SMALL AREA ESTIMATION
USING ROBUST EXTENSIONS TO AREA LEVEL MODELS

THEORY, IMPLEMENTATION AND SIMULATION STUDIES

SEBASTIAN WARNHOLZ

Submitted in fulfillment of
the requirements for the degree
Dr. rer. pol.
School of Business & Economics
Freie Universität Berlin

April, 2016

SUPERVISORS:
Prof. Dr. Timo Schmid
Prof. Dr. Nikos Tzavidis

LOCATION:
Berlin

TIME FRAME:
April, 2016

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

I have gone so far as to say if you (or someone else)
aren't writing software for your methods,
they don't really exist.

— Leek (2013)

1.1 MOTIVATION

The demand for reliable small area statistics from sample surveys has grown substantially over the past decades due to their growing use in public and private sectors. The importance of the field of Small Area Estimation (SAE) can be explained by the increasing need for reliable estimates by policy makers and in official statistics. Results may be used for fund allocation, health programs, agriculture, or poverty mapping to name only a few of the fields of application. Traditionally such estimates have relied on survey data; but as the target domains become more diverse, reliable estimates call for an increasing demand for sampled units within the domains. The conflict between the demand for more diverse domains and the cost and feasibility of larger samples is the factor that stimulates the progress within the field; it supplies the mechanism for optimising the ratio between sampled units and the reliability of estimates.

It is the main endeavour of Small Area Estimation to produce *reliable predictions* of a *target statistic* for *small domains*. A *target statistic* can be a statistic such as a mean, a count, or quantiles; but it can take other forms: e.g. an inequality measure like the Gini coefficient for poverty mapping. Such statistics are produced for *small domains* where domain refers to specific groups, e.g. an industry sector, or groups defined by socio-economic characteristics. Because of its frequent application to administrative data, domains are often defined by areas as a geographical unit. They are small in the sense that they comprise few or no sampled units within these domains. This has the effect that a direct estimation, i.e. an estimation which only relies on the information available within domains, tends to be unreliable. *Reliability* is here measured either by the variance or mean squared error of the predictions (MSPE).

Small Area Estimation tries to improve such domain predictions – often in terms of mean squared error – by *borrowing strength* from other domains. This can happen in taking additional information from other data sources into account, like census and register inform-

ation. Also structures in the data, like spatial or temporal correlation, can be exploited to improve predictions.

Since applications in the SAE field are often related to official statistics, the incorporation of spatial and temporal correlation structures is highly relevant. Such applications have often a geographical dimension since predictions by official statistics are often produced on a regional basis. The Nomenclature of units for territorial statistics (NUTS) (Eurostat, 2015) is a hierarchical system to define regions in the European Union in different granularity. Since such classifications are based on a geographical dimension we may take advantage of the fact that neighbouring districts may have similarities – hence the incorporation of spatial information.

Official statistics is often interested in producing up-to-date results and this can mean that information over a time frame is available since these predictions may be produced, for example, on an annual basis. Such information can be exploited and may have a significant and overall positive impact on the precision of predictions.

Pratesi and Salvati (2008) exemplify this with predictions based on the Survey on Life Conditions in Tuscany, Italy. They aim at predicting the mean per capita income within municipalities. They found that incorporating the unobserved spatial correlation between domains into their predictions improved the precision of these quantities. A combination of incorporating spatial as well as temporal correlation can be found in Marhuenda et al. (2013). They make predictions for two poverty indicators based on the Survey on Income and Living Conditions for Spain. Their results suggest that the incorporation especially of the correlation over time presents a beneficial effect in terms of precision.

Methods using spatial and temporal correlation structures are being used more and more routinely in applications. This is also indicated by the availability of their software implementations which can, for example, be found as packages in the R-language (R Core Team, 2015). Molina and Marhuenda (2015) provide implementations for commonly used methods in the SAE field including models incorporating spatial and temporal correlation.

These methods often rely on strong distributional assumptions which provide the additional advantage of a gain in precision; however such methods can easily be influenced by single observations. Such observations may be framed as *outlying*. Hence it can be advantageous to use methods which can be assumed to be more robust against the influence of such outlying observations. In this context this Thesis aims at combining robust estimation methodology with the use of various spatial and temporal correlation structures. The remaining part of this Introduction provides an overview of how this Thesis is framed within the SAE field – Section 1.2 – and presents more precisely its mechanics – Section 1.3.

1.2 LOCATING THE CONTENT WITHIN THE FIELD OF SMALL AREA ESTIMATION

In the following I give a general overview of the field of Small Area Estimation to the extent necessary to accommodate this Thesis within the SAE field. For a general overview of the field, Rao (2003) as well as Rao and Molina (2015) give comprehensive overviews of the established methods and the published research. Ghosh and Rao (1994), Rao (1999), Pfeffermann (2002), and Pfeffermann (2013) focus on the status quo and main lines of discussion within the field at their respective point in time.

In general small area methods may be divided into two categories: design based and model based methods. This classification may not be distinct but provides a frame for discussion. Design based methods can be considered the traditional methodology for analysing survey data; a comprehensive overview of these methods for SAE can be found in Lehtonen and Veijanen (2009). Design based methods comprise different direct and indirect techniques. For example the Horvitz-Thompson (HT) estimator of Horvitz and Thompson (1952) uses only sampled units within domains; synthetic regression estimates and model assisted methods like generalized regression (GREG) estimators are other examples of such estimators – see Särndal et al. (1992) for a discussion of these methods. These methods have in common that they incorporate information of the sampling design into the estimation.

Conceptually design based and model based methods differ in that design based methods are used to optimally estimate a target parameter of a fixed and finite population. Model based methods rely instead on the idea that an observed sample is drawn from a population which is but one possible realisation of a *superpopulation* model, and it is the parameters of that superpopulation which are targeted. This difference leads to a trade-off when choosing between methods: model based methods can improve domain predictions in terms of variance even with small samples; however they cannot be considered design unbiased. Design based methods on the other hand are design unbiased but have larger and possibly unacceptably high variances for small samples – see Lehtonen and Veijanen (2009).

Model based methods can be further divided into area and unit level models. Observations which can be associated with a specific domain are referred to as units. These can be companies within an industry sector or individuals within a municipality. The area level describes models which use information on area level, i.e. direct estimates for domains. A situation in which these models are considered is when data can only be provided as aggregates due to factors such as confidentiality. Also such methods may be useful in situations in

which the computational effort is high – e.g. when complex variance structures are combined with large data sets.

One class of models in particular is favoured in different variations: the mixed models. The Fay-Herriot (FH) model introduced by Fay and Herriot (1979) and the Battese-Harter-Fuller (BHF) model which was introduced by Battese et al. (1988) are the two basic models which are used respectively for area and unit level estimations. Underlying this is the idea to use auxiliary information in a regression to estimate a global conditional mean and add an extra component to capture the domain specific difference from that global mean. This general idea can be found in combination with different estimation methodologies, i.e. general linear mixed models which are typically associated with best linear unbiased predictors (BLUPs), empirical Bayes, and hierarchical Bayes. Although these different frameworks for estimation differ in respect of optimality criteria, the equivalence of the derived estimators can be shown for special cases. A more general discussion of mixed models in SAE can be found in Jiang and Lahiri (2006). Rao (2003) and Rao and Molina (2015) provide a comprehensive overview and comparison of the different frameworks.

A general property of model based methods is that a lot of their benefits in terms of efficiency rely on strong distributional assumptions. Hence it is not only in the field of SAE that robust methods have been exploited to reduce the negative effect of a potential violation of these assumptions. The general problem here is that single observations can have unwanted and overly large impact on results. Such observations are typically called outliers. Chambers (1986) uses the term *representative* outliers to describe observations which are correctly recorded and can not be assumed to be unique in the population. Non-representative outliers, on the other hand, may be best described as *not correctly recorded* and should be imputed or generally dealt with during the editing process of survey data.

To summarise robust methods in SAE, it is necessary to distinguish between three different lines of discussion. Firstly, if the distributional assumption – often a Gaussian distribution – appears to be implausible then intuition demands that it be replaced. This often leads to the use of non-symmetric or heavy-tailed distributions for the model error or the random effect. Due to their flexibility Bayesian modelling strategies are often used in this context; see for example Datta and Lahiri (1995) and Bell and Huang (2006). Secondly, methods are applied which are *naturally* more robust against outlying observations. Chambers and Tzavidis (2006) and Tzavidis, Marchetti et al. (2010) model a global conditional median, or more generally a quantile, instead of a mean. The third approach is to remain with the original model or method and *robustify* the estimation equations. In this context Sinha and Rao (2009) develop a robust EBLUP; Beaumont and Rivest (2009) refer to a winzorisation of the Horvitz-Thompson

estimator; and Beaumont (2004) introduces a robust extension to generalised regression estimation.

Given this background, in this Thesis I introduce extensions to the Fay-Herriot area level model. More precisely an EBLUP based approach is taken to derive predictions in a way that makes it possible to model spatial and temporal covariance structures in the random effects. These are the methods introduced by Rao and Yu (1994) for incorporating temporal correlation; and by Pratesi and Salvati (2008) for incorporating spatial correlation; and by Marhuenda et al. (2013) who introduce the combination of these two. In contrast to these methods, in this Thesis the estimation procedure is based on robust methodology for deriving area level robust predictions. The approach is based on the methodology of Sinha and Rao (2009) who have extended the approach of Richardson and Welsh (1995) to the SAE field. This will lead (i) to an area level robust EBLUP (REBLUP) which is simply a special case of using the results of Sinha and Rao (2009); (ii) to an area level spatial REBLUP (SREBLUP) which – to contrast it from Schmid (2011) – is based on an area level model instead of a unit level model; (iii) to a temporal REBLUP (TREBLUP); and (iv) to a spatio-temporal REBLUP (STREBLUP).

