginal densities and specification of the dependence structure (joint behaviour) of the variables. For further details refer to Piantadosi et al. (2012) and AghaKouchak (2014). There are few limitations with this approach. As with most methods that use copulas, the choice of an appropriate copula is ambiguous. Additionally, there appear to be no analytical solutions for many types of entropy copula models.

The proposed approach in this chapter aims to address some of the issues with the existing methods. The emphasis is on being able to adequately handle multiple variables as well as higher order moment conditions (well beyond the bivariate cases mentioned in most of the literature). This approach has the additional benefit of being able to select conditions (constraints) prior to the optimisation process. The resulting density possess an analytical form and is a natural extension of the univariate case.

The second contribution of this chapter is to allow the parameters of the multivariate MED to be functions of conditioning variables or covariates. Specifically, the MED parameters are functions of exogenous variables and the resulting MED is conditional on these variables i.e. the shape, scale and location of the MED is dependant on the exogenous variables. Based on this, the proposed model is quite different from the varying parameter model developed by Chan (2009). Furthermore, there is an alternative way to incorporate exogenous variables. This consists of including these variables in the moment constraints itself. Thus solving the optimisation problem will lead to the variables being directly present in the density itself. However, there exists a non-linear relationship between the parameters and the moment constraints (Chan (2009)). As such, it can be argued that allowing the parameters itself to be functions of exogenous variables produces a similar effect compared to including them in the moment constraints.

Hence, these variables can affect the value of the MED parameter which further affects the final structure of the resulting MED. Under this set up one is interested in the coefficients (marginal effects) of the covariates. The proposed framework provides considerable flexibility with regard to allocating covariates to each of the MED parameters. The statistical properties (consistency and asymptotic normality) are shown in this chapter. The implicit relationship between the moments of the MED and the covariates is derived i.e. the marginal change for a specified moment given a change in a covariate value. This allows the modeller to assess the marginal change in the MED resulting from a marginal change in one of the covariates. Finally, the conditioning framework is used to model the distribution of body mass index of an individual given their social and demographic characteristics (covariates).

In section 5.2 the proposed multivariate framework is presented along with the statistical properties of the resulting multivariate MED. Section 5.3 introduces the notion of body mass index and relevant literature on modelling this measure. Subsequent sections consist of applying a special case of the model developed in this chapter to the HILDA data set. The results are discussed in detail for the benefit of health policy professionals.

5.2 Multivariate Framework

The entropy maximisation framework for a multivariate density $f()$ can be written as:

$$\text{max.} \; -\int f(\boldsymbol{x}) \log f(\boldsymbol{x}) \, d\boldsymbol{x} \tag{5.1}$$

subject to

$$\begin{aligned} &\int_{\mathbf{A}} f(\boldsymbol{x}) = 1 \\ &\int_{\mathbf{A}} \boldsymbol{x}^{\otimes \ell} f(\boldsymbol{x}) \, d\boldsymbol{x} = \mu_\ell \text{ for } \ell = 1, \ldots, k. \end{aligned}$$

Here $\boldsymbol{x}$ represents a vector of random variables and $\otimes$ denotes the Kronecker product. This product is applied to the vector of random variables using an index. For example, $\boldsymbol{x}^{\otimes 1} = \boldsymbol{x}$, $\boldsymbol{x}^{\otimes 2} = \boldsymbol{x} \otimes \boldsymbol{x}$, $\boldsymbol{x}^{\otimes 3} = \boldsymbol{x} \otimes \boldsymbol{x} \otimes \boldsymbol{x}$ and so on. This allows the framework to generate all the possible combinations of random variables so that the k^{th} order moments (and cross moments) can be defined.

The first constraint ensures that the multivariate density integrates to 1 over a specified region i.e. set $\mathbf{A}$. In the second constraint, the Kronecker product is applied to $\boldsymbol{x}$ in order to compute the moment vector μ_ℓ, corresponding to each moment ℓ. Note that these vectors also contain cross moments which capture the dependence structure between the variables.

Next, the details of the above setup are presented. Starting with

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

the Kronecker product is applied in order to generate the kth order moments (and cross moments). This can be represented in the following way.

$$\boldsymbol{x}^{(k)} = \begin{bmatrix} \boldsymbol{x}^{\otimes 1} \\ \boldsymbol{x}^{\otimes 2} \\ \vdots \\ \boldsymbol{x}^{\otimes k} \end{bmatrix}$$

Given that $\boldsymbol{x}$ is $n \times 1$ and the dimension of $\boldsymbol{x}^{\otimes \ell}$ is $n^{\ell} \times 1$, the vector $\boldsymbol{x}^{(k)}$ will contain

$$\sum_{\ell=1}^{k} n^{\ell} = \frac{n}{n-1}(n^k - 1) = N_k$$

elements. A closer look at $\boldsymbol{x}^{(k)}$ indicates that it contains some duplicate elements. In order decrease the computational burden, the duplicate entries need to be removed. In order to do this, a selection[1] matrix consisting of zeros and ones is used (Magnus and Neudecker (1999)). This matrix selects unique entries (or removes duplicate entries).

Given the n variables and k moment conditions, the total number of unique moments is given by

$$\sum_{\ell=1}^{k} \binom{\ell + n - 1}{n - 1} = M_k.$$

Hence, knowing the number of unique moments allows one to construct a suitable selection matrix. Additionally, each unique moment corresponds to a MED parameter. As such the dimension of the parameter vector is $M_k \times 1$. Solving the optimisation problem yields the following multivariate MED,

$$f(\boldsymbol{x}) = Q^{-1} \exp\left(\boldsymbol{\lambda}' \boldsymbol{S} \boldsymbol{x}^{(k)}\right). \tag{5.2}$$

Here $\boldsymbol{\lambda}$ denotes the parameter vector and $\boldsymbol{S}$ is the selection matrix of size $M_k \times N_k$. As in

[1] also known as elimination matrix

the univariate case, Q denotes the normalising constant given by

$$Q = \int_{\mathbf{A}} \exp\,(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)})\, d\boldsymbol{x}. \tag{5.3}$$

The resulting density is a generalised multivariate exponential density. For a given sample of data, it is possible to estimate the parameters ($\boldsymbol{\lambda}$) of this density. As in the univariate case, the parameters values control the shape, scale and location of the density. Given this, it may useful to let λ be a function of one or more exogenous variables/ covariates. As such the resulting MED is a conditional density i.e. depends on the values of these variables/ covariates. In addition to this, the proposed setup allows the modeller to assign a variable to any of the parameters. This way the modeller can control (to an extent) which of the variables control the different aspects (moments) of the resulting MED. Let each parameter be a function of p covariates i.e.

$$\lambda_m = \beta_{m1}z_1 + \beta_{m2}z_2 + \cdots + \beta_{mp}z_p.$$

Given k moments and p covariates,

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{M_k} \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1p} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{M_k1} & \beta_{M_k2} & \dots & \beta_{M_kp} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{bmatrix}.$$

The above expression can be written using matrix notation as

$$\boldsymbol{\lambda} = \boldsymbol{\beta z}. \tag{5.4}$$

Note that some of the values for β can be set to zero. This allows the modeller the flexibility to allocate different covariates across MED parameters if required. The conditional density can be expressed as

$$f(\boldsymbol{x}|\boldsymbol{z}) = Q_z^{-1}\exp\left[(\boldsymbol{\beta z})'\,\boldsymbol{x}^{(k)}\right] \tag{5.5}$$

where

$$Q_z = \int_{\mathbf{A}} \exp\left[(\boldsymbol{\beta z})'\,\boldsymbol{x}^{(k)}\right] d\boldsymbol{x}. \tag{5.6}$$

Note that the normalising constant is dependent on the covariate values. For a given set of $\boldsymbol{x}^{(k)}$ and covariates, one can use the method of maximum likelihood estimation to estimate the values of $\boldsymbol{\beta}$. Hence, the optimal parameter values are the ones which maximise the log-likelihood function. The estimator is defined as

$$\log L_T(\hat{\boldsymbol{\beta}}_T) = \max_{\beta \in \Theta} \log L_T(\boldsymbol{\beta}) \tag{5.7}$$

where L_T denotes the likelihood function and Θ represent the parameter space.

Next, the properties of the above estimator are shown. Specifically, the focus is on the consistency and asymptotic normality. Proving these conditions allows one to use this framework in an empirical setting.

Proposition 1. *The estimator as defined in equation 5.7 is consistent:* $\hat{\boldsymbol{\beta}}_{\boldsymbol{T}} \xrightarrow{p} \boldsymbol{\beta}_{\mathbf{0}}$.

Proof. Refer to appendix section A.3.1. □

Proposition 2. *The maximum likelihood estimator is asymptotically normal.* $\sqrt{T}\left(\hat{\boldsymbol{\beta}}_{\boldsymbol{T}} - \boldsymbol{\beta}_{\mathbf{0}}\right) \sim N(\mathbf{0},\ \boldsymbol{B}(\boldsymbol{\beta}_{\mathbf{0}})^{-1}\ \boldsymbol{C}(\boldsymbol{\beta}_{\mathbf{0}})\ \boldsymbol{B}(\boldsymbol{\beta}_{\mathbf{0}})^{-1})$ *where* $\boldsymbol{B}$ *denotes the matrix of second order derivative of the log likelihood function and* **C** *denotes a matrix of the product of first order derivative of the log likelihood function.*

Proof. Refer to appendix section A.3.1. □

This completes the presentation of the multivariate framework. In the next section, this framework is applied to a well known problem in Health Economics.

5.3 Body Mass Index

The objective of this study is to model the distribution of Body Mass Index (BMI) using a set of covariates. BMI is one of the leading indicators of an individual's health. Specifically, it estimates the amount of body fat of an individual. This is done by dividing an individual's mass (kilogram) by the square of their height (metre). Differences in BMI across adults are generally due to the amount of body fat. As such, this metric is used as a comparison tool across individuals. As per the Australian Institute of Health and Welfare (AIHW), a BMI value under 18 is classified as underweight, values from 18 to 25 (inclusive) are considered normal, values from 25 to 30 are considered overweight and values over 30 are classified obese. According to the AIHW in 2012, 63% of Australian adults and 25% of children were overweight or obese. The AIHW claimed that being overweight and obese is the second highest contributor to the burden of disease[2] and obesity rates have doubled since the 1980's in Australia. This is consistent with the other developed nations around the globe according to the World Health Organisation (WHO). Aside from the health implications of being obese, there are also economic consequences such as loss of productivity arising due to employee absenteeism.

Given these reasons, it is not surprising that many researchers are investigating obesity rates using measures such as BMI. As such, there exists a vast amount of academic literature on BMI. However, for the purpose of this study, the focus is on a particular subset of this literature. This study classifies BMI research into two board categories. The first category consists of studies which attempt to fit a density function to an empirical distribution of BMI. An example of this study is Flegal and Troiano (2000) which uses graphical methods (mean difference plots) to describe changes in the distribution of BMI for both adults and children in the US. Another paper by Penman and Johnson (2006) proposes the log-normal distribution to estimate BMI for a given population. A comprehensive paper by Lin et al. (2007) estimates the BMI distribution using a finite mixture model of Normal, skew Normal, Student t and skew Student t distributions as defined in Azzalini (1985), Azzalini (1986) and Azzalini and Capitaino (2003). Lin et al. (2007) found that a finite mixture of skew student t distribution provided a better fit compared to normal mixtures. The paper by Contoyannis and Wildman (2007) estimates the BMI distribution of two different countries using non-parametric techniques. Once these distributions are constructed, a range of measures are used to examine the differences in the modelled distributions. Houle (2010) uses similar methods as Contoyannis and Wildman (2007) to study differences in BMI distributions across gender and education. For more recent study and review, refer to Bann et al. (2018) and Yu et al. (2018).

The second category consists of studies which attempt to model the conditional mean or median of an individual's BMI using a set of covariates. An example of this kind of study is Beyerlein et al. (2008) where three different regression approaches- Generalized Linear Models (GLMs), Quantile Regression and Generalized Additive Models for Location, Scale and Shape (GAMLSS) were employed to model a child's BMI. The major finding of their paper was that GAMLSS and Quantile regression provided a much better fit compared to GLMs for a given set of risk factors. Another more recent paper by Bottai et al. (2014) examined associations among age, physical activity and birth cohort on BMI

[2] after dietary risks and before smoking

percentiles in men using Quantile regression. The paper concluded that Quantile regression allow one to examine how various covariates affected BMI at different percentiles of the estimated BMI distribution.

Based on this classification, this study attempts to combine the objectives from both categories. In other words, this study attempts to model the distribution of an individual's BMI using covariates (risk factors, attributes). This framework will allow different covariates to influence different aspects (moments) of the individual's estimated BMI distribution. This is a point that studies based on Quantile regression do claim. However, the impact of the covariate is usually measured at specific percentiles such as the 90% or 95% percentiles rather than the entire distribution.

Another paper which attempts to combine objectives from both categories is the paper by Brown et al. (2014). It proposes a statistical model (Normal distribution) to model the BMI distribution of an unobserved (latent) class of individuals within a population. It is expected that a finite mixture of these models will provide a good fit for the overall BMI distribution. The weight of the each model is determined using the covariates (individual attributes) in the class and these covariates are same for each class. As a result of this differing values of the same covariate determine the weights for distribution of each class. Hence, the paper has been able to model the distribution of BMI for a given population using information from the covariates.

There are however, a number of factors that one needs to consider when implementing such an approach. One such factor is the number of distributions/classes one should use. Especially since this choice affects the level of complexity in the estimation i.e. as the number of distributions/classes increases the estimation procedure may result in non-convergence. Secondly, each model (normal distribution) as well as the resulting final mixture have infinite support. Whilst this may be desirable for certain applications, it is not the case for BMI. Negative or zero BMI values are nonsensical. The interpretation of the results can be complex. The weights for the each class specify the probability of an individual (based on covariates) falling into a that class. Hence, the covariates affect the weight assigned to the distribution rather than drive any changes in the distribution itself.

In comparison, using the MED approach with covariate information, a single density is produced for a given set of covariate values. The estimated density is constructed over a closed interval. In this case, a set of plausible BMI values. For a given set of optimal parameters, the covariate values determine different moments of the estimated density. This provides a much more intuitive explanation. Although the idea of using measures pertaining to entropy is not entirely new to BMI studies (Contoyannis and Wildman (2007) and Houle (2010)), the application of MED to model the distribution of BMI using covariates is indeed novel.

Despite the vast amount of literature on BMI, it is important to address the limitations that some health professionals have identified. Given the definition of BMI, it is easier to interpret a change in BMI when only the mass of an individual changes. In most cases, this is associated with an increase in body fat Ranasinghe et al. (2013). This is generally the case with adults. However, with children both height and mass can vary and as a result it is more difficult to interpret the change in BMI levels. This is also the case in adults who may have increased their muscle mass i.e. the extra mass does not consist entirely of body fat. Given both these cases, it is possible to exclude children (under 16) from the study and if possible also athletes provided they can be identified. There are nonetheless studies which solely focus on studying the BMI levels in both these groups (Walsh et al. (2011), Ortlepp et al. (2003) and Beyerlein et al. (2008)). Lastly, health professional have introduced a new measure in 2012 appropriately named Body Shape Index (BSI). The definition of BSI contains the waist circumference, height and BMI itself. It is claimed that this new measure can capture a wider array of health risks compared to BMI. However, almost all national and global health institutions as well as medical personnel continue to use and report BMI statistics for the general population.

The rest of the sections are organized as follows: Section 5.3.1 contains the details of the model specification and estimation. Most importantly, it proposes a method of incorporating covariates into the MED framework and contains the estimation methodology for the proposed model. Section 5.3.2 provides a brief introduction to the data set used in this study. Section 5.3.3 contains the estimated model along with some discussion of the results. Lastly, section 5.3.5 summarizes the major results of the paper with some points on the future direction of the study.

5.3.1 Specification and Estimation

The random variable of interest in this study is BMI. This study aims to estimate the distribution of BMI for an individual using their attributes/characteristics. These attributes form the conditioning variables (covariates) used in the estimation procedure. Let y denote the variable BMI and z_i is a vector of covariates pertaining to individual i. The MED for this application can be expressed as

$$f(y|z_i) = Q_i^{-1} \exp\left[(\boldsymbol{\beta} z_i)' \, \boldsymbol{y}\right] \tag{5.8}$$

where

$$Q_i = \int_{\mathbf{A}} \exp\left[(\boldsymbol{\beta} z_i)' \, \boldsymbol{y}\right] \, dy. \tag{5.9}$$

The above specification is a special case ($n = 1$) of the multivariate framework with conditioning presented in the previous section. Note that the normalizing constant is dependent on a given individual's covariate values and is calculated by integrating the density over a set of all possible BMI values (set **A**).

The values of $\boldsymbol{\beta}$ and covariates govern the shape, scale and location of the BMI distribution. More specifically for fixed $\boldsymbol{\beta}$ values, changing the covariates will result in changes in the BMI distribution. This specification offers flexibility with regard to how the covariates affect different moments of the BMI distribution. For example, covariates may be transformed and/or combined with other covariates and/or with intercept terms to model their impact. The resulting conditional density is a generalized exponential distribution which has infinite support. One could rightly argue that BMI values are restricted to a subset of positive values and hence this specification may not be accurate. However, the normalizing constant (equation 5.9) is obtained by integrating the density over of a set of plausible BMI values. As a result of this, there is a zero probability of obtaining a BMI value outside this set of plausible BMI values.[3]

Given the specification for the conditional distribution of BMI, the next step is to estimate the parameters of this distribution i.e. the values of $\boldsymbol{\beta}$. The method of Maximum Likelihood can be used to carry out this estimation. For a sample of size n individuals, the log-likelihood function for equation 5.8 can be expressed as

$$\log L(\boldsymbol{\beta}; z_i, y_i) = \sum_{i=1}^{n} \log f(y_i|z_i) = -\sum_{i=1}^{n} \log Q_i + \sum_{i=1}^{n} (\boldsymbol{\beta} z_i)' \, \boldsymbol{y}_i \tag{5.10}$$

Here $\boldsymbol{y}_i$ is a vector containing individual i's BMI raised to a power (1 to k) and z_i is a matrix consisting of an individual's attributes (p covariates). Note that the normalizing constant Q_i (equation 5.9) differs across individuals. The values of parameters that maximize the above log likelihood function are considered to be optimal estimates. Hence, one needs to maximize equation 5.10 over a set of all possible parameters values. Nu-

[3] Plausible BMI values for this study range from 9 to 100

merical optimization and integration procedures are used to achieve this since no closed form expressions exist when $k > 2$ (Rockinger and Jondeau (2002)). For computational convenience, the first order derivatives are derived (section A.3.2) and included in the optimization routine.

Prior to the estimating the parameters, it is desirable to ensure that the estimator is consistent and asymptotically normal. These properties were proven for the multivariate framework with conditioning in the previous section. Given that the proposed BMI model is a special case (univariate framework with conditioning), the existing proofs ensure that these properties are true for this case.

It is also worth investigating how the changes in covariate values affect the resulting BMI distribution. More specifically, how changes in covariate values produce changes in the moments of the distribution.

Proposition 3. *The marginal change in a moment given a change in the value of a covariate can be expressed as*

$$\frac{\partial \boldsymbol{\mu}}{\partial z} = \boldsymbol{\beta}' \mathbf{M} \tag{5.11}$$

where $\mathbf{M} = (\boldsymbol{\Omega} - \boldsymbol{\mu}\boldsymbol{\mu}')$ *and is a symmetric matrix.*

Proof. From equation A.8, focus on the derivative with respect to z. □

Remark 1. *From Proposition 3, the marginal change in the* ℓ^{th} *moment with respect to change in the* j^{th} *variable can be written as*

$$\frac{\partial \mu_\ell}{\partial z_j} = \beta_{1j}(\mu_{\ell+1} - \mu_\ell\mu_1) + \beta_{2j}(\mu_{\ell+2} - \mu_\ell\mu_2) + \cdots + \beta_{\ell j}(\mu_{\ell+\ell} - \mu_\ell\mu_\ell).$$

Using the above expression one can measure the change in the moment given the change in the value of a covariate. Changes in moment will result in changes in the final distribution. For demonstration purposes, the HILDA dataset (see section below) is used to examine the changes in the BMI distribution given a change in the value of a covariate. All other covariate values remain unchanged.

5.3.2 Data

The BMI data used in this study has been sourced from the Household, Income and Labour Dynamics in Australia (HILDA) survey. This is a household based panel study which began in 2001 and collects information about all individuals in a household. This includes family attributes, economic well being, labour market information, health and subjective well-being, education status as well as a variety of other household and individual variables. Physical attributes such as height and weight have been captured since 2006. This data set is particularly suited to this study because it is the only national level household panel data set available in Australia. As such, it is representative of the Australian population. This paper focuses on survey results for the year 2012. This survey had approximately 25,000 individuals from 10,000 households.

Figure 5.1 contains the histogram of the BMI values for all individuals older than 16 years of age in the 2012 survey. This figure illustrates the stylized facts of BMI distributions. These include the fact that they are prominently uni-modal and are skewed to the right. Table 5.1 provides the summary statistics for BMI values. Both the histogram as well the summary statistic table show the range of BMI values in the survey. It is worth noting that both the median and mean values are in the overweight category (greater than 25 and less then 30) as defined by AIHW.

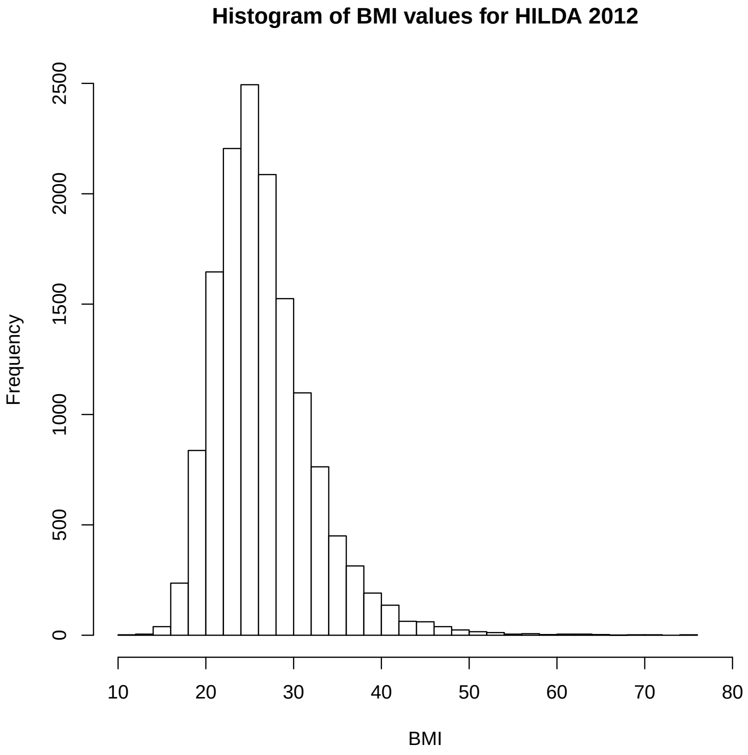


Figure 5.1: BMI Histogram

Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
11.7	22.9	25.8	26.7	29.4	75.8

Table 5.1: Summary Statistics: BMI

5.3.3 Results

In order to estimate the MED, a value for k (equation 5.5) is required. This value along with the number of covariates in the model specifies the number of MED parameters (β) to be estimated. The value of k is chosen based on the existence of moments in the data. Hence, the first step is to estimate the number of moments available in the data. The Hill estimator (Hill (1975)) is used to estimate the tail index ($\hat{\alpha}$) of the BMI distribution. This can then be used to estimate the highest moment available in the BMI distribution. The results indicate that the sixth moment exists. Given this, this study conservatively sets the value of k to 4 based on the paper by Wu (2003). The results in Wu (2003) provide an insight on the effect of sequentially updating the moment conditions i.e. iteratively including one moment condition at a time in the optimization process. The results show that there is only a marginal improvement in AIC and BIC measures when increasing the values of k from 4 to 12. Additionally, the interpretation could be an issue with moments higher than 4. Next the conditions for the existence of the MED are verified (chapter 2). This is done by computing the determinant of the Hankel matrix and ensuring that it is positive.

Finally, covariates are selected from the survey data. This selection process takes into account different type of covariates which may potentially impact BMI levels in individuals. These covariates include examples of physical, economic and social attributes of individuals. This consistent with the approach used in the literature. For instance, a study by Zhang and Wang (2004) examines the relationship between BMI and Gender, Socio-economic inequality, Age and Ethnicity. Similarly, Houle (2010) investigates the effect of Gender, Ethnicity and Education on BMI. A study by Bottai et al. (2014) considers the impact of physical activity on BMI. Section A.3.3 provides summary statistics for the covariates considered in this study.

Next, a number of model specifications consisting of different covariates affecting different parameters are attempted. Expert opinion is used to guide this process. The final model is selected based the BIC value. The resulting model specification is given by:

$$
\begin{aligned}
\lambda_1 &= \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 \\
\lambda_2 &= \beta_4 z_4 + \beta_5 z_5 + \beta_6 z_6 \\
\lambda_3 &= \beta_7 z_7 \\
\lambda_4 &= \beta_8 z_8.
\end{aligned}
$$

Table 5.2 contains the final covariates and their corresponding estimates. The estimates are significant at 5% level.

Covariate	Estimate $\hat{\beta}$
(Log of) Age (z_1)	0.238735
Active (z_2)	0.100000
Married (z_3)	0.000012
Male (z_4)	-0.000036
(Log of) Household Income (z_5)	-0.001385
Number of Children (z_6)	-0.000557
Employment (z_7)	0.000017
University Education (z_8)	-0.000003

Table 5.2: Final Model

Inputting these estimates along with a set of covariates pertaining to an individual into equation 5.5 produces a BMI distribution for that individual. As expected the final model does contain covariates which are significant in other BMI studies. For example, in the study by Brown et al. (2014) all of above covariates were used in their analysis. However, in their study the household income and number of children were not significant. One possible reason for this may be that their approach tracks changes in the mean value of BMI for a given covariate. On the other hand, the approach used in this paper is able to track changes in moments other than the mean. Hence, the results in this paper show that the household income affects more than just the mean value of the distribution i.e. variance or higher moments. Similarly, the number of children may not impact the mean value of BMI distribution as shown in Brown et al. (2014), but does affect higher moments of the BMI distribution.

5.3.4 Discussion

Given the attributes of an individual (covariates), the proposed model can estimate the BMI distribution for that individual. More specifically, this model can be used to determine the marginal effect/ sensitivity that each covariate has on the distribution of BMI. Furthermore, it can be shown how the broader BMI categories (as defined by AIHW) change with respect changes in a given covariate. In order to clearly show the marginal effect of a covariate, a *base case* BMI distribution is used to benchmark the change in distribution. The base case is chosen to ensure that the distribution of the broad BMI categories is consistent with the estimates produced by AIHW.

Figure 5.2 shows the marginal effect of age on the BMI distribution and table 5.3 translates this change in terms of the broader BMI categories. In this instance, the base case age has been increased by 10 years with the remaining covariates left unchanged. It can be seen that this change shifts the base case distribution to the right. Hence the marginal effect of age results in a changing the mean of the BMI distribution. In terms of the BMI categories, this marginal effect produces changes in all of the BMI categories. These changes represent changes in probability of an individual belonging to a particular category. For example, in the base case the probability of an individual being in the underweight category is 6.66%. Increasing the age of the individual by 10 years (all other covariate staying the same), the probability of the individual belonging to the underweight category drops to 3.09%. A similar interpretation can be applied to movements in the other categories. Note that the overweight category does not change much relative to the other categories.

	Underweight	Normal	Overweight	Obese
Base Case	6.66%	32.56%	33.17%	27.62%
Age	3.09%	22.89%	32.82%	41.20%

Table 5.3: Age Effect - Change in BMI categories

Figure 5.3 shows the marginal effect of household income on the BMI distribution and table 5.4 translates this change in terms of the broader BMI categories. In this instance, the base case household income has been reduced by 50% with the remaining covariates left unchanged. From figure 5.3, it can be seen that the resulting distribution differs from the base case considerably. Not only does the mean of the modified distribution shift, the variance and other moments also change. Hence the marginal effect of household income results in changing more than the mean of the BMI distribution.

Figure 5.4 shows the marginal effect of employment on the BMI distribution and table 5.5 translates this change in terms of the broader BMI categories. In this instance, the base case has been changed from an individual being employed to being unemployed

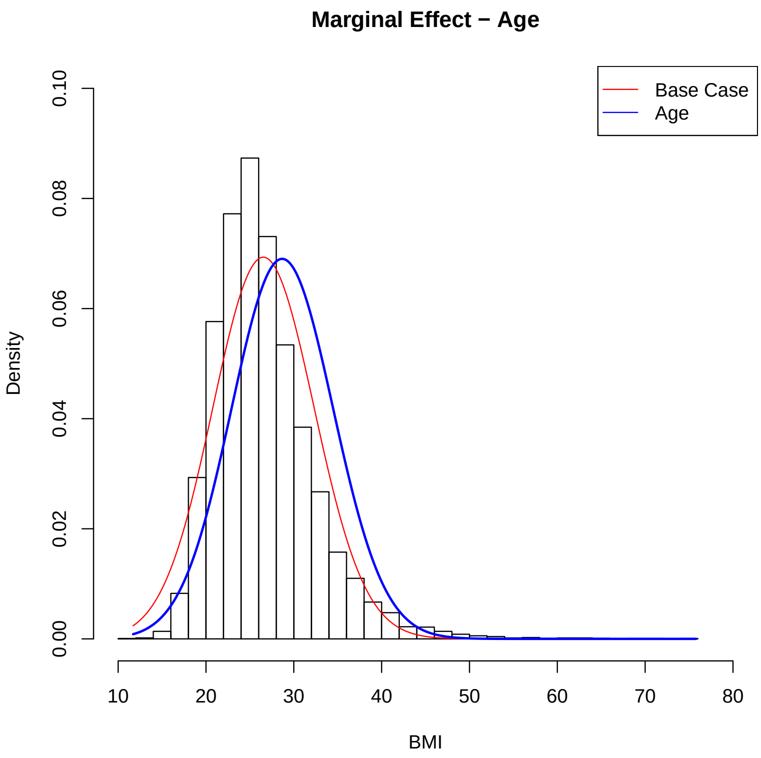


Figure 5.2: Marginal Effect of Age

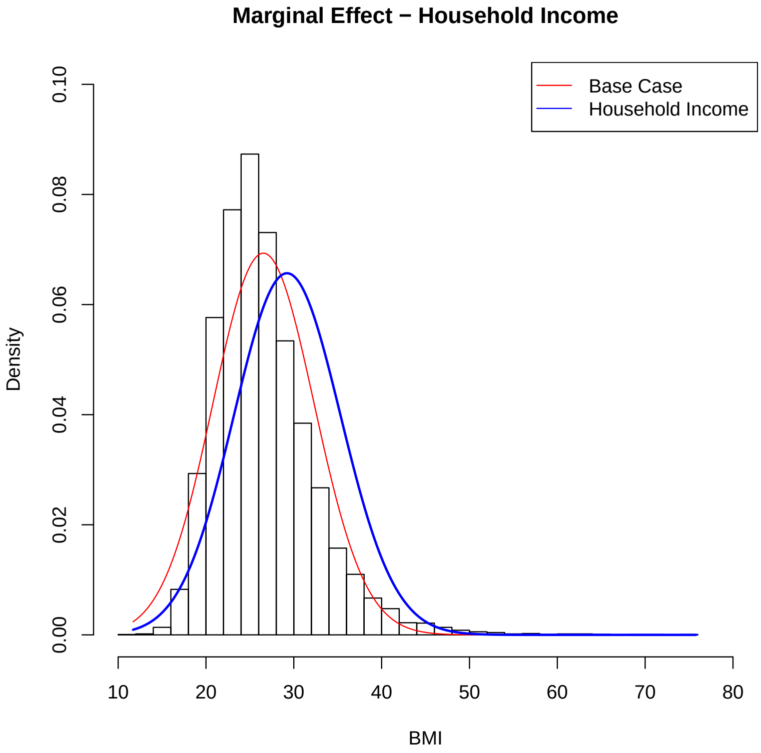


Figure 5.3: Marginal Effect of Household Income

with the remaining covariates left unchanged. It can be seen that the resulting distribution producing a shift in the mean as well as a possible decrease in the overall variance. In

	Underweight	Normal	Overweight	Obese
Base Case	6.66%	32.56%	33.17%	27.62%
Household Income	3.05%	20.96%	30.74%	45.25%

Table 5.4: Household Income: Change in BMI categories

this scenario, the individual's BMI results shift towards a more healthier range. The probabilities of falling in the overweight and obese category drop and as a result the probability of falling in the normal and underweight categories increase.

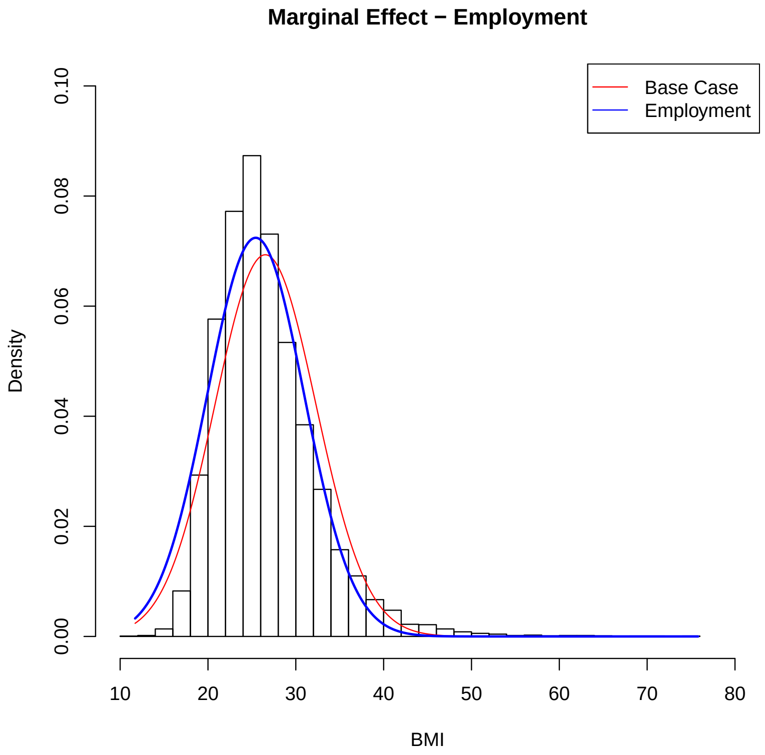


Figure 5.4: Marginal Effect of Employment

	Underweight	Normal	Overweight	Obese
Base Case	6.66%	32.56%	33.17%	27.62%
Employment	8.76%	38.02%	32.79%	20.43%

Table 5.5: Employment: Change in BMI categories

Figure 5.5 shows the marginal effect of employment on the BMI distribution and table 5.6 shows effects of this change in terms of the broader BMI categories. Here, the base case is changed from an individual not having a university education to having one. All other covariates remain unchanged. Similar to the employment effect, the resulting BMI distribution has a slightly lower mean and notable difference in the variance. In particular, the tail of the resulting distribution exhibits different behaviour. The changes in the broader categories highlight the positive impact that a university education has on BMI.

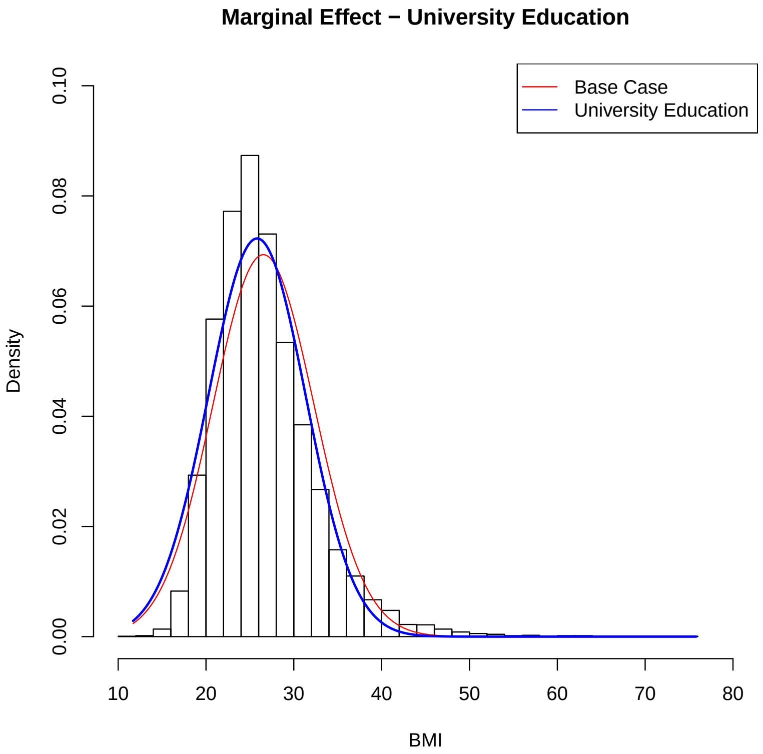


Figure 5.5: Marginal Effect of a University Education

	Underweight	Normal	Overweight	Obese
Base Case	6.66%	32.56%	33.17%	27.62%
University Education	7.87%	36.34%	33.47%	22.32%

Table 5.6: University Education: Change in BMI categories

5.3.5 Conclusion

This study has proposed a method to model the distribution of an individual's BMI using information from the covariates. The application of the MED framework as well the incorporation of covariates into this framework presents a novel approach with regard to BMI modelling. The results clearly show how different covariates affect different aspects (moments) of the BMI distribution. Furthermore, the results also show the shift in the broad BMI categories caused by the marginal changes in the covariates. Additionally, the consistency and asymptotic normality of the estimator have been shown. In terms of future direction, it is expected that this methodology can be extended to accommodate a panel data set up. This would allow one to assess changes in the distribution of BMI over time.