Some extensions introduced in the literature around REBLUPs have focused on unit level models; thus results especially for MSPE estimation and bias correction are extended to area level models. In this regard the parametric bootstrap by Sinha and Rao (2009) is here adapted to obtain estimates for the domain specific MSPE. Furthermore an analytical solution is presented which is based on a pseudolinear form of the area level REBLUPs. This approach extends the results of Chambers, Chandra and Tzavidis (2011) to area level models and combines them with the results of Chambers, Chandra, Salvati et al. (2014) for robust predictions. In addition I present a simple correction for the bias associated with robust predictions. This correction is based on the *limited translation estimator* of Efron and Morris (1972) and has already been used by Fay and Herriot (1979) with a somewhat similar goal.

1.3 OUTLINE

The overall structure of Thesis is divided into three parts. Part i comprises the methodological foundation underpinning the proposed REBLUPs. In Part ii I will introduce two software packages which have been developed alongside this Thesis. Then in Part iii the statistical properties of the predictions based on the proposed REBLUPs are investigated in several model and design based simulation studies.

Part i includes a literature review in Chapter 2 which gives the methodological background underpinning the results and extensions around the robust extensions to area level models in Chapter 3. The

review includes the basic foundation in the form of direct domain estimation in Section 2.1 and predictions using linear mixed models in Section 2.2. Framed as linear mixed models the FH model and the spatial and temporal extensions are reviewed in Section 2.3. Then the robust methodology around robust predictions under linear mixed models in the SAE field is reviewed in Section 2.5. These results are often based on unit level models; hence the BHF model and unit level notation is introduced in 2.4.

Based on these results I establish a coherent framework for robust predictions using spatial and temporal area level models in Chapter 3. These predictions are framed as area level robust best linear unbiased predictions. The models under consideration are merely special cases in this framework and results are given in general notation for linear mixed models where possible. This includes the framework presented in Section 3.2, the bias correction in Section 3.4, and the MSPE estimators in Section 3.5. The main motivation to consider robust area level models are outlying areas; however different kinds of outlying observations can be defined. A discussion of what *outliers* are with respect to area level models is given in Seciton 3.1 as it presents the main motivation for the sebsequent presented framework.

In Part ii I continue the presentation of robust area liner mixed models; however the presentation now changes to the implementation side of these methods. The introduced methods of Chapter 3 are implemented in the R-package `saeRobust` (Warnholz, 2016) which is provided as supplementary material to this Thesis. In Chapter 4 the numerical properties of the proposed estimators are investigated. Furthermore code examples are given to illustrate the possibility to use these methods in a data analysis. The results in Part iii are based on simulation studies and to that extent a simulation framework and its implementation are presented in Chapter 5. Here a simulation study is framed as a sequence of data manipulation steps and this idea is implemented in the R-package `saeSim` (Warnholz and Schmid, 2016). The aim of the package is to provide tools and most importantly a convention how to conduct simulations in the SAE field. In that I hope to promote the publication of source code and simplify the process of understanding it – at least with respect to simulation studies in this field.

In the final Part iii model and design based simulation studies are conducted. The model based studies aim at explaining how the robust predictions differ from their non-robust counterparts – Chapter 6. Here different settings are investigated: Section 6.1 investigates the performace of the robust predictions in area level scenarios, and in Section 6.2 the performance of the MSPE estimators is the subject matter. Both of which rely on the generation of area level data, thus in this setting only area level outliers are considered. This perspective is then changed in Section 6.3 where the data is simulated beginning

with a unit level population model. This opens up the possibility to incorporate also unit level outliers into the discussion. The discussion is then continued in Chapter 7 by providing results for a design based simulation. Here the target is the tax-turnover in 20 industry domains in the Netherlands. This analysis is based on a synthetic unit level population which now presents the interesting combination of an informative sampling design, unit level outliers, and area level outliers. A summary of the main findings and contributions of this Thesis as well as avenues for further research are then presented in Chapter 8.

1.4 REPRODUCING THE RESULTS

Some effort went into the possibility to reproduce the results presented in this Thesis. All outcome of the simulation studies are produced using the R-language (R Core Team, 2015) which is freely available. Two R-packages are provided as supplementary material to this Thesis: `saeRobust` and `saeSim` which are subject to discussion later in this Thesis. The package `saeRobust` implements the statistical methods here developed and `saeSim` provides a simulation framework which is utilised for implementing the studies below.

To provide the option to reconcile the results with their source code, the files implementing the simulation studies are also part of the supplementary material. Running these files is possible however the computations should expected to take a considerable amount of time. This can be speeded up by lowering the number of iterations in simulations in order to make the computation more transparent. To set up the computational environment inside the R-language the file `00Dependencies.R` can be used; in that file all installation instructions necessary are listed since various additional packages have been used – e.g. for data manipulation and graphics.

It should be possible to reproduce all results related to model based simulations. For reproducing the design based simulation study in Chapter 7 it would be necessary to deliver the synthetic population here used; however the copyright does not grand me permission in order to do so. The source file implementing the simulation is included nonetheless.

Part I

THEORY

2 SMALL AREA ESTIMATION

In 2002, small area estimation (SAE) was flourishing both in research and applications, but my own feeling then was that the topic has been more or less exhausted in terms of research and that it will just turn into a routine application in sample survey practice. As the past 9 years show, I was completely wrong.

— Pfeffermann (2013)

2.1 DIRECT ESTIMATORS

In this Chapter a general overview of the SAE field is provided. Here I begin with a simple direct estimator and notation which is then used throughout the Thesis. The area models later introduced are all based on some direct estimator. This direct estimator can be any statistic; however I implicitly often refer to it as an estimator for the population mean. This may be plausible when we keep in mind that the context of this Thesis is to study outlier robust methods; and in this context it may be most intuitive when we consider the population mean and a simple estimator as a starting point. Lehtonen and Veijanen (2009) give a more comprehesive overview of direct estimators and design based methods.

Let a population be denoted by U, consisting of N units, which can be devided into D distinct domains or areas $U_1 \cup \cdots \cup U_D$. Each area is of size N_i, with $i = 1, \ldots, D$ denoting area i, such that $N = \sum_{i=1}^{D} N_i$. S denotes a sample from U which in turn can be devided into $S_1 \cup \cdots \cup S_D$, where n is the sample size of S and n_i of S_i such that $n = \sum_{i=1}^{D} n_i$.

Let $\mathbf{y}$ define the characteristic of interest and y_{ij} the response value for unit j with $j = 1, \ldots, N_i$ in area i. Furthermore let the target quantity be the population mean in area U_i defined by $\theta_i = \bar{Y}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} y_{ij}$. Note that the target quantity is not necessarily a mean but can be any statistic in the population.

Given that the sample is drawn with simple random sampling without replacement (SRSWOR), a design unbiased estimator is given by

$$\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij} \tag{2.1}$$

with its design variance $\mathbb{V}(\bar{y}_i | n_i) = (S_i^2/n_i)\,(1 - n_i/N_i)$ where $S_i^2 = 1/(N_i - 1) \sum_{j=1}^{N_i} \left(y_{ij} - \bar{Y}_i\right)^2$. To estimate the sampling variance, S_i^2

can be replaced by s_i^2 being an estimator using the sampled data, i.e. $s_i^2 = 1/(n_i - 1) \sum_{j=1}^{n_i} \left(y_{ij} - \bar{y}_i\right)^2$. Note that n_i is random under SRSWOR such that specific areas can have very few or even no sampled units.

2.2 SMALL AREA PREDICTION USING LINEAR MIXED MODELS

This Section gives a general overview of linear mixed models and the best linear unbiased prediction (BLUP) and empirical BLUP (EBLUP). The unit and area level models introduced later in this Chapter are based on this class of models and the robust methodology is based on a *robustified* EBLUP. The original model dates back to Henderson (1950) and an early comprehensive overview can be found in Searle (1971). Jiang and Lahiri (2006) review liner mixed models in the context of SAE and the main results can also be found in Rao (2003) and Rao and Molina (2015).

2.2.1 *Linear Mixed Models*

Following Rao and Molina (2015, 98 ff) a linear mixed model can be expressed by:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \tag{2.2}$$

where $\mathbf{y}$ is the $(n \times 1)$ vector of response values; $\mathbf{X}$ is a $(n \times P)$ matrix containing deterministic auxiliary information; $\boldsymbol{\beta}$ is the $(P \times 1)$ vector of regression coefficients; $\mathbf{Z}$ is a known matrix and $\mathbf{u}$ is a vector of random effects, such that $\mathbf{Z}\mathbf{u}$ is of dimension $(n \times 1)$; $\mathbf{e}$ is the $(n \times 1)$ vector of model errors. Note that $\mathbf{u}$ and $\mathbf{e}$ are both random variables where the basic assumption is that both have mean zero and finite variances. Furthermore they are assumed to be independent.

If, in addition, $\mathbf{u}$ and $\mathbf{e}$ are assumed to follow a normal distribution, the model is called a Gaussian Linear Mixed Model (Jiang and Lahiri, 2006). The distribution of $\mathbf{y}$ can then be derived as a multivariate normal of the form:

$$\begin{aligned} \mathbf{y} &\sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \mathbf{V}) \\ \mathbf{y}|\mathbf{u} &\sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}, \mathbf{V}_e) \end{aligned}$$

where $\mathbf{V} = \mathbf{Z}\mathbf{V}_u\mathbf{Z}^\top + \mathbf{V}_e$ with $\mathbf{V}_u$ and $\mathbf{V}_e$ being the variance matrices of $\mathbf{u}$ and $\mathbf{e}$, respectively. Such variance structures typically depend on some unknown dispersion parameters. To be more precise: $\mathbf{V}_u = \mathbf{V}_u(\boldsymbol{\delta}_u)$ and $\mathbf{V}_e = \mathbf{V}_e(\boldsymbol{\delta}_e)$ such that $\mathbf{V} = \mathbf{V}(\boldsymbol{\delta})$ with $\boldsymbol{\delta} = (\boldsymbol{\delta}_u, \boldsymbol{\delta}_e)$.