Chapter 6

Modelling Populations in Remote Communities

6.1 Introduction

There is a substantial and well-documented disparity in outcomes for Aboriginal and Torres Strait Islander and other Australians across virtually all mainstream indicators of social and economic well-being. *Closing the gap* is the broad label for the Council of Australian Governments' National Indigenous Reform Agenda aiming to reduce this statistical inequity faced by Aboriginal and Torres Strait Islander Australians.

The role of remoteness in this narrative is both pivotal and controversial. Relative to other Australians, a far higher proportion of Aboriginal and Torres Strait Islanders are found in remote areas of the continent. Based on data from the 2011 Census, 21.4% of Aboriginal and Torres Strait Islander Australians live in areas classified as 'remote' or 'very remote' compared to just 1.7% of other Australians. Only 34% of Aboriginal and Torres Strait Islanders live in a major city, compared to 71% of other Australians. While the large population centres generally offer greater access to services and infrastructure and better labour market opportunities, for many Aboriginal and Torres Strait Islander Australians those centres cannot offer the connections to homelands, culture and kinship networks that are intrinsic to their well-being.

Australia's current policy discourse expresses reservations over the viability, or sustainability, of remote Aboriginal and Torres Strait Islander communities. Implicitly, and in some cases explicitly, it seems governments would prefer a rationalisation of these communities, such that there was a migration out of the smaller and more remote communities, and their eventual disappearance. Examples of this policy mindset can be seen in the Northern Territory Government's 2009 *Working Future* policy statement (NT-Government, 2009), which identified 20 'Territory Growth Towns' as part of a hub-and-spoke model of service delivery; and more recently the Western Australian Government foreshadowing the withdrawal of services to up to 150 of the smaller and more remote communities in that state (WAToday, 2014a,b). The Western Australian Government's strategy set out in the recently formulated 'roadmap' for the reform of service delivery in remote communities is to focus efforts and investment on communities that have 'significant educational and employment opportunities', explicitly acknowledging that: *In concentrating on towns and larger communities, the State Government expects to support fewer communities over time, particularly as migration away from small outstations continues. (Regional-Services-Reform-Unit, 2016) page 12'.*

If some Aboriginal or Torres Strait Islander communities are seen to be unviable then, broadly speaking, there are two policy approaches that could be taken. One is to promote economic development and employment in those communities, to make them more self-sufficient. The second is to promote adjustment that would see people move out of those communities. Such adjustment may be achieved by pull factors, which provide positive incentives for out-migration or reduce the cost associated with moving; push factors, such as withdrawing services, activity testing and welfare quarantining that make life in those communities more difficult; or some combination of the two. However, Australia's history is littered with examples of failed attempts to manipulate the geography and mobility of Aboriginal and Torres Strait Islander people, stemming largely from the failure to recognise the importance of cultural drivers to their well-being (Dockery, 2016).

Clearly the effectiveness of policy and service funding allocations for remote communities will be highly sensitive to the degree of consistency between assumed and actual population trends. That is, there is a paramount need for separate estimates for Aboriginal and Torres Strait Islander population projections, a result of the vastly differing demographies, service requirements and migration patterns from other populations (Taylor et al., 2006; Wilson, 2009). In addition to the total population of remote communities, projections of the age structure of individual communities are extremely important in determining the likely mix of services required, such as the number of school places, available health services and aged care facilities. Such forecasting exercises are inherently difficult for Aboriginal and Torres Strait Islander populations in remote Australia. This study proposes an empirical approach to dealing with these challenges and reports results from one of a suite of models being developed for forecasting Aboriginal and Torres Strait Islander population changes at the community level by age group and gender. The focus here is on broad population trends from 2011 to 2016 by remoteness and community size. The forecasted release of the actual 2016 Census data (expected in late 2017, at the time of writing), will not only provide actual data by which to evaluate our predictions, but will also provide additional data which can be used to improve the models thus far developed. This will also provide a basis to to extend detailed projections out to 2021 and 2026.

The following section provides background on existing approaches to projecting the Aboriginal and Torres Strait Islander population. Sections 6.3 and 6.4 contain details on the data set and the methodology used in the study. Section 6.5 reports the results from the modelling of 2011 populations using the 2006 data, and Section 6.6 presents the 2016 projections that are obtained by applying the modelling results to the 2011 data. Projections by individual community are provided in the Appendix. The final section discusses scope for further development and application of the method.

6.2 Background

The Census of Population and Housing, conducted every five years by the Australian Bureau of Statistics (ABS), is the principal source of estimates of Australia's population and its demographic composition. It is also the main means by which the Aboriginal and Torres Strait Islander population is counted and, since 1971, the intention has been for a full enumeration of that population (Wilson, 2009). The ABS 2011 Census initially recorded 548,000 people of Aboriginal or Torres Strait Islander descent, representing 2.5% of the total population. The ABS also produces estimates of the Aboriginal and Torres Strait Islander resident population based on a post-enumeration survey and statistical methods to adjust for the net undercount and the substantial number of people for whom Aboriginal or Torres Strait Islander status is unknown. Table 6.1 shows how the Aboriginal and Torres Strait Islander share of the population increases with remoteness according to the original 2011 Census estimates. Only 34% of Aboriginal and Torres Strait Islander Australians live in Major cities but 14% live in areas classified as Very remote; 71% of other Australians live in Major cities and less than 1% live in areas classified as Very remote. So while Aboriginal and Torres Strait Islander Australians make up around 2.5% of the overall population, they represent almost 41% of people living in 'very remote' Australia.

One of the main approaches to population projection is the cohort-component method

Table 6.1: Australian population estimates by Aboriginal and Torres Strait Islander status and remoteness, 2011 Census

	Aboriginal and Torres Strait Islander		Non-Aboriginal and Torres Strait Islander		Total		
Remoteness Area (ARIA level)	People	%	People	%	People	%	Aboriginal and Torres Strait Islanders Share
Major cities (1)	188,537	34.4	14,094,903	70.8	15,006,519	69.8	1.3
Inner regional (2)	121,301	22.1	3,695,423	18.6	3,998,424	18.6	3
Outer regional (3)	118,491	21.6	1,735,627	8.7	1,963,404	9.1	6
Remote (4)	39,751	7.2	234,833	1.2	300,107	1.4	13.2
Very remote (5)	77,493	14.1	98,564	0.5	190,266	0.9	40.7
Total	548,366	100	19,900,767	100	21,507,719	100	2.5

Source: ABS CData online Table Builder facility. Note: columns sum to less than the total as the categories of 'Migratory – offshore – shipping' and 'No usual address' have not been reported. Rows do not sum, as people for whom Aboriginal and Torres Strait Islander status was 'not stated' have not been reported.

(see Booth (2006) for a review). For given age cohorts in a base year, assumptions regarding deaths, births, immigration and emigration are applied to arrive at future projections for that group at a given point in time. For a range of reasons, applying this method to project populations of Aboriginal and Torres Strait Islander people living in remote communities, which is done in this study is fraught with additional complications that are likely to compound projection errors.

First, the accuracy of even the baseline data is questionable. The ABS post-enumeration estimate of the total Aboriginal and Torres Strait Islander population was 21% higher than the original census estimate as a result of both under-counting of people and non-identification of people as being of Aboriginal or Torres Strait Islander descent. There is evidence that earlier censuses under-counted young children and young to middle-aged adults, with inaccuracies more pronounced in remote areas (Taylor, 1997). Despite considerable efforts on the part of the ABS, such problems persisted for more recent censuses (Wilson and Barnes, 2007). One issue is that kinship structures and mobility among Aboriginal and Torres Strait Islander families do not match assumptions underlying the enumeration strategies that individuals can be traced to a unique household and reported upon by a unique reference person within that household. In the 2011 Census, there were twice as many people for whom Aboriginal or Torres Strait Islander status was not stated as there were people who identified as being of Aboriginal or Torres Strait Islander descent, meaning error margins around those population estimates are very large.

In addition to its effect on baseline estimates, the issue of identification means there is a further source of population change within a cohort in addition to the three standard components of mortality, fertility and migration. For cohort-component models that attempt to derive separate projections of the Aboriginal and Torres Strait Islander and other populations, this can include the question of how children from mixed families are likely to be identified, or identify themselves, in future censuses and the impact of policy changes on the propensity to identify (Biddle and Wilson, 2013; Taylor, 1997; Wilson, 2009). These challenges are specific to the enumeration of Aboriginal and Torres Strait Islander Australians. On top of these, Taylor (2014) notes general problems associated with projecting populations for sparsely populated areas, including that they are more vulnerable to vagaries of exogenous impacts that may impact upon them, such as weather events and policy changes; data collection is more resource-intensive; and proportional errors in projections tend to increase the smaller the population size of the units being analysed and if there is rapid change occurring in the period in which the baseline data is compiled.

Compounding the above issues, most projection methodologies are based on the assumption of large population counts and cannot be applied to small populations. Taylor et al. (2006) suggest a regional population size of around 10,000 people is required to meaningfully apply age-conditional mortality analyses. Wilson (2011) suggests that exponential models have favourable properties over linear ones. However, if there are zero counts in the component categories (such as in certain gender-by-age categories) exponential models cannot be used, and inferences can be distorted by extreme values in terms of percentage change where counts are small.

An example of how these difficulties impact upon population forecasts for sparsely populated areas is provided by Taylor (2014) in the assessment of the accuracy of ABS projections for the Northern Territory (from the 1970s through to 2012). Even at this territory level, mean absolute percentage errors in the ABS projections are far higher than for Australia as a whole. Several relatively naive models, based on simple extrapolation of growth trends, out-performed the more sophisticated ABS cohort-replacement model.

Wilson (2009, p. 232) compiles a summary table of studies providing projections for the Aboriginal and Torres Strait Islander population. Of the 13 publications included, 11 produced estimates at the national or state/territory level, with Taylor and Bell (2002) and Taylor (2003) being the exceptions. Using a composite approach that combines ABS census population estimates with other data sources, such as health clinic registrations and school enrolments, Taylor and Bell (2002) produced projections to 2021 for Cape York Peninsula (then the Peninsula ATSIC region), commencing from an adjusted baseline population in 1986 of 6,500 Aboriginal and Torres Strait Islander people. Taylor (2003) produced projections for 2006 to 2021 from the 2001 baseline of 37,000 Aboriginal and

Torres Strait Islander people in the Australian desert region, which focused on the arid zone covering much of Western Australia, the Northern Territory, South Australia and parts of Queensland and New South Wales.

Following the same approach, Taylor et al. (2006) later generated separate projections for the arid, semi-arid and savanna-biogeographical zones to 2021, with the 2001 baseline Aboriginal and Torres Strait Islander populations in the semi-arid and savanna zones put at 51,000 and 84,000, respectively. More recently, Biddle and Taylor (2009) applied the cohort-component approach to generate projections of the Aboriginal and Torres Strait Islander population from 2006 to 2031 for each of 37 ABS defined 'Indigenous Regions'. The smallest of those regions in terms of the Aboriginal and Torres Strait Islander population was Ceduna, with 2,248 people.

Taylor et al. (2006, p. 3), nominate methodological developments in the treatment of small areas and subsequent small number analysis as a key imperative to improving demographic information for Aboriginal and Torres Strait Islanders in desert Australia. Of the studies discussed above, only the Biddle and Taylor (2009) report, and possibly Taylor and Bell (2002) projections for the Cape York Peninsula, commencing from a baseline Aboriginal and Torres Strait Islander population of 6,500, could be considered as dealing with a small area or small population. In contrast, this study produces projections for communities with as few as nine Aboriginal and Torres Strait Islander inhabitants. In addition, existing exercises often attempt long projection horizons. While some projections for Aboriginal and Torres Strait Islander populations are for five or 10 years beyond the baseline, typically they are for 20, 30 or even 50 years forward, well beyond the planning and policy cycles that Aboriginal and Torres Strait Islander communities are subject to.

6.3 Data

To explore demographic changes in remote Aboriginal and Torres Strait Islander communities in more detail, this study models intercensal changes based on the ABS defined geography of 'Indigenous Locations': *Indigenous Locations (ILOCs) are aggregates of one or more Statistical Areas (Level 1). ILOCs generally represent small Aboriginal and Torres Strait Islander communities with a minimum population of 90 Aboriginal and Torres Strait Islander usual residents. An ILOC is an area designed to allow the production of census statistics relating to Aboriginal and Torres Strait Islander people with a high level of spatial accuracy while maintaining the confidentiality of individuals. For the 2011*

Census, 1,116 ILOCs have been defined to cover the whole of geographic Australia.There are non-spatial ILOCs for Migratory - Offshore - Shipping and No Usual Address in each state and territory (S/T). (ABS, 2011).

The approach is to develop an empirical model of the 2011 Aboriginal and Torres Strait Islander population counts for communities based on the 2006 Census counts and selected characteristics of the communities. It utilises the fact that with the census undertaken every five years, and the data being available in five-year ranges, population changes for each community can be derived by age group. Treating those age-specific changes as multiple observations for a community, this study adopts a multilevel modelling approach. Using this framework, it is possible to estimate community-specific effects to identify growth communities from those in decline, conditional upon factors such as their remoteness and initial size. Projections for the 2016 population are then generated by applying the coefficients from the model to the 2011 data as the base year, under the assumption that the 2006–2011 trends identified in the model continue. Potentially, population projections for 2021 and beyond could be generated treating the 2016 forecasts as the base year, and so on.

Census population estimates were downloaded from the ABS Table Builder online facility for all 1,098 spatial ILOCs defined in 2011 (*i.e.*, excluding the non-spatial ILOCs) and for all 838 spatial ILOCs that were defined for 2006. The data extracted for each ILOC include the population in five-year age groups by sex and Aboriginal or Torres Strait Islander status. The status variable has five categories: non-Indigenous, Aboriginal, Torres Strait Islander, both Aboriginal and Torres Strait Islander, and not stated. There have been considerable changes to the ILOC geography since 2006. The ABS provided a concordance of 838 2006 ILOCs to the 2011 ILOCs, which reports estimates of the percentage of the population within each 2006 ILOC that would correspond to a 2011 ILOC, with the proportions weighted according to the respective Aboriginal and Torres Strait Islander populations. These were used to generate 2006 population estimates corresponding to the 2011 geography. For example, the 2011 ILOC of Amoonguna corresponds to the 2006 ILOCs of Amoonguna (100%) and some of Sandover and Outstations (10%). Hence for each age group by gender by Indigenous status' cell, 2006 estimates are computed as a weighted sum: 1 times the Amoonguna estimate + 0.1 of the Sandover and Outstations estimate. These concordances will be less appropriate for the non–Aboriginal and Torres Strait Islander cells, but this will be of little consequence since the analyses concentrates upon estimates of the Aboriginal and Torres Strait Islander population.

The resulting data include 2006 and 2011 population estimates for 1,098 ILOCS that are geographically comparable between the two years. As described below, a subset of 618 ILOCs in ‘outer regional’, ‘remote’ and ’very remote’ Australia was utilised in the modelling.

The census counts of the number of Aboriginal, Torres Strait Islander and both Aboriginal and Torres Strait Islander people were aggregated to a single Aboriginal and Torres Strait Islander population estimate. For each ILOC in each census year, population data are available for 42 gender-by-age cohorts: 21 age groups (0–4, 5–9, 10–14 ... 90–94, 95–99, 100+) each for males and females. In the older cohorts, a large proportion of these cells have either zero or very small population counts and this posed a problem for a number of the models that were considered for this study. Consequently, the older age cohorts were aggregrated as shown in Figure 6.1.

A key concern of the current paper is to generate projections for remote Aboriginal and Torres Strait Islander communities and, given that population changes in the major cities and regional centres are likely to be driven by markedly different processes, only ILOCs in ’outer regional’, ’remote’ and ’very remote’ Australia were included in the sample for estimation (corresponding to ARIA levels of 3, 4 and 5). A small number of other ILOCS were excluded due to the fact that there were almost no Aboriginal and Torres Strait Islanders present at the time of the census. These exclusions were:

- Lord Howe Island, off New South Wales. The 2006 Census recorded no Aboriginal or Torres Strait Islander usual residents of Lord Howe Island
- The external territories of Christmas Island and the Cocos Islands (in part to enable state dummies to be included among the dependent variables)
- The Northern Territory ILOCs of Apatula (Finke) Homelands and Walungurru Outstations. The data for these ILOCs had zero counts in the vast majority of age-by-gender cells.

The final sample is based on data from 618 ILOCs, generating 19,776 observations for the regression analysis (32 observed cohort changes per ILOC). Also included in the data set is the natural logarithm of the total population of the ILOC (including non–Aboriginal and Torres Strait Islander people) to capture differences by community size; dummy variables for cohort age, gender, ARIA and state/territory; the gender and age-specific five-year survival rate for each cohort; and interaction terms between the age group and ILOC size, and between age group and ARIA. The interaction terms are included to allow for possible differential effects of age in smaller *versus* larger communities and in less remote versus more remote communities. For example, previous studies have identified trends in which younger Aboriginal and Torres Strait Islander people tend to move away from smaller and more remote communities into larger regional centres, while older people tend to move back out into the country. A dummy variable was also included to indicate whether the community was nominated as a Territory Growth Town under the Northern Territory Government's 2009 Working Future policy (see (Sanders, 2010)). Table 6.2 provides definitions of the variables in the data sets. Descriptive statistics for these, along with the dependent variable, are provided in the section A.4.1 of the appendix.

6.4 Methods

In order to assess changes in the population of an ILOC across the two time periods, there are two board approaches that this study employs. The first approach is an aggregated approach, whereby the change in the total population of an ILOC across the two time periods is considered. The second approach considers these changes for each age group within an ILOC across the two time periods. The methods for both approaches are discussed in this section.

The data available provides the number of individuals in each age group (and gender) across all ILOCs for the years 2006 and 2011. From this, one can build an *age group* distribution of the overall population by aggregating each age group category for all ILOCs. This is carried out for both the 2006 and 2011 census years. This provides a reasonable starting point for measuring the changes in the overall age group distribution between 2006 and 2011. Similarly, by aggregating all age groups for a each ILOC, an *ILOC group* distribution can be constructed for each time period. As before, these distributions can be compared across the two time periods in order to assess if there has been a change in the overall population.

One measure that can be used to assess the *change* in such a distribution is the *Kullback-Leibler (KL)* divergence measure, also known as relative entropy (Kullback and Leibler, 1951). This is a measure of how a *proposed* distribution diverges from a *benchmark* distribution. In this context, the benchmark distribution is the 2006 age group distribution and the proposed distribution is the 2011 age group distribution. A significant advantage of this measure is that it does not require any distributional assumptions. Belov and

Table 6.2: Definitions of variables in the data set

Variable(s)	**Description**
ILOC size	The natural logarithm of the resident total population of the ILOC in 2006 (including Aboriginal and Torres Strait Islander people, other Australians and those for whom Aboriginal or Torres Strait Islander status is not stated).
remote, Very remote	Two mutually exclusive dummy variables indicating whether the ILOC is in ARIA level 4 (remote) or ARIA level 5 (Very remote). The omitted or 'reference' category is Outer regional (ARIA level 3).
Victoria, Queensland, South Australia, Western Australia, Tasmania, Northern Territory	Six mutually exclusive dummy variables indicating the state or territory of the ILOC. New South Wales is the omitted category. There are no Outer regional, remote or Very remote ILOCs in the Australian Capital Territory.
Female	Dummy variable equal to 1 if the observation is for a female cohort, and equal to 0 if it relates to a male cohort.
Growth town	Dummy variable taking on a value of 1 if the ILOC contains or corresponds to one of the communities nominated as growth towns under the Northern Territory Government's *Working Future* policy announced in 2009. While the policy named 20 towns, one of these (Daguragu-Kalkarindji) falls across two ILOCs (Daguragu and Kalkarindji), meaning there are 21 ILOCs coded with a value of 1.
Age 10–14; Age 15–19, ... Age 75–79, Age 80+	Sixteen mutually exclusive dummy variables indicating the age of the cohort in 2011. The omitted category is Age 5–9.
Survival rate	Based on ABS Catalogue 3238.0 – *Estimated and projected Aboriginal and Torres Strait Islander population Series B for Australia* (ABS 2014). The ratio of the estimated population in each age cohort i in 2011 to the estimated population of age cohort $i-1$ in 2006. This gives an age-specific *apparent survival rate* and is calculated separately for males and females. The survival rates are close to unity for younger cohorts and decline to under 0.7 for cohorts beyond the age of 70 years.
ILOC size*age interaction terms	Fifteen separate variables are generated by interacting the continuous ILOC size variable with the age group dummies. The omitted age category is Age 5–9. The coefficients on these variables indicate whether, within each specific age group, there is any further effect of community size in addition to the average effect of community size observed across all age cohorts.
Outer regional*age and remote*age interaction terms	Thirty separate dummy variables generated by interacting the Outer regional dummy variable (ARIA level 3) with age cohort and the remote dummy variable (ARIA level 4) with age cohort. The omitted categories are Very remote (ARIA level 5) and Age 5–9. The coefficients on these variables indicate whether, within each specific age group, there is any further effect of remoteness in addition to the average effects observed across all age cohorts.

D Armstrong (2011) use this methodology to identify differences in performance across a range of tests for a given individual.

Let $p_{i,2006}$ be the probability of individuals in age group i for all ILOCs in 2006. Let $p_{i,2011}$ be the probability of individuals in age group i for all ILOCs in 2011. Then, the KL divergence from $p_{i,2006}$ to $p_{i,2011}$ is given by

$$KL = \sum_{i=1}^{N} p_{i,2006} \log \left(\frac{p_{i,2006}}{p_{i,2011}} \right), \tag{6.1}$$

where N is the total number of age-groups. From the equation above, it can see that if the two distributions are similar, then the KL divergence would be close to zero. If this is not the case, then the resulting divergence would be significantly different from 0.[1]. This measure allows to one to answer the question whether or not there has been an overall change in the age group distribution across the five year period. This method can be applied to various other distributions computed from the data. For example, one can focus on the assessing the change in distribution for across all ILOCs conditioning on gender or level of remoteness.

In order to compare the results produced by the KL divergence measure, one can use the chi-squared statistic and the Kolmogrov-Smirnov (KS) statistic. The chi-squared statistic is a second order accurate approximation of twice the KL divergence. This statistic is given by:

$$\chi^2 = \frac{\sum_{i=1}^{N} (p_{i,2011} - p_{i,2006})^2}{p_{i,2006}} \tag{6.2}$$

Another measure to compare distributions is the two sample KS test statistic. This measure can be used to test if two samples are drawn from the same distribution. This is achieved by computing the maximum distance between two the empirical cumulative distributions. Let $F_{i,2006}$ and $F_{i,2011}$ denote the empirical age group distributions for years 2006 and 2011 respectively. The KS test statistic is defined as

$$KS = \sup_i |F_{i,2006} - F_{i,2011}| \tag{6.3}$$

This quantity is then compared against the critical values obtained from the KS distribution. If this quantity is significant, then the two samples are not drawn from the same distribution. The KS test statistic is a non-parametric test and similar to the KL divergence does not require any distributional assumptions. All of the above measures indicate

[1] Critical values for the KL divergence can be obtained from the double F-distribution under normality

Census	Age					
j	1	2	…	15	16	17
2006	0–4 ↘	5–9 ↘	…	70–74 ↘	75+ ↘	
2011	0–4	5–9	…	70–74	75–79	80+

Figure 6.1: Cohort Structure of the Census population data

whether or not there has been a change in the age group distribution across both years. Furthermore, these measures are also applied to other distributions based on gender and remoteness.

Thus, this first approach allows one to test whether the overall distribution has changed over time. Changes in the overall distribution would imply that some aspect of the population has indeed changed over time. However, in order to assess what is happening within the ILOC, a different approach is required. This second approach aims to study the changes in population for each of the age groups within a given ILOC. This approach is more widely known as the cohort-component model. Using this approach, the number of Aboriginal and Torres Strait Islanders aged 5–9 in a given ILOC in 2011 will be the number who were aged 0–4 in 2006 minus deaths, net migration and net changes associated with identification of Aboriginal or Torres Strait Islander status. The data set-up is demonstrated in Figure 6.1, where the arrows trace the cohort movement through time.

Using j to index the age group categories, it can be seen that there are 16 such cohort progressions or flows in which the population (P) in 2011 (t) can be related to the population in the earlier age category in 2006($t-1$):

$$P_{j-1,t-1} \rightarrow P_{jt}. \tag{6.4}$$

That is, ignoring extraneous factors, the age group in say 0-4 simply ages to that in 5-9, between censuses.

With population data for males and females, this gives 32 flows observed for each ILOC. This can be treated as a multi-level model framework in which there are 32 observations for each ILOC. Following the convention in the econometric multi-level modelling literature (Rabe-Hesketh and Skrondal, 2008, p. 65), the entity or the ‘cluster’ is denoted by subscript i and the ‘occasions’ providing repeat observations for that cluster by subscript j. Hence, in this case communities are denoted by subscript i (i = 1 to 1,098) and the 2011 age groups by subscript j (j = 2 to 17). For convenience, the gender distinction is ignored for the purposes of setting out the model. From 6.4, a modelling framework can be developed which incorporates clustering at the ILOC level,

$$P_{ijt} = f(P_{i,j-1,t-1}). \tag{6.5}$$

This study does not impose estimates of the components of migration, deaths or changes in identification. Rather, the functional relationship between P_{ijt} and $P_{i,j-1,t-1}$ is estimated by taking account of observable ILOC characteristics and unobservable ILOC-specific effects. A range of options are available for modelling the changes in cohort populations between 2006 and 2011. Exploration with these options indicated that it is preferable to model the change in population in linear terms (or 'levels') rather than in growth terms. Models based on the proportionate change, such as the growth rate, are highly sensitive to extreme values that mostly arise where cell counts are small. For example, percentage changes in some age categories in smaller communities were in the thousands, and are likely to be affected by concordance and enumeration issues. Modelling changes in growth rates also means omitting observations for which the base number is zero, even though that is a legitimate value (that is, a gender/age cohort/ILOC observation with a currently zero population, in this context 'could not' be allowed to grow). Indeed, these issues are certainly all very pertinent when dealing with small area counts of Aboriginal and Torres Strait Islander people by gender and age.

An alternative approach is to model the population in 2011 as a direct function of the corresponding lagged (2006) population values. Explaining current population levels with past ones, especially over a short number of time observations, raises several econometric concerns, such as endogeneity, non-stationarity and the strong possibility of being adversely affected by the spurious regression problem. Indeed, in a simple regression, the lagged dependent variable clearly exhibited signs of these issues, with the estimated coefficient being very close to unity in value and with an extremely high t-statistic. Further, cells with zero counts and very small counts continue to pose a problem. Thus the preferred strategy adopted here is to model the change in population for each cohort. That is, the dependent variable becomes:

$$c_{ijt} = P_{ijt} - P_{i,j-1,t-1}. \tag{6.6}$$

As such there are 32 observations of c_{ijt} for each ILOC ($i = 2$ to 17 for males and females). Referring back to Figure 6.1 above, the model estimates changes in the population of people aged 5-9 in 2011 from the number aged 0-4 in 2006; of people aged 10-14 in 2011 from the number aged 5-9 in 2006; and so on. However, it cannot provide estimates for the population aged 0-4 in 2011, as there is no younger cohort in 2006 to use as the baseline. To enable projections for the total populations by ILOC, a separate fertility model is developed to generate estimates of the number of males and females aged 0-4 in 2016 (see 6.5).

Stochastically then, population changes can be modelled as:

$$c_{ijt} = \alpha_i + \mathbf{x}'_{i,j,t-1}\boldsymbol{\beta} + \varepsilon_{ijt} \text{ where } \varepsilon_{ijt} \sim N(0, \sigma^2) \qquad (6.7)$$

and $\mathbf{x}_{i,j,t-1}$ represents a vector containing the independent variables (all defined in the base year 2006; $\boldsymbol{\beta}$ is a vector of coefficients to be estimated; α_i consists of unobservable community-level (ILOC) effects and ε_{ijt} denotes the errors in the model. For most applied work, there are two major approaches to estimating unobserved effects: fixed and random-effects. The primary consideration in choosing between them involves the assumption regarding the relationship between the α_i's and the x's. Unlike the random-effects approach, the fixed-effects one assumes an arbitrary correlation between these.

The above model (equation 6.7) can be estimated by simple linear regression. However, this is not appropriate for the 'small numbers' model we are dealing with here, as the date will be necessary truncated in many instances as the dependent variable has an effective lower bound. That is, any population cannot decrease by more than the starting value. For example, consider an ILOC with five male Aboriginal and Torres Strait Islander people aged 20–24 in 2006. The change in the population from 2006 to 2011 can only range from -5 upwards, and hence the population of males aged 25–29 in 2011 cannot be less than zero.

Note also that if the population of males aged 25–29 in 2011 is also five, then there has been no change and the dependent variable c_{ijt} equals zero. This zero is a legitimate value indicating no change in the population, and indeed any estimation the expected value of c_{ijt} would (should) accordingly be close to zero. However, consider the case in which there are no individuals in a particular age-by-gender category in 2006, as is common for older age cohorts. In this case the population can increase ($c_{ijt} > 0$), but has a lower bound of zero, and the probability of observing zero change is much higher than for observations with a positive initial population. In such situations where there is a 'latent' potential change in the population that cannot be observed because of the effective lower bound, linear regression models will produce biased and inconsistent results, which will worsen with the extent of such censoring/*boundary* observations (Amemiya, 1984; Greene, 2003).

To alleviate any such potential issues in estimation, a preferable approach is to implement a Tobit model with varying censoring limits. As noted above, in the current context the limit is equal to the 2006 population for a specific gender-by-age group at a given ILOC. This differs from the usual Tobit model set-up, where lower (and/or upper) limits are usually fixed (for example, commonly at zero) and the same for all observations in the sample.

The previous model (equation 6.7) can be updated to reflect this possible truncation as

$$c^*_{ijt} = \alpha_i + \mathbf{x}'_{i,j,t-1}\boldsymbol{\beta} + \varepsilon_{ijt} \text{ where } \varepsilon_{ijt} \sim N\left(0, \sigma^2\right) \tag{6.8}$$

and c^*_{ijt} now denotes the *latent* underlying change in ILOC i for age group j at time t (2011). However, this cannot be fully observed due to the fact that the change cannot be less than the current population. In other words, there is lower tail censoring such that only c_{ijt} is observed. Hence

$$\text{if } c^*_{ijt} < -P_{i,j-1,t-1} \text{ then } c_{ijt} = -P_{i,j-1,t-1}. \tag{6.9}$$

Thus, as noted, in contrast to the standard Tobit model in which the lower (and/or upper) limit is assumed to be a fixed value(s) for all i and j, the proposed framework contains the 2006 population as a varying lower limit for each observation. Like the linear model, the Tobit model can be estimated assuming either random or fixed effects, although the latter will suffer from the well-known incidental parameters problem, if the dimension over which these are constant is ‘small’ (Greene, 2012). However, as robustness checks, extensive modelling was undertaken using both the linear regression and Tobit models, both of which with fixed and random unobserved effects. In light of these, it was found that the Tobit Model with random effects provided a superior fit compared to other models.

Currently the data contains only one observation on each c_{ijt} i.e. namely the change from 2006 ($t - 1$) to 2011 (t). In this sense, the model is cross-sectional rather than longitudinal and as such the time subscript can be omitted. The release of 2016 Census data will provide a second observation, which will result in a true multi-level panel structure. Observations on changes across multiple time periods will provide more rigorous estimation of community-level unobserved effects plus scope for further development. As a result, the random-effects Tobit model can be expressed as

$$c_{ij}^{*} = \mathbf{x}_{ij}^{'}\boldsymbol{\beta} + \varepsilon_{ij} + u_i, \tag{6.10}$$

where ε_{ij} and u_i are both normally distributed, with zero means and variances of σ^2 and ω^2, respectively. The data is observed as $c_{ij} = \max(L_{ij}, c_{ij}^{*})$ where $L_{ij} = -P_{i,j-1}$. In this context, this is an example of lower tail censoring: the change in population in the next time period cannot be less than the number of people (in an age group) currently residing in the ILOC. As per usual, the random effect is assumed to be the same for each time period and the ε_{ij} is uncorrelated across all ILOCS. To derive the log likelihood function, the focus here is on the conditional distribution of $f(c_{ij}|u_i)$ (Greene, 2012).

Let the dummy variable, $d_{ij} = 1$ indicate that $c_{ij} > L_{ij}$. This is the uncensored case and $d_{ij} = 0$ for censored cases. The conditional density of c_{ij} can then be expressed as

$$f(c_{ij}|u_i,\ d_{ij} = 0) = P(c_{ij}^{*} \leq L_{ij}|u_i) = \Phi\left(\frac{L_{ij} - \mathbf{x}_{ij}^{'}\boldsymbol{\beta} - u_i}{\sigma}\right)$$

for censored cases and

$$f(c_{ij}|u_i,\ d_{ij} = 1) = \frac{1}{\sigma}\phi\left(\frac{c_{ij} - \mathbf{x}_{ij}^{'}\boldsymbol{\beta} - u_i}{\sigma}\right)$$

for uncensored cases; where Φ and ϕ, respectively denote the *c.d.f* and *p.d.f* of the standardised normal distribution. Combining the two cases above

$$f(c_{ij}|u_i) = [f(c_{ij}|u_i,\ d_{ij} = 0)]^{1-d_{ij}} \text{ x } [f(c_{ij}|u_i, d_{ij} = 1)]^{d_{ij}}.$$

Assuming independence, the joint density of all observations in a group can be expressed as

$$f(c_{i1}, c_{i2}, \ldots, c_{iT_j}|u_i) = \prod_{j=1}^{T_j} f(c_{ij}|u_i).$$

Based on the results above, the log likelihood function of this model can be written as

$$\log L = \sum_{i=1}^{n} \log\left\{\int_{-\infty}^{\infty} \frac{1}{\omega\sqrt{2\pi}}\exp\left(-\frac{u_i}{2\omega^2}\right)\prod_{j=1}^{T_j}\left[\Phi\left(\frac{L_{ij} - \mathbf{x}_{ij}^{'}\boldsymbol{\beta} - u_i}{\sigma}\right)\right]^{1-d_{ij}}\left[\frac{1}{\sigma}\phi\left(\frac{c_{ij} - \mathbf{x}_{ij}^{'}\boldsymbol{\beta} - u_i}{\sigma}\right)\right]^{d_{ij}} du_i\right\}$$

Lastly, find values of β, σ and ω such that this function is maximised. The integrals can be computed using Gauss-Hermite quadrature (or by simulation) and the function can be maximised using standard non-linear optimisation methods.

6.5 Results

As mentioned above, we first compare age group and ILOC group distributions across the two time periods. In addition to this, distributional comparisons are also carried out for these distributions conditioned on gender and remoteness (ARIA level). Table 6.3 shows the results of this analysis. The KL divergence measure indicates that the ILOC group distribution has not changed over time. Although the divergence measure increases marginally for ILOC distributions conditioned on gender, it is still a relatively low value and as such indicates that the distribution has not changed over time. The divergence measure drops further for ILOC distributions conditioned on remoteness. Similarly, the KL divergence is quite low for the age group distribution. This indicates there is no significant change in these distributions across the two time periods. This is also true age group distributions conditioned of gender and remoteness.

The results for the KL divergence are consistent with both the Chi-squared and KS test statistic. Hence, at an aggregate level, there seems to be no change in both ILOC and age group distributions across the two periods. This is to be expected, since significant changes in population do not typically occur in such a short time frame (one census period). From a policy standpoint, this does not support the idea of removing services to regional and remote areas due to declining populations.

Thus, although at an aggregate level there appears to be no significant change in distributionS across the two time periods, there are instances where the changes in population haVE been quite extreme at the ILOC level. This may be due factors specific to the ILOC. For example, the community of Kargaru (ILOC number 1095) decreased by 83% from 2006 to 2011. Similarly, the communities of Umoona (ILOC Number 640) and Coonana (ILOC number 693) decreased by 69% and 62% respectively. On the other hand, there were communities such as Tiwi Islands (a 8 fold increase) and Broome (Surrounds) (a four fold increase) which experienced significant growth in population.

As per the methodology described above, population change is defined using the cohort-component approach. The objective here is then to model this change (using using ILOC specific factors. Initially linear regression models with both random and fixed effects were fitted [2]. At first impression, these models seems to be performing reasonably

[2] The results for these models are available upon request

Table 6.3: Measuring change in distribution

Group	KL Divergence	$\frac{1}{2}$ Chi Sq	KS Statistic	KS p-val
ILOC group	0.023	0.024	0.065	0.15
ILOC group males	0.027	0.027	0.065	0.149
ILOC group females	0.024	0.025	0.063	0.17
ILOC group aria3	0.014	0.015	0.102	0.156
ILOC group aria4	0.014	0.015	0.074	0.959
ILOC group aria5	0.035	0.036	0.076	0.406
age group	0.003	0.003	0.143	0.987
age group males	0.003	0.003	0.15	0.983
age group females	0.003	0.003	0.19	0.853
age group aria3	0.004	0.004	0.2	0.832
age group aria4	0.005	0.005	0.15	0.983
age group aria5	0.002	0.002	0.05	1.000

well with the fixed effect model performing marginally better than the random effects model. However for some of the observations, the predicted change value is much smaller than the initial (2006) population value itself. For example, the predicted change for age group 9 for a given ILOC is -10 when the initial (2006) population for that age group is 8. This is clearly nonsensical given that the final projected population value is negative (-2).

To rectify this issue the Tobit model with varying lower limits was proposed (Section 6.4). The introduction of this censoring point ensures that the expected values of the age group/ gender/ILOC population change observation cannot breach their respective lower limit (and hence eliminates the possibility of a negative population projections). The results (model coefficients and significance) of the Tobit model are provided in Table 6.4. Note that these results pertain to the pooled version of the model. Having conditioned on all the available information (independent variables) no residual unobserved heterogeneity remains and as such a pooled version of the Tobit model is suitable, and statistically preferable.

Comparing the fitted values of the Tobit model with the linear regression model, the correlation measure (fitted versus actual values) improves significantly from 23% (linear regression model) to 33% (Tobit). It appears that the linear regression model underestimates change for 15% of the observations. Approximately 29% of the observations are censored cases *i.e.*, cases where the value of change is equal to the lower limit ($L_{i,j}$). The linear regression model under-predicts change for 42% of these censored cases. Furthermore, the correlation between actual change and predicted change is approximately 3% for the linear regression model compared to 61% for the Tobit model for these cases. Based on these metrics, the Tobit model clearly provides a significantly better fit.

The projected population numbers can be computed by simply summing the fitted values from the model to the initial (2006) population values. This produces a final estimate for each age group category across all ILOCs. Figure 6.2 shows the actual versus predicted values produced by the Tobit model.

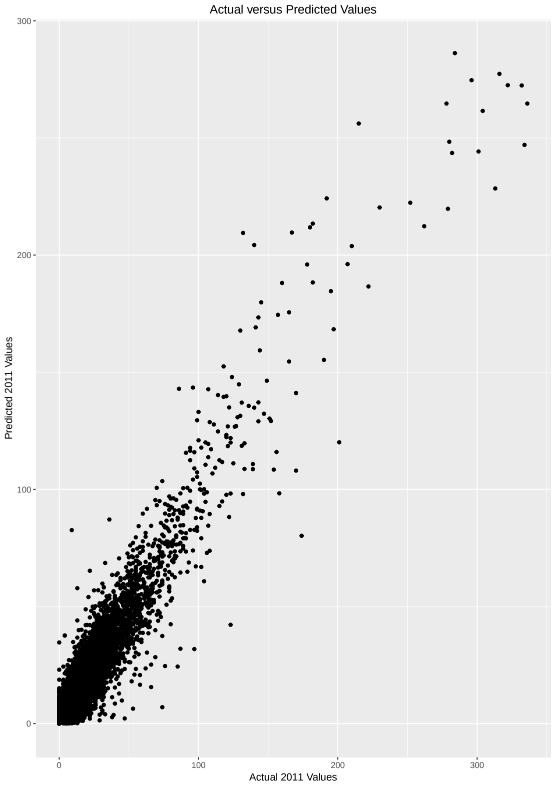


Figure 6.2: Actual 2011 values versus Predicted 2011 Values

Apparent Fertility Rates

To develop a model to predict the Aboriginal and Torres Strait Islander population aged 0–4 in each ILOC, a linear regression model was estimated across ILOCs with the 2011 Aboriginal and Torres Strait Islander population aged 0–4 as the dependent variable. Models were tested with a variety of specifications that included summations of the male and female Aboriginal and Torres Strait Islander populations in the ILOC in 2006, initially focusing on what were considered to be adults of child-bearing age. However, experimenting with which age groups to include and allowing differential effects by age and by region to maximise the model fit returned the relatively simple model set out in equation (6.11) below. A model that included the number of Aboriginal and Torres Strait Islander children aged 0–14 in the ILOC in 2006 interacted with ARIA dummies was found to have the best predictive capability. For the full sample of ILOCs across all of

Table 6.4: Model Coefficients and Significance

Variable	Coefficient	Standard Error
ILOC Size	2.313***	0.45
ARIA		
Outer Regional	—	
Remote	0.062	1.03
Very Remote	3.209***	1.02
State/Territory		
New South Wales	—	
Victoria	-0.786	0.75
Queensland	0.102	0.39
South Australia	-0.602	0.42
Western Australia	0.266	0.36
Tasmania	-0.587	0.42
Northern Territory	-0.052	0.35
Female	0.044	0.17
Growth Town	-0.685	0.91
Cohort age (2011)		
Age 5-9	—	
Age 10-14	4.851***	1.66
Age 15-19	10.965***	2.23
Age 19-24	13.38***	2.70
Age 25-29	7.256***	2.63
Age 30-34	5.931**	2.58
Age 35-39	6.318**	2.46
Age 40-44	6.566**	2.62
Age 45-49	4.513*	2.31
Age 50-54	5.695**	2.39
Age 55-59	4.764*	2.68
Age 60-64	1.303	2.69
Age 65-69	-2.835	2.59
Age 70-74	-1.526	2.95
Age 75-79	-5.648*	3.06
Age 80+	-5.262	3.69
Survival Rate	8.972*	5.26
ILOC size by Age Interaction terms		
ILOC Size*Age 10-14	-1.409***	0.32
ILOC Size*Age 15-19	-2.382***	0.39
ILOC Size*Age 19-24	-2.623***	0.47
ILOC Size*Age 25-29	-1.642***	0.46
ILOC Size*Age 30-34	-1.449***	0.45
ILOC Size*Age 35-39	-1.532***	0.43
ILOC Size*Age 40-44	-1.537***	0.46
ILOC Size*Age 45-49	-1.248***	0.40
ILOC Size*Age 50-54	-1.437***	0.42
ILOC Size*Age 55-59	-1.434***	0.47
ILOC Size*Age 60-64	-0.859*	0.46
ILOC Size*Age 65-69	-0.418	0.43
ILOC Size*Age 70-74	-0.744	0.47
ILOC Size*Age 75-79	-0.296	0.45
ILOC Size*Age 80+	-0.640	0.58

Variable	Coefficient	Standard Error
ARIA by Age interaction terms		
Outer Regional*Age 10-14	6.272***	1.18
Outer Regional*Age 15-19	3.621***	1.03
Outer Regional*Age 19-24	1.532	1.10
Outer Regional*Age 25-29	3.266***	1.01
Outer Regional*Age 30-34	3.007***	0.97
Outer Regional*Age 35-39	3.746***	1.04
Outer Regional*Age 40-44	3.270***	1.03
Outer Regional*Age 45-49	2.901***	0.97
Outer Regional*Age 50-54	3.458***	1.01
Outer Regional*Age 55-59	3.443***	1.14
Outer Regional*Age 60-64	1.420	1.13
Outer Regional*Age 65-69	1.113	1.18
Outer Regional*Age 70-74	1.324	1.24
Outer Regional*Age 75-79	0.878	1.33
Outer Regional*Age 80+	0.128	1.56
Remote*Age 10-14	3.119***	1.08
Remote*Age 15-19	2.572**	1.10
Remote*Age 19-24	1.653	1.14
Remote*Age 25-29	2.324**	0.98
Remote*Age 30-34	2.318**	0.95
Remote*Age 35-39	3.075***	1.02
Remote*Age 40-44	3.508***	0.97
Remote*Age 45-49	2.917***	0.95
Remote*Age 50-54	2.953***	0.97
Remote*Age 55-59	2.913***	1.08
Remote*Age 60-64	1.717	1.08
Remote*Age 65-69	2.405**	1.12
Remote*Age 70-74	2.544**	1.19
Remote*Age 75-79	1.440	1.30
Remote*Age 80+	2.177	1.48
Constant	-24.267***	6.2

Australia, the model returned an R-squared of 0.96.

The number of Aboriginal and Torres Strait Islander males aged 15–34 marginally added to the predictive power, but this component of the model applied only to projections for ILOCs in the major cities (*i.e.*, ARIA level 1). For ILOCs in outer regional, remote and very remote Australia, the predicted number of Aboriginal and Torres Strait Islander people aged 0–4 years is given by:

$$P_{1,j,t} = -2.37 + 0.42 * OREG_j \sum_{i=1}^{3} P_{i,j,t-1} + 0.35 * REM_j \sum_{i=1}^{3} P_{i,j,t-1} + 0.35 * VREM_j \sum_{i=1}^{3} P_{i,j,t-1} \tag{6.11}$$

where P_1, P_2 and P_3 denote the Aboriginal and Torres Strait Islander population aged $0-4$, $5-9$ and $10-14$ respectively. OREG, REM and VREM are dummy variables indicating that the ILOC is in outer regional, remote and very remote Australia. In the estimated model of the 2011 population on 2006 values, each of the estimated coefficients (including the intercept term) was significant at the 1% level.

The coefficients were applied to the 2011 data to obtain projections for the population aged $0-4$ in 2016 for each ILOC in our analysis sample. The projected population was allocated as 50% boys and 50% girls. The cases where the predicted population was negative, the projection was set to zero.

6.6 Projections

By applying the estimated coefficients from the cohort and fertility models to the 2011 data, including the estimated ILOC-specific effect (where appropriate), projections of the 2016 populations for each ILOC by gender and five-year age group were generated. In absolute terms, projected indigenous population changes range from an increase of 888 individuals for Thuringowa in Queensland to a decline of 38 individuals in the remote community of Miali Brumby – Warlpiri, a town camp on the outskirts of Katherine in the Northern Territory. Outer regional Davenport, suburb on the north eastern outskirts of Port Augusta, is projected to experience the most rapid decline in terms of percentage change (-15.7%), relating to a fall in population of around 27. The neighbouring ILOC of Port Augusta – Surrounds, in South Australia, is projected to have the highest growth rate (364%), albeit from a population base of just 14 in 2011. A map of the percentage changes by ILOC is shown in Figure 6.3, where the dots indicate the centroid of the ILOC region. Section A.4.2 of the appendix contains the projections (and percentage change) for all the ILOCs.

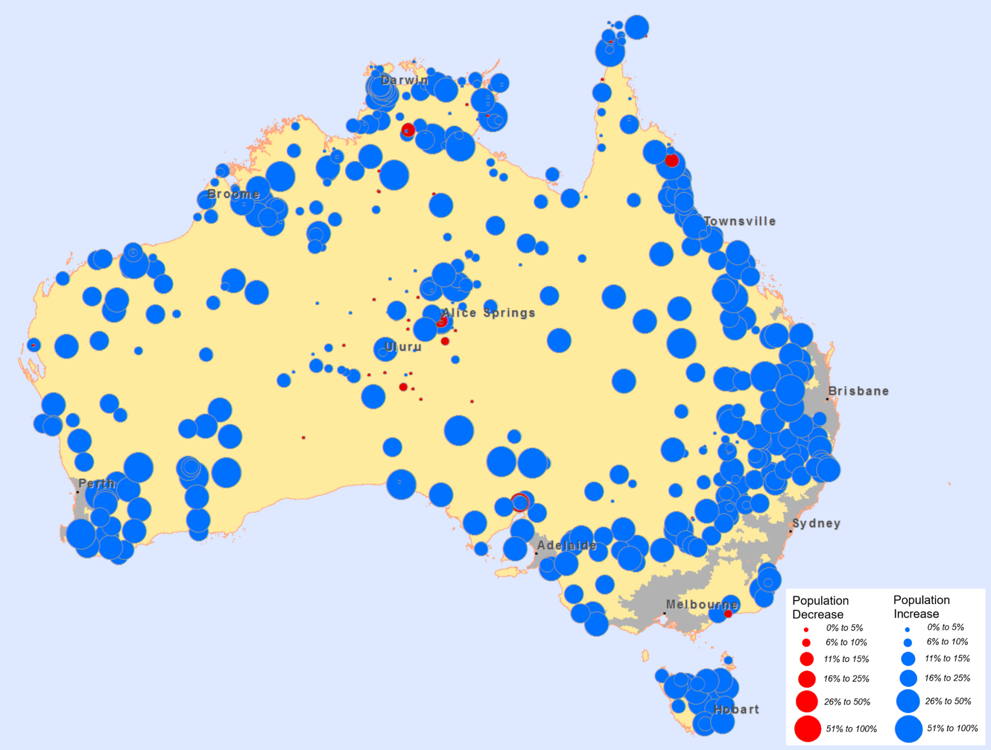


Figure 6.3: Projected changes in Aboriginal and Torres Strait Islander populations by ILOC: 2011–2016

For the many reasons stressed above, estimates for individual communities should be treated with caution. Certainly the use of projections based on modelling results for individual communities in any decision-making – particularly where they indicate unusually high or low growth – should only be undertaken following robust verification of trends through alternative data sources, local knowledge and local consultation. More confidence can be placed in results pertaining to trends at more aggregated spatial and demographic levels, as these will be less subject to idiosyncratic factors relating to a specific community and point in time. Selected aggregated results of potential policy relevance are presented below, commencing with projected changes in the Aboriginal and Torres Strait Islander population by remoteness.

Overall, the Aboriginal and Torres Strait Islander population living in Outer regional, remote and Very remote Australia is projected to increase by 34,400 people between 2011 and 2016, or growth of 14.8%. This represents a larger increase than that recorded by the census between 2006 and 2011 in both absolute terms (27,476 persons) and growth (13.4%). As shown in Table 6.5, populations in each of the three ARIA categories are projected to increase, but with growth strongest in Outer regional Australia (21.4%).

The projections suggest increases in the Aboriginal and Torres Strait Islander popu-

Table 6.5: Aboriginal and Torres Strait Islander populations: 2006 and 2011 census estimates and 2016 projections, by remoteness

		Population			Percent Change	
ARIA	Number of ILOCs	2006	2011	2016	2006–2011	2011–2016
3 – Outer regional	245	97929	116891	139953	19.4	19.7
4 - Remote	95	35537	38930	43073	9.5	10.6
5 – Very remote	278	72067	77188	84384	7.1	9.3
Total	618	205533	233009	267409	13.4	14.8

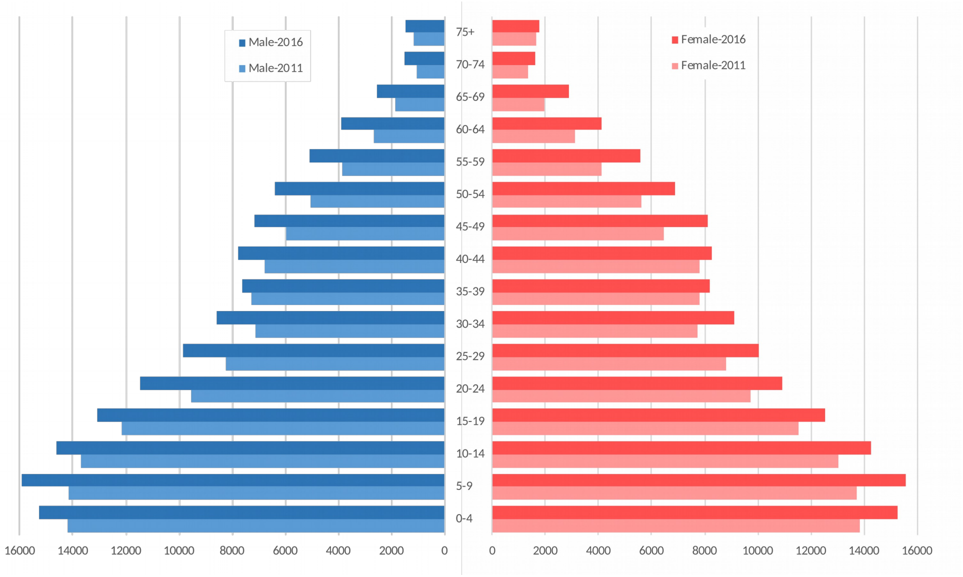


Figure 6.4: Age structure of the Aboriginal and Torres Strait Islander population in outer regional, remote and very remote Australia: 2011 (actual) and 2016 (projected)

Table 6.6: Projected Aboriginal and Torres Strait Islander population growth rates, by age and remoteness (per cent change from 2011 to 2016)

	Age (years)			
ARIA	0–19	20–59	60	Total
3 – Outer regional	15	21.7	40.5	19.7
4 – Remote	3.1	13.9	35.3	10.6
5 – Very remote	4	12.4	22.5	9.3
Total	9.6	17.1	34	14.8

Table 6.7: Projected Aboriginal and Torres Strait Islander population growth rates, by age and remoteness (per cent change from 2011 to 2016)

		Population		Per cent change	
ARIA/ILOC size in 2011	Number of ILOCs	2011	2016	2006-2011	2011-2016
Outer Regional					
Small (0–2,750 people)	75	16306	19131	0.7	17.3
Medium (2,751–6,500 people)	75	26453	32109	18.9	21.4
Large (6,501+ people)	95	74132	88713	24.6	19.7
Remote					
Small (0–700 people)	35	5438	5495	-3.8	1.1
Medium (701–2,000 people)	21	8046	8626	12.2	7.2
Large (2,001+ people)	39	25446	28951	12	13.8
Very Remote					
Small (0–200 people)	99	10319	11703	0.3	13.4
Medium (201–500 people)	94	23975	25196	7.7	5.1
Large (501+ people)	85	42894	47485	8.5	10.7
Total	618	233009	267409	13.4	14.8

lation in all five-year age groups for both males and females. The detailed comparative population pyramids for the 2011 population as recorded by the census and the projected 2016 population are shown in Figure 6.4.

This cohort analysis suggests a rapid ageing of the regional Aboriginal and Torres Strait Islander population, with the number of people aged 60 and over projected to increase by 34% between 2011 and 2016 (Table 6.7). This is on top of a 28% increase in the previous intercensal period. Within each ARIA region, ILOCs were classified as relatively small, medium or large on the basis of their 2011 Census total population estimates (that is, including the non–Aboriginal and Torres Strait Islander population), with the bands chosen such that roughly one-third of ILOCs are in each category. Growth in the Aboriginal and Torres Strait Islander population is projected to be relatively even across communities of different size in Outer regional Australia (Table 6.7). In remote Australia, population growth between 2011 and 2016 is projected to be markedly lower in the smaller communities, almost stagnant in communities that had overall populations of less than 700. Contrary to assumptions underpinning a range of government policies, Indigenous populations are projected to grow substantially in small, very remote communities.

Finally, Table 6.8 shows projected changes in the regional Aboriginal and Torres Strait Islander populations by state and territory. The growth rate of the Aboriginal and Torres Strait Islander population is projected to be lowest in the Northern Territory, at just 9.5% between 2011 and 2016. With this exception, projected growth rates are relatively uniform across jurisdictions, ranging from 15.0% for Western Australia to 21.1% for Victoria. Projected growth rates are also quite uniform in Outer regional areas across the states and territories. The projections for remote and very remote areas are more variable. In terms of the absolute number of Aboriginal and Torres Islander people, the largest increase is projected to occur in Queensland with an additional 12,101 people, well above New South Wales (6,327 additional Aboriginal and Torres Strait Islander people) and Western Australia (up by 3,135 people).

An important policy variable that was included in the analysis was that indicating whether the ILOC included a town that had been nominated as a Northern Territory Growth Town. As reported in Table 6.4, the coefficient on this variable was negative and highly significant, meaning that these towns were in fact associated with lower population increases between 2006 and 2011 than would have been expected given their characteristics. All the Growth Towns are in ILOCs classified as very remote. The projections to 2016 are for the populations of those ILOCs to increase by 6.4%, marginally below the 7.2% overall growth projected for the Aboriginal and Torres Strait Islander population in very remote Northern Territory.

Robustness checks

As noted previously, a range of specifications were experimented with regard to estimation techniques. The preferred specification was the pooled Tobit model with varying lower limits, with robust standard errors, clustered at the ILOC level. There was little evidence of unobserved heterogeneity, such that the pooled version was statistically preferred. Convergence issues were encountered with the fixed effects Tobit specification (and moreover, there are few *a priori* reasons to expect correlations with observed and unobserved heterogeneity terms here, as this is not individual-level data). There was little scope in experimenting with the set of explanatory variables, due to data limitations. However, all variants of linear regression models (pooled and fixed and random effects) yielded essentially similar results to those presented above, although would not be preferred for reasons previously discussed (such as potentially forecasting negative population levels).

Table 6.8: Projected Aboriginal and Torres Strait Islander populations by state/territory and remoteness

	Outer regional	Remote	Very Remote	Total
	(a) Projected 2016 population			
New South Wales	35853	4828	2936	43616
Victoria	6149	n.a.	n.a.	6149
Queensland	54515	12061	21194	87771
South Australia	8718	1344	4289	14352
Western Australia	12287	12238	19084	43609
Tasmania	9146	539	134	9820
Northern Territory	13285	12062	36746	62093
Total	139953	43073	84384	267409
	(b) Projected change from 2011 (people)			
New South Wales	5681	417	230	6327
Victoria	1070	n.a.	n.a.	1070
Queensland	8801	1377	1922	12101
South Australia	1440	228	205	1874
Western Australia	2427	1343	2365	6135
Tasmania	1443	72	9	1525
Northern Territory	2200	705	2464	5369
hline Total	23062	4143	7196	34400
	(c) Per cent change 2011–2016			
New South Wales	18.8	9.5	8.5	17
Victoria	21.1	n.a.	n.a.	21.1
Queensland	19.3	12.9	10	16
South Australia	19.8	20.4	5	15
Western Australia	24.6	12.3	14.1	16.4
Tasmania	18.7	15.5	7.3	18.4
Northern Territory	19.8	6.2	7.2	9.5
Total	19.7	10.6	9.3	14.8

A further interesting feature of the data set are instances where there are no individuals for an observation across both time periods. These *zero-zero* cases are a still classified as censored cases with the censoring point being zero. These cases do deserve special attention since predicting a zero change given that the censoring point is also zero would be challenging. These cases form approximately 10% of the data and 36% of all censored cases. That is, a zero change observation, where there were zero observations in both time-periods could be considered distinctly different from an observationally equivalent outcome where all of the existing population left (and there were no net gains from other sources).

Since the zero-zero values do not affect the absolute size of the population, removing them does not change the population values. Superficially, this would appear to be sensible. Hence, if the model is re-estimated excluding these observations, and the results remain relatively unchanged. However, it should be noted that selecting a sample on the basis of the dependent variable can cause sample selection biases, and is therefore not advocated.

Thus although these particular observations did not appear to be adversely affecting our results, given their particularly unusual status, it may be beneficial to understand the factors that drive them. A probit model can be used for this purpose. Based on the results of this model, it appears that the older age groups and remoteness levels, as well as interactions between these two variables, are the main factors that determine these zero-zero cases. This is to be expected as mortality rates for older age groups start to fall and this seems to be more prevalent in remote/regional areas where the number of individuals in these age groups is relatively small.

6.7 Conclusion

This study has attempted to generate projections of Aboriginal and Torres Strait Islander populations in remote communities. There is an extremely pressing need for reliable estimates of such. Existing projections of the Aboriginal and Torres Strait Islander population disaggregated by detailed age group are available on a regular basis only at the state/territory level. As noted, the literature has stressed the need for such small area estimates for practical purposes of policy and planning for remote Aboriginal communities and the more general need for methodological advances in small number analysis to meet the demographic informational needs of Aboriginal and Torres Strait Islanders in desert Australia ((Taylor et al., 2006)). As noted by (Wilson, 2011): *Probably the logical starting point in the design of any projection system is to consider its purpose ... what question or problems does it need to solve? What projections are required to solve these problems? And of the required outputs, what is the most important and should be prioritised given the resources available?*

The approach taken here has been specifically designed for handling small population numbers while still providing projections by age categories that are critical for decisions relating to service demand, such as education and health. As such the model focuses on changes in population levels by cohort, thereby negating the difficulties associated with models based on growth rates when handling small populations. A second motivation is to generate policy-relevant information on what communities are shrinking and growing, and moreover what are the characteristics that distinguish these. Here the multi-level modelling approach allows estimation of unobservable community-specific effects, as well as statistical estimates of effects such as remoteness, community size and state/territory.

Another critical difference is with respect to the projection horizon. While demographic projections are often prepared with projections 20, 30 or even 50 years into the future, this paper has focused on generating projections five years ahead of a baseline census: namely projecting the 2016 populations based on 2011 data. With the release of the 2016 Census data by ILOC, expected early in latter part of 2017, it will be possible to generate projections to 2021. It is relatively straightforward to use the method to project further ahead; however, projection errors will increase, and a projection horizon of around five years seems more in line with the typical policy cycles impacting upon Aboriginal and Torres Strait Islanders in remote Australia. The approach is very economical in its resource requirements, with almost all input data coming from the ABS census and being publicly accessible from the ABS website. Neither of the two variables used that were not derived from census data were significant. These are the age-specific survival rate (also readily accessible from the ABS) and our one experiment with a policy variable capturing Northern Territory Growth Town status.

The need for caution in the use of estimates must be reiterated, particularly with respect to estimates for individual communities. They should be considered experimental given that the model fit has only been tested ‘within sample’. The real test of such models is their ability to predict out of sample, and it will only be possible to assess the performance of the models out of sample when the 2016 Census data become available. With these reservations in mind, it would be unwise to dwell on the implications of the projections for policy and planning. It is to be noted that at a broad level, the Aboriginal and Torres Strait Islander populations in both remote and very remote Australia are projected to increase substantially. However, there is variation across communities by size. The Aboriginal and Torres Strait Islander population in Very remote Australia is projected to grow in small communities (populations of less than 200 people), and considerably faster in medium-sized communities (201 – 500 people) and larger communities. Hence the projections do not suggest that Aboriginal and Torres Strait Islander people are moving out of smaller communities in very remote Australia. Only in remote Australia do the projections imply some rationalisation of smaller communities. Strong growth in Outer regional Australia of around 18% between 2011 and 2016 is projected irrespective of community size. Given limitations of the census estimates that form our baseline data, it could be said that in effect we are projecting census estimates of the population, rather than actual populations. Be that as it may, this is still important information as those census estimates are used extensively in decision-making.

Only a very limited range of explanatory variables have been tested in the model, and this could be expanded considerably, even with the existing census data. For example, information on the number of people for whom Aboriginal or Torres Strait Islander status is not stated could be incorporated to capture identification effects. The proportion of each ILOC's population that are identified as Aboriginal and Torres Strait Islander could be added as a further community-level characteristic.

Deviations of the age structure from that for the total Aboriginal and Torres Strait Islander population could be added as an ILOC and cohort-specific variable that might also partially capture identification effects.

The analysis here has included only ILOCs in outer regional, remote and very remote Australia due to our focus on regional and remote communities and likely differences in trends influencing demographic change in remote and non-remote Australia. However, ILOCs are defined to spatially cover all of Australia, and those in inner regional and major capital cities can be included in the sample. Using the full sample would mean that the sum of projections would constitute a projection of the total Aboriginal and Torres Strait Islander population for Australia. This would allow for 'top down' adjustment of the estimates to match existing projections of the total Aboriginal and Torres Strait Islander population if that was considered desirable.

The release of the 2016 Census data will mean there will be observations on population changes for each ILOC-by-gender-by-age cohort for two time periods, providing a true multi-level, panel structure for the modelling. The availability of repeat observations should improve capacity of the modelling to identify unobserved community-specific effects; the effects of observable community characteristics, such as remoteness and size; and to uncover various time trends in the population data. It is hoped that the ABS' area definitions under the 'Indigenous location' geography will be more stable between 2011 and 2016, so that fewer baseline observations need to be imputed by a concordance matrix than was the case in the current sample.

Upon release of the 2016 data, there is a plan to re-assess the preferred specification for modelling population changes and then apply the coefficients from that model to forecast populations for 2021. The process can readily be repeated using the 2021 projections as baseline data to generate projections for 2026. The projections will be used in a range of exercises for the CRC-REP's Population Mobility and Labour Markets project, such as mapping labour supply, for forecasting service demand within remote communities and modelling traffic volumes for the road network in remote Australia.

Chapter 7

Conclusion

This thesis has applied the principle of maximum entropy to selected problems in Economics and Finance. In doing so, it has attempted to address some of the challenging issues/problems in both these disciplines as well as demonstrate the applicability of the maximum entropy method.

In chapter 3, a method for forecasting liquidation discount rate was introduced. This process consisted of constructing a time series of liquidation discount rate for two time segments (morning and afternoon) and three portfolio sizes (small, medium and large) using real market data. Analysing the properties of these time series revealed that they possessed a long memory property. This finding implies that current liquidation rates are affected by past liquidation discount rates. More specifically, the influence of previous discount rates decayed slowly over time. Hence, an ARFIMA-GARCH method was used to model the time series across both time segments and the three portfolio sizes. An exhaustive list of different lag combinations and specifications were considered in the estimation process prior to settling on a parsimonious model. The MED was used to construct a liquidation discount-at-risk measure. This measure provided the likelihood of expected future level of discount. This allowed portfolio managers to budget for the future cost of liquidity for a chosen liquidation horizon and confidence level. The results for the 5-day liquidation horizon with a 99% confidence level were provided across both time periods and all portfolio sizes.

Chapter 4 introduced a method for detecting seasonal patterns in financial returns. This was achieved by estimating a MED for different time segments (weekdays and time of the day) extracted from the returns data. The estimated MEDs were compared pairwise within each time segment. Differences in the MEDs (parameters) indicated the presence of seasonal behaviour. This approach of comparing densities (parameters) allows for a richer comparison relative to traditional measures such as mean and/or variance. Based on this approach, one can check for differences in higher moments. This is especially relevant when seasonality only exists in higher moments i.e. differences in lower moments are not significant. This is precisely one of the results found in this chapter. This methodology was applied to detect seasonality in returns of foreign exchange rates. The final results indicate that returns for Wednesday are significantly different from the rest of the weekdays. In other words, some of the parameters of the MED estimated for the Wednesday segment are significantly different from the rest of the weekdays. Similarly, the returns pertaining to 12 p.m. to 2 p.m. interval were significantly different from the rest of the time of day intervals. This difference corresponded to a difference in higher moments. This result strengthens the evidence of the lunch time effect in the greater finance literature.

Chapter 5 formulated the principle of maximum entropy for a multivariate distribution. The proposed formulation is intuitive and can accommodate large number of variables as well as constraints with relative ease compared to the approaches used in the existing literature. Furthermore, the parameters of the resulting multivariate MED were allowed to be functions of the exogenous variables. Hence, these variables can affect the shape, scale and location of the multivariate density. This proposed framework provides ample flexibility to the modeller. Finally, the estimation framework for the resulting multivariate MED was provided. This consisted of proving the consistency and asymptotic normality of the estimator. An empirical example was used to demonstrate the applicability of the proposed framework. This consisted of modelling the distribution of BMI for an individual given their socioeconomic attributes or risk factors. Here, these risk factors were the exogenous variables present the framework. The resulting density was affected by the values of these risk factors. In fact, the results indicated that different factors affect different moments of the resulting BMI distribution. Of particular importance were the risk factors that greatly affect the right hand tail probabilities of the distribution since these indicate the likelihood of obesity. The results illustrate the changes in distribution produced by changes in selected risk factors. This allows policy makers to measure the impact a risk factor has on the likelihood of obesity.

Chapter 6 developed a method of modelling change in an age group for given population over time. It was based on the cohort-component model whereby individuals in a given age group transition to the next age group. The variable of interest, change is defined as the change in number of individuals in a given age group across two time periods. The proposed methodology consisted of three stages. In the first stage, the Kullback-Leibler divergence measure was used to assess if the overall population distribution had changed over time. Given that it had changed, the second stage involved modelling this change. This was done by using a varying limit censoring regression model. This was a notable contribution. To date, there appears to be no application of censoring models to examine population changes. Finally, the resulting model is used to produce population forecasts. This method was applied to model indigenous populations/communities in regional and remote Australia. The results indicate that the overall population distribution had changed over time. The model predicted an increase of approximately 15% in the overall indigenous population from 2011 to 2016 in regional and remote Australia. The population forecast for each location/community are shown in the thesis. In absence of population forecasts for regional and remote locations, these results bring direct benefits to researchers and planning agencies. It is expected that these forecasts will aid to allocating resources/services such as housing, health and infrastructure for each community.

With regard to future direction, there are couple of technical developments which would make the application of maximum entropy even more attractive. One such possibility is to address the issue of moment selection. Given a set of moment constraints, which of these constraints would be most suitable? In other words, since the resulting density is affected by the choice of constraints, how does one go about choosing the *optimal* ones? Another technical development which greatly assist in the applicability of this method would be derive the existence and uniqueness of solution for the multivariate MED case. Currently, most researchers are using the bivariate maximum entropy density for a number of applications even though there are no results on the existence and/or uniqueness of the solution.

Bibliography

ABS (2011, September). Australian Statistical Geography Standards (ASGS). http://www.abs.gov.au/ausstats/abs@.nsf/0/540CFB81D35853FECA2579100014B035?opendocument.

AghaKouchak, A. (2014). Entropy copula in hydrology and climatology. *Journal of Hydrometeorology 15*(6), 2176–2189.

Aitken, M. and C. Comerton-Forde (2003). How should liquidity be measured? *Pacific-Basin Finance Journal 11*(1), 45–59.

Amemiya, T. (1984). Tobit models: A survey. *Journal of Econometrics 24*(1), 3 – 61.

Amemiya, T. (1985). *Advanced econometrics / Takeshi Amemiya*. Blackwell Oxford, UK.

Armstrong, N., G. J. Sutton, and D. B. Hibbert (2019, jan). Estimating probability density functions using a combined maximum entropy moments and bayesian method. theory and numerical examples. *Metrologia 56*(1), 015019.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics 12*, 171–178.

Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones. *Statistica 46*, 199–208.

Azzalini, A. and A. Capitaino (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of Royal Statistical Society, Series B 65*, 367–389.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics 73*(1), 5–59.

Baillie, R. T. and T. Bollerslev (1992). Prediction in dynamic models with time-dependent conditional variances. *Journal of Econometrics 52*(1–2), 91–113.

Baillie, R. T., C.-F. Chung, and M. A. Tieslau (1996). Analysing inflation by the fractionally integrated ARFIMA-GARCH model. *Journal of Applied Econometrics 11*(1), 23–40.

Baillie, R. T., Y. Han, and T. Kwon (2002). Further long memory properties of inflationary shocks. *Southern Economic Journal 68*, 496–510.

Bank for International Settlements (2001). *Final Report of Multidisciplinary Working Group on Enhanced Disclosure*. Bank for International Settlements.

Bann, D., W. Johnson, L. Li, D. Kuh, and R. Hardy (2018). Socioeconomic inequalities in childhood and adolescent body-mass index, weight, and height from 1953 to 2015: an analysis of four longitudinal, observational, british birth cohort studies. *The Lancet Public Health 3*(4), e194 – e203.

Barndorff-Nielsen, O. (2014). *Information and exponential families: in statistical theory*. John Wiley & Sons.

Belov, D. and R. D Armstrong (2011, 05). Distributions of the kullback-leibler divergence with applications. *The British journal of mathematical and statistical psychology 64*, 291–309.

Berkes, I., L. Horváth, and P. Kokoszka (2003). Estimation of the maximal moment exponent of a GARCH(1,1) sequence. *Econometric Theory null*(04), 565–586.

Bertsimas, D. and A. W. Lo (1998). Optimal control of execution costs. *Journal of Financial Markets 1*(1), 1–50.

Beyerlein, A., L. Fahrmeir, U. Mansmann, and A. M. Toschke (2008, Sep). Alternative regression models to assess increase in childhood bmi. *BMC Medical Research Methodology 8*(1), 59.

Biddle, N. and J. Taylor (2009). Indigenous population projections, 2006–31: Planning for growth. Technical report, Centre for Aboriginal Economic Policy Research, Australian National University, Canberra.

Biddle, N. and T. Wilson (2013). Indigenous Australian population projections: Problems and prospects. *Journal of Population Research 30*(2), 101–116.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting 22*, 547–581.

Bottai, M., E. A. Frongillo, X. Sui, J. R. O'Neill, R. E. McKeown, T. L. Burns, A. D. Liese, S. N. Blair, and R. R. Pate (2014). Use of quantile regression to investigate the longitudinal association between physical activity and body mass index. *Obesity 22*(5), E149–E156.

Brown, S., W. Greene, and M. N. Harris (2014, Apr). A New Formulation for Latent Class Models. Working Papers 2014006, The University of Sheffield, Department of Economics.

Brunnermeier, M. and M. Pedersen (2009). Market liquidity and funding liquidity. *Review of Financial Studies 22*, 2201–2238.

Cao, C., . Hansch, and X. Wang (2009). The information content of an open limit-order book. *Journal of Futures Markets 29*, 16–41.

Chan, F. (2009). Modelling time-varying higher moments with maximum entropy density. *Mathematics and Computers in Simulation 79*(9), 2767–2778.

Chan, L. K. C. and J. Lakonishok (1993). Institutional trades and intraday stock price behavior. *Journal of Financial Economics 33*(2), 173–199.

Chan, L. K. C. and J. Lakonishok (1995). The Behavior of Stock Prices Around Institutional Trades. *The Journal of Finance 50*(4), 1147–1174.

Contoyannis, P. and J. Wildman (2007). Using relative distributions to investigate the body mass index in england and canada. *Health Economics 16*(9), 929–944.