2.2.2 *Best Linear Unbiased Prediction*

Here I also follow the results given by Rao and Molina (2015, 98 ff). Given the model (2.2) above in SAE problems we are generally interested in estimating the expected value of $\mathbf{y}$ given $\mathbf{u}$:

$$\boldsymbol{\mu} = \mathbf{l}^\top \boldsymbol{\beta} + \mathbf{m}^\top \mathbf{u}$$

for specified values of $\mathbf{l}$ and $\mathbf{m}$. An estimator for $\boldsymbol{\mu}$ can be obtained by replacing $\boldsymbol{\beta}$ and $\mathbf{u}$ with suitable estimators. For known variance components, $\boldsymbol{\delta}$, the best linear unbiased estimator (BLUE) is given by:

$$\tilde{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}}(\boldsymbol{\delta}) = \left(\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{V}^{-1} \mathbf{y} \tag{2.3}$$

and the BLUP for $\mathbf{u}$ by:

$$\tilde{\mathbf{u}} = \tilde{\mathbf{u}}(\boldsymbol{\delta}) = \mathbf{V}_u \mathbf{Z}^\top \mathbf{V}^{-1} \left(\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}\right) \tag{2.4}$$

such that the BLUP estimator for $\boldsymbol{\mu}$ can be stated as:

$$\tilde{\boldsymbol{\mu}} = \tilde{\boldsymbol{\mu}}(\boldsymbol{\delta}) = \mathbf{l}^\top \tilde{\boldsymbol{\beta}} + \mathbf{m}^\top \tilde{\mathbf{u}}. \tag{2.5}$$

The BLUP (2.5) of $\boldsymbol{\mu}$ depends on known variance components $\boldsymbol{\delta}$. These values are typically unknown in applications and are themselves subject to estimation. If we replace $\boldsymbol{\delta}$ with a suitable estimator, $\hat{\boldsymbol{\delta}}$, the empirical BLUP (EBLUP) obtained is:

$$\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}(\hat{\boldsymbol{\delta}}) = \mathbf{l}^\top \hat{\boldsymbol{\beta}} + \mathbf{m}^\top \hat{\mathbf{u}} \tag{2.6}$$

where $\hat{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}}(\hat{\boldsymbol{\delta}})$ and $\hat{\mathbf{u}} = \tilde{\mathbf{u}}(\hat{\boldsymbol{\delta}})$. To estimate $\boldsymbol{\delta}$ a variety of estimators have been proposed. Commonly used estimators are based on maximum likelihood (ML) and restricted maximum likelihood (REML). For a detailed discussion of the estimation of the variance parameters see also Jiang and Lahiri (2006) and the literature quoted there.

2.2.3 *Mean Squared Prediction Error*

One of the main reasons for relying on small area methods is to reduce the mean squared error of domain predictions. Since domain predictions under a linear mixed model are derived as the EBLUP, we are generally interested in the mean squared prediction error (MSPE) of the EBLUP. Note that here I present results directly for the EBLUP instead of the BLUP since the latter has little practical relevance and the results are mainly needed to give a comprehensive context in which the existing literature can be extended. In general the estimation of the MSPE can be identified as one of the challenging problems in model based SAE (Pfeffermann, 2013). Two approaches are taken

in the literature: The analytical identification of the MSPE and different resampling strategies. More detailed reviews of these strategies can be found in Rao (2003, p. 95 ff), Jiang and Lahiri (2006), and in Datta (2009).

Early results can be found in Kacker and Harville (1984) who propose an approximation to the MSPE of the EBLUP of a Gaussian linear mixed model. In particular they show that:

$$\mathrm{MSPE}(\hat{\boldsymbol{\mu}}) = \mathrm{MSPE}(\boldsymbol{\mu}) + \mathbb{E}\left(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}\right)^2. \tag{2.7}$$

The MSPE of μ can be decomposed such that

$$\mathrm{MSPE}(\boldsymbol{\mu}) = g_1(\boldsymbol{\delta}) + g_2(\boldsymbol{\delta})$$

where

$$g_1(\boldsymbol{\delta}) = \mathbf{m}^\top \left(\mathbf{V}_u - \mathbf{V}_u \mathbf{Z}^\top \mathbf{V}^{-1} \mathbf{Z} \mathbf{V}_u\right) \mathbf{m}$$
$$g_2(\boldsymbol{\delta}) = \mathbf{d}^\top \left(\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X}\right) \mathbf{d}$$

with $\mathbf{d}^\top = \mathbf{l}^\top - \mathbf{b}^\top \mathbf{X}$ and $\mathbf{b}^\top = \mathbf{m}^\top \mathbf{V}_u \mathbf{Z}^\top \mathbf{V}^{-1}$ (Rao, 2003, 98 ff). The second term in equation (2.7) was approximated by Kacker and Harville (1984) using a Taylor series approximation. Prasad and Rao (1990) have used instead the approximation by

$$\mathbb{E}\left(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}\right)^2 \approx g_3(\boldsymbol{\delta})$$

where

$$g_3(\boldsymbol{\delta}) = \mathrm{tr}\left(\frac{\partial \mathbf{b}^\top}{\partial \boldsymbol{\delta}} \mathbf{V} \left(\frac{\partial \mathbf{b}^\top}{\partial \boldsymbol{\delta}}\right) \mathbf{V}_{\hat{\delta}}\right)$$

with $\mathbf{V}_{\hat{\delta}}$ being the asymptotic covariance matrix of $\hat{\boldsymbol{\delta}}$. The derived MSPE depends on the unknown parameter vector $\boldsymbol{\delta}$. Prasad and Rao (1990) used a moment estimator for $\boldsymbol{\delta}$ and replaced the above formulae such that an estimator of the MSPE can be defined as $\widehat{\mathrm{MSPE}}(\hat{\boldsymbol{\mu}}) = g_1(\hat{\boldsymbol{\delta}}) + g_2(\hat{\boldsymbol{\delta}}) + g_3(\hat{\boldsymbol{\delta}})$. Datta and Lahiri (2000) extended this approach for a wider range of models in SAE including ML and REML estimators. An overview and comparison of these methods can be found in Datta, Rao et al. (2005); their presentation focuses on MSPE estimators using area level models.

A different approach has been taken by Chambers, Chandra and Tzavidis (2011) which is to define the EBLUP as a weighted sum of the sampled values and to derive an MSPE estimator under the assumption of independence between weights and sampled values. An advantage of this method is its wide applicability as the approach is not restricted to predictions under linear mixed models but extends to any predictor which can be represented as a weighted sum

of the sampled values. The approach has been extended in Chambers, Chandra, Salvati et al. (2014) to robust methods in SAE and is of special interest in deriving an MSPE estimator for the methods proposed in this Thesis. Hence these results are reviewed in more detail in Section 2.5.3.2.

As an alternative approach a wide range of different resampling strategies have been proposed. Jiang, Lahiri and Wan (2002) introduce a jackknife method to estimate the MSPE in the context of longitudinal linear mixed models; this was modified by Lohr and Rao (2009) into a simpler form. Also important is the proposed double bootstrap method of Hall and Maiti (2006) and the block bootstrap of Chambers and Chandra (2013) for mixed linear models. However of special interest are the methods in the context of robust predictions under linear mixed models which is why the bootstrap methods of Sinha and Rao (2009), Jiongo and Nguimkeu (2014), and Mokhtarian and Chambers (2013) are reviewed in more detail in Section 2.5.3.1.

2.3 AREA LEVEL MODELS

In SAE mixed models are generally devided into area and unit level models. In this Section a review of some of the results is given with respect to area level models. Section 2.4 then presents the basic unit level model.

2.3.1 *The Fay-Herriot Model*

The basic area level model was introduced by Fay and Herriot (1979) and has been used to predict the mean income of small areas using census data. The general backdrop is that only information on the area level is available, i.e. the direct estimates for the domains. The model is then built around two stages. The first stage is the sampling model:

$$\tilde{y}_i = \theta_i + e_i$$

where $\tilde{y}_i$ is a direct estimator for a statistic of interest, θ_i, for an area i with $i = 1, \dots, D$. The sampling error e_i is assumed to be independent and normally distributed with known variances σ^2_{ei}, i.e. $e_i \sim \mathcal{N}(0, \sigma^2_{ei})$. The model is modified with a second stage, the linking model, by assuming a linear relationship between the true area statistic, θ_i, and some deterministic auxiliary variables $\mathbf{x}_i$:

$$\theta_i = \mathbf{x}_i^\top \boldsymbol{\beta} + u_i$$

where $\mathbf{x}_i$ is a $(P \times 1)$ vector containing area level information for P variables and $\boldsymbol{\beta}$ is a $(P \times 1)$ vector of regression coefficients. The model errors u_i are assumed to be independent and identically distributed following a normal distribution: $u_i \sim \mathcal{N}(0, \sigma^2_u)$. Furthermore

e_i and u_i are assumed to be independent. Combining the sampling and the linking model thus leads to:

$$\tilde{y}_i = \mathbf{x}_i^\top \boldsymbol{\beta} + u_i + e_i. \tag{2.8}$$

2.3.1.1 *Best Linear Unbiased Prediction*

Obtaining small area predictions under the model in (2.8) can be defined as a linear mixed model from which a BLUP and EBLUP can be derived. Basically the model in (2.8) can be viewed directly as a linear mixed model as it was introduced in equation (2.2) where $\mathbf{Z} = \mathbf{I}_D$, $\mathbf{V}_u = \sigma_u^2 \mathbf{I}_D$ and $\mathbf{V}_e = \text{diag}(\{\sigma_{ei}^2\}_{i=1}^D)$ with $\mathbf{I}_D$ being a $(D \times D)$ identity matrix. The vector of unknown variance parameters in this case is a scalar, such that $\boldsymbol{\delta} = \sigma_u^2$ since σ_{ei}^2 is assumed to be known. The BLUP under the Fay-Herriot model, defined in equation (2.5), can then be obtained by setting $\tilde{\mu}_i = \tilde{\theta}_i^{FH}$, $\mathbf{l}_i^\top = \mathbf{x}_i^\top$, $\mathbf{m}_i^\top = 1$ and $y_i = \tilde{y}_i$:

$$\begin{aligned}
\tilde{\theta}_i^{FH} = \tilde{\theta}_i^{FH}(\sigma_u^2) &= \mathbf{x}_i^\top \tilde{\boldsymbol{\beta}} + \tilde{u}_i \\
&= \mathbf{x}_i^\top \tilde{\boldsymbol{\beta}} + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_{ei}^2} \left(\tilde{y}_i - \mathbf{x}_i^\top \tilde{\boldsymbol{\beta}} \right) \\
&= \gamma_i \tilde{y}_i + (1 - \gamma_i) \mathbf{x}_i^\top \tilde{\boldsymbol{\beta}}
\end{aligned}$$

with $\gamma_i = \sigma_u^2 / (\sigma_u^2 + \sigma_{ei}^2)$. The BLUP depends on the variance parameter of the random effects, σ_u^2, which is unknown. To obtain the EBLUP under the Fay-Herriot model we can replace the unknown parameter with an estimate, leading to:

$$\hat{\theta}_i^{FH} = \hat{\gamma}_i \tilde{y}_i + (1 - \hat{\gamma}_i) \, \mathbf{x}_i^\top \hat{\boldsymbol{\beta}} \tag{2.9}$$

where $\hat{\gamma}_i = \hat{\sigma}_u^2 / (\hat{\sigma}_u^2 + \sigma_{ei}^2)$. Note that the regression parameters are still estimated using the weighted least squares estimator of equation (2.3) with $\delta = \hat{\sigma}_u^2$.

For the estimation of σ_u^2 different approaches exist. Fay and Herriot (1979) propose a moment estimator from which they derive an algorithm to estimate σ_u^2. Rao (2003, pp. 118-119) reviews several other ideas. Also based on a moment estimator he derives the estimator also used in Prasad and Rao (1990). Both moment estimators have the property that they do not rely on a normal distribution, which is also true for the estimation of the regression coefficients. Alternatively σ_u^2 can be estimated using maximum likelihood or restricted maximum likelihood, which, in contrast, relies on the distributional assumptions. For details see also Datta and Lahiri (2000).

2.3.1.2 *Mean Squared Prediction Error*

The MSPE of the EBLUP under the Fay-Herriot model, $\hat{\theta}_i^{FH}$, is the subject of several studies. However it needs to be noted that interestingly Fay and Herriot (1979) did not assess the quantification of

uncertainty associated with their predictions. In principle the MSPE can be defined for an EBLUP as it was discussed in Section 2.2.3. Datta, Rao et al. (2005) study the MSPE estimation using the results of Prasad and Rao (1990) for different estimations of the variance component σ_u^2. They compare the method of moment estimator of Prasad and Rao (1990) with the original estimator of Fay and Herriot (1979) and the maximum likelihood estimator proposed by Datta and Lahiri (2000). The main finding is that for the MSPE estimation using the predictor by Fay and Herriot (1979) performs the best overall.

A jackknife MSPE estimator for linear mixed models was introduced by Jiang, Lahiri and Wan (2002) and later subjected to several refinements; see for example Chen and Lahiri (2003). Chen and Lahiri (2008) then introduced a jackknife estimator based on the results of Jiang, Lahiri and Wan (2002) and explicitly targeted the MSPE estimation of the prediction under a Fay-Herriot model. They found satisfying results for the MSPE estimation using their method; however the conclusion with respect to the estimation of σ_u^2 is not as clear as in Datta, Rao et al. (2005), where they compare a method of moments estimator to the method proposed by Fay and Herriot (1979).

A different line of discussion is stimulated by the fact that it is assumed that the sampling variances, σ_{ei}^2, are known parameters. In practice this is not the case and these parameters are estimated using the sample data. This can mean that they are themselves direct estimators; but if a direct mean is considered unreliable then its variance estimation cannot be considered reliable. Fay and Herriot (1979) use generalised variance functions – see Wolter (2007, pp. 272 ff) for a discussion of these methods – instead of direct estimators. Maiti et al. (2014) suggest instead to shrink both means and variances to account for the possibility of unstable direct variance estimates and also provide an estimator of the MSPE of the predictions. This approach is based on a Bayesian modeling strategy. You and Chapman (2006) provide results for the case that direct variance estimates are used in a hierarchical Bayes approach and can account for that extra variability in the MSPE estimation. MSPE estimators for an EBLUP based prediction using estimated sampling variances can be found in Wang and Fuller (2003) and in Rivest and Vandal (2003). Wang and Fuller (2003) derive an MSPE using asymptotic properties of the EBLUP. Rivest and Vandal (2003) on the other hand extend the results of Prasad and Rao (1990) and add an extra term to the MSPE estimator to account for the additional variability associated with the estimation of direct sampling variances.

2.3.1.3 *Discussion*

From a practical point of view the assumption of known sampling variances under the model is not plausible. Here these variances, though subject to estimation, are treated as known constants. Some

approaches to deal with this problem have been reviewed in Section 2.3.1.2 because in principle this relates mostly to an underestimation of the true uncertainty of the predictions. Another dimension which was for example addressed by Maiti et al. (2014) is the instability of predictions when very heterogenous sampling variances are observed, and to deal with this by shrinking both means and variances. A further approach is to stabilise the sampling variances by using generalised variance functions or other smoothing techniques. Albeit these parameters are assumed to be known they can have a large impact on the validity of domain predictions. In Section 3.1 I show how the robust FH model relates to this discussion.

The response variable, $\tilde{y}_i$, denotes a direct estimator. This is, of course, not necessarily the sample mean but can be any other statistic. An important feature of this statistic is that it is design unbiased. So in principle it can be a direct design based estimator such as the HT estimator. However it is assumed that the sampling errors are independent. This can be a plausible assumption under simple random sampling but is not necessarily valid under an informative sampling design. With respect to the response variable it also needs to be noted that quite often it is not the direct estimator itself but a suitable transformation that is used. Fay and Herriot (1979) log-transform the direct estimator and suggest the use of a transformation such that a normal distribution is plausible, i.e. is supported by the observed data. Hence several suggestions have been made as to how to optimally transform the response variable and how to assess the estimation of the MSPE of the back transformed domain prediction; see for example Slud and Maiti (2006). Sugasawa and Kubokawa (2015) review and introduce several parametric transformations for the FH model and also discuss the possibility of MSPE estimation.

One of the main motivations to consider area level models is based on the availability of data. Especially with census or administrative data it may not be possible to give unit level information directly to the analyst due to reasons of confidentiality. Thus only aggregates, i.e. direct domain predictions, are available. Even with the availability of information there are other reasons to consider. One is the integration of sampling weights which in general is not directly feasible in model based methodology. Area level models present a way to at least incorporate design weights into the direct estimation and then have a design unbiased estimator on the area level. Other reasons can be practical considerations, e.g. the reduction of computational demands because area level data is often connected with a dramatic reduction in the number of observations hence more complex variance structures can be modeled with less computational effort. Other reasons are discussed in more detail in Namazi-Rad and Steel (2015) who consider different scenarios for the availability of auxiliary information, e.g. unit and area level variables and contextual variables.

Their findings show that overall unit level models have more potential to reduce the MSPE of domain predictions; this is not surprising as parameter estimation under a unit level model uses more information and thus is more precise. This may leave the availability of data as the main reason to consider area level models.

Area level models in Small Area Estimation have been used in many applications and the Fay-Herriot model is subject to numerous extensions. See for example Clement (2014) for a review in disease mapping; Guadarrama et al. (2015) for a review of methods used for poverty mapping; and Benavent and Morales (2016) for an EBLUP under a multivariate FH model. A comprehensive review of extensions to the FH model can be found in Rao (2003, pp. 153 ff) and Rao and Molina (2015, pp 81 ff). The following Section reviews some advances for incorporating structures in space and time into the estimation process. Some of these results in addition to a robust extension is then the subject matter of Chapter 3.

2.3.2 *Spatial and Temporal Fay-Harriot Models*

This Section reviews spatial and temporal extensions to the FH model. These extensions are again the subject of Chapter 3 where they are then combined with robust estimation methodology which is reviewed in Section 2.5. In principle it is intuitive that if historical data, e.g. annually repeated surveys, are available they should be exploited; and the same is true of spatial structures. From a mixed model perspective two components can be modified in order to take such information into account: the random effect and the model error term. The use of correlated random effects to allow for spatial, temporal, or spatio-temporal effects may also be beneficial with respect to domain predictions. In constrast to modifications to the model error this may add to the improvement of domain predictions in terms of MSPE and is looked upon favourably in the literature. Modifying the model error can be motivated to account for correlated sampling errors of the direct estimators. This may be of special interest if, for example, the area level information is based on a panel survey.

2.3.2.1 *Spatial Extensions*

The standard FH model (2.8) uses a random effect to capture unobserved variation between areas. However it ignores unobserved spatial patterns which may be present when areas refer to geographical units. In the FH model the area specific random effects are independent; when spatial correlation is an issue then the assumption of correlated random effects appear instead to be more plausible since neighbouring areas may be similar. If no covariates are present to capture such effects Molina, Salvati et al. (2009) show that taking spatial correlation into account can be beneficial for domain predictions.