Coval, J. and E. Stafford (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics 86*, 479–512.

Dockery, M. (2016). A wellbeing approach to mobility and its application to Aboriginal and Torres Strait Islander Australians. *Social Indicators Research 125(1)*, 243–255.

Doyle, J. R. and C. H. Chen (2009, aug). The wandering weekday effect in major stock markets. *Journal of Banking & Finance 33*(8), 1388–1399.

Engel, R., R. Ferstenburg, and J. Rusell (2012). Measuring and modeling execution cost and risk. *Journal of Portfolio Management 38*, 14–28.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of {United Kingdom} inflation. *Econometrica 50*(4), pp. 987–1007.

Engle, R. F. (1996, November). The econometrics of ultra-high frequency data. Working Paper 5816, National Bureau of Economic Research.

Flegal, F. and R. Troiano (2000, Jul). Changes in the distribution of body mass index of adults and children in the us population. *International journal of obesity and related metabolic disorders 24*(7), 807–18.

French, K. R. (1980). Stock returns and the weekend effect. *Journal of Financial Economics 8*(1), 55–69.

Frontini, M. and A. Tagliani (1997). Entropy-convergence in Stieltjes and Hamburger moment problem. *Applied Mathematics and Computation 88*(1), 39–51.

Gao, L. and D. Han (2019). Methods of moment and maximum entropy for solving nonlinear expectation. *Mathematics 7*(1).

Ghalanos, A. (2012). {rugarch}: Univariate GARCH models.

Gibbons, M. R. and P. Hess (1981). Day of the Week Effects and Asset Returns. *The Journal of Business 54*(4), pp. 579–596.

Golan, A. and E. Maasoumi (2008). Information Theoretic And Entropy Methods: An Overview. *Econometric Reviews 27*(4-6), 317–328.

Golan, A. and J. M. Perloff (2002). Comparison of maximum entropy and higher-order entropy estimators. *Journal of Econometrics 107*(12), 195–211.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics 92*(2), 153–181.

Granger, C. W. J. and R. Joyeux (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis 1*(1), 15–29.

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

Greene, W. H. (2012). *Econometric Modelling Guide*. Econometric Software Inc.

Hameed, A., W. Kang, and S. Viswanathan (2010). Stock market declines and liquidity. *Journal of Finance 55*, 257–293.

Hamid, S. (2018, May). Seasonality in the monthly returns of large and small stocks. *Journal of Accounting and Finance 18*(2).

Handley, W. and M. Millea (2019). Maximum-entropy priors with derived parameters in a specified distribution. *Entropy 21*(3).

Harris, L. (1990). *Liquidity, trading rules and electronic trading systems*. New York University Salomon Center Monograph Series in Finance and Economics, Monograph 1990-4.

Hill, B. M. (1975, 09). A simple general approach to inference about the tail of a distribution. *Ann. Statist. 3*(5), 1163–1174.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika 68*(1), 165–176.

Houle, B. C. (2010). Measuring distributional inequality: Relative body mass index distributions by gender, race/ethnicty, and education, united states 1999–2006. *Journal of Obesity 2010*.

Jaynes, E. T. (1957, March). Information Theory and Statistical Mechanics. *Phys. Rev. 106*(4), 620–630.

Joe, H. (1997). *Multivariate models and dependence concepts*. Chapman & Hall.

Kapur, J. N. (1989). *Maximum-entropy models in science and engineering*. John Wiley & Sons.

Kullback, S. and R. A. Leibler (1951). On Information and Sufficiency. *Annals of Mathematical Statistics 22*(1), pp. 79–86.

Lillo, F. and J. Farmer (2004). The long memory of the efficient market. *Studies in Nonlinear Dynamics and Econometrics 8*(3).

Lin, T. I., J. C. Lee, and W. J. Hsieh (2007, Jun). Robust mixture modeling using the skew t distribution. *Statistics and Computing 17*(2), 81–92.

Ling, S. (2003). Adaptive estimators and tests of stationary and nonstationary short- and long-Memory {ARFIMA-GARCH} models. *Journal of the American Statistical Association 98*(464), 955–967.

Maasoumi, E. (1993). A compendium to information theory in economics and econometrics. *Econometric Reviews 12*(2), 137–181.

Maasoumi, E. and J. Racine (2002). Entropy and predictability of stock market returns. *Journal of Econometrics 107*(1-2), 291–312.

Magnus, J. R. and H. Neudecker (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics* (Second ed.). John Wiley.

Mead, L. R. and N. Papanicolaou (1984). Maximum entropy in the problem of moments. *Journal of Mathematical Physics 25*(8), 2404–2417.

NT-Government (2009). A working future: real towns, real jobs, real opportunities. Technical report, Northern Territory Government.

O'Hara, M. (1997). *Market Microstructure Theory*. Blackwell Publishing.

Ortlepp, J., J. Metrikat, M. Albrecht, P. Maya-Pelzer, H. Pongratz, and R. Hoffmann (2003, 09). Relation of body mass index, physical fitness, and the cardiovascular risk profile in 3127 young normal weight men with an apparently optimal lifestyle. *International Journal of Obesity 27*, 979–82.

Park, S. Y. and A. K. Bera (2009). Maximum entropy autoregressive conditional heteroskedasticity model. *Journal of Econometrics 150*(2), 219–230.

Penman, A. and W. Johnson (2006, Jul). The changing shape of the body mass index distribution curve in the population: Implications for public health policy to reduce the prevalence of adult obesity. *Preventing Chronic Disease 3*(3).

Piantadosi, J., P. Howlett, J. Borwein, J. Henstridge, et al. (2012). Maximum entropy methods for generating simulated rainfall.

Rabe-Hesketh, S. and A. Skrondal (2008). *Multilevel and longitudinal modeling using Stata* (Second Edi ed.). Stata Press, Texas.

Ranasinghe, C., P. Gamage, P. Katulanda, N. Andraweera, S. Thilakarathne, and P. Tharanga (2013). Relationship between body mass index (bmi) and body fat percentage, estimated by bioelectrical impedance, in a group of sri lankan adults: a cross sectional study. *BMC Public Health 13*, 797.

Regional-Services-Reform-Unit (2016). Resilient families, strong communities: A roadmap for regional and remote Aboriginal communities. Technical report, Department of Regional Development, State of Western Australia, Perth.

Rockinger, M. and E. Jondeau (2002). Entropy densities with an application to autoregressive conditional skewness and kurtosis. *Journal of Econometrics 106*(1), 119–142.

Rogalski, R. J. (1984). New Findings Regarding Day-of-the-Week Returns over Trading and Non-Trading Periods: A Note. *The Journal of Finance 39*(5), pp. 1603–1614.

Sanders, W. (2010). Working Future: A critique of policy by numbers. Technical report, Centre for Aboriginal Economic Policy Research, Australian National University, Canberra.

Sarr, A. and T. Lybek (2002). Measuring liquidity in Financial Markets. Technical report, International Monetary Fund.

Scholes, M. (1972). The market for corporate securities: Substitution versus price pressure and the effects of information in stock prices. *Journal of Business 45*, 179–211.

Shannon, C. (1948). The mathematical theory of communication. *Bell Systems Technical Journal 27*, 349–423.

Shleifer, A. and R. Vishny (2011). Fire sales in finance and macroeconomics. *Journal of Economic Perspectives 25*, 29–48.

Smirlock, M. and L. Starks (1986). Day-of-the-week and intraday effects in stock returns. *Journal of Financial Economics 17*(1), 197–210.

Taylor, J. (1997). The contemporary demography of Indigenous Australians. *Journal of Australian Population Association 14(1)*, 77–114.

Taylor, J. (2003). Population futures in the Australian desert, 2001–16. *Australian Geographer 34*, 355–370.

Taylor, J. (2014). Population projections for sparsely populated areas: Reconciling ‘error’ and ‘context’. *International Journal of Population Research 2014*, 9.

Taylor, J. and M. Bell (2002). The Indigenous population of Cape York Peninsula, 2001–2016. Technical report, Centre for Aboriginal Economic Policy Research, Australian National University, Canberra.

Taylor, J., D. Brown, and M. Bell (2006). Population dynamics and Demographic accounting in arid and savannah Australia: Methods, issues and outcomes. Technical report, Desert Knowledge Cooperative Research Centre. Alice Springs. DKCRC Research Report 16.

Tse, Y. (2018). Return seasonality in the foreign exchange market. *Applied Economics Letters 25*(1), 5–8.

Walsh, J., M. Climstein, H. I.T., B. S., K. J., A. K., and D. M. (2011). The loess regression relationship between age and bmi for both sydney world masters games athletes and the australian national population. *International journal of biological and medical sciences 1*(1).

WAToday (2014a, Nov.). Colin Barnett expected flak over Aboriginal community closures. http://www.watoday.com.au/wa-news/colin-barnett-expected-flak-over-aboriginal-community-closures-20141114-11mybe.html.

WAToday (2014b, Nov.). WA's remote communities plan condemned. http://www.watoday.com.au/wa-news/was-remote-communities-plan-condemned-20141113-11ltpt.html.

Wilson, T. (2009). A multistate model for projecting regional populations by Indigenous status: An application to the Northern Territory, Australia. *Environment and Planning 41*, 230–249.

Wilson, T. (2011). Modelling with NEWDSS: Producing state, regional and local area population projections for New South Wales. In J. Stillwell and M. Clarke (Eds.), *Population dynamics and projection methods.*, pp. 61–97. Springer Netherlands.

Wilson, T. and T. Barnes (2007). Continuing challenges in attempting to measure the size, and changing size, of Australia's Indigenous population. *People and Place 15(3)*, 12–21.

Wu, X. (2003). Calculation of maximum entropy densities with application to income distribution. *Journal of Econometrics 115*(2), 347–354.

Yu, K., X. Liu, R. Alhamzawi, F. Becker, and J. Lord (2018). Statistical methods for body mass index: A selective review. *Statistical Methods in Medical Research 27*(3), 798–811.

Zellner, A. and R. A. Highfield (1988, feb). Calculation of maximum entropy distributions and approximation of marginal posterior distributions. *Journal of Econometrics 37*(2), 195–209.

Zhang, Q. and Y. Wang (2004). Socioeconomic inequality of obesity in the united states: do gender, age, and ethnicity matter? *Social Science and Medicine 58*(6), 1171 – 1180.

Appendix A

Appendix

A.1 Appendix - Time Series Properties of Liquidation Discount

Table A.1: Morning and afternoon log liquidation discount time series unit root test statistics and autocorrelation for portfolio size factor γ ∈{0.0001%, 0.0004%, 0.0008%}. For each of the tests, the null hypothesis states that a unit root exists.

			Unit root test statistic (* denotes rejection of null at 1% significance level)			Autocorrelation			
Portfolio detail		Obs	Augmented Dickey-Fuller	KPSS	Phillips-Peron	Lag 1	Lag 5	Lag 10	Lag 20
γ= 0.0001%	morn	1,229	-12.36*	2.459	-974.7*	0.48	0.48	0.44	0.40
	aftn	1,218	-11.73*	2.390	-656.5*	0.60	0.55	0.52	0.51
γ= 0.0004%	morn	1,229	-11.89*	2.632	-849.4*	0.53	0.50	0.48	0.41
	aftn	1,218	-11.09*	3.035	-616.9*	0.62	0.59	0.57	0.54
γ= 0.0008%	morn	1,229	-11.48*	2.607	-712.0*	0.58	0.54	0.52	0.44
	aftn	1,218	-10.31*	3.835	-507.9*	0.66	0.65	0.63	0.59

A.2 Appendix - Detecting Intra-Daily Seasonality in Returns Data

A.2.1 Plots - Foreign Exchange (FX) Returns

A.2.2 Summary Statistics

Weekdays

Table A.2: AUD FX weekday summary statistics

Weekday	Minimum	Q1	Median	Mean	Q3	Maximum
Monday	-0.966500	-0.019700	0.000000	0.000024	0.019640	0.612200
Tuesday	-0.6609000	-0.0204200	0.0000000	0.0002009	0.0209800	1.1370000
Wednesday	-0.818700	-0.020880	0.000000	-0.000001	0.020840	2.550000
Thursday	-0.8401000	-0.0214500	0.0000000	-0.0001516	0.0210000	0.7477000
Friday	-0.6655000	-0.0205800	0.0000000	0.0001821	0.0208400	0.8563000

Table A.3: EUR FX weekday summary statistics

Weekday	Minimum	Q1	Median	Mean	Q3	Maximum
Monday	-2.327e-01	-1.440e-02	0.000e+00	-4.928e-05	1.437e-02	2.215e-01
Tuesday	-0.2157000	-0.0148500	0.0000000	-0.0000206	0.0148200	0.3376000
Wednesday	-0.4500000	-0.0149800	0.0000000	-0.0001411	0.0148300	0.5997000
Thursday	-0.4442000	-0.0152000	0.0000000	-0.0001736	0.0150200	0.5462000
Friday	-0.4832000	-0.0150200	0.0000000	-0.0002144	0.0149000	0.4479000

Table A.4: EURAUD FX weekday summary statistics

Weekday	Minimum	Q1	Median	Mean	Q3	Maximum
Monday	-11.430000	-0.022410	0.000000	-0.000076	0.022440	11.690000
Tuesday	-0.7921000	-0.0229400	0.0000000	-0.0002248	0.0228200	0.8273000
Wednesday	-2.7380000	-0.0231800	0.0000000	-0.0001352	0.0230200	0.8866000
Thursday	-0.8159000	-0.0232800	0.0000000	-0.0000178	0.0233900	0.8547000
Friday	-0.736600	-0.023010	0.000000	-0.000393	0.022810	0.655100

Table A.5: EURGBP FX weekday summary statistics

Weekday	Minimum	Q1	Median	Mean	Q3	Maximum
Monday	-0.3213000	-0.0120300	0.0000000	0.0000003	0.0120300	0.5239000
Tuesday	-0.2642000	-0.0121100	0.0000000	-0.0000862	0.0120900	0.3290000
Wednesday	-0.3251000	-0.0122200	0.0000000	-0.0001346	0.0121500	0.3027000
Thursday	-0.8214000	-0.0125600	0.0000000	-0.0000446	0.0125300	0.8764000
Friday	-0.4571000	-0.0125500	0.0000000	-0.0003128	0.0123500	0.3018000

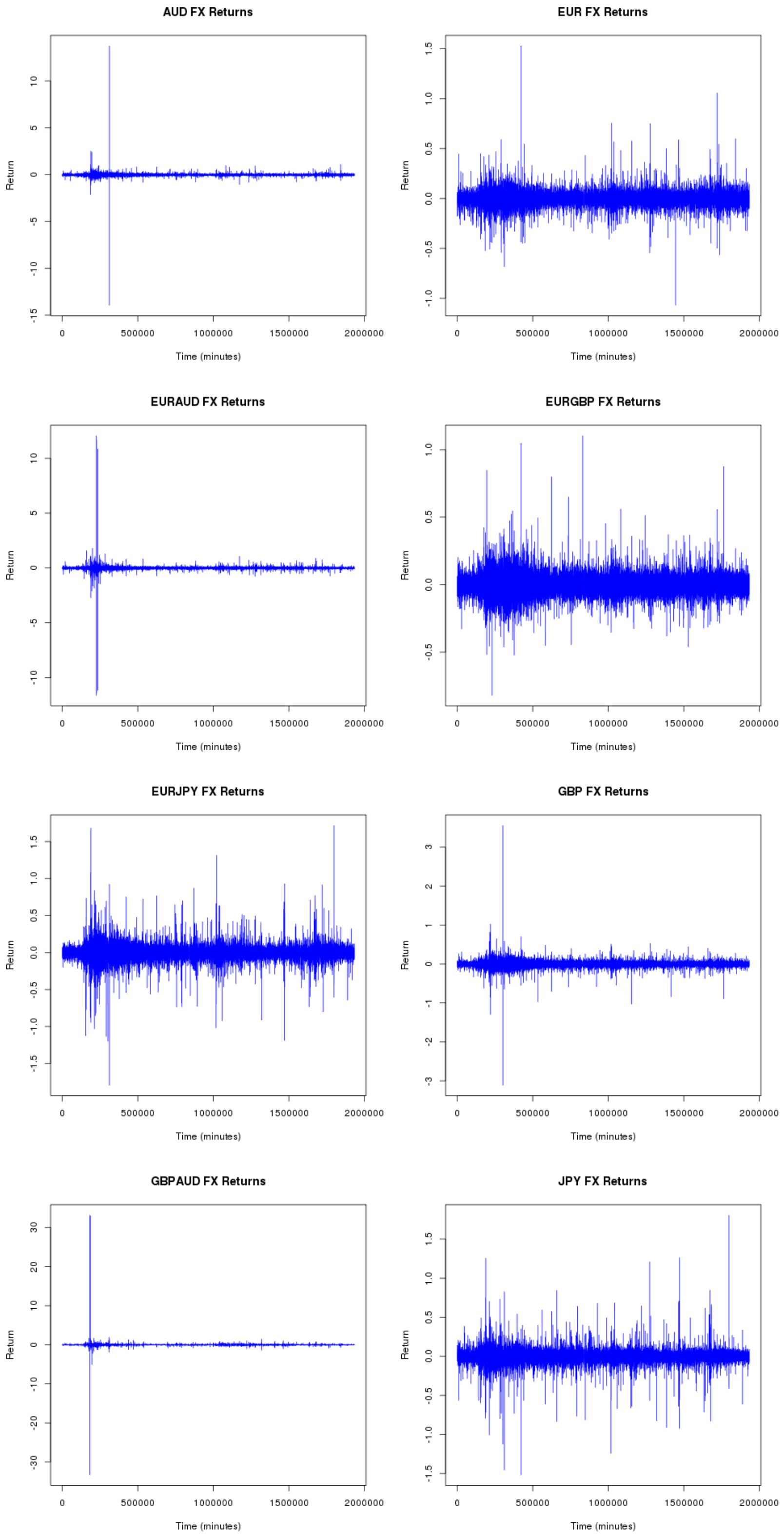


Figure A.1: Plots - FX Returns

Table A.6: EURJPY FX weekday summary statistics

Weekday	Minimum	Q1	Median	Mean	Q3	Maximum
Monday	-0.8802000	-0.0174400	0.0000000	-0.0001467	0.0173800	0.4718000
Tuesday	-0.4036000	-0.0178900	0.0000000	0.0000498	0.0179000	1.0850000
Wednesday	-0.9489000	-0.0179400	0.0000000	-0.0000487	0.0178500	1.6850000
Thursday	-0.6270000	-0.0181500	0.0000000	-0.0000979	0.0182100	0.6491000
Friday	-0.9254000	-0.0180700	0.0000000	-0.0000682	0.0180800	0.8425000

Table A.7: GBP FX weekday summary statistics

Weekday	Minimum	Q1	Median	Mean	Q3	Maximum
Monday	-0.5934000	-0.0126400	0.0000000	-0.0000461	0.0126900	0.2868000
Tuesday	-0.3666000	-0.0131700	0.0000000	0.0000731	0.0130200	0.2941000
Wednesday	-0.5037000	-0.0131900	0.0000000	-0.0000021	0.0130500	0.3549000
Thursday	-0.9772000	-0.0135800	0.0000000	-0.0001181	0.0132500	0.4874000
Friday	-0.3542000	-0.0131700	0.0000000	0.0001044	0.0132100	0.5857000

Table A.8: GBPAUD FX weekday summary statistics

Weekday	Minimum	Q1	Median	Mean	Q3	Maximum
Monday	-0.6635000	-0.0196300	0.0000000	-0.0000422	0.0196100	0.7731000
Tuesday	-0.9633000	-0.0198000	0.0000000	-0.0001005	0.0196700	0.9405000
Wednesday	-1.4070000	-0.0202200	0.0000000	0.0000219	0.0202800	0.9026000
Thursday	-1.3400000	-0.0207000	0.0000000	0.0000558	0.0219900	0.8063000
Friday	-0.8588000	-0.0213300	0.0000000	-0.0000545	0.0203500	0.7890000

Table A.9: JPY FX weekday summary statistics

Weekday	Minimum	Q1	Median	Mean	Q3	Maximum
Monday	-0.7946000	-0.0128300	0.0000000	-0.0000933	0.0128300	0.4062000
Tuesday	-0.3270000	-0.0130100	0.0000000	0.0000686	0.0130300	0.7538000
Wednesday	-0.7171000	-0.0130200	0.0000000	0.0000924	0.0130400	1.2580000
Thursday	-0.4688000	-0.0130200	0.0000000	0.0000651	0.0130600	0.4399000
Friday	-0.8275000	-0.0130200	0.0000000	0.0001411	0.0130300	0.6794000

Timeslots

Table A.10: AUD FX timeslot summary statistics

Timeslot	Minimum	Q1	Median	Mean	Q3	Maximum
2.10-12	-0.8401000	-0.0190800	0.0000000	0.0001493	0.0192300	2.5500000
3.12-14	-0.8187000	-0.0208400	0.0000000	0.0000855	0.0208300	1.1380000
4.14-16	-0.9665000	-0.0220100	0.0000000	-0.0000811	0.0220800	0.6122000

Table A.11: EUR FX timeslot summary statistics

Timeslot	Minimum	Q1	Median	Mean	Q3	Maximum
2.10-12	-0.2776000	-0.0143100	0.0000000	-0.0001569	0.0142000	0.5462000
3.12-14	-0.4832000	-0.0151000	0.0000000	-0.0000902	0.0150100	0.5997000
4.14-16	-0.4500000	-0.0154000	0.0000000	-0.0001112	0.0153100	0.3476000

Table A.12: EURAUD FX timeslot summary statistics

Timeslot	Minimum	Q1	Median	Mean	Q3	Maximum
2.10-12	-2.7380000	-0.0224300	0.0000000	-0.0003006	0.0222000	0.8866000
3.12-14	-1.1200000	-0.0229400	0.0000000	-0.0001608	0.0230600	1.0660000
4.14-16	-11.430000	-0.023700	0.000000	-0.000047	0.023690	11.690000

Table A.13: EURGBP FX timeslot summary statistics

Timeslot	Minimum	Q1	Median	Mean	Q3	Maximum
2.10-12	-0.4571000	-0.0120900	0.0000000	-0.0001978	0.0120300	0.8764000
3.12-14	-0.8214000	-0.0123300	0.0000000	-0.0000789	0.0123100	0.3837000
4.14-16	-0.3213000	-0.0124100	0.0000000	-0.0000695	0.0122300	0.3027000

Table A.14: EURJPY FX timeslot summary statistics

Timeslot	Minimum	Q1	Median	Mean	Q3	Maximum
2.10-12	-0.948900	-0.017250	0.000000	-0.000034	0.017090	1.685000
3.12-14	-0.9254000	-0.0180600	0.0000000	-0.0000629	0.0180600	1.0850000
4.14-16	-0.8802000	-0.0185000	0.0000000	-0.0000896	0.0185700	0.4718000

Table A.15: GBP FX timeslot summary statistics

Timeslot	Minimum	Q1	Median	Mean	Q3	Maximum
2.10-12	-0.9772000	-0.0127200	0.0000000	0.0000558	0.0126700	0.5857000
3.12-14	-0.5583000	-0.0132500	0.0000000	-0.0000099	0.0131900	0.3992000
4.14-16	-0.5037000	-0.0137100	0.0000000	-0.0000387	0.0135400	0.3217000

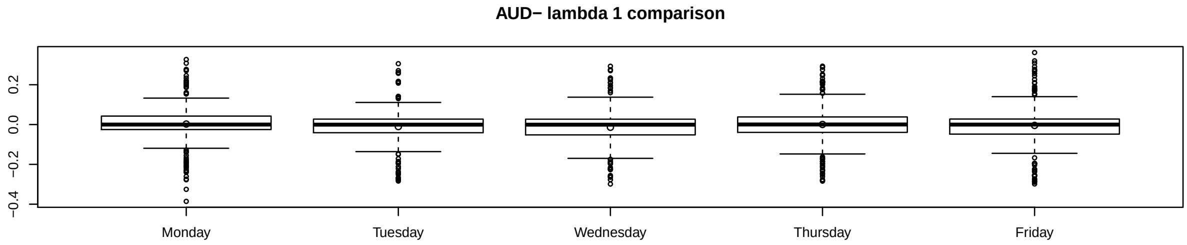
Table A.16: GBPAUD FX timeslot summary statistics

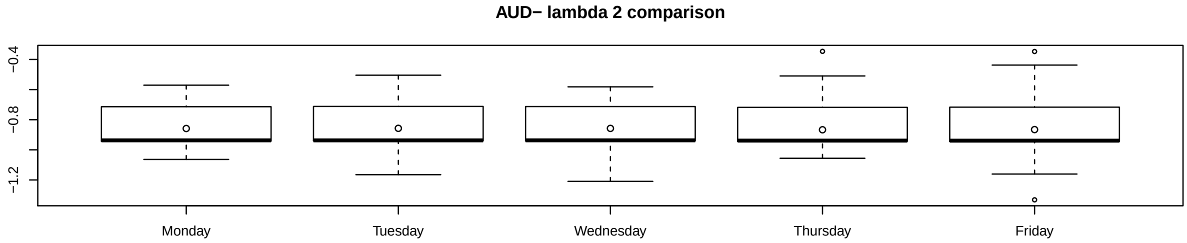
Timeslot	Minimum	Q1	Median	Mean	Q3	Maximum
2.10-12	-1.4070000	-0.0197800	0.0000000	-0.0000386	0.0197800	0.9405000
3.12-14	-1.3400000	-0.0199600	0.0000000	-0.0000821	0.0199900	0.9026000
4.14-16	-0.6635000	-0.0208800	0.0000000	0.0000484	0.0206400	0.7731000

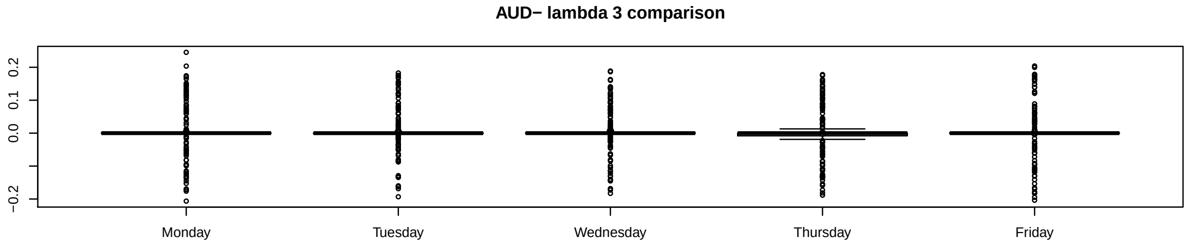
Table A.17: JPY FX timeslot summary statistics

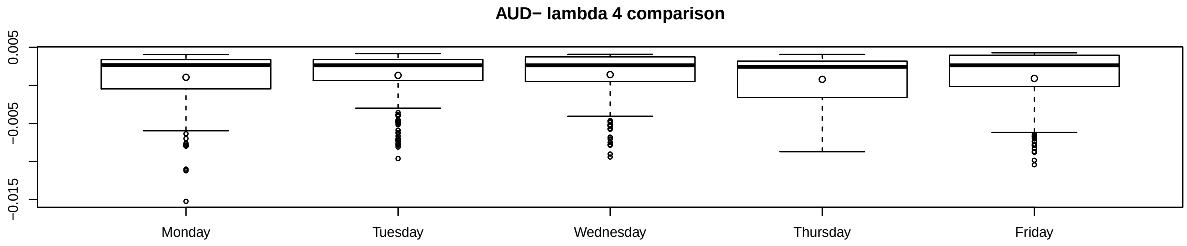
Timeslot	Minimum	Q1	Median	Mean	Q3	Maximum
2.10-12	-0.7171000	-0.0124700	0.0000000	0.0001108	0.0126600	1.2580000
3.12-14	-0.8275000	-0.0130400	0.0000000	0.0000194	0.0130500	0.7538000
4.14-16	-0.7946000	-0.0131100	0.0000000	0.0000334	0.0182300	0.4294000

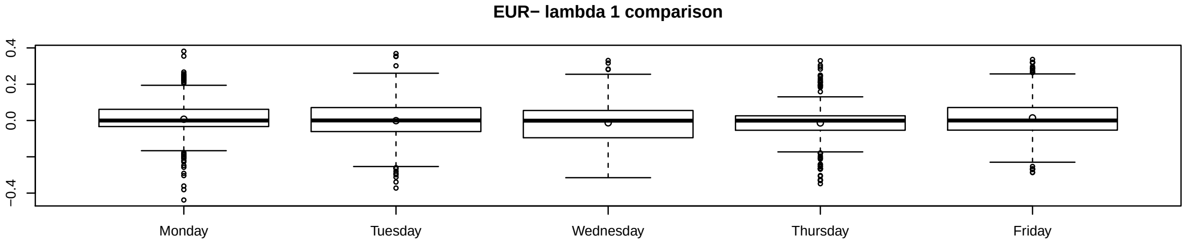
Boxplots

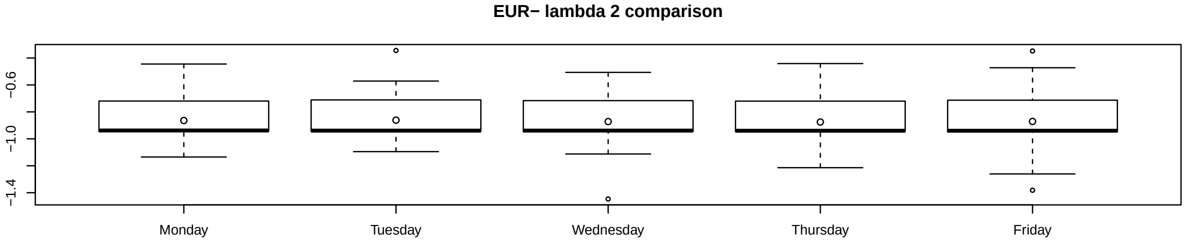


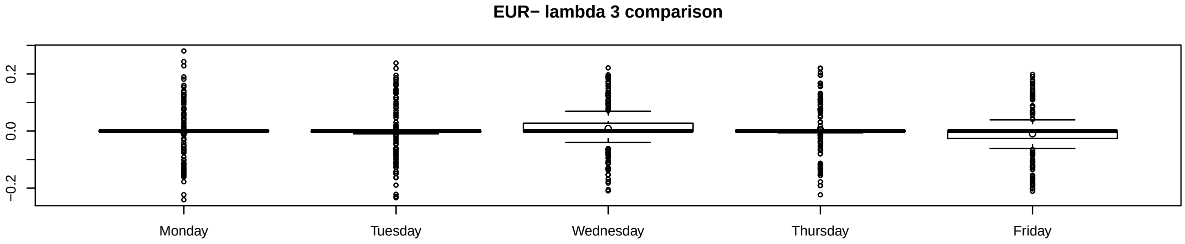


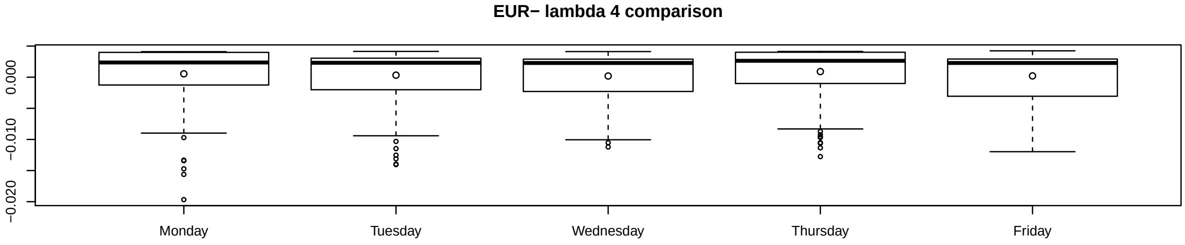


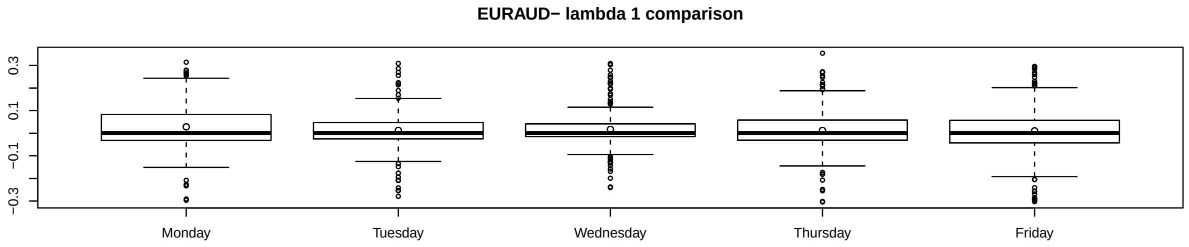


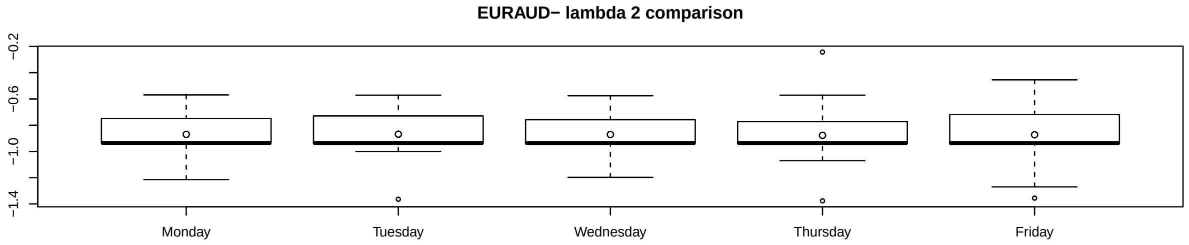

EUR− lambda 1 comparison
0.4
0.2
0.0
−0.4
Monday
Tuesday
Wednesday
Thursday
Friday

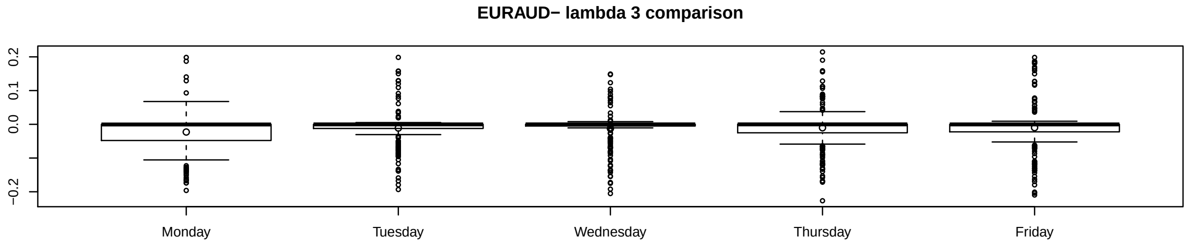

EUR− lambda 2 comparison
−0.6
−1.0
−1.4
Monday
Tuesday
Wednesday
Thursday
Friday

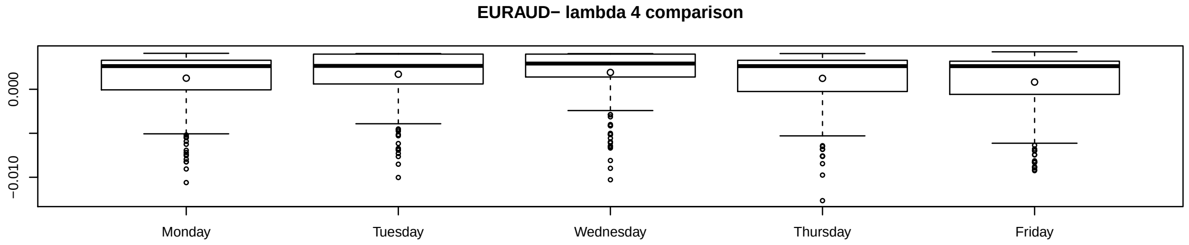

EUR− lambda 3 comparison
0.2
0.0
−0.2
Monday
Tuesday
Wednesday
Thursday
Friday

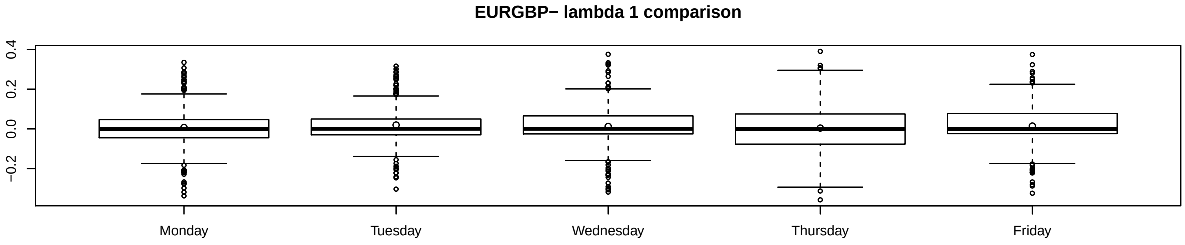

EUR− lambda 4 comparison
0.000
−0.010
−0.020
Monday
Tuesday
Wednesday
Thursday
Friday