Singh, Shukla et al. (2005), Petrucci and Salvati (2006), and Pratesi and Salvati (2008) investigated the possibility of incorporating a simultaneous autoregressive process (SAR) into the domain predictions. In principle there are two options to model spatial correlation: conditional autoregressive models (CAR) and SAR. The main difference is that CAR models are based on a Markov field which implies that spatial correlation exists only locally between two or more neighbours but does not affect other non neighbouring units, i.e. such a process has no memory across space. SAR processes do not have this restriction and are more useful in describing a global correlation structure. For a comprehensive overview of the different approaches see Cressie (1993). In order to incorporate spatial correlation we can modify model (2.8) such that:

$$\tilde{y}_i = \mathbf{x}_i^\top \boldsymbol{\beta} + u_{1i} + e_i \tag{2.10}$$

where the only difference to model (2.8) is the replacement of u_i with u_{1i} to represent the random effect. This is the model used by Pratesi and Salvati (2008). In contrast to the formulation in model (2.8) u_{1i} now follows a SAR(1), i.e. a simultaneous autoregressive process of order one, which is defined by:

$$u_{1i} = \rho_1 \sum_{l \neq i} w_{il} u_{1l} + \epsilon_{1i}$$

where $|\rho_1| < 1$ and $\epsilon_{1i} \sim \mathcal{N}(0, \sigma_1^2)$ are i.i.d. with $i = 1, \ldots, D$. w_{il} are the elements of $\mathbf{W}$ which is the row standardised transformation of the proximity matrix $\mathbf{W}^0$. The elements in $\mathbf{W}^0$ are equal to one where areas are neighboured and zero otherwise, it then follows that the dimension of $\mathbf{W}^0$ is $D \times D$. Note that this definition of $\mathbf{W}^0$ is but one possibility. In general the choice of $\mathbf{W}$ must lead to a nonsingular form of $\mathbf{I}_D - \rho_1 \mathbf{W}$, where $\mathbf{I}_D$ denotes an identity matrix of order D, such that the variance structure can be identified (Rao and Molina, 2015, p. 87). Using the methodology of Section 2.2 the BLUP can be stated as:

$$\tilde{\theta}_i^{SFH} = \mathbf{x}_i^\top \tilde{\boldsymbol{\beta}} + \tilde{u}_{1i}$$

where $\tilde{\theta}_i^{SFH}$ depends on the variance parameters $\boldsymbol{\delta} = (\rho_1, \sigma_1^2)$. Pratesi and Salvati (2008) use a maximum likelihood estimator for both parameters. Replacing the unknown parameters with their respective estimates the EBLUP can be obtained by:

$$\hat{\theta}_i^{SFH} = \mathbf{x}_i^\top \hat{\boldsymbol{\beta}} + \hat{u}_{1i}. \tag{2.11}$$

This model is again the subject matter of Section 3.2.2 where the variance structure is reviewed in more detail. The MSPE for the EBLUP under model (2.10) is developed by Pratesi and Salvati (2008) who extend the results from Prasad and Rao (1990).

A different approach to incorporate spatial correlation structures, but for unit level models, can be found in Chandra, Salvati, Chambers and Tzavidis (2012) who use a geographically weighted regression model and derive the EBLUP under this model. Salvati, Tzavidis et al. (2012) use a geographically weighted M-quantile regression model for small area predictions. Porter et al. (2014) is an example how a CAR process may be used in the context of the FH model, which is, additionally, extended to allow for functional covariates. The model in (2.10) assumes spatial stationarity of the regression parameters; that is that they do not vary spatially. Hence Chandra, Salvati and Chambers (2015) extend this approach for spatial nonstationarity and derive a nonstationary EBLUP.

2.3.2.2 *Temporal Extensions*

An early extension by Choudhry and Rao (1989) allows for an autoregressive process in the error term, which, in the context of the FH model, is one way to account for correlated sampling errors. Rao and Yu (1994) then introduced a model to use autocorrelated random effects to borrow strength for domain predictions:

$$\tilde{y}_{it} = \mathbf{x}_{it}^{\top}\boldsymbol{\beta} + u_i + u_{2it} + e_{it} \tag{2.12}$$

where $i = 1, \dots, D$ and $t = 1, \dots, T$ where D and T are the total number of areas and time periods, respectively. Here $e_{it} \sim \mathcal{N}(0, \sigma^2_{eit})$ is independent with known variances σ^2_{eit}. u_i corresponds to the random effect in model (2.8) and follows a normal distribution with zero mean and variance σ^2_u. Correlation over time is now incorporated by adding u_{2it} which is a correlated random effect following an AR(1), i.e. an autoregressive process of order one:

$$u_{2it} = \rho_2 u_{2i,t-1} + \epsilon_{2it}$$

where ρ_2 is the autocorrelation coefficient with $|\rho_2| < 1$ and $\epsilon_{2it} \sim \mathcal{N}(0, \sigma^2_2)$ being i.i.d. with $i = 1, \dots, D$ and $t = 1, \dots, T$. The BLUP under model (2.12) can then be defined as:

$$\tilde{\theta}^{TFH}_{it} = \mathbf{x}_{it}^{\top}\tilde{\boldsymbol{\beta}} + \tilde{u}_i + \tilde{u}_{2it}$$

where the variance parameters $\boldsymbol{\delta} = (\sigma^2_u, \rho_2, \sigma^2_2)$ are assumed to be known. Replacing these parameters with their respective estimates the EBLUP can be obtained:

$$\hat{\theta}^{TFH}_{it} = \mathbf{x}_{it}^{\top}\hat{\boldsymbol{\beta}} + \hat{u}_i + \hat{u}_{2it}. \tag{2.13}$$

Rao and Yu (1994) used a method of moments estimator for the elements in $\boldsymbol{\delta}$. Extensions to this model have been made by Datta, Lahiri and Maiti (2002) who replace the AR(1) process with a random walk. Singh, Mantel et al. (1991) use a random slope model instead of correlated random effects, but also use an AR(1) to describe the variation of the random regression coefficients.

2.3.2.3 *Spatio Temporal Extensions*

The subject matter of this Section is the combination of the spatial model (2.10) and temporal model (2.12); this combination was introduced by Marhuenda et al. (2013). A similar approach but for a multinomial response can be found in López-Vizcaíno et al. (2015). Singh, Shukla et al. (2005) use a similar approach with respect to spatial autocorrelation but they use a State Space Model with a Kalman filter to take advantage of time series data.

Following Marhuenda et al. (2013) a spatio temporal FH model can be formulated as:

$$\tilde{y}_{it} = \mathbf{x}_{it}^{\top}\boldsymbol{\beta} + u_{1i} + u_{2it} + e_{it} \tag{2.14}$$

where in contrast to model (2.12) the first random effect component has been replaced with a correlated random effects component following a SAR(1). To summarise this model, we have u_{1i} following a SAR(1), u_{2it} following an AR(1), and e_{it} are i.i.d.; furthermore u_{1i}, u_{2it} and e_{it} are assumed to be pairwise independent. Following the presentation of the previous sections, the spatio temporal BLUP under model (2.14) can be defined by:

$$\tilde{\theta}_{it}^{STFH} = \mathbf{x}_{it}^{\top}\tilde{\boldsymbol{\beta}} + \tilde{u}_{1i} + \tilde{u}_{2it}.$$

Marhuenda et al. (2013) have used a REML estimator for the unknown variance components $\boldsymbol{\delta} = (\rho_1, \sigma_1^2, \rho_2, \sigma_2^2)$. Replacing the unknown components with their respective estimates leads then to the EBLUP under model (2.14):

$$\hat{\theta}_{it}^{STFH} = \mathbf{x}_{it}^{\top}\hat{\boldsymbol{\beta}} + \hat{u}_{1i} + \hat{u}_{2it}. \tag{2.15}$$

In contrast to Singh, Shukla et al. (2005) who derive an analytical MSPE for spatio-temporal domain predictions and Pratesi and Salvati (2008) who derive an MSPE for spatial predictions, Marhuenda et al. (2013) propose to use a parametric bootstrap for the estimation of the MSPE.

2.3.2.4 *Discussion*

The use of correlation across space to improve domain predictions is in principle a promising approach for applications. However an important aspect and result of the literature reviewed is that modeling spatial autocorrelation in terms of a random effect is useful only when this structure cannot be captured by auxiliary information, i.e. in the fixed effects part of a mixed model. This means that such models can be used to capture *unobserved* spatial correlation. Also of interest is that although the proximity matrix $\mathbf{W}^0$ is introduced to represent neighbouring units in a geographical sense it can refer to more abstract relationships. Neighbouring units may also be defined, for

example, by capturing the structure between industry sectors which are inter-related. Such structures, in contrast to geographical units, may be defined through domain expertise. Thus this strategy can be generally useful for domain predictions instead of being restricted to the geographical sense of area predictions.

The use of information over time on the other hand is quite different. Marhuenda et al. (2013) note that in practice we may be interested in making predictions for the current time period – and not the past – and use historic information as additional data. The use of historic information may lead to an improvement of parameter estimates due to the increased sample size. The temporal random effect in model (2.12) can additionally be of use in improving domain predictions. Singh, Shukla et al. (2005) note that this is especially the case if the historic information, often itself a prediction, is more reliable than the available information of the current time period.

The spatial, temporal, and spatio-temporal FH models are again the subject matter of Chapter 3 where they are combined with robust estimation methodology. Note that the review in this Chapter may not be sufficient to deduce the representation of these models as linear mixed models; this will be discussed in more detail in Section 3.2.

2.4 UNIT LEVEL MODELS

The Battese-Harter-Fuller (BHF) model, which can also be regarded as a linear mixed model, was introduced by Battese et al. (1988). The main focus of the present Thesis is the extension of area level models. However some recent advances in the field, especially with respect to robust estimation methodology, have been made for unit level models. In Chapter 3 these results are applied to the area level; here the basic unit level model and unit level notation is reviewed preparatory to more details in Section 2.5 below. A review of unit level models in the context of linear mixed models and EBLUP based predictions can be found in Jiang and Lahiri (2006), Rao (2003), and in Rao and Molina (2015) who also review a variety of extensions.

The unit level model of Battese et al. (1988) can be expressed as linear mixed model by:

$$y_{ij} = \mathbf{x}_{ij}^{\top}\boldsymbol{\beta} + u_i + e_{ij} \tag{2.16}$$

where y_{ij} denotes the response of unit $j = 1, \dots, N_i$ in area $i = 1, \dots, D$. The random effects, u_i, are independent and identically distributed following a normal distribution such that $u_i \sim \mathcal{N}(0, \sigma_u^2)$. The error terms e_{ij} are i.i.d. following a normal distribution with zero mean and variance σ_e^2. Furthermore the random effect and model error are independent. $\mathbf{x}_{ij}$ denote the auxiliary variables for the jth unit in the ith area.

Following Rao and Molina (2015, pp. 78 ff) this model can be split into sampled and non-sampled units. Let the set of n sampled and $N-n$ non-sampled units be denoted by S and R respectivey. The sets can furthermore be split into the distinct $S_1 \cup \cdots \cup S_D$ and $R_1 \cup \cdots \cup R_D$, where S_i and R_i denote the sampled and non sampled units in area $i = 1, \ldots, D$, respectively. Then (2.16) can be written in matrix notation as

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_S \\ \mathbf{y}_R \end{pmatrix} = \begin{pmatrix} \mathbf{X}_S \\ \mathbf{X}_R \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} \mathbf{Z}_S \\ \mathbf{Z}_R \end{pmatrix} \mathbf{u} + \begin{pmatrix} \mathbf{e}_S \\ \mathbf{e}_R \end{pmatrix}$$

where the subscripts R and S denote respectively the sampled and non-sampled part of the model. The fundamental assumption of this representation is that the sampled units follow the population model. This is plausible under simple random sampling within each domain and, as Rao and Molina (2015, p. 79) note, also when $\mathbf{X}_S$ contains the variables used for sampling.

These results can now be used to derive an EBLUP for the target statistic. In principle this also happens according to Section 2.2; however we are generally not interested in a prediction at unit level but for the area level target statistic. Following Rao and Molina (2015, pp. 174 ff) an EBLUP under model (2.16) for the ith area mean can be derived as

$$\hat{\theta}_i^{BHF} = N_i^{-1} \left(\sum_{j \in S_i} y_{ij} + \sum_{j \in R_i} (\mathbf{x}_{ij}^\top \hat{\boldsymbol{\beta}} + \hat{u}_i) \right) \tag{2.17}$$

where $\hat{\boldsymbol{\beta}}$ denotes the BLUE of $\boldsymbol{\beta}$ which is estimated using the sampled information. Note that this predictor is based on the auxiliary information for each unit in the non-sampled part of the population. It may be unrealistic to have such information for each observation. This assumption can be relaxed by using the area means of the population to make domain predictions; however this information needs to be available to the analyst. The unknown parameters of the model, i.e. $\boldsymbol{\beta}$, σ_u^2 and σ_e^2, can be estimated using standard ML and REML estimation techniques and are based on the sampled data.

A wide range of MSPE estimators have been proposed for basic unit level small area methods – a comprehensive review of these methods can be found in Rao and Molina (2015, 179 ff). Among these methods is an unconditional MSPE estimator which has been proposed by Datta and Lahiri (2000). For the robust extension of (2.17), which is introduced in the following Section, a number of methods are discussed in more depth in Section 2.5.3.

2.5 ROBUST METHODS IN SMALL AREA ESTIMATION

The quality of predictions using model based methods is strongly dependent upon the distributional assumptions. Serious problems with

these methods can arise already when single outlying observations are present in the data. Hence discussion and some progress has been made in connection with methods robust against outliers. Outliers are here defined as representative outliers as in Chambers (1986). A thorough discussion of what exactly outlying observations on the area level are will be given in Section 3.1 ahead.

In the broader context of this subfield within SAE the present Thesis relies strongly on the results of Sinha and Rao (2009) who propose robust estimation equations for linear mixed models. Their proposition focuses on the unit level model (2.16); but the methodology builds on linear mixed models where (2.16) can be derived as a special case of these models. Also the basic area level model (2.8) can be framed within this context; and the estimation equations can be adapted accordingly. This is also pointed out in Sinha and Rao (2008) and in Rao and Molina (2015, pp. 146 ff). The use of these results for area level models is again addressed in Chapter 3 where robust extensions to the FH model are proposed.

In what follows the main results around the methodology introduced by Sinha and Rao (2009) are reviewed. Together with the models reviewed in Section 2.3 these results are the basis for the extensions proposed in this Thesis. The Section is structured as follows: Section 2.5.1 reviews the methodology around robust estimation equations; in Section 2.5.2 these results are extended to allow for a correction of the bias associated with this type of robust predictions; then in Section 2.5.3 different possibilities to estimate the MSPE of the area level predictions are described; and in Section 2.5.4 these results are embedded in a broader context of proposed robust methods in the SAE field.

2.5.1 *Robust Estimation Equations*

Section 2.2 reviewed the domain prediction under linear mixed models. The key to these predictions are the BLUE given by (2.3) and the BLUP given by (2.4). These estimators can be derived based on the log-likelihood of the joint density of $\mathbf{y}$ and $\mathbf{u}$ of (2.2)

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

when we assume normality in the distribution of $\mathbf{u}$ and $\mathbf{e}$. Note that the model can now be stated as $\mathbf{y} \sim \mathcal{N}(\mathbf{X}\beta, \mathbf{V})$ where $\mathbf{V} = \mathbf{Z}\mathbf{V}_u\mathbf{Z}^\top + \mathbf{V}_e$. The first derivatives of the log-likelihood with respect to $\boldsymbol{\beta}$ and $\mathbf{u}$ for given variance parameters $\boldsymbol{\delta}$ lead then to the so-called *mixed model equations*:

$$\mathbf{X}^\top \mathbf{V}_e^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) = 0 \tag{2.18}$$

$$\mathbf{Z}^\top \mathbf{V}_e^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) - \mathbf{V}_u^{-1}\mathbf{u} = 0. \tag{2.19}$$

Henderson (1963) has shown that the solutions to (2.18) and (2.19) are identical to the BLUE and BLUP as defined in (2.3) and (2.4), respectively.

Fellner (1986) studied the robust estimation of $\boldsymbol{\beta}$ and $\mathbf{u}$ by modifying (2.18) and (2.19) to restrict the influence of outlying observations. He suggests the use of an influence function to restrict the impact of the residuals in (2.18) and (2.19), transforming them into:

$$\mathbf{X}^\top \mathbf{V}_e^{-\frac{1}{2}} \psi\left(\mathbf{V}_e^{-\frac{1}{2}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u})\right) = 0 \tag{2.20}$$

and

$$\mathbf{Z}^\top \mathbf{V}_e^{-\frac{1}{2}} \psi\left(\mathbf{V}_e^{-\frac{1}{2}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u})\right) - \mathbf{V}_u^{-\frac{1}{2}} \psi\left(\mathbf{V}_u^{-\frac{1}{2}}\mathbf{u}\right) = 0 \tag{2.21}$$

where $\psi(\cdot)$ is a bounded monotonous function. Fellner (1986) suggests to use Huber's influence function (Huber, 1964) where $\psi_b(x) = x \min\left(1, \frac{b}{|x|}\right)$ for a given tuning constant b where a common choice for b is 1.345. Rao (2003, p. 102) suggests to use a robust version of Henderson's (Henderson, 1950) estimation equations for the unknown parameters in $\boldsymbol{\delta}$; remember that $\mathbf{V} = \mathbf{V}(\boldsymbol{\delta})$ where $\boldsymbol{\delta}$ denotes the vector of unknown variance parameters.

Building on these results Sinha and Rao (2009) propose a similar approach in that they use *robustified* estimation equations. In contrast to the approach by Fellner (1986) these equations are derived from the marginal model of (2.2) and specifically for $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$; thus the first derivatives of the marginal log-likelihood of $\mathbf{y}$ with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ are:

$$\mathbf{X}^\top \mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = 0$$

and

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta_l} \mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) - \operatorname{tr}\left(\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \delta_l}\right) = 0$$

for $l = 1, \ldots, Q$ where Q denotes the total number of variance parameters in $\boldsymbol{\delta}$. These estimation equations are now modified in order to restrict the influence of outlying observations similar to (2.20) in that the residuals have bounded influence using an influence function denoted by $\psi(\cdot)$:

$$\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \psi(\mathbf{r}) = 0 \tag{2.22}$$

$$\psi(\mathbf{r})^\top \mathbf{U}^{\frac{1}{2}} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta_l} \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \psi(\mathbf{r}) - \operatorname{tr}\left(\mathbf{K}\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \delta_l}\right) = 0 \tag{2.23}$$

where $\mathbf{K} = \mathbb{E}(\psi_b^2(z))\mathbf{I}_n$ is a diagonal matrix of the same order as $\mathbf{V}$ with z following a standard normal distribution; and $\mathbf{r} = \mathbf{r}(\boldsymbol{\beta}) = \mathbf{U}^{-\frac{1}{2}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ denotes the vector of the standardised residuals and $\mathbf{U} = \mathbf{U}(\boldsymbol{\delta})$ is the matrix containing the diagonal elements of $\mathbf{V}$.

In order to find solutions for β and δ based on the robust estimation equations (2.22) and (2.23) Sinha and Rao (2009) propose Newton-Raphson algorithms based on a Taylor series expansion. Schmid (2011) and Schoch (2012) have reported numerical problems with this approach; and different alternative solutions have been investigated. Schmid (2011) proposes to find the minimum of the squared estimation equations in (2.23) for all variance parameters; and this was further optimised by Auerbach (2015) giving stable and satisfying results. Schoch (2012) uses instead an iteratively re-weighted least squares (IRWLS) algorithm for a solution of (2.22) which also yields satisfactory results. A further proposition has been offered by Chatrchi (2012) which is to derive a system of fixed point equations based on (2.23): she derives a stable algorithm – using a fixed point algorithm – for δ.

The main advantage of these alternatives may be that they do not rely on the derivatives of the estimation equations, which was already a known concern in the work of Sinha and Rao (2009). See Kennedy Jr and James (1980, Chapter 11) and Thisted (1988, Chapter 4) for comprehensive overviews for solving ML based robust estimation equations, they also suggest the IRWLS algorithm as an alternative to the NR algorithm for this very reason.

With given robust estimates for β and δ, Sinha and Rao (2009) solve the robust mixed model equation (2.21) proposed by Fellner (1986) in order to derive a robust EBLUP (REBLUP). Although they also suggest to use a NR algorithm which is developed using a Taylor series expansion, the numerical solution appears to be less problematic. On the other hand Schoch (2012) suggests finding solutions on the basis of a robust version of a method of moments estimator; and a similar suggestion can be found in Rao and Molina (2015, p. 196): they propose the use of a robust version of (2.4):

$$\hat{\mathbf{u}}^{\psi} = \mathbf{V}_u \mathbf{Z}^{\top} \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \psi(\mathbf{r}) \tag{2.24}$$

where the notation is the same as given before with the only difference being that $\hat{\mathbf{u}}^{\psi}$ now depends on the robust parameter estimates $\hat{\beta}^{\psi}$ and $\hat{\delta}^{\psi}$. From a computational point of view this has the obvious advantage that no iterative algorithm is needed. Rao and Molina (2015, p. 196) note that a disadvantage here may be that (2.24) depends on the composite error – $\mathbf{Zu} + \mathbf{e}$ – of (2.2) whereas in (2.21) the influence function is applied only to $\mathbf{e}$. However, an empirical investigation of this issue seems not to be available in the literature at this time.