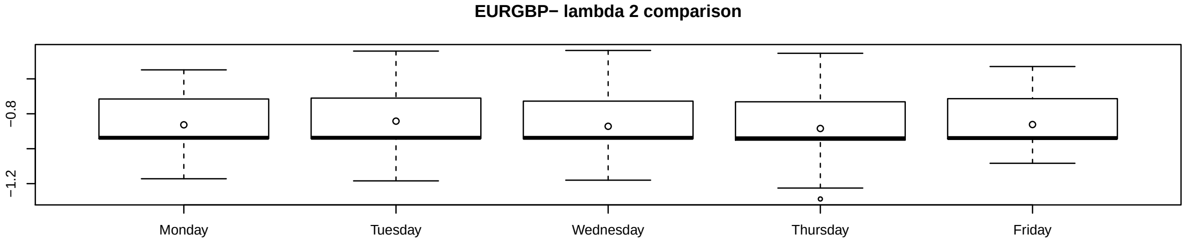

EURAUD− lambda 1 comparison
0.3
0.1
−0.1
−0.3
Monday
Tuesday
Wednesday
Thursday
Friday

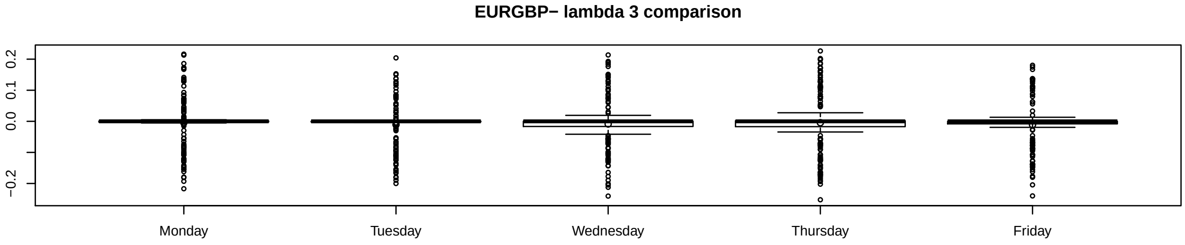

EURAUD− lambda 2 comparison
−0.2
−0.6
−1.0
−1.4
Monday
Tuesday
Wednesday
Thursday
Friday

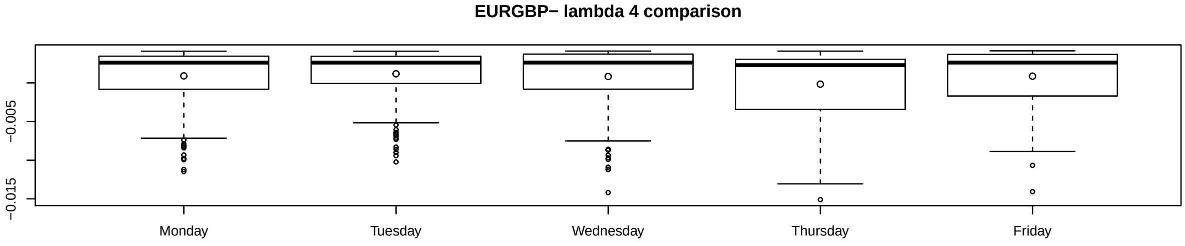

EURAUD− lambda 3 comparison
0.2
0.1
0.0
−0.2
Monday
Tuesday
Wednesday
Thursday
Friday

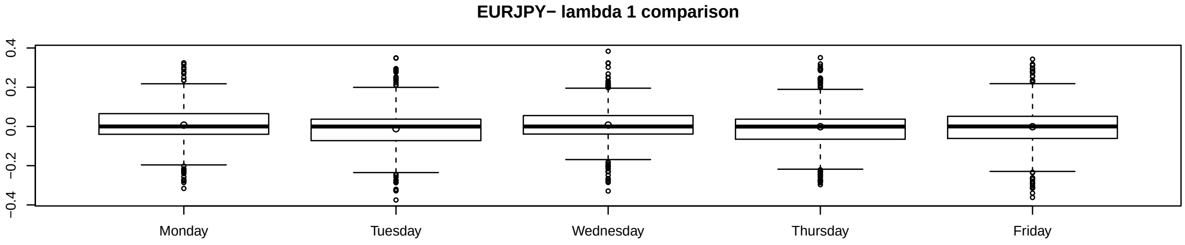

EURAUD− lambda 4 comparison
0.000
−0.010
Monday
Tuesday
Wednesday
Thursday
Friday

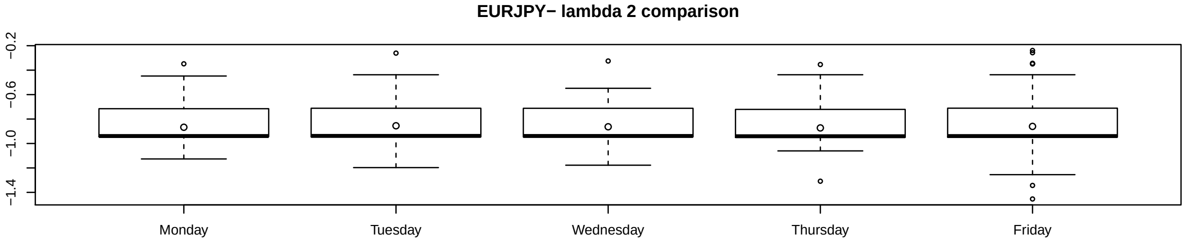

EURGBP− lambda 1 comparison
0.4
0.2
0.0
−0.2
Monday
Tuesday
Wednesday
Thursday
Friday

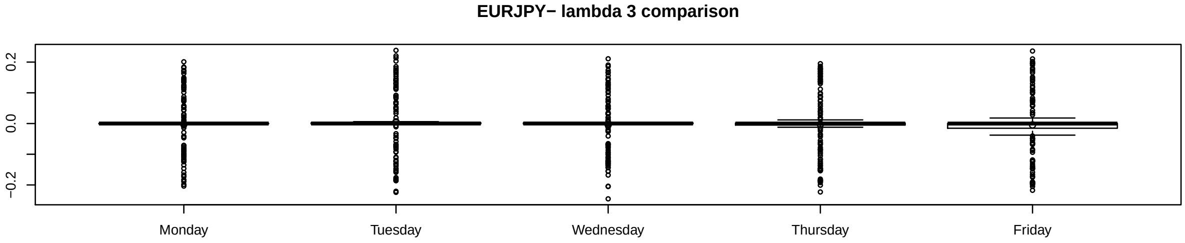

EURGBP− lambda 2 comparison
−0.8
−1.2
Monday
Tuesday
Wednesday
Thursday
Friday

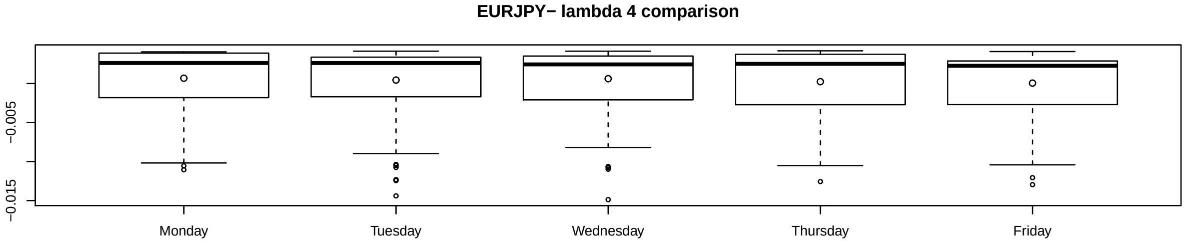

EURGBP− lambda 3 comparison
0.2
0.1
0.0
−0.2
Monday
Tuesday
Wednesday
Thursday
Friday

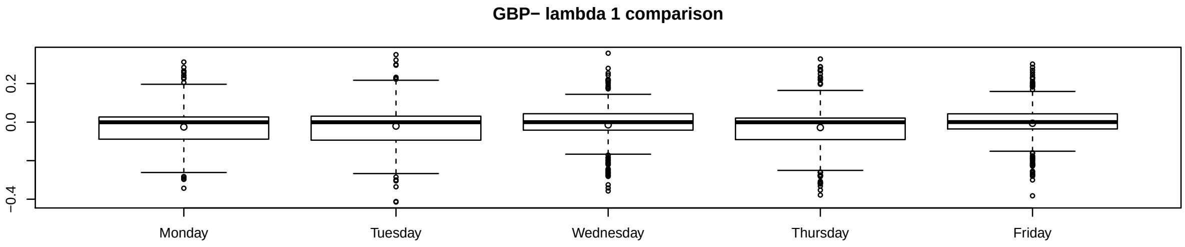

EURGBP− lambda 4 comparison
−0.005
−0.015
Monday
Tuesday
Wednesday
Thursday
Friday

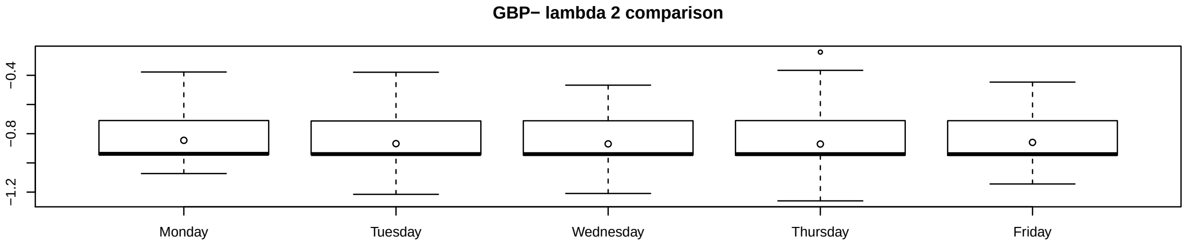

EURJPY− lambda 1 comparison
−0.4
−0.2
0.0
0.2
0.4
Monday
Tuesday
Wednesday
Thursday
Friday

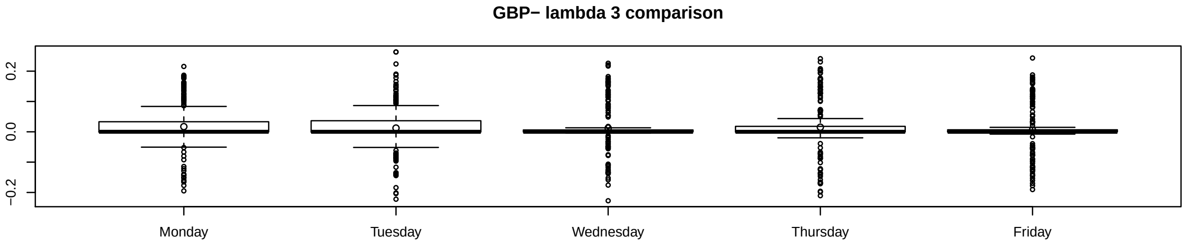

EURJPY− lambda 2 comparison
−1.4
−1.0
−0.6
−0.2
Monday
Tuesday
Wednesday
Thursday
Friday

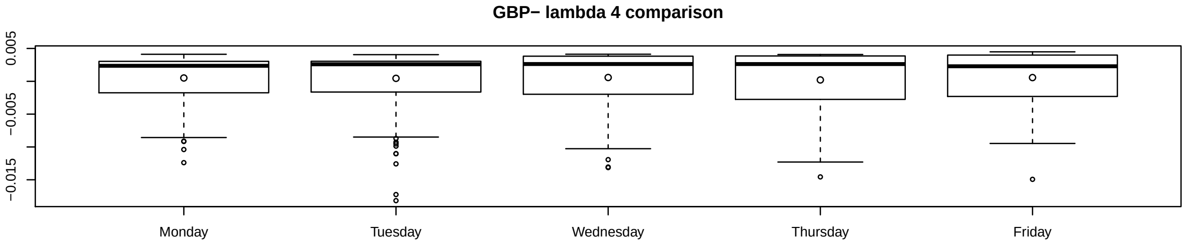

EURJPY− lambda 3 comparison
−0.2
0.0
0.2
Monday
Tuesday
Wednesday
Thursday
Friday

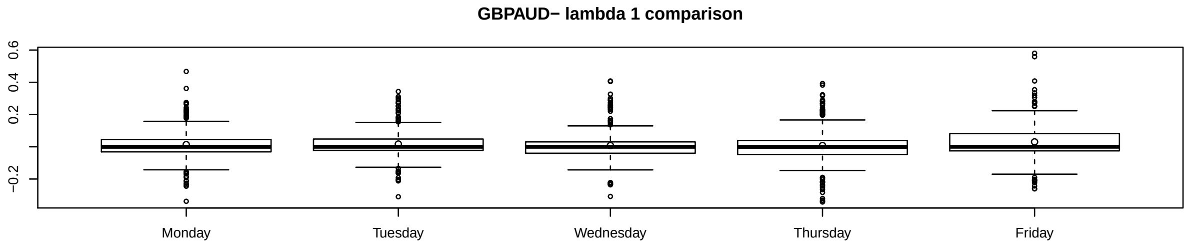

EURJPY− lambda 4 comparison
−0.015
−0.005
Monday
Tuesday
Wednesday
Thursday
Friday

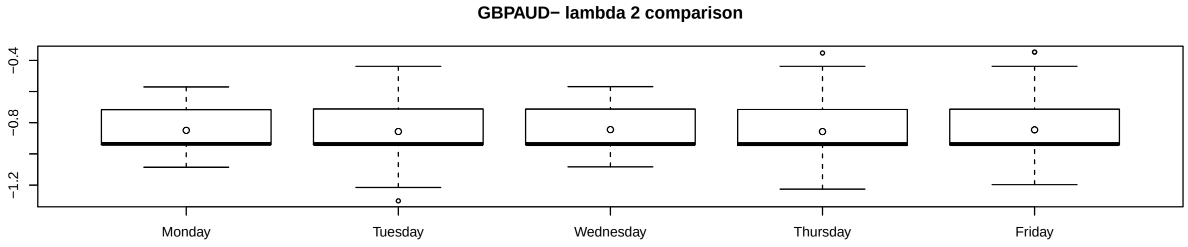

GBP− lambda 1 comparison
0.2
0.0
−0.4
Monday
Tuesday
Wednesday
Thursday
Friday

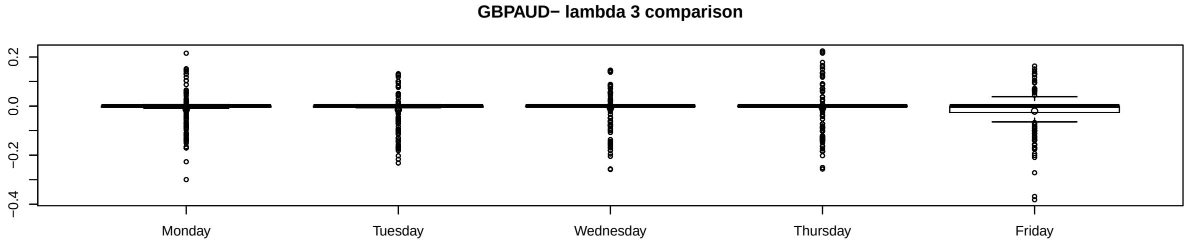

GBP− lambda 2 comparison
−0.4
−0.8
−1.2
Monday
Tuesday
Wednesday
Thursday
Friday

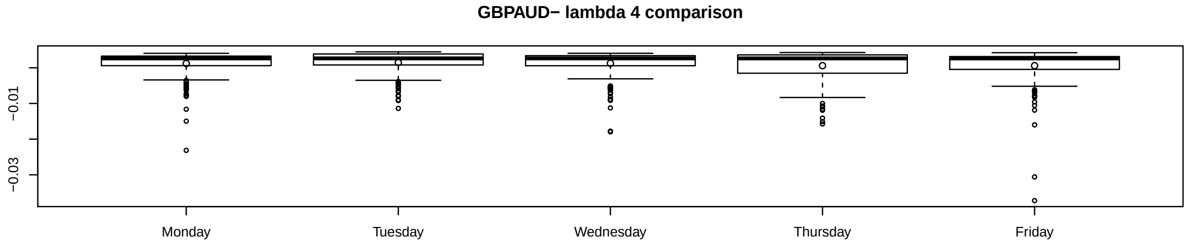

GBP− lambda 3 comparison
0.2
0.0
−0.2
Monday
Tuesday
Wednesday
Thursday
Friday

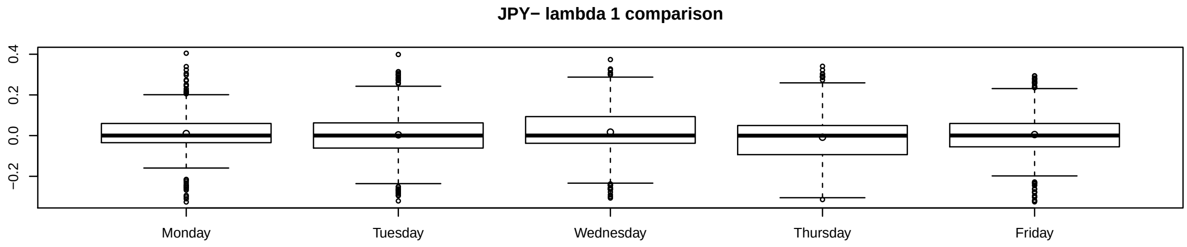

GBP− lambda 4 comparison
0.005
−0.005
−0.015
Monday
Tuesday
Wednesday
Thursday
Friday

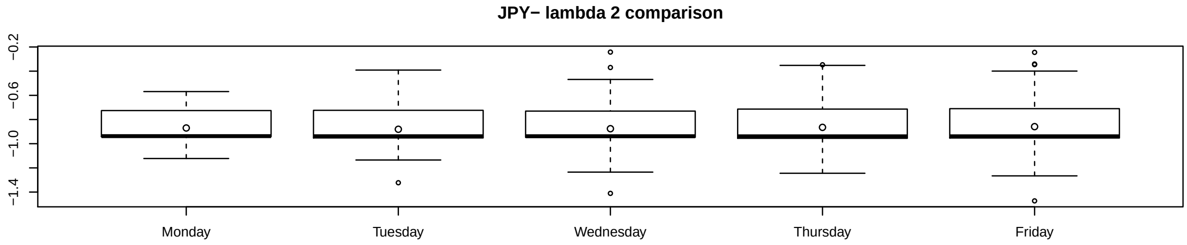

GBPAUD− lambda 1 comparison
0.6
0.4
0.2
−0.2
Monday
Tuesday
Wednesday
Thursday
Friday

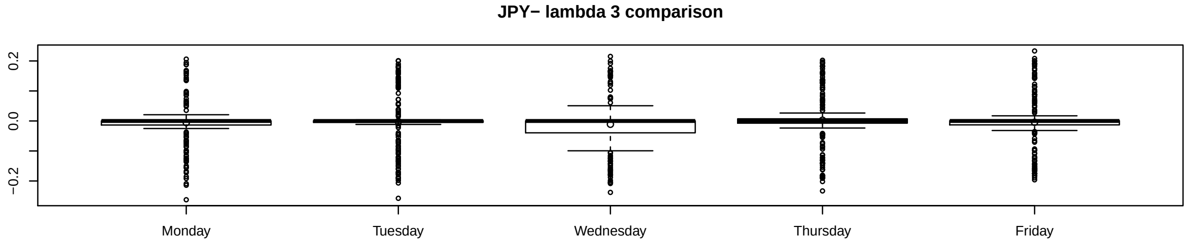

GBPAUD− lambda 2 comparison
−0.4
−0.8
−1.2
Monday
Tuesday
Wednesday
Thursday
Friday

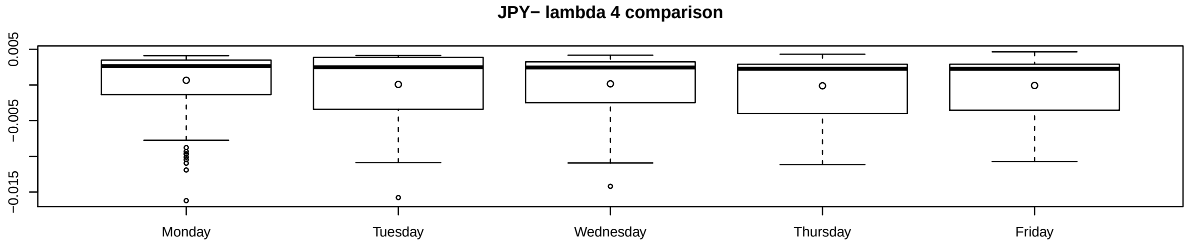

GBPAUD− lambda 3 comparison
0.2
0.0
−0.2
−0.4
Monday
Tuesday
Wednesday
Thursday
Friday


GBPAUD− lambda 4 comparison
−0.01
−0.03
Monday
Tuesday
Wednesday
Thursday
Friday


JPY− lambda 1 comparison
0.4
0.2
0.0
−0.2
Monday
Tuesday
Wednesday
Thursday
Friday


JPY− lambda 2 comparison
−0.2
−0.6
−1.0
−1.4
Monday
Tuesday
Wednesday
Thursday
Friday


JPY− lambda 3 comparison
0.2
0.0
−0.2
Monday
Tuesday
Wednesday
Thursday
Friday


JPY− lambda 4 comparison
0.005
−0.005
−0.015
Monday
Tuesday
Wednesday
Thursday
Friday

Significance of means - MED parameters

Weekdays	Mondays		Tuesday		Wednesday		Thursday		Friday	
AUD	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0027	0.3180	−0.0103	−1.3698	−0.0135	−1.8057	0.0005	0.0599	−0.0043	−0.4770
$\bar{\ell}_2$	−0.8584	−99.5065	−0.8572	−90.7101	−0.8576	−94.5376	−0.8667	−89.0032	−0.8656	−79.1640
$\bar{\ell}_3$	0.0023	0.4677	0.0036	0.8105	0.0061	1.4131	−0.0003	−0.0660	0.0012	0.2201
$\bar{\ell}_4$	0.0011	3.9371	0.0013	5.5545	0.0014	6.1655	0.0008	3.1053	0.0009	3.3209

Weekdays	Mondays		Tuesday		Wednesday		Thursday		Friday	
AUDJPY	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0035	0.4400	−0.0153	−1.6191	0.0144	1.4100	−0.0015	−0.1431	0.0067	0.6019
$\bar{\ell}_2$	−0.8665	−88.2205	−0.8611	−84.1053	−0.8498	−79.4751	−0.8862	−79.0901	−0.8730	−64.7712
$\bar{\ell}_3$	−0.0024	−0.4946	0.0086	1.5186	−0.0055	−0.8974	0.0014	0.2154	−0.0016	−0.2286
$\bar{\ell}_4$	0.0013	4.9474	0.0005	1.5712	0.0004	0.9166	0.0001	0.1718	−0.0003	−0.8286

Weekdays	Mondays		Tuesday		Wednesday		Thursday		Friday	
EUR	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0072	0.7649	−0.0009	−0.0896	−0.0126	−1.3119	−0.0134	−1.4609	0.0145	1.4836
$\bar{\ell}_2$	−0.8642	−86.0274	−0.8613	−85.2540	−0.8718	−86.1999	−0.8754	−88.2742	−0.8709	−77.3681
$\bar{\ell}_3$	−0.0040	−0.7206	0.0011	0.1964	0.0088	1.5217	0.0052	0.9780	−0.0098	−1.6410
$\bar{\ell}_4$	0.0005	1.7418	0.0003	1.0740	0.0002	0.6635	0.0009	3.1542	0.0002	0.6941

Weekdays	Mondays		Tuesday		Wednesday		Thursday		Friday	
EURAUD	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0282	3.7827	0.0126	1.9211	0.0169	2.6417	0.0123	1.6316	0.0106	1.2412
$\bar{\ell}_2$	−0.8698	−110.4415	−0.8684	−103.8229	−0.8709	−115.2259	−0.8770	−99.8758	−0.8727	−91.4518
$\bar{\ell}_3$	−0.0228	−5.3859	−0.0111	−2.8697	−0.0125	−3.3917	−0.0102	−2.2637	−0.0098	−1.8697
$\bar{\ell}_4$	0.0013	5.2608	0.0017	8.0457	0.0019	9.0784	0.0012	5.2585	0.0008	3.0344

Weekdays	Mondays		Tuesday		Wednesday		Thursday		Friday	
EURGBP	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0075	0.8549	0.0196	2.4859	0.0125	1.4459	0.0043	0.4054	0.0144	1.6464
$\bar{\ell}_2$	−0.8631	-93.6781	−0.8421	-82.2203	−0.8717	-93.6157	−0.8842	-87.5416	−0.8616	-92.3317
$\bar{\ell}_3$	−0.0044	−0.8783	−0.0108	−2.3486	−0.0087	−1.6665	−0.0043	−0.6691	−0.0110	−2.1757
$\bar{\ell}_4$	0.0009	3.3986	0.0012	4.8276	0.0008	2.9465	−0.0002	−0.4953	0.0009	3.2383

Weekdays	Mondays		Tuesday		Wednesday		Thursday		Friday	
EURJPY	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0070	0.7914	−0.0113	−1.1648	0.0072	0.7883	−0.0013	−0.1303	−0.0014	−0.1368
$\bar{\ell}_2$	−0.8672	-91.1346	−0.8549	-77.1994	−0.8632	-82.7670	−0.8725	-79.4449	−0.8602	-63.0802
$\bar{\ell}_3$	−0.0028	−0.5287	0.0052	0.9052	−0.0040	−0.7442	−0.0026	−0.4167	−0.0044	−0.6970
$\bar{\ell}_4$	0.0007	2.4987	0.0004	1.4664	0.0006	2.2125	0.0002	0.7692	0.0001	0.1821

Weekdays	Mondays		Tuesday		Wednesday		Thursday		Friday	
GBP	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	−0.0252	−2.7815	−0.0206	−2.2601	−0.0143	−1.5467	−0.0281	−2.7561	−0.0060	−0.6579
$\bar{\ell}_2$	−0.8453	−78.3991	−0.8680	−89.6813	−0.8695	−92.0307	−0.8707	−78.6564	−0.8595	−78.5562
$\bar{\ell}_3$	0.0171	3.1995	0.0124	2.2404	0.0123	2.2284	0.0152	2.4635	0.0059	1.0725
$\bar{\ell}_4$	0.0005	1.7874	0.0005	1.5019	0.0006	1.9600	0.0002	0.6271	0.0006	2.0238

Weekdays	Mondays		Tuesday		Wednesday		Thursday		Friday	
GBPAUD	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0134	1.6812	0.0172	2.2820	0.0077	0.9369	0.0076	0.7890	0.0305	3.1892
$\bar{\ell}_2$	−0.8484	−96.8598	−0.8561	−87.0078	−0.8440	−92.3869	−0.8565	−78.2681	−0.8453	−76.8665
$\bar{\ell}_3$	−0.0143	−3.0990	−0.0152	−3.5476	−0.0088	−1.8996	−0.0084	−1.4756	−0.0203	−3.6863
$\bar{\ell}_4$	0.0012	4.3798	0.0015	6.2543	0.0013	4.6027	0.0006	1.8938	0.0006	1.5908

Weekdays	Mondays		Tuesday		Wednesday		Thursday		Friday	
JPY	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0103	1.1194	0.0038	0.3660	0.0174	1.7526	−0.0079	−0.7471	0.0053	0.5072
$\bar{\ell}_2$	−0.8699	−93.3671	−0.8805	−84.5935	−0.8759	−78.5359	−0.8648	−74.1277	−0.8588	−62.9030
$\bar{\ell}_3$	−0.0057	−1.0438	−0.0023	−0.3602	−0.0110	−1.8161	0.0025	0.3894	−0.0044	−0.6796
$\bar{\ell}_4$	0.0007	2.2431	0.0001	0.2345	0.0002	0.5216	−0.0001	−0.3625	−0.0001	−0.2322

Difference in means - MED Parameters

Comparisons AUD	Mon vs Tue t statistic	Mon vs Wed t statistic	Mon vs Thu t statistic	Mon vs Fri t statistic	Tue vs Wed t statistic	Tue vs Thu t statistic	Tue vs Fri t statistic	Wed vs Thu t statistic	Wed vs Fri t statistic	Thu vs Fri t statistic
$\Delta\ \bar{\ell}_1$	1.1549	1.4432	0.1816	0.5657	0.2998	−0.9603	−0.5181	−1.2468	−0.7928	0.3889
$\Delta\ \bar{\ell}_2$	−0.0901	−0.0613	0.6421	0.5157	0.0295	0.7006	0.5768	0.6853	0.5596	−0.0800
$\Delta\ \bar{\ell}_3$	−0.1963	−0.5759	0.3728	0.1595	−0.3978	0.5833	0.3562	0.9618	0.7242	−0.2040
$\Delta\ \bar{\ell}_4$	−0.7134	−0.9932	0.7277	0.3789	−0.2889	1.5112	1.1062	1.8131	1.3881	−0.3300

Comparisons AUDJPY	Mon vs Tue t statistic	Mon vs Wed t statistic	Mon vs Thu t statistic	Mon vs Fri t statistic	Tue vs Wed t statistic	Tue vs Thu t statistic	Tue vs Fri t statistic	Wed vs Thu t statistic	Wed vs Fri t statistic	Thu vs Fri t statistic
$\Delta\ \bar{\ell}_1$	1.5183	−0.8361	0.3823	−0.2318	−2.1345	−0.9826	−1.5049	1.0902	0.5071	−0.5381
$\Delta\ \bar{\ell}_2$	−0.3791	−1.1525	1.3200	0.3913	−0.7670	1.6502	0.7033	2.3504	1.3518	−0.7499
$\Delta\ \bar{\ell}_3$	−1.4784	0.3994	−0.4675	−0.0925	1.6911	0.8374	1.1360	−0.7719	−0.4195	0.3141
$\Delta\ \bar{\ell}_4$	2.1125	1.8610	2.9249	3.5871	0.1605	0.9130	1.6377	0.6084	1.2353	0.7242

Comparisons EUR	Mon vs Tue t statistic	Mon vs Wed t statistic	Mon vs Thu t statistic	Mon vs Fri t statistic	Tue vs Wed t statistic	Tue vs Thu t statistic	Tue vs Fri t statistic	Wed vs Thu t statistic	Wed vs Fri t statistic	Thu vs Fri t statistic
$\Delta\ \bar{\ell}_1$	0.5966	1.4730	1.5682	−0.5383	0.8605	0.9396	−1.1156	0.0597	−1.9775	−2.0816
$\Delta\ \bar{\ell}_2$	−0.2048	0.5338	0.7934	0.4435	0.7365	0.9972	0.6353	0.2534	−0.0607	−0.3005
$\Delta\ \bar{\ell}_3$	−0.6417	−1.5950	−1.1957	0.7033	−0.9362	−0.5149	1.3133	0.4621	2.2374	1.8763
$\Delta\ \bar{\ell}_4$	0.4917	0.8346	−0.8527	0.7737	0.3307	−1.3740	0.2804	−1.7637	−0.0444	1.6771

Comparisons EURAUD	Mon vs Tue t statistic	Mon vs Wed t statistic	Mon vs Thu t statistic	Mon vs Fri t statistic	Tue vs Wed t statistic	Tue vs Thu t statistic	Tue vs Fri t statistic	Wed vs Thu t statistic	Wed vs Fri t statistic	Thu vs Fri t statistic
$\Delta\,\bar{\ell}_1$	1.5757	1.1482	1.5075	1.5474	−0.4747	0.0321	0.1800	0.4728	0.5878	0.1422
$\Delta\,\bar{\ell}_2$	−0.1262	0.0995	0.6083	0.2334	0.2250	0.7113	0.3419	0.5256	0.1481	−0.3306
$\Delta\,\bar{\ell}_3$	−2.0450	−1.8326	−2.0453	−1.9376	0.2674	−0.1534	−0.2020	−0.4020	−0.4285	−0.0585
$\Delta\,\bar{\ell}_4$	−1.3874	−2.0357	0.0570	1.2343	−0.6879	1.4593	2.5948	2.1133	3.2032	1.1888

Comparisons EURGBP	Mon vs Tue t statistic	Mon vs Wed t statistic	Mon vs Thu t statistic	Mon vs Fri t statistic	Tue vs Wed t statistic	Tue vs Thu t statistic	Tue vs Fri t statistic	Wed vs Thu t statistic	Wed vs Fri t statistic	Thu vs Fri t statistic
$\Delta\,\bar{\ell}_1$	−1.0278	−0.4083	0.2350	−0.5622	0.6051	1.1653	0.4354	0.6049	−0.1578	−0.7432
$\Delta\,\bar{\ell}_2$	−1.5257	0.6524	1.5429	−0.1175	2.1359	2.9275	1.4058	0.9134	−0.7652	−1.6460
$\Delta\,\bar{\ell}_3$	0.9272	0.5845	−0.0132	0.9196	−0.3003	−0.8114	0.0371	−0.5237	0.3217	0.8138
$\Delta\,\bar{\ell}_4$	−0.7563	0.2055	2.5485	0.0922	0.9490	3.3121	0.8487	2.3085	−0.1148	−2.4559

Comparisons EURJPY	Mon vs Tue t statistic	Mon vs Wed t statistic	Mon vs Thu t statistic	Mon vs Fri t statistic	Tue vs Wed t statistic	Tue vs Thu t statistic	Tue vs Fri t statistic	Wed vs Thu t statistic	Wed vs Fri t statistic	Thu vs Fri t statistic
$\Delta\,\bar{\ell}_1$	1.3935	−0.0170	0.6204	0.6179	−1.3877	−0.7126	−0.6943	0.6269	0.6245	0.0073
$\Delta\,\bar{\ell}_2$	−0.8401	−0.2832	0.3665	−0.4194	0.5435	1.1279	0.3012	0.6156	−0.1733	−0.7025
$\Delta\,\bar{\ell}_3$	−1.0242	0.1623	−0.0317	0.1943	1.1693	0.9243	1.1250	−0.1821	0.0446	0.2114
$\Delta\,\bar{\ell}_4$	0.5740	0.1609	1.0562	1.4961	−0.4156	0.4680	0.8875	0.8955	1.3310	0.4139

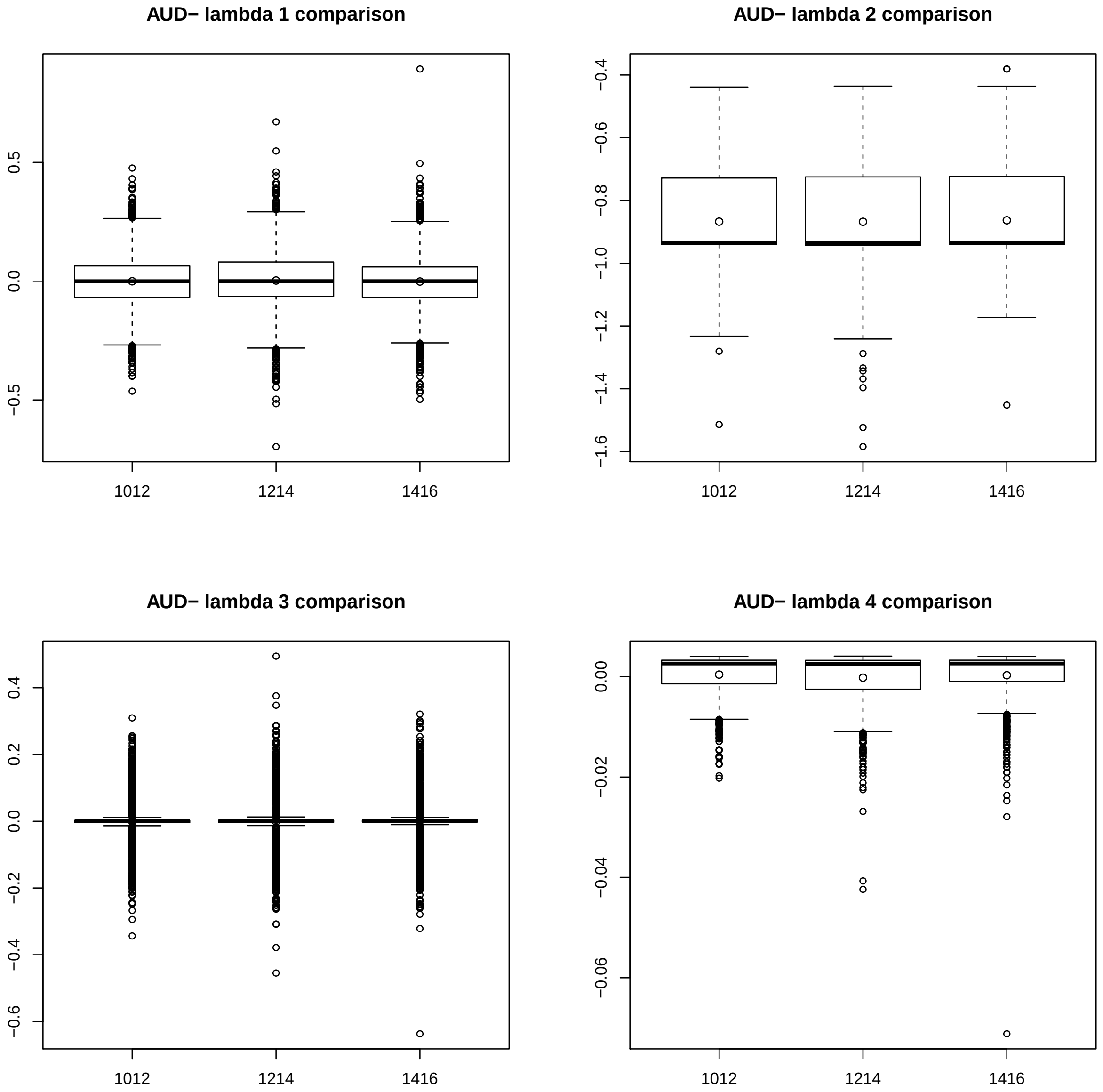
Comparisons GBP	Mon vs Tue t statistic	Mon vs Wed t statistic	Mon vs Thu t statistic	Mon vs Fri t statistic	Tue vs Wed t statistic	Tue vs Thu t statistic	Tue vs Fri t statistic	Wed vs Thu t statistic	Wed vs Fri t statistic	Thu vs Fri t statistic
$\Delta\ \bar{\ell}_1$	−0.3609	−0.8461	0.2107	−1.4986	−0.4865	0.5496	−1.1349	1.0051	−0.6403	−1.6192
$\Delta\ \bar{\ell}_2$	1.5631	1.6852	1.6418	0.9204	0.1117	0.1852	−0.5826	0.0833	−0.6932	−0.7218
$\Delta\ \bar{\ell}_3$	0.6107	0.6243	0.2318	1.4731	0.0126	−0.3384	0.8430	−0.3507	0.8314	1.1359
$\Delta\ \bar{\ell}_4$	0.1353	−0.1713	0.6859	−0.1496	−0.2994	0.5440	−0.2813	0.8330	0.0257	−0.8271