Given the robust estimates $\hat{\beta}^{\psi}$, $\hat{\delta}^{\psi}$, and $\hat{\mathbf{u}}^{\psi}$, Sinha and Rao (2009) derive the REBLUP for area level means under the BHF model (2.16).

The REBLUP under the BHF model is then a robust variation of (2.17) and can be stated as:

$$\hat{\theta}_i^{RBHF} = N_i^{-1}\left(\sum_{j\in S_i} y_{ij} + \sum_{j\in R_i}\left(\mathbf{x}_{ij}^{\top}\hat{\boldsymbol{\beta}}^{\psi} + \hat{u}_i^{\psi}\right)\right) \tag{2.25}$$

2.5.2 *Bias Correction*

Predictions using a REBLUP in the form of (2.25) can result in the introduction of a bias. This can also happen in the case of different outlier robust estimation approaches as already noted, for example by Tzavidis and Chambers (2005), in the context of M-quantile regression. In general a bias can be introduced if outliers in $\mathbf{u}$ or $\mathbf{e}$ are present but are not drawn from a distribution with zero mean or are from an asymmetric distribution.

This may be viewed as a contrast to the situation say in which we have a well behaved population model following, for example, the Gaussian Linear Mixed Model; in this situation outliers can occur randomly due to sampling but should be of no concern under repeated sampling. This effectively means that realisations of outliers are driven purely by chance. In many real world scenarios outliers have a direction: for instance outliers for income data are typically high incomes. The same holds true for example for the revenue of a company. This is modeled in the robust SAE literature as either a mixture distribution in which outlying observations are believed to follow a different population model with a shifted location than the rest of the observations; or as a non-symmetric, heavy tailed distribution which is better suited for describing the occurrence of outlying observations than a symmetric distribution.

Where we in fact suspect a mixture or non-symmetric population model, then weighting down outlying observations effectively means that we tend to ignore the outlier population model or the inherint asymmetry. The robust estimation of location parameters using the techniques described above are essentially a trade-off between predicting the mean and median which, with respect to bias, introduces a problem with an asymmetric population model.

This reasoning explains why some advances have been attempted to control a potential bias introduced by robust estimation techniques. Chambers, Chandra, Salvati et al. (2014) have proposed a correction of REBLUPs in the form of (2.25) by adding an area specific estimator of the potential bias:

$$\hat{\theta}_i^{RBHF-BC} = \hat{\theta}_i^{RBHF} + \left(\frac{1}{n_i} - \frac{1}{N_i}\right)\sum_{j\in S_i}\phi_i\psi_c\left(\frac{\hat{e}_{ij}^{\psi}}{\phi_i}\right) \tag{2.26}$$

where $\hat{e}_{ij}^{\psi} = y_{ij} - \mathbf{x}_{ij}^{\top}\hat{\boldsymbol{\beta}}^{\psi} - \hat{u}_i^{\psi}$ is the unit level prediction error and ϕ_i with $i = 1, \ldots, D$ is the median absolute deviation of $\hat{e}_{ij}^{\psi}$ within area i. Note that here ψ_c denotes typically the same influence function as in the robust estimation equations, ψ_b, with the important restriction that $c > b$. This means that ψ_c is still bounded but more liberal than ψ_b; this introduces the possibility to compute a potential prediction bias for each area.

An extension of this approach was introduced by Jiongo, Haziza et al. (2013) whose main concern has been that the correction in (2.26) only depends on the sampled units within domains. Hence they propose two alternatives. First they propose a bias correction based on the results of Chambers (1986) which, in contrast to (2.26), relies on S instead of S_i – i.e. the whole sample and not just the sampled units within domains – and they call this a fully-bias correction. The second approach uses a conditional bias to account for units in the population. This effectively reduces to the different treatment of the robust predictions of the random effects in their two approaches. In their simulation study they find an improvement in terms of bias for all three bias corrections for the REBLUP in (2.25). However, especially in terms of stability measured by relative efficiency – relative to the EBLUP in (2.17) – the two approaches by Jiongo, Haziza et al. (2013) show promising results in their simulation study.

At this point it is important to note that for area level robust predictions the application of the results of Chambers, Chandra, Salvati et al. (2014) and Jiongo, Haziza et al. (2013) are not immediately obvious. All three approaches are based on some form of mean of the unit level prediction error which is not available at the area level. However the underlying problem of a potential bias continues to persist – also at the area level. A solution for robust area level predictions is discussed in Section 3.4.

2.5.3 *Mean Squared Prediction Error*

One of the challenging problems in this discipline is the estimation of the MSPE in model based SAE (Pfeffermann, 2013). This is only aggravated with the combination of robust estimation methodology so that Sinha and Rao (2009) initially proposed a bootstrap instead of an analytical solution. In Section 2.5.3.1 below existing bootstrap methods in this context are reviewed. Then in Section 2.5.3.2 one analytical solution is presented based on a pseudolinear representation of the REBLUP which also allows the estimation of the MSPE of the bias corrected REBLUP.

2.5.3.1 *Bootstrap Methods*

In order to estimate the MSPE for the REBLUP in (2.25) Sinha and Rao (2009) propose the use of a parametric bootstrap. With the robust parameter estimates $\hat{\boldsymbol{\beta}}^{\Psi}$ and $\hat{\boldsymbol{\delta}}^{\Psi} = \left(\hat{\sigma}_u^{2\Psi}, \hat{\sigma}_e^{2\Psi}\right)$ samples are generated from the following model:

$$y_{ij}^* = \mathbf{x}_{ij}^\top \hat{\boldsymbol{\beta}}^{\Psi} + u_i^* + e_{ij}^*$$

with $j = 1, \ldots, n_i$ and $i = 1, \ldots, D$ where $u_i^* \sim \mathcal{N}(0, \hat{\sigma}_u^{2\Psi})$ and $e_{ij}^* \sim \mathcal{N}(0, \hat{\sigma}_e^{2\Psi})$. In each repetition, i.e. for each bootstrap sample, the bootstrap population mean is computed as:

$$\bar{Y}_i^* = N_i^{-1} \left(\sum_{j \in S_i} y_{ij}^* + \sum_{j \in R_i} \left(\mathbf{x}_{ij}^\top \hat{\boldsymbol{\beta}}^{\Psi} + (1 - n_i N_i^{-1})(u_i^* + \bar{e}_i^*) \right) \right)$$

where $\bar{e}_i^* \sim \mathcal{N}(0, (N_i - n_i)^{-1} \hat{\sigma}_e^{2\Psi})$. Using each bootstrap sample, the REBLUP (2.25) is computed and the suggested bootstrap estimator of the MSPE with B repetitions is then defined by:

$$\widehat{\mathrm{MSPE}}\left(\hat{\theta}_i^{\mathrm{RBHF}}\right) = \frac{1}{B} \sum_{b=1}^{B} \left(\hat{\theta}_i^{\mathrm{RBHF}(b)} - \bar{Y}_i^{*(b)} \right)^2$$

where $\hat{\theta}_i^{\mathrm{RBHF}(b)}$ and $\bar{Y}_i^{*(b)}$ denote respectively the REBLUP and the bootstrap population mean of the bth bootstrap sample. Jiongo, Haziza et al. (2013) use a very similar approach with the important difference that their model – from which the bootstrap samples are drawn – depends on the non-robust parameter estimates $\hat{\boldsymbol{\delta}} = (\hat{\sigma}_u^2, \hat{\sigma}_e^2)$. Sinha and Rao (2009) argue that they are interested in resampling the non-contaminated part of the observations because the MSPE of their robust estimator should be uneffected by outliers. Jiongo, Haziza et al. (2013) found that the method used by Sinha and Rao (2009) lead to poor coverage rates of constructed confidence intervals and they suspect that this is because the robust parameter estimates do not reflect a sufficiently large part of the sample.

Jiongo and Nguimkeu (2014) propose a different bootstrap method. Their non-parametric approach is different from the above in that the random effects, u_i^*, and the error term, e_{ij}^*, are drawn from a transformed set of the estimated $\{\hat{u}_i\}_{i=1}^D$ and $\{\{\hat{e}_{ij}\}_{j=1}^{N_i}\}_{i=1}^D$ under a non-robust BHF model to generate bootstrap populations. From these generated populations samples are drawn using the same sampling scheme as for the realised sample at hand. With these differences they provide preliminary results which show very good properties in terms of relative bias and relative root mean squared error of their bootstrap estimator. Albeit their results are very promising, it is not entirely clear how to adapt this approach in the case of mixed linear

models with correlated random effects which would be needed for the spatial and temporal extensions reviewed in Section 2.3.2.

Mokhtarian and Chambers (2013) proposed a further extension which is to use a robust block bootstrap. Similar to the approach taken by Jiongo and Nguimkeu (2014) they also construct a set of errors and random effects from which they draw their bootsrtap samples. However instead of using a robust estimation method, they construct outlier robust samples; this means that the way in which the bootstrap samples are constructed is robust against outliers. On each sample then a non-robust small area estimator can be used. This approach has the interesting effect that a MSPE estimator for robust domain predictions as well as the robust predictions themselfes can be produced. Although the results look promising it is at this time, again, not clear how to extend their method for models with correlated random effects.

2.5.3.2 *Pseudolinearisation-based Approach*

Parallel to the development of the bootstrap methods discussed above Chambers, Chandra and Tzavidis (2011) provide an analytical solution for a bias-robust MSPE estimator for small area predictions. The appeal of their method lies in their available applicability for a wide range of predictors. Their method can be used for any predictor which can be represented as a weighted sum of the sampled response vector. Chambers, Chandra, Salvati et al. (2014) then adapted this approach for the REBLUP in (2.25) and the bias corrected version in (2.26). They derive an analytical estimator to the conditional MSPE – conditional on the set of random effects – based on a pseudolinear form of the REBLUP and extend this approach to a linearisation-based approach.