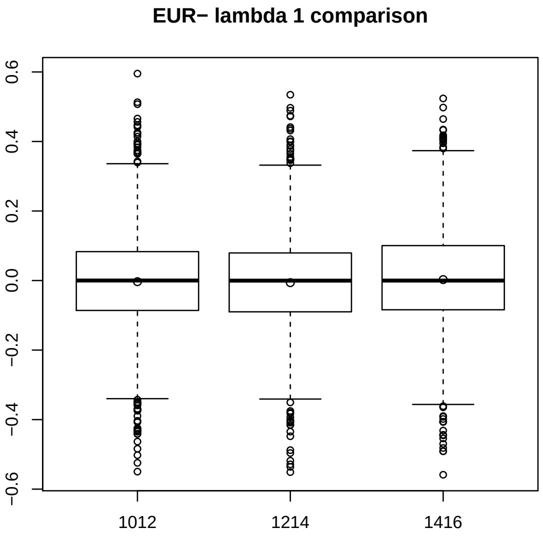
Comparisons GBPAUD	Mon vs Tue t statistic	Mon vs Wed t statistic	Mon vs Thu t statistic	Mon vs Fri t statistic	Tue vs Wed t statistic	Tue vs Thu t statistic	Tue vs Fri t statistic	Wed vs Thu t statistic	Wed vs Fri t statistic	Thu vs Fri t statistic
$\Delta\ \bar{\ell}_1$	−0.3377	0.5067	0.4624	−1.3660	0.8557	0.7762	−1.0951	0.0011	−1.8148	−1.6782
$\Delta\ \bar{\ell}_2$	0.5872	−0.3484	0.5766	−0.2211	−0.9045	0.0235	−0.7349	0.8762	0.0910	−0.7213
$\Delta\ \bar{\ell}_3$	0.1402	−0.8319	−0.8039	0.8404	−1.0018	−0.9510	0.7394	−0.0598	1.5924	1.5052
$\Delta\ \bar{\ell}_4$	−0.7632	−0.1328	1.4569	1.3360	0.6242	2.2494	2.0325	1.5854	1.4516	0.0148

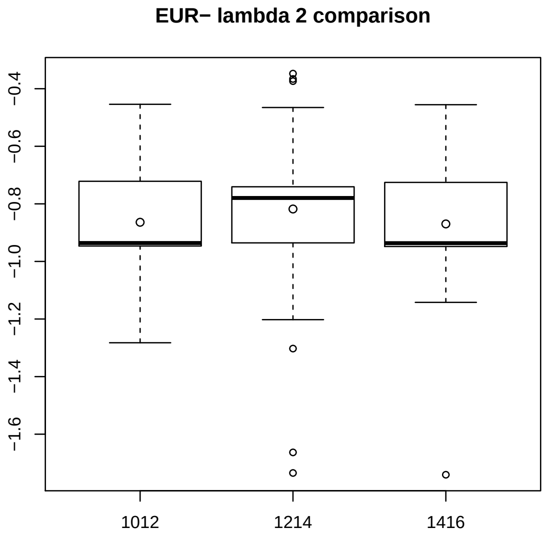
Comparisons JPY	Mon vs Tue t statistic	Mon vs Wed t statistic	Mon vs Thu t statistic	Mon vs Fri t statistic	Tue vs Wed t statistic	Tue vs Thu t statistic	Tue vs Fri t statistic	Wed vs Thu t statistic	Wed vs Fri t statistic	Thu vs Fri t statistic
$\Delta\ \bar{\ell}_1$	0.4614	−0.5275	1.2976	0.3564	−0.9389	0.7883	−0.0991	1.7438	0.8381	−0.8879
$\Delta\ \bar{\ell}_2$	0.7587	0.4157	−0.3415	−0.6692	−0.2987	−1.0040	−1.2617	−0.6902	−0.9702	−0.3321
$\Delta\ \bar{\ell}_3$	−0.4141	0.6486	−0.9747	−0.1603	1.0019	−0.5302	0.2336	−1.5323	−0.7532	0.7563
$\Delta\ \bar{\ell}_4$	1.3355	1.1641	1.7945	1.7122	−0.1919	0.4210	0.3300	0.6242	0.5331	−0.0959

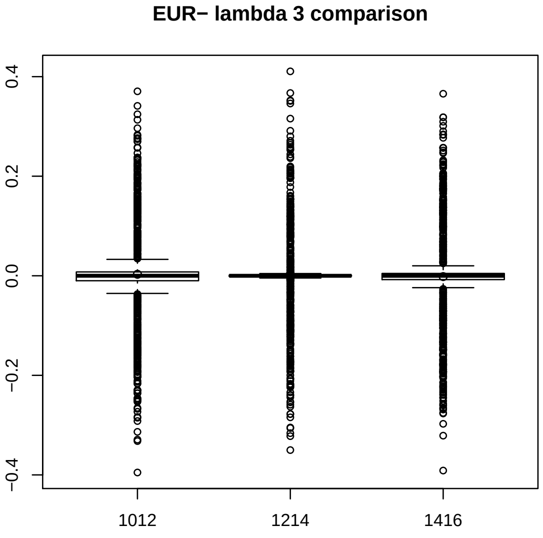
A.2.3 Results - Time of Day Effect

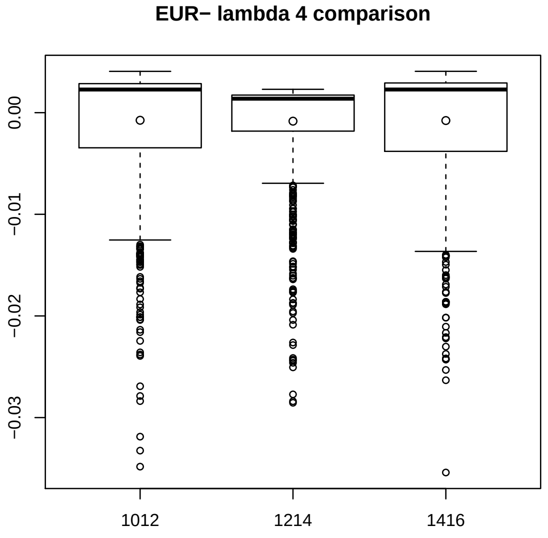
Boxplots

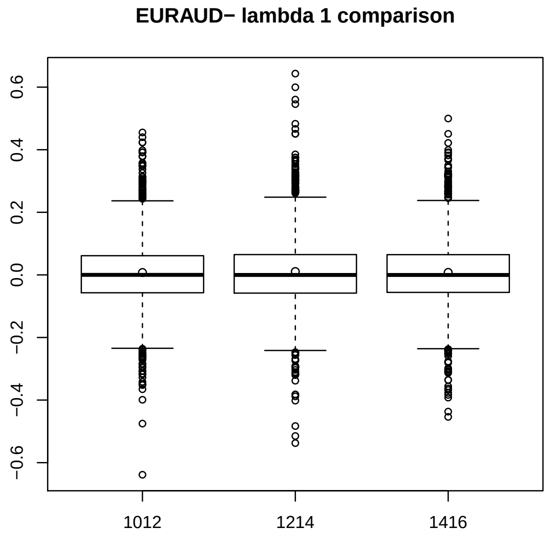


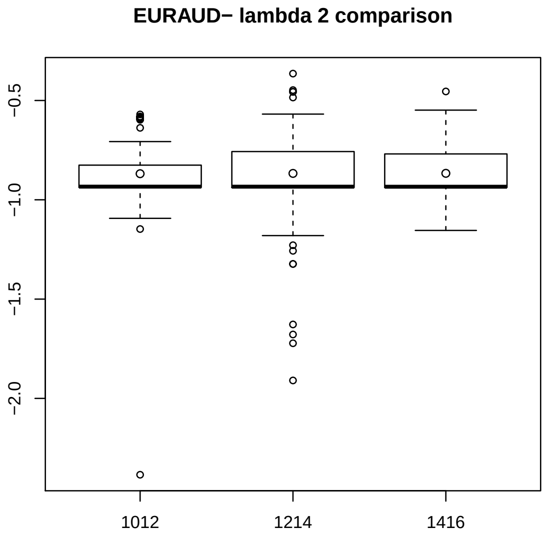

EUR− lambda 1 comparison
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1012
1214
1416

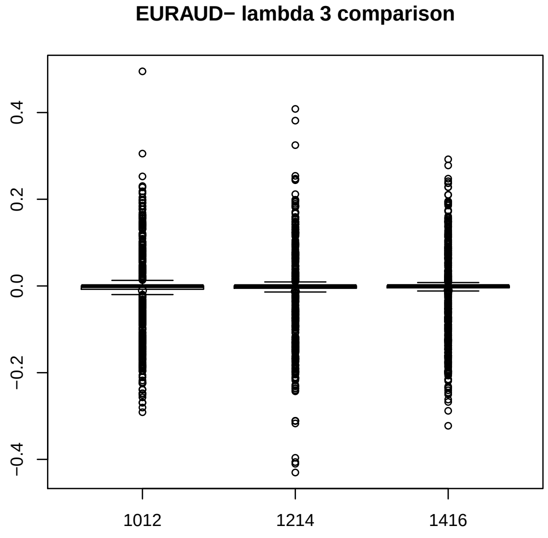

EUR− lambda 2 comparison
−0.4
−0.6
−0.8
−1.0
−1.2
−1.4
−1.6
1012
1214
1416

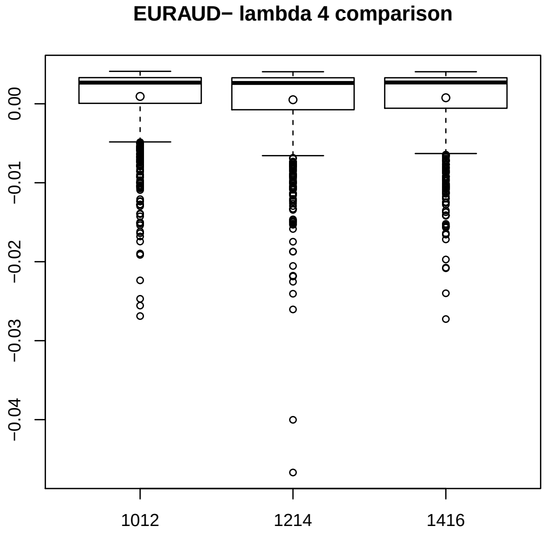

EUR− lambda 3 comparison
0.4
0.2
0.0
−0.2
−0.4
1012
1214
1416


EUR− lambda 4 comparison
0.00
−0.01
−0.02
−0.03
1012
1214
1416

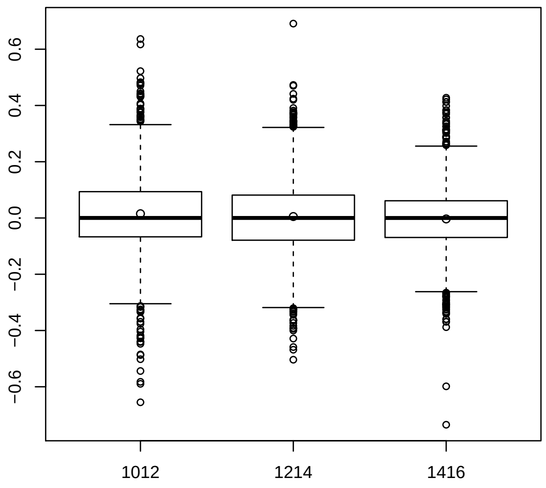

EURAUD− lambda 1 comparison


EURAUD− lambda 2 comparison

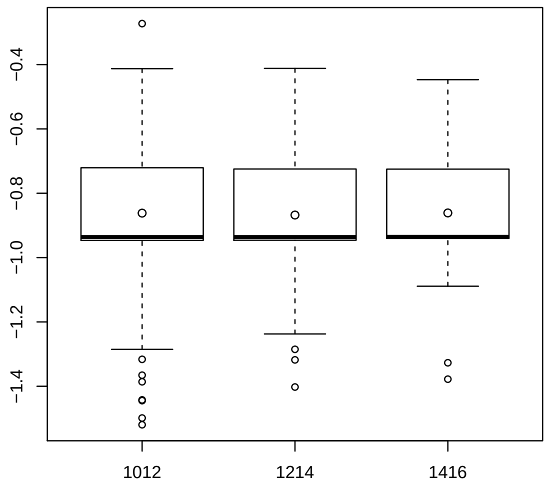

EURAUD− lambda 3 comparison


EURAUD− lambda 4 comparison

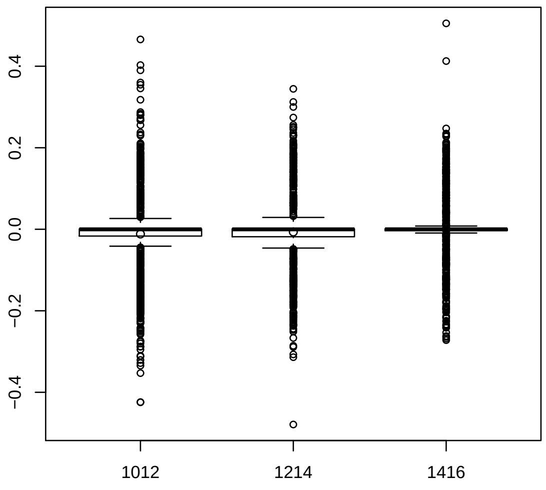
EURGBP− lambda 1 comparison



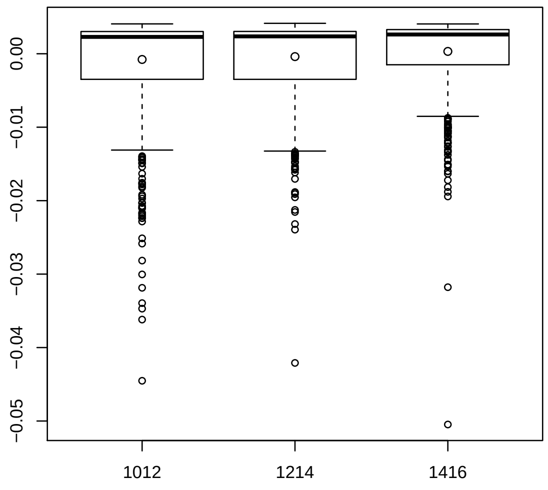
EURGBP− lambda 2 comparison



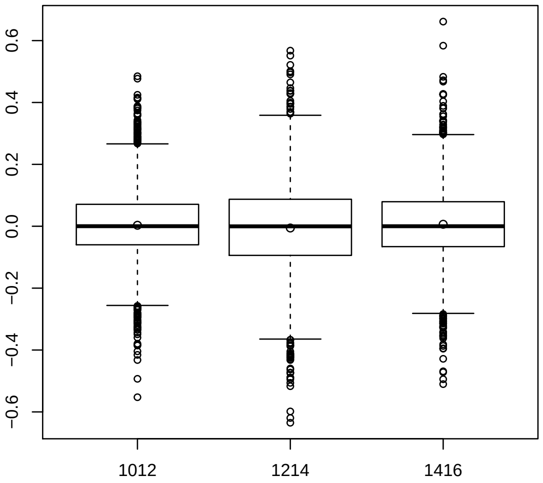
EURGBP− lambda 3 comparison



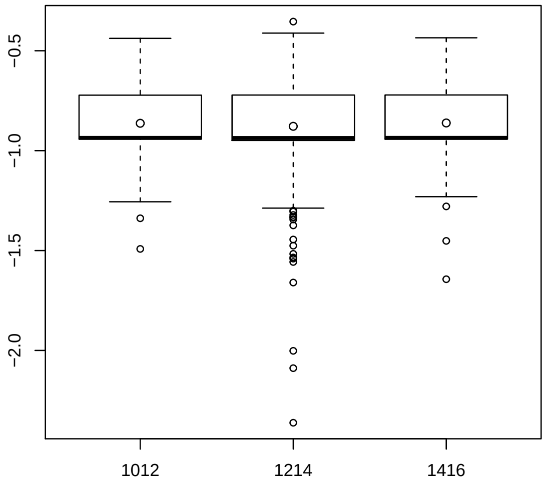
EURGBP− lambda 4 comparison



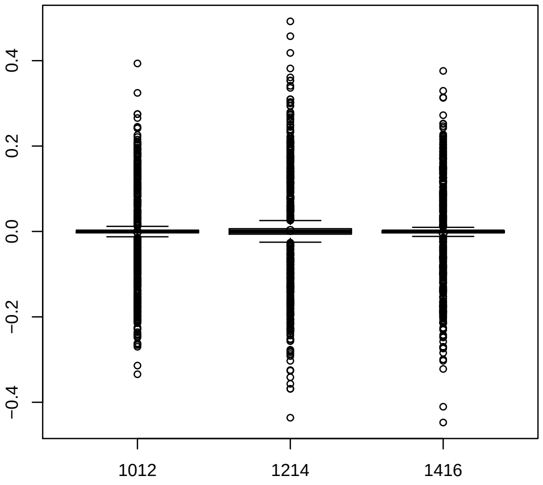
EURJPY− lambda 1 comparison



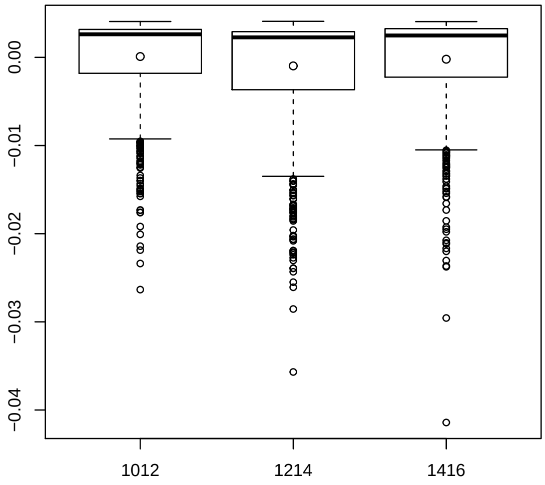
EURJPY− lambda 2 comparison

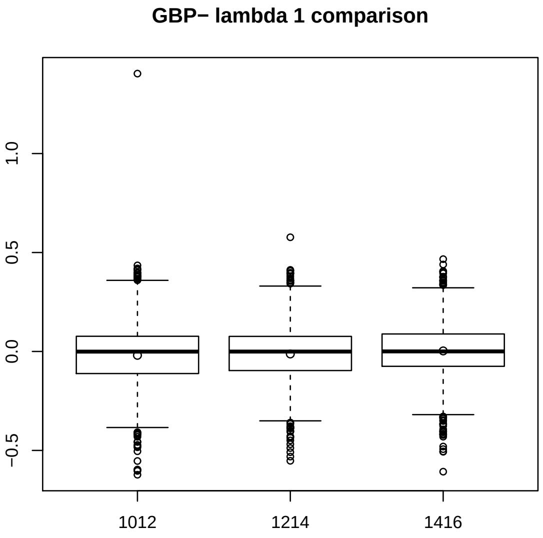


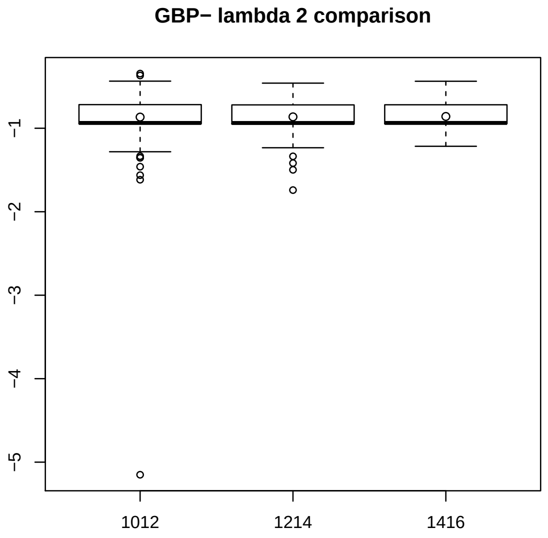
EURJPY− lambda 3 comparison

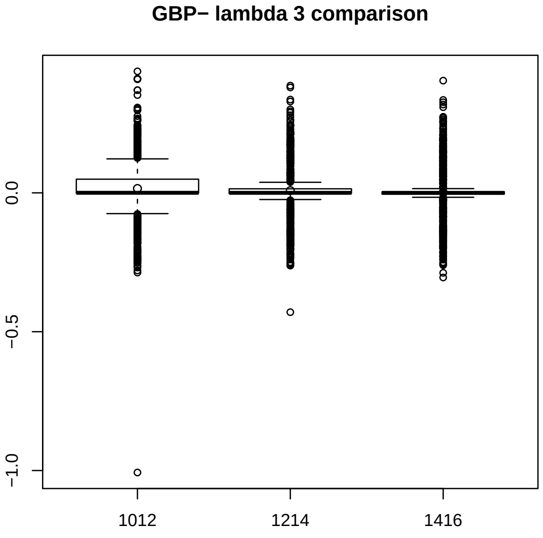


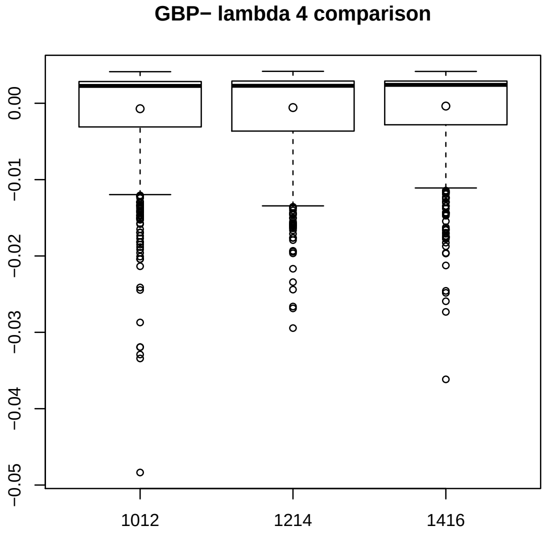
EURJPY− lambda 4 comparison

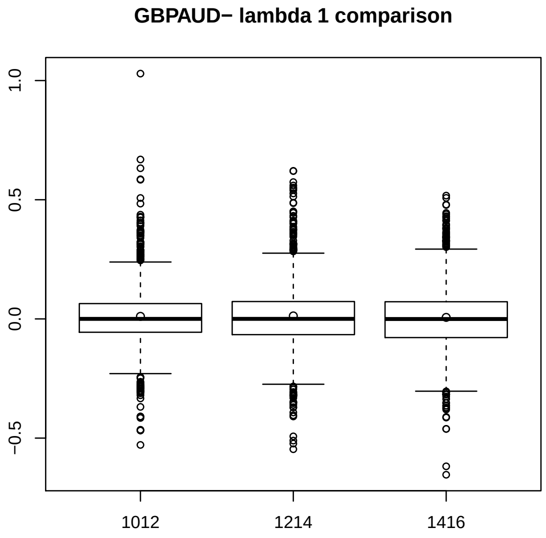



GBP− lambda 1 comparison

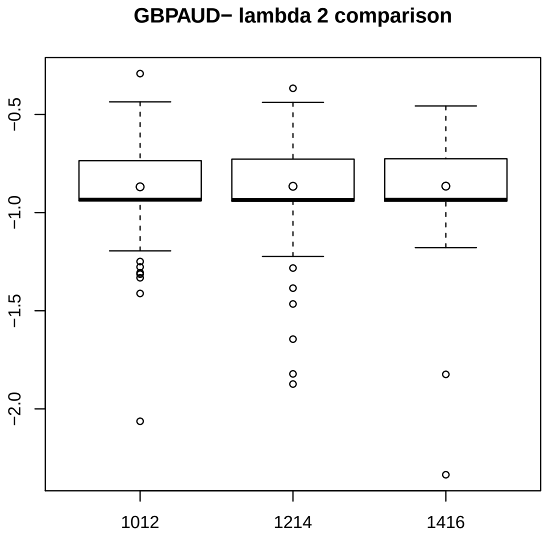

GBP− lambda 2 comparison


GBP− lambda 3 comparison

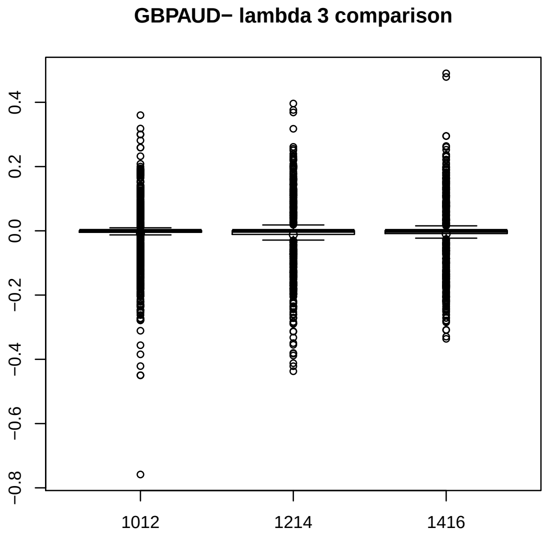

GBP− lambda 4 comparison



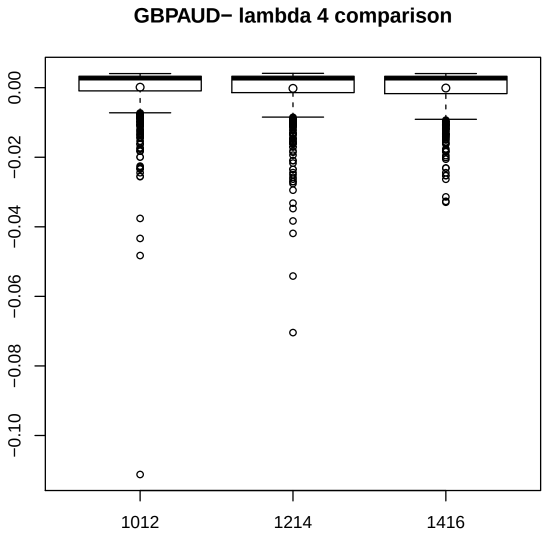
GBPAUD− lambda 1 comparison



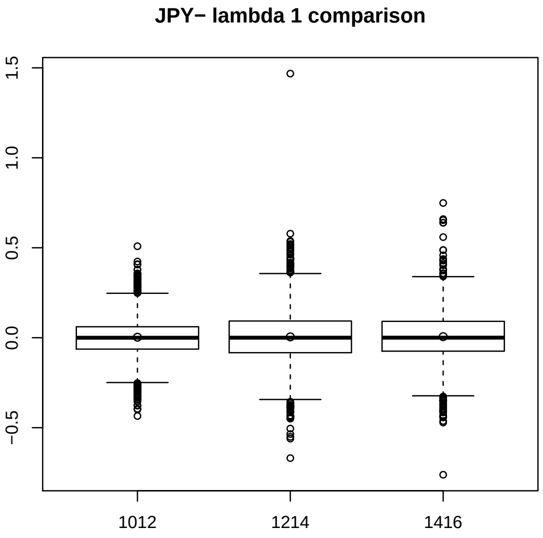
GBPAUD− lambda 2 comparison



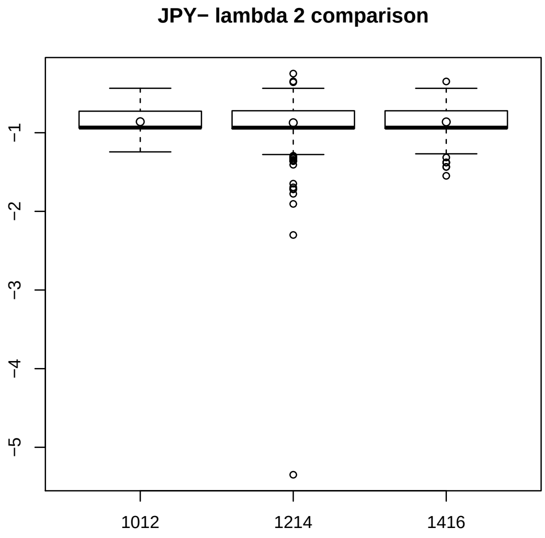
GBPAUD− lambda 3 comparison



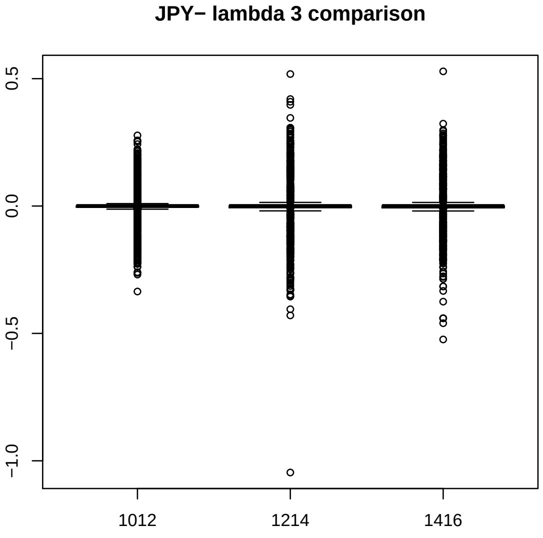
GBPAUD− lambda 4 comparison



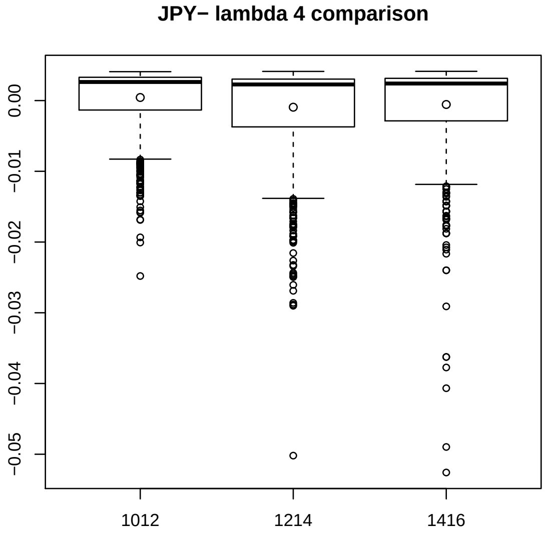
JPY− lambda 1 comparison



JPY− lambda 2 comparison



JPY− lambda 3 comparison



JPY− lambda 4 comparison

Significance of Means - MED Parameters

TimeSlots	10-12		12-14		14-16	
AUD	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	−0.0003	−0.0699	0.0029	0.5723	−0.0015	−0.3204
$\bar{\ell}_2$	−0.8673	−221.5744	−0.8677	−198.1286	−0.8629	−214.5959
$\bar{\ell}_3$	−0.0001	−0.0230	−0.0010	−0.3281	0.0019	0.6897
$\bar{\ell}_4$	0.0004	2.9996	−0.0002	−1.0884	0.0003	1.7276

TimeSlots	10-12		12-14		14-16	
AUDJPY	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	−0.0033	−0.6801	−0.0045	−0.7456	−0.0124	−2.3900
$\bar{\ell}_2$	−0.8675	−198.1785	−0.8799	−110.6193	−0.8766	−197.7718
$\bar{\ell}_3$	0.0048	1.6581	0.0020	0.5196	0.0069	2.2235
$\bar{\ell}_4$	0.0001	0.5027	−0.0010	−4.4280	−0.0003	−1.8458

TimeSlots	10-12		12-14		14-16	
EUR	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	−0.0032	−0.5999	−0.0061	−1.1796	0.0027	0.5061
$\bar{\ell}_2$	−0.8638	−196.7022	−0.8178	−161.3425	−0.8697	−202.9937
$\bar{\ell}_3$	0.0030	0.9368	0.0019	0.6587	−0.0016	−0.5079
$\bar{\ell}_4$	−0.0007	−3.8580	−0.0008	−5.2821	−0.0008	−4.1254

TimeSlots	10-12		12-14		14-16	
EURAUD	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0076	1.8690	0.0107	2.3813	0.0080	1.9244
$\bar{\ell}_2$	−0.8686	−238.2774	−0.8671	−206.8385	−0.8666	−249.2099
$\bar{\ell}_3$	−0.0105	−4.5045	−0.0125	−4.7193	−0.0101	−4.2007
$\bar{\ell}_4$	0.0009	6.5590	0.0005	3.1221	0.0008	5.2369

TimeSlots	10-12		12-14		14-16	
EURGBP	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0152	2.7762	0.0052	1.0263	−0.0033	−0.7357
$\bar{\ell}_2$	−0.8617	−175.7317	−0.8677	−201.8855	−0.8612	−215.8698
$\bar{\ell}_3$	−0.0117	−3.5371	−0.0066	−2.1804	−0.0002	−0.0817
$\bar{\ell}_4$	−0.0008	−3.7943	−0.0004	−2.2772	0.0003	2.0525

TimeSlots	10-12		12-14		14-16	
EURJPY	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0032	0.6851	−0.0056	−0.9758	0.0064	1.2944
$\bar{\ell}_2$	−0.8629	−209.7009	−0.8781	−154.0951	−0.8614	−191.8692
$\bar{\ell}_3$	−0.0006	−0.2168	0.0024	0.6758	−0.0038	−1.2641
$\bar{\ell}_4$	0.0001	0.6199	−0.0010	−4.7570	−0.0002	−1.1816

TimeSlots	10-12		12-14		14-16	
GBP	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	−0.0194	−3.4804	−0.0125	−2.3752	0.0029	0.5658
$\bar{\ell}_2$	−0.8661	−130.8422	−0.8632	−183.9386	−0.8586	−195.5184
$\bar{\ell}_3$	0.0158	4.7073	0.0081	2.5823	0.0015	0.4872
$\bar{\ell}_4$	−0.0007	−3.6455	−0.0006	−3.1346	−0.0004	−2.0890

TimeSlots	10-12		12-14		14-16	
GBPAUD	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0108	2.1769	0.0127	2.5539	0.0064	1.2805
$\bar{\ell}_2$	−0.8685	−198.8939	−0.8658	−189.2006	−0.8650	−198.4785
$\bar{\ell}_3$	−0.0089	−2.9321	−0.0122	−4.0670	−0.0097	−3.3107
$\bar{\ell}_4$	0.0001	0.6277	−0.0002	−0.9271	−0.0001	−0.5400

TimeSlots	10-12		12-14		14-16	
JPY	estimate	t statistic	estimate	t statistic	estimate	t statistic
$\bar{\ell}_1$	0.0027	0.6092	0.0048	0.8272	0.0059	1.0949
$\bar{\ell}_2$	−0.8599	−219.9647	−0.8729	−116.2722	−0.8624	−179.1072
$\bar{\ell}_3$	−0.0042	−1.6322	−0.0032	−0.8711	−0.0047	−1.4440
$\bar{\ell}_4$	0.0004	2.9790	−0.0009	−4.4864	−0.0006	−2.7026

Differences in Means - MED Parameters

Comparisons AUD	1012 vs 1214 t statistic	1012 vs 1416 t statistic	1214 vs 1416 t statistic
$\Delta \bar{\ell}_1$	−0.4774	0.1834	0.6388
$\Delta \bar{\ell}_2$	0.0814	−0.7725	−0.8096
$\Delta \bar{\ell}_3$	0.2346	−0.5170	−0.7057
$\Delta \bar{\ell}_4$	2.7337	0.5765	−1.9869

Comparisons AUDJPY	1012 vs 1214 t statistic	1012 vs 1416 t statistic	1214 vs 1416 t statistic
$\Delta \bar{\ell}_1$	0.1543	1.2720	0.9845
$\Delta \bar{\ell}_2$	1.3656	1.4630	−0.3607
$\Delta \bar{\ell}_3$	0.5942	−0.4900	−1.0039
$\Delta \bar{\ell}_4$	3.8102	1.6888	−2.2484

Comparisons EUR	1012 vs 1214 t statistic	1012 vs 1416 t statistic	1214 vs 1416 t statistic
$\Delta \bar{\ell}_1$	0.3860	−0.7816	−1.1791
$\Delta \bar{\ell}_2$	−6.8668	0.9501	7.8174
$\Delta \bar{\ell}_3$	0.2704	1.0210	0.8158
$\Delta \bar{\ell}_4$	0.3539	0.1013	−0.2492

Comparisons EURAUD	1012 vs 1214 t statistic	1012 vs 1416 t statistic	1214 vs 1416 t statistic
$\Delta \bar{\ell}_1$	−0.5188	−0.0760	0.4404
$\Delta \bar{\ell}_2$	−0.2756	−0.3848	−0.0749
$\Delta \bar{\ell}_3$	0.5655	−0.1220	−0.6723
$\Delta \bar{\ell}_4$	1.9079	0.8188	−1.1287

Comparisons EURGBP	1012 vs 1214 t statistic	1012 vs 1416 t statistic	1214 vs 1416 t statistic
$\Delta \bar{\ell}_1$	1.3444	2.6189	1.2565
$\Delta \bar{\ell}_2$	0.9103	−0.0888	−1.1080
$\Delta \bar{\ell}_3$	−1.1365	−2.7100	−1.5875
$\Delta \bar{\ell}_4$	−1.4543	−4.2630	−3.0657

Comparisons	1012 vs 1214	1012 vs 1416	1214 vs 1416
EURJPY	t statistic	t statistic	t statistic
$\Delta\ \bar{\ell}_1$	1.1893	−0.4730	−1.5847
$\Delta\ \bar{\ell}_2$	2.1601	−0.2430	−2.2969
$\Delta\ \bar{\ell}_3$	−0.6650	0.7776	1.3326
$\Delta\ \bar{\ell}_4$	4.1301	1.2928	−2.8195

Comparisons	1012 vs 1214	1012 vs 1416	1214 vs 1416
GBP	t statistic	t statistic	t statistic
$\Delta\ \bar{\ell}_1$	−0.8970	−2.9479	−2.0992
$\Delta\ \bar{\ell}_2$	−0.3572	−0.9522	−0.7259
$\Delta\ \bar{\ell}_3$	1.6637	3.1725	1.5269
$\Delta\ \bar{\ell}_4$	−0.5631	−1.3088	−0.7747

Comparisons	1012 vs 1214	1012 vs 1416	1214 vs 1416
GBPAUD	t statistic	t statistic	t statistic
$\Delta\ \bar{\ell}_1$	−0.2763	0.6177	0.8904
$\Delta\ \bar{\ell}_2$	−0.4371	−0.5678	−0.1168
$\Delta\ \bar{\ell}_3$	0.7801	0.2024	−0.5883
$\Delta\ \bar{\ell}_4$	1.0911	0.8273	−0.3571

Comparisons	1012 vs 1214	1012 vs 1416	1214 vs 1416
JPY	t statistic	t statistic	t statistic
$\Delta\ \bar{\ell}_1$	−0.2901	−0.4550	−0.1310
$\Delta\ \bar{\ell}_2$	1.5308	0.3957	−1.1776
$\Delta\ \bar{\ell}_3$	−0.2297	0.1106	0.3044
$\Delta\ \bar{\ell}_4$	5.3833	3.9340	−1.3166

A.3 Appendix - Modelling the distribution of Body Mass Index

A.3.1 Proving consistency and asymptotic normality of the estimator

In order to prove that estimator is consistent and asymptotically normal, it is neccessary to make some assumptions:

Assumption 1. *Existence of moments*
Each of the n random variables in the multivariate framework have finite moments up the $2k^{th}$ order i.e. $\mathbb{E}(x_i^{2k}) < \infty$.
All associated cross moments are finite i.e. $\mathbb{E}(x_1^{p_1}\ x_2^{p_2} \ldots x_n^{p_n})$ where $\sum_{i=1}^{n} p_i = 2k$.

Assumption 2. *The second order moment of every conditioning variable/covariate exits i.e. $\mathbb{E}(zz') < \infty$.*

Assumption 3. *Sample moments (and cross moments) for all n variables converge in probability to their corresponding population moments.*

Proposition 4. *The estimator as defined in equation 5.7 is consistent: $\hat{\boldsymbol{\beta}}_T \xrightarrow{p} \boldsymbol{\beta}_0$.*

The outline of the proof is as follows:

1. First, the relationship between $\boldsymbol{\mu}$ and $\boldsymbol{\beta}$ is derived. Specifically, how the moments (and cross moments) change with respect to the parameters i.e. a derivative. This is done using the differentiating $\boldsymbol{\mu}$ which is a vector that contains the moments (and cross moments). Matrix calculus methods are used to carry this out. For further details see Magnus and Neudecker (1999).

2. Based on assumptions 1 and 2, the resulting derivative exists. Hence, by the implicit function theorem the function $\boldsymbol{\beta}(\boldsymbol{\mu})$ also exists.

3. Based on assumption 3, $\hat{\boldsymbol{\mu}}_T \xrightarrow{p} \boldsymbol{\mu}$ as $T \to \infty$. Given that this occurs, then by the continuous mapping theorem (Amemiya (1985)), $\hat{\boldsymbol{\beta}}_T(\hat{\boldsymbol{\mu}}_T) \xrightarrow{p} \boldsymbol{\beta}_0(\boldsymbol{\mu})$ as $T \to \infty$.

Proof. Let

$$\boldsymbol{\mu} = \mathbb{E}\left(\boldsymbol{S}\boldsymbol{x}^{(k)}\right) = \int \boldsymbol{S}\boldsymbol{x}^{(k)}\ f(\boldsymbol{x})\, d\boldsymbol{x} \tag{A.1}$$

and

$$\boldsymbol{\Omega} = \mathbb{E}\left(\boldsymbol{S}\boldsymbol{x}^{(k)}(\boldsymbol{S}\boldsymbol{x}^{(k)})'\right) = \int \boldsymbol{S}\boldsymbol{x}^{(k)}\left(\boldsymbol{S}\boldsymbol{x}^{(k)}\right)'\ f(\boldsymbol{x})\,d\boldsymbol{x}. \tag{A.2}$$

where $f(\boldsymbol{x})$ denotes the multivariate MED. Note that $\boldsymbol{\Omega}$ is a symmetric matrix of dimension $M_k \times M_k$.

Next, differentiating the equation A.1 gives

$$\begin{aligned}
\boldsymbol{\mu} &= Q^{-1}\int \boldsymbol{S}\boldsymbol{x}^{(k)} \exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\ d\boldsymbol{x}\\
d\boldsymbol{\mu} &= dQ^{-1}\int \boldsymbol{S}\boldsymbol{x}^{(k)} \exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\ d\boldsymbol{x} + Q^{-1}\int d\left[\boldsymbol{S}\boldsymbol{x}^{(k)} \exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\right]\ d\boldsymbol{x}\\
d\boldsymbol{\mu} &= -Q^{-2}dQ\int \boldsymbol{S}\boldsymbol{x}^{(k)} \exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\ d\boldsymbol{x} + \int Q^{-1}\boldsymbol{S}\ d\boldsymbol{x}^{(k)} \exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\ d\boldsymbol{x}\\
&\quad + \int Q^{-1}\boldsymbol{S}\boldsymbol{x}^{(k)} \exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\ d\boldsymbol{\lambda}'\ \boldsymbol{S}\boldsymbol{x}^{(k)}\ d\boldsymbol{x}\\
&\quad + \int Q^{-1}\boldsymbol{S}\boldsymbol{x}^{(k)} \exp\ (\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)})\ \boldsymbol{\lambda}'\ \boldsymbol{S}\ d\boldsymbol{x}^{(k)}\ d\boldsymbol{x}\\
d\boldsymbol{\mu} &= -Q^{-1}\ \boldsymbol{\mu}\ dQ + \boldsymbol{S}\ d\boldsymbol{x}^{(k)} + \int \boldsymbol{S}\boldsymbol{x}^{(k)}\left(\boldsymbol{S}\boldsymbol{x}^{(k)}\right)'\ Q^{-1}\exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\ d\boldsymbol{x}\ d\boldsymbol{\lambda}\\
&\quad + \int \boldsymbol{S}\boldsymbol{x}^{(k)}Q^{-1}\exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\ d\boldsymbol{x}\ \boldsymbol{\lambda}'\ \boldsymbol{S}\ d\boldsymbol{x}^{(k)}\\
d\boldsymbol{\mu} &= -Q^{-1}\ \boldsymbol{\mu}\ dQ + \boldsymbol{S}\ d\boldsymbol{x}^{(k)} + \boldsymbol{\Omega}\ d\boldsymbol{\lambda} + \boldsymbol{\mu}\ \boldsymbol{\lambda}'\boldsymbol{S}\ d\boldsymbol{x}^{(k)}.
\end{aligned} \tag{A.3}$$

Now, differentiate equation 5.3 in order to derive dQ in the above equation.

$$\begin{aligned}
dQ &= \int d\left[\exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\right]d\boldsymbol{x}\\
dQ &= \int \left[\exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\ d\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\right]d\boldsymbol{x}\\
dQ &= \int \left[\exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\ \left(d\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)} + \boldsymbol{\lambda}'\boldsymbol{S}\ d\boldsymbol{x}^{(k)}\right)\right]d\boldsymbol{x}\\
dQ &= \int \left(\boldsymbol{S}\boldsymbol{x}^{(k)}\right)'\exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)d\boldsymbol{x}\ d\boldsymbol{\lambda} + \int \exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right)\ d\boldsymbol{x}\ \boldsymbol{\lambda}'\ \boldsymbol{S}\ d\boldsymbol{x}^{(k)}\\
dQ &= Q\ \boldsymbol{\mu}'\ d\boldsymbol{\lambda} + Q\ \boldsymbol{\lambda}'\ \boldsymbol{S}\ d\boldsymbol{x}^{(k)}
\end{aligned} \tag{A.4}$$

Substitute dQ back into equation A.3

$$\begin{aligned}
d\boldsymbol{\mu} &= -Q^{-1}\ \boldsymbol{\mu}\left(Q\ \boldsymbol{\mu}'\ d\boldsymbol{\lambda} + Q\ \boldsymbol{\lambda}'\ \boldsymbol{S}\ d\boldsymbol{x}^{(k)}\right) + \boldsymbol{S}\ d\boldsymbol{x}^{(k)} + \boldsymbol{\Omega}\ d\boldsymbol{\lambda} + \boldsymbol{\mu}\ \boldsymbol{\lambda}'\boldsymbol{S}\ d\boldsymbol{x}^{(k)}\\
d\boldsymbol{\mu} &= -\boldsymbol{\mu}\boldsymbol{\mu}'\ d\boldsymbol{\lambda} - \boldsymbol{\mu}\boldsymbol{\lambda}'\boldsymbol{S}\ d\boldsymbol{x}^{(k)} + \boldsymbol{S}\ d\boldsymbol{x}^{(k)} + \boldsymbol{\Omega}\ d\boldsymbol{\lambda} + \boldsymbol{\mu}\ \boldsymbol{\lambda}'\boldsymbol{S}\ d\boldsymbol{x}^{(k)}\\
d\boldsymbol{\mu} &= (\boldsymbol{\Omega} - \boldsymbol{\mu}\boldsymbol{\mu}')\,d\boldsymbol{\lambda} + \boldsymbol{S}\ d\boldsymbol{x}^{(k)}
\end{aligned} \tag{A.5}$$

Let $\boldsymbol{M} = (\boldsymbol{\Omega} - \boldsymbol{\mu}\boldsymbol{\mu}')$ and hence,

$$d\boldsymbol{\mu} = \boldsymbol{M}\, d\boldsymbol{\lambda} + \boldsymbol{S}\, d\boldsymbol{x}^{(k)}. \tag{A.6}$$

Note that $\boldsymbol{M}$ is a symmetric matrix of dimension $M_k \times M_k$. Since $\boldsymbol{\lambda}$ is a function of $\boldsymbol{z}$ (equation 5.4), the differential $d\boldsymbol{\lambda}$ is equal to

$$d\boldsymbol{\lambda} = d(\boldsymbol{\beta}\boldsymbol{z}) = d\boldsymbol{\beta}\, \boldsymbol{z} + \boldsymbol{\beta} d\boldsymbol{z}. \tag{A.7}$$

Next, substituting equation A.7 into A.6 gives

$$\begin{aligned} d\boldsymbol{\mu} &= \boldsymbol{M}\,(d\boldsymbol{\beta}\, \boldsymbol{z} + \boldsymbol{\beta} d\boldsymbol{z}) + \boldsymbol{S}\, d\boldsymbol{x}^{(k)} \\ d\boldsymbol{\mu} &= \boldsymbol{M} d\boldsymbol{\beta}\, \boldsymbol{z} + \boldsymbol{M}\boldsymbol{\beta} d\boldsymbol{z} + \boldsymbol{S}\, d\boldsymbol{x}^{(k)}. \end{aligned} \tag{A.8}$$

Focusing on the derivative with respect to $\boldsymbol{\beta}$,

$$\begin{aligned} d\boldsymbol{\mu} &= \boldsymbol{M} d\boldsymbol{\beta}\, \boldsymbol{z} \\ d\boldsymbol{\mu} &= vec\,(\boldsymbol{M} d\boldsymbol{\beta}\, \boldsymbol{z}) \\ d\boldsymbol{\mu} &= (\boldsymbol{z}' \otimes \boldsymbol{M})\; vec\; d\boldsymbol{\beta}. \end{aligned}$$

Hence, the partial derivative of $\boldsymbol{\mu}$ with respect to $\boldsymbol{\beta}$ can be written as

$$\frac{\partial \boldsymbol{\mu}}{\partial\; vec\, \boldsymbol{\beta}} = (\boldsymbol{z} \otimes \boldsymbol{M})\,.$$

Given the assumptions, this derivative exists and based on the implicit function theorem the function $\boldsymbol{\beta}(\boldsymbol{\mu})$ also exists. Therefore, for a given sample of T observations, $\hat{\boldsymbol{\beta}}_T(\hat{\boldsymbol{\mu}}_T)$ exists. Based on the law of large numbers ($\hat{\boldsymbol{\mu}}_T \xrightarrow{p} \boldsymbol{\mu}$ as $T \to \infty$) and the continuous mapping theorem (Amemiya (1985)), $\hat{\boldsymbol{\beta}}_T(\hat{\boldsymbol{\mu}}_T) \xrightarrow{p} \boldsymbol{\beta}_0(\boldsymbol{\mu})$ as $T \to \infty$. This completes the proof for consistency of the parameters. □

Proposition 5. *The maximum likelihood estimator is asymptotically normal.* $\sqrt{T}\left(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0\right) \sim N(\boldsymbol{0},\; \boldsymbol{B}(\boldsymbol{\beta}_0)^{-1}\, \boldsymbol{C}(\boldsymbol{\beta}_0)\, \boldsymbol{B}(\boldsymbol{\beta}_0)^{-1})$ *where* $\boldsymbol{B}$ *denotes the matrix of second order derivative of the log likelihood function and* $\boldsymbol{C}$ *denotes a matrix of the product of first order derivative of the log likelihood function.*

Proof. Given that the estimator is consistent, in order to show asymptotic normality the first and second partial derivatives of the log likelihood function must satisfy the following conditions (Amemiya (1985)):

1. $\frac{\partial^2 \log L}{\partial \, vec\boldsymbol{\beta} \, \partial \, vec\boldsymbol{\beta}'}$ exists and is continuous in an open, convex neighbourhood of $\boldsymbol{\beta_0}$.

2. $\frac{1}{T}\frac{\partial^2 \log L}{\partial \, vec\boldsymbol{\beta} \, \partial \, vec\boldsymbol{\beta}'}\Big|_{\hat{\boldsymbol{\beta}}_T}$ converges to a finite non-singular matrix $\boldsymbol{B}(\boldsymbol{\beta_0}) = \lim \mathbb{E}\left[\frac{1}{T}\frac{\partial^2 \log L}{\partial \, vec\boldsymbol{\beta} \, \partial \, vec\boldsymbol{\beta}'}\right]_{\boldsymbol{\beta}_0}$ in probability for any sequence $\hat{\boldsymbol{\beta}}_T$ such that plim $\hat{\boldsymbol{\beta}}_T = \boldsymbol{\beta}_0$.

3. $\frac{1}{\sqrt{T}}\left[\frac{\partial \log L}{\partial \, vec\boldsymbol{\beta}}\right]_{\boldsymbol{\beta}_0} \rightarrow N(0, \boldsymbol{C}(\boldsymbol{\beta}_0))$ where $\boldsymbol{C}(\boldsymbol{\beta}_0) = \lim \mathbb{E}\left[\frac{1}{T}\left[\frac{\partial \log L}{\partial \, vec\boldsymbol{\beta}}\right]_{\boldsymbol{\beta}_0}\left[\frac{\partial \log L}{\partial \, vec\boldsymbol{\beta}'}\right]_{\boldsymbol{\beta}_0}\right]$

Given the density function

$$f(\boldsymbol{x}) = Q^{-1}\exp\left(\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}^{(k)}\right),$$

the log likelihood function can be written as

$$\log L(\boldsymbol{\lambda}|\boldsymbol{x}_t^{(k)}) = \sum_{t=1}^{T} \ell(\boldsymbol{\lambda}|\boldsymbol{x}_t^{(k)})$$

where $\ell(\boldsymbol{\lambda}|\boldsymbol{x}_t^{(k)}) = -\log Q + \boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}_t^{(k)}$. Using differentials, the derivative of ℓ_t with respect to the parameters can be derived.

$$\begin{aligned}
d\ell_t &= d[-\log Q] + d\left[\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}_t^{(k)}\right] \\
d\ell_t &= -\frac{1}{Q}\, dQ + d\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}_t^{(k)} + \boldsymbol{\lambda}'\boldsymbol{S}\, d\boldsymbol{x}_t^{(k)} \\
d\ell_t &= -\frac{1}{Q}\left(Q\,\boldsymbol{\mu}'\, d\boldsymbol{\lambda} + Q\,\boldsymbol{\lambda}'\,\boldsymbol{S}\, d\boldsymbol{x}_t^{(k)}\right) + d\boldsymbol{\lambda}'\boldsymbol{S}\boldsymbol{x}_t^{(k)} + \boldsymbol{\lambda}'\boldsymbol{S}\, d\boldsymbol{x}_t^{(k)} \\
d\ell_t &= d\boldsymbol{\lambda}'\left(\boldsymbol{S}\boldsymbol{x}_t^{(k)} - \boldsymbol{\mu}\right). \qquad \text{(A.9)}
\end{aligned}$$

Apply the differential operator again in order to derive the second order derivatives,

$$d^2\ell_t = d^2\boldsymbol{\lambda}'\left(\boldsymbol{S}\boldsymbol{x}_t^{(k)} - \boldsymbol{\mu}\right) + d\boldsymbol{\lambda}'\, d\left(\boldsymbol{S}\boldsymbol{x}_t^{(k)} - \boldsymbol{\mu}\right).$$

Note that the $d^2\boldsymbol{\lambda}'$ is equal to zero and hence,

$$d^2\ell_t = d\boldsymbol{\lambda}'\left(\boldsymbol{S}d\boldsymbol{x}_t^{(k)} - d\boldsymbol{\mu}\right).$$

Substituting equation A.8 yields

$$d^2\ell_t = d\boldsymbol{\lambda}'\boldsymbol{S}d\boldsymbol{x}_t^{(k)} - d\boldsymbol{\lambda}'\boldsymbol{M}d\boldsymbol{\beta}\, z_t - d\boldsymbol{\lambda}'\boldsymbol{M\beta}dz_t - d\boldsymbol{\lambda}'\boldsymbol{S}d\boldsymbol{x}_t^{(k)} \tag{A.10}$$

As per the specification, the covariate values are given as $z_t = [z_{1,t}, z_{2,t} \ldots, z_{p,t}]'$. From equation A.10, focus on the derivative with respect to β (second term) and vectorize it,

$$d^2\ell_t = -d\boldsymbol{\lambda}'(z_t' \otimes \boldsymbol{M})\; vec\; d\boldsymbol{\beta}.$$

Next, substitute the transpose of $d\boldsymbol{\lambda}$ (equation A.7) into the above

$$d^2\ell_t = -\left[(d\boldsymbol{\beta}\, z_t)' + (\boldsymbol{\beta}\, dz_t)'\right]\; (z_t' \otimes \boldsymbol{M})\; vec\; d\boldsymbol{\beta}$$
$$d^2\ell_t = -(d\boldsymbol{\beta}\, z_t)'\; (z_t' \otimes \boldsymbol{M})\; vec\; d\boldsymbol{\beta} + (\boldsymbol{\beta}\, dz_t)'\; (z_t' \otimes \boldsymbol{M})\; vec\; d\boldsymbol{\beta}$$

Focus on the derivative with respect to β i.e. the first term

$$d^2\ell_t = -(d\boldsymbol{\beta}\, z_t)'\; (z_t' \otimes \boldsymbol{M})\; vec\; d\boldsymbol{\beta} \tag{A.11}$$

Note that

$$(d\boldsymbol{\beta}\, z_t)' = ((z_t' \otimes I)\; vec\; d\boldsymbol{\beta})' = vec\; d\boldsymbol{\beta}'\; (z_t \otimes I). \tag{A.12}$$

Substituting equation A.12 into A.11 yields

$$d^2\ell_t = -vec\; d\boldsymbol{\beta}'\; (z_t \otimes I)(z_t' \otimes \boldsymbol{M})\; vec\; d\boldsymbol{\beta} = -vec\; d\boldsymbol{\beta}'(z_t z_t' \otimes \boldsymbol{M})\; vec\; d\boldsymbol{\beta}.$$

Hence the second order partial derivative with respect to β is equal to

$$\frac{\partial^2 \log\, \ell_t}{\partial\; vec\boldsymbol{\beta}\; \partial\; vec\boldsymbol{\beta}'} = -(z_t z_t' \otimes \boldsymbol{M}).$$

In terms of the log likelihood function,

$$\frac{\partial^2 \log\, L}{\partial\; vec\boldsymbol{\beta}\; \partial\; vec\boldsymbol{\beta}'} = \sum_{t=1}^{T} \frac{\partial^2 \log\, \ell_t}{\partial\; vec\boldsymbol{\beta}\; \partial\; vec\boldsymbol{\beta}'} = -\sum_{t=1}^{T}(z_t z_t' \otimes \boldsymbol{M}).$$

Given assumptions 1 and 3 the second order partial derivative with respect to β exits i.e. $\frac{\partial^2 \log\, L}{\partial\; vec\boldsymbol{\beta}\; \partial\; vec\boldsymbol{\beta}'} < \infty$. As such, the first condition is satisfied.

The second condition can be rewritten as

$$\begin{aligned}
&\frac{1}{T}\frac{\partial^2 \log L}{\partial\, vec\boldsymbol{\beta}\,\partial\, vec\boldsymbol{\beta}'}\bigg|_{\hat{\boldsymbol{\beta}}_T} - \boldsymbol{B}(\boldsymbol{\beta_0}) &&= 0\\
&\frac{1}{T}\frac{\partial^2 \log L}{\partial\, vec\boldsymbol{\beta}\,\partial\, vec\boldsymbol{\beta}'}\bigg|_{\hat{\boldsymbol{\beta}}_T} - \mathbb{E}\left[\frac{1}{T}\frac{\partial^2 \log L}{\partial\, vec\boldsymbol{\beta}\,\partial\, vec\boldsymbol{\beta}'}\right]_{\boldsymbol{\beta}_0} &&= 0\\
&-\frac{1}{T}\sum_{t=1}^{T}(z_t z_t' \otimes \boldsymbol{M})\bigg|_{\hat{\boldsymbol{\beta}}_T} + \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}(z_t z_t' \otimes \boldsymbol{M})\right]_{\boldsymbol{\beta}_0} &&= 0\\
&-\frac{\sum_{t=1}^{T}(z_t z_t')}{T}\otimes \boldsymbol{M}\bigg|_{\hat{\boldsymbol{\beta}}_T} + \left[\mathbb{E}(z_t z_t')\otimes \boldsymbol{M}\right]_{\boldsymbol{\beta}_0} &&= 0 \qquad \text{(A.13)}
\end{aligned}$$

Based on the assumption of the law of large numbers $\frac{\sum_{t=1}^{T}(z_t z_t')}{T} \to \mathbb{E}(z_t z_t')$ as $T \to \infty$ and it is assumed that $\mathbb{E}(z_t z_t')$ exists (assumption 3). Since $\boldsymbol{M} = (\boldsymbol{\Omega} - \boldsymbol{\mu\mu}')$ is a function of $\boldsymbol{\beta}$ and given that the parameters are consistent (proposition 1), then by the continuous mapping theorem, $\boldsymbol{M}(\boldsymbol{\beta}_0) - \boldsymbol{M}(\hat{\boldsymbol{\beta}}_T) \to 0$ as $T \to \infty$. This proves the second condition.

In order to prove the third condition, start with equation A.9

$$d\ell_t = d\boldsymbol{\lambda}'\left(\boldsymbol{S}\boldsymbol{x}_t^{(k)} - \boldsymbol{\mu}\right).$$

Next, substitute the transpose of $d\boldsymbol{\lambda}$ (equation A.7) into the above

$$d\ell_t = (dz_t'\boldsymbol{\beta}' + z_t' d\boldsymbol{\beta}')\left(\boldsymbol{S}\boldsymbol{x}_t^{(k)} - \boldsymbol{\mu}\right).$$

Focus on the derivative with respect to β

$$d\ell_t = \left((\boldsymbol{S}\boldsymbol{x}_t^{(k)} - \boldsymbol{\mu}) \otimes z_t\right)' vec\; d\boldsymbol{\beta}'.$$

Hence, the partial derivative with respect to β is equal to

$$\frac{\partial \ell}{\partial\, vec\boldsymbol{\beta}} = (\boldsymbol{S}\boldsymbol{x}_t^{(k)} - \boldsymbol{\mu}) \otimes z_t.$$

Based on this,

$$\frac{\partial \log L}{\partial\, vec\boldsymbol{\beta}} = \sum_{t=1}^{T}\frac{\partial \ell}{\partial\, vec\boldsymbol{\beta}} = \sum_{t=1}^{T}\left((\boldsymbol{S}\boldsymbol{x}_t^{(k)} - \boldsymbol{\mu}) \otimes z_t\right). \qquad \text{(A.14)}$$

The above equation represents a sum of random variables. The expected value of this random variable is equal to 0 i.e.

$$\mathbb{E}\left((\boldsymbol{Sx}_t^{(k)} - \boldsymbol{\mu}) \otimes z_t\right) = 0$$

since by definition $\mathbb{E}\left(\boldsymbol{Sx}_t^{(k)}\right) = \boldsymbol{\mu}$. Let $\boldsymbol{D} = (\boldsymbol{Sx}_t^{(k)} - \boldsymbol{\mu})$ then the variance of this random variable is given by

$$\begin{aligned} Var\left((\boldsymbol{Sx}_t^{(k)} - \boldsymbol{\mu}) \otimes z_t\right) &= \mathbb{E}\left((\boldsymbol{D} \otimes z_t)(\boldsymbol{D} \otimes z_t)'\right) \\ &= \mathbb{E}\left(\boldsymbol{DD}' \otimes z_t z_t'\right). \end{aligned} \tag{A.15}$$

Here $\mathbb{E}\left(\boldsymbol{DD}'\right)$ is equal to

$$\begin{aligned} &= \mathbb{E}\left(\boldsymbol{Sx}_t^{(k)} - \boldsymbol{\mu}\right)\left(\boldsymbol{Sx}_t^{(k)} - \boldsymbol{\mu}\right)' \\ &= \mathbb{E}\left(\boldsymbol{Sx}_t^{(k)}\boldsymbol{x}_t^{(k)\prime}\boldsymbol{S}' + \boldsymbol{\mu}\boldsymbol{\mu}' - 2\boldsymbol{Sx}_t^{(k)}\boldsymbol{\mu}'\right) \\ &= (\boldsymbol{\Omega} - \boldsymbol{\mu}\boldsymbol{\mu}') = \boldsymbol{M}. \end{aligned}$$

Substituting this result into equation A.15,

$$Var\left((\boldsymbol{Sx}_t^{(k)} - \boldsymbol{\mu}) \otimes z_t\right) = \boldsymbol{M} \otimes \mathbb{E}(z_t z_t').$$

Based on assumptions 1 and 3, the resulting variance is finite i.e. $\boldsymbol{M} \otimes \mathbb{E}(z_t z_t') < \infty$.

Given these results the Lindeberg-Feller central limit theorem states that

$$\frac{\partial \log L}{\partial\, vec\beta} \sim N(\boldsymbol{0}, T\left(\boldsymbol{M} \otimes \mathbb{E}(z_t z_t')\right))$$

Multiplying by $\frac{1}{\sqrt{T}}$ yields,

$$\frac{1}{\sqrt{T}} \frac{\partial \log L}{\partial\, vec\beta} \sim N(\boldsymbol{0}, (\boldsymbol{M} \otimes \mathbb{E}(z_t z_t')))$$

where $\boldsymbol{C}(\boldsymbol{\beta}_0) = (\boldsymbol{M} \otimes \mathbb{E}(z_t z_t')) = \lim \mathbb{E}\left[\frac{1}{T}\left[\frac{\partial \log L}{\partial\, vec\beta}\right]_{\beta_0}\left[\frac{\partial \log L}{\partial\, vec\beta'}\right]_{\beta_0}\right]$. This proves the third and final condition.

As stated in the proposition, $\sqrt{T}\left(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0\right) \sim N(\boldsymbol{0}, \boldsymbol{B}(\boldsymbol{\beta}_0)^{-1}\, \boldsymbol{C}(\boldsymbol{\beta}_0)\, \boldsymbol{B}(\boldsymbol{\beta}_0)^{-1})$. By applying the Moore-Penrose inverse, this simplifies to $\sqrt{T}\left(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}_0\right) \sim N(\boldsymbol{0}, \boldsymbol{B}(\boldsymbol{\beta}_0)^{-1})$ where $\boldsymbol{B}(\boldsymbol{\beta}_0) = \mathbb{E}(z_t z_t') \otimes \boldsymbol{M}$. Hence, the maximum likelihood estimator is asymptotically normal. □

A.3.2 Analytical derivatives of the log-likelihood function

The conditional distribution of BMI for an individual i is expressed as

$$f(y|z_i) = Q_i^{-1}\exp\left[(\boldsymbol{\beta} z_i)'\,\boldsymbol{y}\right]$$

where

$$Q_i = \int_{\mathbf{A}} \exp\left[(\boldsymbol{\beta} z_i)'\,\boldsymbol{y}\right] dy.$$

For a sample of n individuals, the log-likelihood function for the conditional distribution is

$$\begin{aligned}
\log L(\boldsymbol{\beta}; z_i, y_i) &= \sum_{i=1}^{n} \log f(y_i|z_i) \\
&= \sum_{i=1}^{n} \log\left[Q_i^{-1}\exp\left((\boldsymbol{\beta} z_i)'\,\boldsymbol{y}_i\right)\right] \\
&= \sum_{i=1}^{n} \left[-\log Q_i + (\boldsymbol{\beta} z_i)'\,\boldsymbol{y}_i\right] \\
&= -\sum_{i=1}^{n} \log Q_i + \sum_{i=1}^{n} (\boldsymbol{\beta} z_i)'\,\boldsymbol{y}_i
\end{aligned}$$

Given z_i and y_i, find $\boldsymbol{\beta}$ such that log $L(\boldsymbol{\beta}; z_i, \boldsymbol{y}_i)$ is maximized. It is often useful to incorporate derivatives in the optimization routine. The first order partial derivative of the log likelihood function with respect to $\boldsymbol{\beta}$ is given equation A.14. The elements of this derivative matrix can be written as

$$\begin{aligned}
\frac{\partial \log L}{\partial \beta_{\ell j}} &= -\sum_{i=1}^{n} \frac{\partial \log L}{\partial Q_i}\,\frac{\partial Q_i}{\partial \beta_{\ell j}} + \sum_{i=1}^{n} z_{ji}\, y_i^{\ell} \\
&= -\sum_{i=1}^{n} \frac{1}{Q_i}\,\frac{\partial}{\partial \beta_{\ell j}} \int_{\mathbf{A}} \exp\left[(\boldsymbol{\beta} \boldsymbol{z}_i)'\, \boldsymbol{y}\right] dy + \sum_{i=1}^{n} z_{ji}\, y_i^{\ell} \\
&= -\sum_{i=1}^{n} \frac{1}{Q_i} \int_{\mathbf{A}} \frac{\partial}{\partial \beta_{\ell j}} \exp\left((\boldsymbol{\beta} \boldsymbol{z}_i)'\, \boldsymbol{y}\right) dy + \sum_{i=1}^{n} z_{ji}\, y_i^{\ell} \\
&= -\sum_{i=1}^{n} \frac{1}{Q_i} \int_{\mathbf{A}} z_{ji}\, y^{j} \exp\Big((\boldsymbol{\beta} \boldsymbol{z}_i)^{T}\, \boldsymbol{y}\Big) dy + \sum_{i=1}^{n} z_{ji}\, y_i^{\ell} \\
&= -\sum_{i=1}^{n} \frac{z_{ji}}{Q_i} \int_{\mathbf{A}} y^{j} \exp\Big((\boldsymbol{\beta} \boldsymbol{z}_i)^{T}\, \boldsymbol{y}\Big) dy + \sum_{i=1}^{n} z_{ji}\, y_i^{\ell} \\
&= -\sum_{i=1}^{n} \frac{z_{ji}}{Q_i}\, \mu_{\ell i}\, Q_i + \sum_{i=1}^{n} z_{ji}\, y_i^{\ell} \\
&= -\sum_{i=1}^{n} z_{ji}\, \mu_{\ell i} + \sum_{i=1}^{n} z_{ji}\, y_i^{\ell}.
\end{aligned}$$

Note that there are k moments ($\ell = 1, 2, \ldots, k$) and p covariates ($j = 1, 2, \ldots, p$).

A.3.3 Summary Statistics - Covariates

```
male
Min.    1st Qu.  Median  Mean    3rd Qu.  Max.
0.0000  0.0000   0.0000  0.4726  1.0000   1.0000

university education
Min.    1st Qu.  Median  Mean    3rd Qu.  Max.
0.0000  0.0000   0.0000  0.2488  0.0000   1.0000

certificate or diploma
Min.    1st Qu.  Median  Mean    3rd Qu.  Max.
0.0000  0.0000   0.0000  0.3045  1.0000   1.0000

employed
Min.    1st Qu.  Median  Mean    3rd Qu.  Max.
0.0000  0.0000   1.0000  0.6422  1.0000   1.0000

not in labour force
Min.    1st Qu.  Median  Mean    3rd Qu.  Max.
0.0000  0.0000   0.0000  0.3214  1.0000   1.0000

unemployed
Min.     1st Qu.  Median   Mean     3rd Qu.  Max.
0.00000  0.00000  0.00000  0.03635  0.00000  1.00000

married
Min.    1st Qu.  Median  Mean    3rd Qu.  Max.
0.0000  0.0000   1.0000  0.6419  1.0000   1.0000

separate
Min.    1st Qu.  Median  Mean    3rd Qu.  Max.
0.0000  0.0000   0.0000  0.0865  0.0000   1.0000

widow
Min.     1st Qu.  Median   Mean     3rd Qu.  Max.
0.00000  0.00000  0.00000  0.04615  0.00000  1.00000

single
Min.    1st Qu.  Median   Mean    3rd Qu.  Max.
0.0000  0.0000   0.0000   0.2255  0.0000   1.0000

smoker
Min.    1st Qu.  Median   Mean    3rd Qu.  Max.
0.0000  0.0000   0.0000   0.1806  0.0000   1.0000

non-smoker
Min.    1st Qu.  Median   Mean    3rd Qu.  Max.
0.0000  0.0000   1.0000   0.5473  1.0000   1.0000

non-drinker
Min.    1st Qu.  Median   Mean    3rd Qu.  Max.
0.0000  0.0000   0.0000   0.1846  0.0000   1.0000

inactive
Min.    1st Qu.  Median  Mean    3rd Qu.  Max.
0.0000  0.0000   0.0000  0.1112  0.0000   1.0000

household income
Min. 1st Qu.  Median   Mean    3rd Qu.   Max.
52   51690    94260    110300  144300    1079000

age
Min.   1st Qu.  Median  Mean   3rd Qu.  Max.
15.00  29.00    44.00   44.97  59.00    100.00

number of children
Min.    1st Qu.  Median  Mean   3rd Qu. Max.
0.000   0.000    2.000   1.637  3.000   14.000
```

A.4 Appendix - Modelling Populations in Remote Communities

A.4.1 Descriptive statistics for variables included in the regression model

Variable	Mean	Std.Dev.	Minimum	Maximum
Population change (dependent variable)	-0.0295	6.0961	-79.71	100.78
ILOC size (log)	6.9704	1.6393	2.17	11.36
Remote	0.1537	0.3607	0	1
Very remote	0.4498	0.4975	0	1
Victoria	0.0162	0.1262	0	1
Queensland	0.2104	0.4076	0	1
South Australia	0.0744	0.2625	0	1
Western Australia	0.2120	0.4087	0	1
Tasmania	0.0356	0.1853	0	1
Northern Territory	0.2961	0.4566	0	1
Female	0.5000	0.5000	0	1
Growth town	0.0340	0.1812	0	1
Cohort age (2011)				
Age 10–14	0.0625	0.2421	0	1
Age 15–19	0.0625	0.2421	0	1
Age 19–24	0.0625	0.2421	0	1
Age 25–29	0.0625	0.2421	0	1
Age 30–34	0.0625	0.2421	0	1
Age 35–39	0.0625	0.2421	0	1
Age 40–44	0.0625	0.2421	0	1
Age 45–49	0.0625	0.2421	0	1
Age 50–54	0.0625	0.2421	0	1
Age 55–59	0.0625	0.2421	0	1
Age 60–64	0.0625	0.2421	0	1
Age 65–69	0.0625	0.2421	0	1
Age 70–74	0.0625	0.2421	0	1
Age 75–79	0.0625	0.2421	0	1

Age 80+	0.0625	0.2421	0	1
Survival rate	0.9012	0.1252	0.60	1.00
ILOC Size*Age 10–14	0.4356	1.7364	0	11.36
ILOC Size*Age 15–19	0.4356	1.7364	0	11.36
ILOC Size*Age 19–24	0.4356	1.7364	0	11.36
ILOC Size*Age 25–29	0.4356	1.7364	0	11.36
ILOC Size*Age 30–34	0.4356	1.7364	0	11.36
ILOC Size*Age 35–39	0.4356	1.7364	0	11.36
ILOC Size*Age 40–44	0.4356	1.7364	0	11.36
ILOC Size*Age 45–49	0.4356	1.7364	0	11.36
ILOC Size*Age 50–54	0.4356	1.7364	0	11.36
ILOC Size*Age 55–59	0.4356	1.7364	0	11.36
ILOC Size*Age 60–64	0.4356	1.7364	0	11.36
ILOC Size*Age 65–69	0.4356	1.7364	0	11.36
ILOC Size*Age 70–74	0.4356	1.7364	0	11.36
ILOC Size*Age 75–79	0.4356	1.7364	0	11.36
ILOC Size*Age 80+	0.4356	1.7364	0	11.36
Outer Reg*Age 10–14	0.0248	0.1555	0	1
Outer Reg*Age 15–19	0.0248	0.1555	0	1
Outer Reg*Age 19–24	0.0248	0.1555	0	1
Outer Reg*Age 25–29	0.0248	0.1555	0	1
Outer Reg*Age 30–34	0.0248	0.1555	0	1
Outer Reg*Age 35–39	0.0248	0.1555	0	1
Outer Reg*Age 40–44	0.0248	0.1555	0	1
Outer Reg*Age 45–49	0.0248	0.1555	0	1
Outer Reg*Age 50–54	0.0248	0.1555	0	1
Outer Reg*Age 55–59	0.0248	0.1555	0	1
Outer Reg*Age 60–64	0.0248	0.1555	0	1
Outer Reg*Age 65–69	0.0248	0.1555	0	1
Outer Reg*Age 70–74	0.0248	0.1555	0	1
Outer Reg*Age 75–79	0.0248	0.1555	0	1
Outer Reg*Age 80+	0.0248	0.1555	0	1
Remote*Age 10–14	0.0096	0.0975	0	1
Remote*Age 15–19	0.0096	0.0975	0	1

Remote*Age 19–24	0.0096	0.0975	0	1
Remote*Age 25–29	0.0096	0.0975	0	1
Remote*Age 30–34	0.0096	0.0975	0	1
Remote*Age 35–39	0.0096	0.0975	0	1
Remote*Age 40–44	0.0096	0.0975	0	1
Remote*Age 45–49	0.0096	0.0975	0	1
Remote*Age 50–54	0.0096	0.0975	0	1
Remote*Age 55–59	0.0096	0.0975	0	1
Remote*Age 60–64	0.0096	0.0975	0	1
Remote*Age 65–69	0.0096	0.0975	0	1
Remote*Age 70–74	0.0096	0.0975	0	1
Remote*Age 75–79	0.0096	0.0975	0	1
Remote*Age 80+	0.0096	0.0975	0	1

Table A.18:

A.4.2 Aboriginal and Torres Strait Islander populations by ILOC: 2011 Census estimates and 2016 projections

Indigenous Location (ILOC)	**State**	**ARIA**	**Population2011**	**Population2016**	**% change**
Balranald	NSW	3	93	103	11.2
Balranald - Wentworth - Surrounds	NSW	3	279	331	18.8
Baryulgil - Malabugilmah	NSW	3	107	115	7.9
Bega	NSW	3	233	286	22.8
Bega - Surrounds	NSW	3	430	520	20.9
Bellingen	NSW	3	375	443	18.1
Bodalla	NSW	3	77	77	0.3
Bogan	NSW	4	415	450	8.5
Boggabilla	NSW	3	369	422	14.4
Bourke	NSW	5	760	831	9.3
Bourke - Surrounds	NSW	5	105	104	-1.2
Bowraville	NSW	3	259	298	14.9
Brewarrina	NSW	5	607	650	7.1
Brewarrina - Surrounds	NSW	5	259	263	1.5

Broken Hill	NSW	3	1398	1630	16.6
Broken Hill - Surrounds	NSW	5	151	146	-3.4
Carrathool - Murrumbidgee - Surrounds	NSW	3	111	141	27.3
Central Murray	NSW	3	494	593	20.1
Clarence Valley	NSW	3	286	349	22.1
Cobar	NSW	4	505	548	8.4
Collarenebri	NSW	4	188	185	-1.6
Condobolin	NSW	3	710	829	16.7
Coolamon - Temora - West Wyalong	NSW	3	510	606	18.9
Coomealla	NSW	3	130	137	5.4
Coonabarabran	NSW	3	430	501	16.6
Coonabarabran - Surrounds	NSW	3	377	457	21.2
Coonamble	NSW	4	903	992	9.9
Coonamble - Surrounds	NSW	4	102	99	-2.7
Dareton	NSW	3	184	203	10.2
Darlington Point	NSW	3	190	215	13.1
Dubbo - Surrounds	NSW	3	345	410	18.9
Eden	NSW	3	226	263	16.2
Eurobodalla	NSW	3	481	563	17.1
Forbes	NSW	3	884	1038	17.4
Gilgandra	NSW	3	533	621	16.5
Gingie Reserve	NSW	4	64	54	-16.2
Glen Innes	NSW	3	496	571	15.2
Goodooga	NSW	5	178	173	-2.9
Griffith	NSW	3	998	1154	15.7
Gulargambone	NSW	3	170	185	9.0
Gunnedah	NSW	3	1099	1293	17.7
Gunnedah - Surrounds	NSW	3	245	284	16.0
Guyra	NSW	3	272	318	17.1
Gwydir	NSW	3	190	228	20.0
Hillston	NSW	4	100	91	-9.2
Inverell	NSW	3	1059	1243	17.4

Jubullum Village	NSW	3	137	148	8.4
Kempsey - Surrounds	NSW	3	878	1010	15.0
Lachlan	NSW	4	135	137	1.5
Lake Cargelligo	NSW	4	229	225	-1.5
Leeton	NSW	3	650	783	20.4
Lightning Ridge	NSW	4	426	455	6.8
Liverpool Plains	NSW	3	145	176	21.4
Macksville	NSW	3	215	256	18.9
Menindee	NSW	5	180	172	-4.4
Mirriwinni Gardens - Bellbrook	NSW	3	152	175	15.1
Moree - North	NSW	3	338	379	12.2
Moree - South	NSW	3	969	1105	14.0
Moree - West	NSW	3	194	220	13.5
Moree Plains	NSW	3	219	262	19.8
Mudgee	NSW	3	963	1127	17.1
Muli Muli - Woodenbong	NSW	3	201	226	12.3
Mungindi	NSW	4	127	123	-3.2
Murrin Bridge	NSW	4	99	81	-18.1
Nambucca - Surrounds	NSW	3	295	351	19.1
Nambucca Heads	NSW	3	578	683	18.1
Namoi Reserve	NSW	4	91	71	-22.0
Narooma	NSW	3	128	153	19.4
Narrabri	NSW	3	634	738	16.4
Narrabri - Surrounds	NSW	3	403	481	19.4
Narrandera	NSW	3	582	687	18.1
Narromine	NSW	3	916	1092	19.3
Narromine - Surrounds	NSW	3	166	195	17.5
Parkes	NSW	3	751	890	18.5
Parkes - Surrounds	NSW	3	220	261	18.7
Peak Hill	NSW	3	205	227	10.8
Quirindi	NSW	3	334	386	15.5
South West Rocks	NSW	3	262	320	22.1
Stanley Village	NSW	3	319	358	12.1
Tamworth - Surrounds	NSW	3	1006	1193	18.5

Tenterfield	NSW	3	310	372	20.1
Tingha	NSW	3	175	204	16.3
Toomelah	NSW	3	230	259	12.8
Trangie	NSW	3	224	256	14.3
Uralla	NSW	3	330	381	15.6
Walcha	NSW	3	225	265	17.9
Walgett	NSW	4	819	885	8.1
Walgett - Surrounds	NSW	4	208	206	-0.9
Walhallow Reserve (Carooma)	NSW	3	87	91	4.3
Wallaga Lake	NSW	3	118	126	6.7
Warren	NSW	3	360	414	15.1
Wee Waa	NSW	3	340	394	15.8
Wellington	NSW	3	1135	1321	16.4
Wellington - Surrounds	NSW	3	600	654	8.9
Wentworth	NSW	3	138	151	9.5
Wilcannia	NSW	5	466	499	7.1
Bairnsdale	Vic	3	538	650	20.9
East Gippsland	Vic	3	677	797	17.8
Glenelg North - Heywood	Vic	3	211	262	24.2
Lake Tyers	Vic	3	124	127	2.2
Mildura	Vic	3	1843	2197	19.2
Portland	Vic	3	184	229	24.2
Swan Hill	Vic	3	432	529	22.5
Swan Hill - Robinvale	Vic	3	293	352	20.1
Swan Hill - Surrounds	Vic	3	155	196	26.1
Wimmera	Vic	3	622	730	17.3
Atherton	Qld	3	763	913	19.6
Aurukun	Qld	5	1201	1325	10.3
Ayr	Qld	3	601	721	19.9
Badu Island	Qld	5	708	797	12.5
Balonne exc. St George and Dirranbandi	Qld	4	131	142	8.4
Bamaga and Surrounds	Qld	5	860	962	11.9
Banana - North	Qld	3	297	367	23.4

Banana - South	Qld	4	43	53	23.0
Barcaldine	Qld	5	200	229	14.3
Barron	Qld	3	1021	1205	18.0
Biloela	Qld	3	227	289	27.3
Blackall - Tambo	Qld	5	88	99	12.3
Blackwater	Qld	3	238	303	27.2
Boigu Island	Qld	5	189	198	4.6
Boulia	Qld	5	185	202	9.4
Bowen (Qld)	Qld	3	873	1023	17.1
Bulloo - Quilpie - Barcoo	Qld	5	185	207	11.7
Burdekin	Qld	3	309	380	23.0
Cairns - City	Qld	3	700	785	12.1
Cairns - North	Qld	3	539	657	21.9
Cairns - Southern Hinterlands	Qld	3	424	504	18.8
Cairns - West	Qld	3	839	986	17.5
Cairns - White Rock - Mt Sheridan	Qld	3	1471	1748	18.8
Camooweal	Qld	5	105	111	5.6
Cape Tribulation - China Camp - Zig Zag	Qld	4	42	40	-4.1
Cape York Wilderness	Qld	4	150	159	6.2
Cardwell	Qld	3	349	422	20.9
Carpentaria exc. Doomadgee	Qld	5	257	285	10.9
Central Highlands	Qld	4	340	396	16.3
Charters Towers	Qld	3	834	988	18.4
Cherbourg	Qld	3	1199	1397	16.5
Chinchilla	Qld	3	238	293	23.0
Cloncurry - McKinlay	Qld	4	744	822	10.5
Coen	Qld	5	271	286	5.5
Cooktown	Qld	4	370	410	10.8
Dauan Island	Qld	5	133	134	0.6
Diamantina	Qld	5	73	70	-4.8
Dirranbandi	Qld	5	111	119	7.0
Doomadgee	Qld	5	1168	1320	13.0

Dysart	Qld	3	80	110	37.0
Eacham	Qld	3	301	376	24.8
Edmonton	Qld	3	2060	2507	21.7
Eidsvold	Qld	3	188	219	16.7
Emerald	Qld	3	448	533	19.0
Erub (Darnley) Island	Qld	5	360	385	7.0
Etheridge Tablelands	Qld	3	1325	1481	11.8
Flinders - Richmond - Dalrymple	Qld	5	290	328	13.0
Gayndah	Qld	3	143	192	34.3
Gladstone - South Coast	Qld	3	141	180	28.0
Goondiwindi	Qld	3	323	404	25.0
Goondiwindi - Surrounds	Qld	3	194	245	26.0
Gordonvale	Qld	3	758	906	19.6
Hammond Island	Qld	5	222	235	5.7
Herberton	Qld	3	200	252	26.0
Herberton Tablelands	Qld	3	344	415	20.6
Hinchinbrook	Qld	4	247	278	12.6
Hope Vale	Qld	4	927	1015	9.5
Horn Island	Qld	5	333	367	10.1
Iama (Yam) Island	Qld	5	305	335	9.8
Ingham	Qld	3	408	483	18.4
Injinoo	Qld	5	461	518	12.3
Innisfail	Qld	3	1348	1560	15.7
Johnstone	Qld	3	540	657	21.7
Jumbun	Qld	4	101	94	-7.0
Kowanyama	Qld	5	940	1028	9.4
Kowrowa - Mantaka - Mona Mona	Qld	3	222	260	17.3
Kubin (Moa Island)	Qld	5	164	171	4.3
Laura	Qld	5	53	47	-11.4
Lockhart River	Qld	5	439	472	7.6
Longreach	Qld	5	283	333	17.7
Mabuiag Island	Qld	5	257	288	11.9
Mackay - Surrounds	Qld	3	422	520	23.3

Manunda - Portsmith	Qld	3	4241	4904	15.6
Mapoon	Qld	5	237	244	2.9
Maranoa exc. Roma and Mitchell	Qld	4	193	215	11.2
Mareeba	Qld	3	1249	1451	16.1
Mer (Murray) Island	Qld	5	356	385	8.2
Millmerran	Qld	3	95	128	35.2
Mirani	Qld	3	237	303	27.7
Mitchell	Qld	5	158	176	11.2
Moranbah	Qld	3	233	293	25.6
Mornington	Qld	5	1002	1132	13.0
Mossman	Qld	3	325	391	20.4
Mossman - Surrounds	Qld	3	453	546	20.4
Mossman Gorge	Qld	3	96	103	7.2
Mount Garnet	Qld	4	87	81	-7.3
Mount Isa exc. Camooweal	Qld	4	3110	3524	13.3
Mount Whitfield	Qld	3	638	760	19.1
Muralag and Inner Islands	Qld	5	50	45	-10.7
Murgon	Qld	3	417	503	20.7
Murilla - Wandoan	Qld	3	115	158	37.3
Murweh	Qld	5	531	604	13.7
Napranum	Qld	5	820	907	10.6
Nebo - Clermont	Qld	4	293	337	14.9
New Mapoon	Qld	5	276	292	5.7
Normanton	Qld	5	670	734	9.6
North Burnett - Rural	Qld	3	275	341	24.0
Palm Island	Qld	4	2209	2451	10.9
Paroo	Qld	5	573	648	13.2
Pormpuraaw	Qld	5	606	653	7.7
Port Kennedy (Thursday Island)	Qld	5	889	987	11.1
Poruma (Coconut) Island	Qld	5	150	152	1.3
Proserpine - Whitsunday	Qld	3	475	584	22.9
Ravenshoe	Qld	3	165	205	24.2
Roma	Qld	3	592	704	18.9

Saibai Island	Qld	5	359	396	10.4
Sarina	Qld	3	645	780	21.0
Seisia	Qld	5	149	146	-2.0
St George	Qld	4	564	641	13.7
St Pauls (Moa Island)	Qld	5	251	260	3.4
Stanthorpe	Qld	3	248	313	26.1
Tara (Qld)	Qld	3	236	288	22.1
Thuringowa	Qld	3	4524	5388	19.1
Townsville	Qld	3	5391	6232	15.6
Townsville - Northern Beaches	Qld	3	203	257	26.4
Townsville - Southern Rangelands	Qld	3	601	651	8.3
TRAWQ (Thursday Island)	Qld	5	817	901	10.3
Trinity	Qld	3	868	1021	17.7
Tully	Qld	3	280	345	23.4
Ugar (Stephens) Island	Qld	5	50	48	-3.8
Umagico	Qld	5	270	292	8.1
Wambo	Qld	3	155	208	34.1
Warraber Island	Qld	5	223	237	6.5
Weipa	Qld	5	630	719	14.1
Winton	Qld	5	123	132	7.1
Wondai	Qld	3	242	307	26.8
Woorabinda	Qld	4	880	979	11.3
Wujal Wujal	Qld	4	253	251	-1.0
Yarrabah	Qld	3	2348	2707	15.3
Yorke Island	Qld	5	241	255	5.8
Amata - Tjurma Homelands	SA	5	451	479	6.2
Anilalya Homelands	SA	5	123	120	-2.1
Barmera	SA	3	203	250	23.2
Berri	SA	3	287	352	22.7
Ceduna	SA	5	473	522	10.4
Coober Pedy	SA	5	254	275	8.4
Copper Coast - Barunga West	SA	3	314	395	26.0
Davenport	SA	3	174	190	9.4

Dunjiba (Oodnadatta)	SA	5	116	114	-1.4
Eyre Peninsula	SA	4	260	300	15.5
Flinders Ranges	SA	3	884	1053	19.1
Iga Warta Homeland	SA	5	61	52	-15.2
Indulkana and Indulkana Homelands	SA	5	299	317	5.9
Kalka and Homelands	SA	5	70	64	-8.7
Kaltjiti (Fregon) and Irintata Homelands	SA	5	240	246	2.4
Kanpi - Nyapari - Angatja	SA	5	110	104	-5.1
Koonibba	SA	5	137	141	3.1
Lake Eyre - Lake Torrens	SA	5	213	228	6.8
Leigh Creek - Copley	SA	5	64	67	4.8
Limestone Coast	SA	3	514	615	19.6
Loxton - Waikerie	SA	3	214	280	30.7
Maralinga Tjarutja	SA	5	67	63	-6.6
Meningie	SA	3	82	106	29.0
Mimili and Mimili Homelands	SA	5	248	257	3.6
Mount Gambier	SA	3	577	710	23.1
Murray Mallee	SA	3	197	252	27.7
Pipalyatjara	SA	5	90	85	-5.3
Point Pearce	SA	3	127	151	19.0
Port Augusta - Central	SA	3	1374	1571	14.3
Port Augusta - Surrounds	SA	3	14	27	92.0
Port Augusta - West	SA	3	611	731	19.6
Port Lincoln	SA	4	781	882	12.9
Pukatja (Ernabella)	SA	5	440	476	8.2
Quorn	SA	3	132	160	21.1
Raukkan	SA	3	95	111	16.9
Renmark Paringa	SA	3	163	218	33.4
Roxby Downs	SA	4	75	83	10.8
South-West Coast	SA	5	53	52	-1.2
Stirling North	SA	3	174	209	20.0

Tjutjunaku Worka Tjuta - Inner Homelands	SA	5	174	184	5.9
Tjutjunaku Worka Tjuta - Outer Homelands	SA	5	74	80	7.5
Umoona	SA	5	22	22	1.1
Watarru and Outstations	SA	5	42	38	-10.5
Whyalla	SA	3	932	1104	18.5
Yalata	SA	5	263	281	7.0
Yorke Peninsula	SA	3	210	270	28.7
Albany - Central	WA	3	597	723	21.0
Albany - Surrounds	WA	3	475	573	20.6
Argyle	WA	5	132	148	12.5
Balgo	WA	5	459	505	10.0
Bardi (One Arm Point)	WA	5	310	351	13.1
Bayulu	WA	5	316	358	13.2
Beagle Bay	WA	5	255	284	11.3
Beverley	WA	3	71	105	47.5
Bidyadanga	WA	5	530	596	12.5
Bridgetown - Scott River East	WA	3	92	135	47.2
Brookton	WA	3	121	163	34.7
Broome - Central	WA	4	1359	1487	9.4
Broome - North	WA	4	312	351	12.7
Broome - Surrounds	WA	5	163	180	10.1
Broome Town Camps	WA	4	240	251	4.6
Burringurrah	WA	5	104	111	6.7
Carnarvon Town exc. Mungullah	WA	4	917	1041	13.6
Carnegie South exc. Mount Magnet	WA	5	202	233	15.6
Cheeditha - Mingullatharndo	WA	4	298	309	3.7
Coolgardie	WA	5	200	223	11.5
Coonana	WA	5	43	43	0.9
Cosmo Newberry	WA	5	66	65	-2.0
Dampier	WA	5	55	74	33.7
Denmark - Plantagenet	WA	3	202	271	34.0

Derby	WA	5	810	911	12.5
Djarindjin - Lombadina	WA	5	208	232	11.4
Djugerari	WA	5	79	81	2.5
Doon Doon	WA	5	89	93	3.9
East Pilbara - Surrounds	WA	5	212	234	10.3
Esperance	WA	4	442	514	16.3
Esperance - Ravensthorpe - Surrounds	WA	4	154	175	13.3
Exmouth - Ashburton - Surrounds	WA	5	360	408	13.3
Fitzroy Crossing	WA	5	212	235	10.9
Fitzroy Crossing - Surrounds	WA	5	50	49	-1.6
Fitzroy River - Surrounds	WA	5	138	144	4.1
Geraldton - Central	WA	3	1409	1676	18.9
Geraldton - North	WA	3	766	920	20.1
Geraldton - Surrounds	WA	3	683	789	15.6
Gnowangerup	WA	4	144	172	19.7
Great Sandy Desert	WA	5	89	92	3.5
Greenough - Chapman Valley	WA	3	42	75	79.0
Halls Creek - Surrounds	WA	5	380	427	12.5
Halls Creek exc.Town Camps	WA	5	613	692	12.9
Injudunna	WA	4	159	161	1.3
Irrunuytju	WA	5	131	135	2.9
Irwin - Morawa	WA	3	337	419	24.3
Jameson	WA	5	111	116	4.8
Jarlmadangah Burru	WA	5	66	69	5.2
Jigalong	WA	5	329	364	10.6
Joy Springs	WA	5	53	53	-0.6
Junjuwa	WA	5	355	394	11.0
Kalgoorlie	WA	3	1995	2342	17.4
Kalgoorlie - Dundas	WA	5	54	65	20.2
Kalumburu	WA	5	408	459	12.5
Kambalda	WA	3	98	143	46.0
Karalundi	WA	5	44	55	25.2

Karmarlinunga - Djimu Nguda - Budulah	WA	5	110	108	-1.9
Karratha	WA	4	874	992	13.5
Katanning	WA	3	384	475	23.8
Kellerberrin	WA	3	115	151	31.1
Kiwirrkurra	WA	5	201	213	5.9
Kojonup	WA	3	278	355	27.6
Kooraby	WA	5	54	57	5.9
Kulin	WA	4	171	203	18.9
Kunawarritji	WA	5	73	73	0.0
Kundat Djaru	WA	5	178	198	11.3
Kununurra exc. Town Camps	WA	4	1048	1194	13.9
Kupungarri	WA	5	57	65	13.7
Kurrawang	WA	5	77	85	10.2
Laverton - Ngaanyatjarraku - Surrounds	WA	5	332	371	11.7
Leonora	WA	5	190	215	12.9
Looma	WA	5	371	405	9.2
Manjimup	WA	3	244	328	34.5
Marble Bar - Mirtunkarra (Goodabinya)	WA	5	126	140	10.8
Mardiwah Loop - Lundja	WA	5	355	393	10.7
Meekatharra exc. Karalundi	WA	5	405	450	11.0
Menzies - Leonora - Surrounds	WA	5	120	140	16.6
Merredin	WA	3	252	328	30.3
Mindi Rardi - Kurnangki	WA	5	184	196	6.3
Mindibungu	WA	5	248	284	14.6
Minyirr - Cable Beach	WA	4	1000	1139	13.9
Mirima	WA	4	156	161	3.0
Moora	WA	3	306	379	23.9
Mount Magnet	WA	5	232	259	11.6
Mount Margaret	WA	5	88	94	6.6
Mowanjum	WA	5	301	331	10.0
Mugarinya (Yandeyarra)	WA	5	97	97	0.4

Mukinbudin	WA	4	92	109	18.4
Mulan	WA	5	137	146	6.7
Mullewa	WA	4	160	173	8.1
Muludja	WA	5	127	133	4.5
Mungullah	WA	4	185	195	5.6
Narrogin - Wagin	WA	3	502	617	23.0
Newman	WA	5	550	639	16.1
Ningia Mia	WA	3	104	122	17.6
Norseman	WA	5	104	120	15.1
Northampton	WA	4	182	210	15.2
North-East Kimberley	WA	4	189	202	7.0
North-West Kimberley	WA	5	157	171	9.2
Nullagine	WA	5	117	128	9.0
Nulleywah	WA	4	130	132	1.9
Onslow	WA	5	176	196	11.3
Outer Derby - West Kimberley	WA	5	131	154	17.5
Pandanus Park	WA	5	129	135	4.7
Papulankutja	WA	5	165	178	8.0
Paraburdoo	WA	5	139	166	19.5
Parnngurr	WA	5	120	126	4.8
Pingelly	WA	3	130	166	27.8
Port Hedland - Surrounds	WA	5	90	106	17.3
Port Hedland exc. Tjalka Brooda	WA	4	192	215	11.9
Punmu	WA	5	141	148	4.9
Quairading	WA	3	136	174	27.8
Roebourne	WA	4	334	352	5.3
Shark Bay - Coral Bay - Upper Gascoyne	WA	5	117	130	11.0
South Hedland	WA	4	1767	1965	11.2
Southern Beaches	WA	3	521	639	22.7
Tjalka Boorda	WA	4	88	83	-5.8
Tjuntjuntjarra	WA	5	181	186	2.6
Tom Price	WA	5	294	348	18.4
Wanarn	WA	5	135	141	4.5

Wangkatjungka	WA	5	172	182	5.6
Warakurna	WA	5	154	166	7.8
Warburton	WA	5	395	433	9.7
Warmun	WA	5	227	248	9.5
Warralong	WA	5	167	178	6.8
Wickham	WA	4	302	346	14.4
Wiluna	WA	5	284	317	11.4
Wyndham	WA	5	411	455	10.7
Yakanarra	WA	5	100	107	6.7
Yardgee - Nicholson Town Camps	WA	5	115	119	3.1
Yungngora	WA	5	259	285	10.0
Break O'Day	Tas	3	232	290	25.0
Burnie	Tas	3	1110	1297	16.9
Campbell Town	Tas	3	105	135	28.9
Central Highlands (Tas.)	Tas	3	110	141	27.8
Circular Head - King Island	Tas	3	968	1126	16.3
Cygnet	Tas	3	183	199	8.6
Dorset	Tas	3	220	269	22.1
Flinders	Tas	5	125	115	-7.8
Geeveston	Tas	3	67	77	15.4
George Town	Tas	3	288	361	25.4
Glamorgan - Spring Day	Tas	4	160	165	3.1
Huonville - South Cape	Tas	3	1055	1226	16.2
Kentish	Tas	3	240	282	17.4
Latrobe - Hawley Beach	Tas	3	397	457	15.1
Meander Valley	Tas	3	261	325	24.7
Southern Midlands	Tas	3	260	311	19.6
Tasman	Tas	3	111	134	21.0
Ulverstone - Penguin	Tas	3	1111	1278	15.0
Ulverstone - Penguin - Surrounds	Tas	3	139	168	21.1
Waratah	Tas	3	211	259	22.6
Wynyard	Tas	3	635	736	15.9
Zeehan - Franklin	Tas	4	307	335	9.2

Acacia-Larrakia	NT	3	83	103	24.5
Adelaide River - Coomalie	NT	4	155	162	4.2
Alawa	NT	3	187	224	19.7
Ali Curung	NT	5	487	481	-1.2
Alpurrurulam	NT	5	416	442	6.2
Alyangula	NT	5	88	100	13.9
Amanbidji (Mialuni)	NT	5	83	79	-4.9
Amoonguna	NT	4	272	278	2.2
Ampilatwatja and Outstations	NT	5	370	402	8.8
Angurugu	NT	5	792	829	4.7
Angurugu Outstations	NT	5	93	87	-6.7
Anmatjere - Surrounds	NT	5	108	98	-9.1
Anmatjere - Ti Tree	NT	5	62	55	-11.4
Anthelk Ewlpaye	NT	4	81	65	-19.4
Anthepe - New Llparpa - Tyeweretye	NT	4	176	170	-3.4
Anula	NT	3	320	385	20.3
Apatula (Finke)	NT	5	145	146	0.6
Areyonga	NT	5	220	220	-0.1
Atitjere	NT	5	167	170	2.1
Atitjere - Akarnenehe Outstations	NT	5	107	103	-3.8
Atneltyey	NT	5	23	18	-22.3
Bagot Community	NT	3	196	216	10.4
Bakewell - Rosebery - Mitchell	NT	3	729	878	20.4
Barkly Tablelands - Outstations	NT	5	109	113	3.9
Barunga	NT	5	291	312	7.1
Batchelor	NT	4	71	64	-10.2
Bees Creek - Virginia	NT	3	139	178	28.1
Belyuen	NT	4	175	171	-2.3
Binjari	NT	4	228	235	3.1
Borroloola exc. Mara - Yanyula	NT	5	359	349	-2.7
Brinkin - Nakara	NT	3	127	158	24.8
Bulla	NT	5	130	130	-0.3

Bulman - Weemol	NT	5	279	296	6.2
Canteen Creek	NT	5	199	210	5.7
Charles	NT	4	709	789	11.2
Cobourg Peninsula - Demed Homelands	NT	5	103	97	-5.5
Coconut Grove - Ludmilla	NT	3	249	291	16.8
Cox - Finniss	NT	4	119	116	-2.9
Daguragu	NT	5	191	167	-12.8
Daguragu Outstations	NT	5	18	17	-6.6
Darwin - Central	NT	3	202	225	11.3
Darwin River - Berry Springs - Southport	NT	3	149	189	26.9
Dhalinybuy	NT	5	115	111	-3.5
Douglas-Daly	NT	5	117	116	-1.2
Driver	NT	3	321	398	24.1
Elliott	NT	5	290	262	-9.5
Elliott Surrounds - Outstations	NT	5	60	57	-5.5
Elsey Roper - Surrounds	NT	5	43	40	-8.1
Emu Point	NT	5	93	94	1.0
Engawala	NT	5	135	129	-4.3
Ewyenper - Ilpeye - Irklancha	NT	4	208	204	-2.0
Fannie Bay - Parap	NT	3	286	314	9.6
Farrar - Durack - Marlow Lagoon	NT	3	340	404	19.0
Flynn	NT	4	697	777	11.4
Galiwinku	NT	5	1895	2091	10.3
Gan Gan	NT	5	69	67	-2.4
Gapuwiyak	NT	5	819	849	3.6
Gapuwiyak Outstations	NT	5	172	173	0.7
Gray	NT	3	520	609	17.2
Gumatj - Surrounds	NT	5	9	11	18.8
Gunbalanya	NT	5	1038	1101	6.1
Gunyangara	NT	5	142	141	-0.6
Haasts Bluff and Outstations	NT	5	148	149	0.7
Heavitree	NT	4	179	173	-3.3

Hermannsburg	NT	5	538	546	1.5
Howard Springs - Gunn Point	NT	3	266	319	19.8
Humpty Doo - Elizabeth Valley	NT	3	613	727	18.6
Illeuwurru - Inkawenyerre	NT	5	68	57	-16.5
Ilparpa community	NT	4	19	15	-21.4
Ilperle-Tyathe - Mt Nancy - Aper-Alwerr	NT	4	172	169	-1.7
Imangara	NT	5	91	88	-3.3
Imanpa (Mount Ebenezer)	NT	5	175	175	0.1
Inarlenge	NT	4	41	31	-24.1
Ingkerreke Outstations	NT	5	68	58	-14.3
Irrultja	NT	5	87	84	-3.4
Jabiru	NT	4	212	222	4.8
Jilkminggan	NT	5	276	303	9.6
Jingili	NT	3	208	253	21.5
Julalikari - Outstations	NT	5	194	197	1.8
Kakadu - Marrakai - Surrounds	NT	4	226	237	5.0
Kalkarindji	NT	5	276	243	-12.0
Kaltukatjara (Docker River)	NT	5	262	267	1.8
Kaltukatjara Outstations	NT	5	69	65	-5.7
Karama	NT	3	790	949	20.1
Kargaru	NT	5	19	16	-14.9
Katherine exc. Town Camps	NT	4	1738	1949	12.1
Knuckey Lagoon	NT	3	91	107	17.7
Kulaluk	NT	3	36	31	-13.5
Lajamanu	NT	5	587	610	3.9
Laramba	NT	5	234	237	1.4
Larapinta	NT	4	957	1078	12.6
Larrakeyah - The Gardens	NT	3	234	283	20.8
Laynhapuy	NT	5	311	332	6.8
Lyons - Lee Point - Leanyer	NT	3	393	475	20.9
Mabunji - Mungoorbada Outstations	NT	5	283	301	6.5
Malak	NT	3	544	654	20.3

Mamadawerre	NT	5	56	49	-12.7
Maningrida	NT	5	2038	2211	8.5
Maningrida Outstations	NT	5	270	284	5.3
Manmoyi	NT	5	61	53	-13.3
Manyallaluk	NT	5	105	101	-4.1
Mara	NT	5	192	199	3.5
Marla Marla - Village Camp	NT	5	127	123	-3.4
Marrara - Winnellie - Berrimah	NT	3	893	952	6.6
Marthakal Homelands exc. Galiwinku	NT	5	404	445	10.1
Mataranka - Mulggan	NT	5	91	86	-5.4
Miali Brumby - Warlpiri	NT	4	290	282	-2.8
Milikapiti	NT	5	406	439	8.1
Milingimbi	NT	5	1020	1082	6.1
Millner	NT	3	228	260	13.9
Milyakburra and Outstations	NT	5	156	150	-3.9
Minjilang	NT	5	270	286	5.8
Minmarama Park	NT	3	139	154	10.9
Minyerri	NT	5	444	496	11.8
Moil	NT	3	237	286	20.8
Moulden	NT	3	675	808	19.7
Mount Johns	NT	4	350	374	7.0
Mount Liebig and Outstations	NT	5	169	170	0.6
Mudginberri	NT	4	60	48	-19.9
Mutitjulu - Uluru	NT	5	249	262	5.3
Nauiyu Nambiyu	NT	5	391	424	8.5
Ngalpa Ngalpa - Wuppa	NT	5	396	416	5.0
Nganmarriyanga (Palumpa)	NT	5	343	381	11.0
Ngukurr	NT	5	973	1036	6.5
Nhulunbuy	NT	5	374	425	13.5
Nightcliff	NT	3	185	221	19.4
Nturiya	NT	5	97	92	-5.2
Numbulwar and Outstations	NT	5	627	632	0.7

Nyewente - Akngwertnarre - Namatjira	NT	4	123	108	-12.2
Nyirripi	NT	5	187	191	2.4
Palmerston Indigenous Village	NT	3	80	79	-1.8
Papunya	NT	5	381	360	-5.4
Papunya Outstations	NT	5	77	73	-5.5
Peppimenarti	NT	5	165	168	1.7
Pigeon Hole	NT	5	120	124	3.1
Pine Creek	NT	4	150	153	1.9
Pirlangimpi	NT	5	333	351	5.3
Pmara Jutunta	NT	5	193	201	4.3
Ramingining	NT	5	775	795	2.5
Ramingining - Milingimbi Outstations	NT	5	219	225	2.8
Rapid Creek	NT	3	185	224	21.2
Rittarangu	NT	5	93	93	0.5
Robinson River (Mungoorbada)	NT	5	243	262	8.0
Rockhole	NT	4	125	122	-2.5
Ross	NT	4	742	815	9.8
Santa Teresa (Ltyentye Purte)	NT	4	504	548	8.7
South MacDonnell Ranges	NT	4	821	835	1.7
Stuart Park - Bayview - Woolner	NT	3	198	237	19.9
Tanami Outstations	NT	5	23	23	-0.2
Tara	NT	5	54	50	-6.7
Tennant Creek exc. Town Camps	NT	5	1040	1144	10.0
Thamarrurr exc. Wadeye	NT	5	189	196	3.9
Timber Creek	NT	5	154	163	5.7
Timber Creek - Surrounds	NT	5	61	55	-10.1
Titjikala	NT	4	189	185	-2.1
Tiwi	NT	3	238	280	17.6
Tiwi Islands - Wilderness	NT	5	180	189	5.2
Tjuwanpa Outstations	NT	5	205	209	1.8
Umbakumba and Outstations	NT	5	429	463	8.0
Utopia - Arawerr - Arlparra	NT	5	483	523	8.3

Wadeye	NT	5	1929	2139	10.9
Wagaman	NT	3	145	179	23.4
Walangeri Outstations	NT	5	48	46	-4.6
Wallace Rockhole	NT	5	65	58	-11.5
Walungurru (Kintore)	NT	5	415	446	7.6
Wanguri	NT	3	155	191	23.4
Warruwi	NT	5	395	439	11.0
Willowra	NT	5	204	214	5.0
Wilora	NT	5	111	105	-5.8
Woodroffe	NT	3	438	532	21.4
Wugular (Beswick)	NT	5	500	540	8.1
Wulagi	NT	3	266	326	22.4
Wurrumiyanga (Nguiu)	NT	4	1358	1435	5.7
Wutunugurra	NT	5	200	214	6.8
Yanyula	NT	5	187	195	4.4
Yarralin	NT	5	252	260	3.2
Yarrenyty-Arltere	NT	4	90	75	-17.0
Yilpara	NT	5	125	119	-4.5
Yirara College and Surrounds	NT	4	120	140	16.4
Yirrkala	NT	5	652	651	-0.2
Yuelamu	NT	5	197	202	2.6
Yuendumu and Outstations	NT	5	587	585	-0.4
Yugul Mangi Outstations	NT	5	29	26	-9.3