Following Chambers, Chandra, Salvati et al. (2014) the RBLUP underlying (2.25) – i.e. with known parameters δ – can be represented as a weigthed sum of the sampled response values:

$$\tilde{\theta}_i^{RBHF} = \sum_{j \in S} \tilde{\mathbf{w}}_{ij}^{RBHF} y_j = \left(\tilde{\mathbf{w}}_{iS}^{RBHF}\right)^\top \mathbf{y}_S$$

where $\mathbf{y}_S$ denotes the vector of sampled response values and $\tilde{\mathbf{w}}_{iS}^{RBHF}$ the vector of weights to produce the ith area prediction. $\tilde{\mathbf{w}}_{iS}^{RBHF}$ is here defined as:

$$\left(\tilde{\mathbf{w}}_{iS}^{RBHF}\right)^\top = N_i^{-1}\left(\mathbf{1}_{iS}^\top + (N_i - n_i)\left(\bar{\mathbf{x}}_{iR}^\top \mathbf{A}_S + \bar{\mathbf{z}}_{iR}^\top \mathbf{B}_S\left(\mathbf{I}_S - \mathbf{X}_S \mathbf{A}_S\right)\right)\right)$$

where in contrast to (2.25) we have substituted $\mathbf{x}_{ij}$ with $\bar{\mathbf{x}}_{iR}$ denoting the known population means of the auxiliary variables and similarly $\bar{\mathbf{z}}_{iR}$ denotes the known vector for selecting the random effects for area i. Here $\mathbf{1}_{iS}$ denotes a vector with n elements which are equal to one if

the jth element is from area i and zero otherwise; and $\mathbf{I}_S$ is a $(n \times n)$ identity matrix. Furthermore

$$\mathbf{A}_S = \left(\mathbf{X}_S^\top \mathbf{V}_S^{-1} \mathbf{U}_S^{\frac{1}{2}} \mathbf{W}_{1S} \mathbf{U}_S^{-\frac{1}{2}} \mathbf{X}_S\right)^{-1} \mathbf{X}_S^\top \mathbf{V}_S^{-1} \mathbf{U}_S^{\frac{1}{2}} \mathbf{W}_{1S} \mathbf{U}_S^{-\frac{1}{2}}$$

where $\mathbf{W}_{1S} = diag\left(\{w_{1j}\}_{j=1}^n\right)_{n\times n}$ is a diagonal matrix with

$$w_{1j} = \psi\left(U_j^{-\frac{1}{2}}\left(y_j - \mathbf{x}_j^\top \tilde{\boldsymbol{\beta}}^\psi\right)\right)\left(U_j^{-\frac{1}{2}}\left(y_j - \mathbf{x}_j^\top \tilde{\boldsymbol{\beta}}^\psi\right)\right)^{-1}$$

as elements. Note that here $U_j^{-\frac{1}{2}}$ denotes the jth diagonal element of the matrix $\mathbf{U}_S^{-\frac{1}{2}}$. Continuing,

$$\mathbf{B}_S = \left(\mathbf{Z}_S^\top \mathbf{V}_{eS}^{-\frac{1}{2}} \mathbf{W}_{2S} \mathbf{V}_{eS}^{-\frac{1}{2}} \mathbf{Z}_S + \mathbf{V}_u^{-\frac{1}{2}} \mathbf{W}_{3S} \mathbf{V}_u^{-\frac{1}{2}}\right)^{-1} \mathbf{Z}_S^\top \mathbf{V}_{eS}^{-\frac{1}{2}} \mathbf{W}_{2S} \mathbf{V}_{eS}^{-\frac{1}{2}}$$

where $\mathbf{W}_{2S} = diag\left(\{w_{2j}\}_{j=1}^n\right)_{n\times n}$ and $\mathbf{W}_{3S} = diag\left(\{w_{3i}\}_{i=1}^D\right)_{D\times D}$ are diagonal matrices where their respective elements are defined by:

$$w_{2j} = \psi\left(\frac{1}{\sigma_e^\psi}\left(y_i - \mathbf{x}_i^\top \tilde{\boldsymbol{\beta}}^\psi - \tilde{u}_i^\psi\right)\right)\left(\frac{1}{\sigma_e^\psi}\left(y_i - \mathbf{x}_i^\top \tilde{\boldsymbol{\beta}}^\psi - \tilde{u}_i^\psi\right)\right)^{-1}$$

and

$$w_{3i} = \psi\left(\frac{\tilde{u}_i^\psi}{\sigma_u^\psi}\right)\frac{\sigma_u^\psi}{\tilde{u}_i^\psi}.$$

The pseudolinear representation for $\tilde{\theta}_i^{RBHF}$ needs to be modified for the REBLUP in that we substitute the unknown variance parameters with their robust estimates, $\hat{\delta}^\psi$. In addition let $\hat{w}_{ij}^{RBHF}$ denote the weight of the REBLUP under the BHF model for the jth observation to predict the ith area mean. Chambers, Chandra, Salvati et al. (2014) then derive the estimator of the MSPE following the results of Chambers, Chandra and Tzavidis (2011) which in turn are based on the results of Royall and Cumberland (1978). Thus an estimator for the MSPE is given by:

$$\widehat{MSPE}\left(\widehat{\theta}_i^{RBHF}\right) = \widehat{\mathbb{V}}\left(\widehat{\theta}_i^{RBHF}\right) + \widehat{\mathbb{B}}\left(\widehat{\theta}_i^{RBHF}\right)^2 \tag{2.27}$$

where

$$\widehat{\mathbb{V}}\left(\widehat{\theta}_i^{RBHF}\right) = N_i^{-2}\sum_{j\in S}\left(a_{ij}^2 + (N_i - n_i)\,n^{-1}\right)\lambda_j^{-1}\left(y_j - \hat{\mu}_j\right)^2$$

is an estimator of the conditional prediction variance with $a_{ij} = N_i w_{ij}^{RBHF} - I(j \in i)$ where $I(j \in i)$ denotes an indicator function for when observation j is in the ith area; and

$$\widehat{\mathbb{B}}\left(\widehat{\theta}_i^{RBHF}\right) = \sum_{j\in S}\hat{w}_{ij}^{RBHF}\hat{\mu}_j - N_i^{-1}\sum_{j\in(R_i\cup S_i)}\hat{\mu}_j$$

is an estimator of the conditional prediction bias. Note that $\hat{\mu}_j$ is an unbiased estimator of the the conditional expectation $\mu_j = \mathbb{E}\left(y_j|\mathbf{x}_j, \mathbf{u}^\psi\right)$ and $\lambda_j = 1 - 2\phi_{jj} + \sum_{k\in S}\phi_{kj}^2$ is a scaling constant.

Due to the shrinkage effect associated with EBLUPs, Chambers, Chandra and Tzavidis (2011) suggest the use of the unshrunken version of the EBLUP under the BHF model for $\hat{\mu}_j$. The approach can be utilised to derive an analytical MSPE estimator for the bias corrected version of the REBLUP. Chambers, Chandra, Salvati et al. (2014) do this by replacing the weights in the pseudolinear representation and argue for the omission of the squared bias term in (2.27) since it is an approximately unbiased estimator of the area mean.

Note that the benefit of this pseudolinear representation has the underlying assumption that the weights and sampled response values are independent; furthermore this estimator of the MSPE neglects the uncertainty associated with the estimation of the variance components. Chambers, Chandra, Salvati et al. (2014) hence note that this is a first order approximation to the actual MSPE estimator of the REBLUP. A second note is that this approach involves an increase in the MSE of the MSPE estimator. However, Chambers, Chandra, Salvati et al. (2014) argue that in realistic applications this approach has good repeated sampling properties – for details see the references in Chambers, Chandra, Salvati et al. (2014).

Chambers, Chandra, Salvati et al. (2014) also provide a linearisation-based approach to the estimation of the conditional MSPE. This approach explicitly aims to incorporate the uncertainty in the estimation of the variance components and hence is a second order approximation. The pseudolinear form has, in principle, the great advantage that it is easily adapted for other small area estimators through replacing the weights. The linearisation-based approach makes it necessary to provide components to capture the uncertainty associated with specific variance components and variance structures. For this reason the pseudolinear form is further explored and adapted for the robust area level estimators in Section 3.3.2 and Section 3.5.2 below.

2.5.4 *Discussion*

The discussion in this Section has been narrowed to the SAE field and specifically around the results by Sinha and Rao (2009). Their results are strongly influenced by Richardson and Welsh (1995) who propose robust estimation equations for linear mixed models and by Huggins (1993) who proposes similar methods. A further extension of these methods to generalised linear mixed models can be found in Yau and Kuk (2002).

Recent extensions to the robust methods by Sinha and Rao (2009) can be found in Schmid and Münnich (2014) who use a unit level model with spatially correlated random effects; they propose a spatial

REBLUP (SREBLUP). Schmid, Tzavidis et al. (2016) then build on the results of Chambers, Chandra, Salvati et al. (2014) and make the necessary extensions for the SREPLUP. Rao, Sinha et al. (2014) have extended the unit level mixed model to allow for P-splines in the fixed part of the model. It is interesting to note that they also report problems with their originally proposed Newton-Raphson algorithm and suggest using a fixed point algorithm for the variance parameters instead.

In general the approach to outlier robust predictions, as it was discussed above, is but one possible approach; it is plausible to apply it in an application when the analyst believes that the population model is correctly specified for a substential part of the population. When the model is incorrect – for example in the case of a mixture distribution or in general in the context of non-symmteric outliers – the proposed methods can easily lead to a prediction bias. However solutions have been suggested to overcome this issue. Similar methods – in that they use robust estimation equations – with respect to survey methods can be found in Beaumont and Rivest (2009) and Hulliger (1999) who recommend for example a robust HT estimator; and Beaumont (2004) has introduced robust extensions to generalised regression estimation.

A different approach to robust small area predictions can be found in Chambers and Tzavidis (2006) and in Tzavidis, Marchetti et al. (2010) which is to model a conditional quantile instead of a conditional mean. This approach also leads to the same problem under non-symmetric outliers in that the predictions can be biased. Tzavidis and Chambers (2005) have suggested a correction term for this bias which was then adapted for the REBLUP by Chambers, Chandra, Salvati et al. (2014).

Furthermore robust small area predictions can be addressed by changing the underlying model assumptions. This is most natural in the situation where the original distributional assumption, for example a normal distribution, is implausible and needs to be replaced. Choices can fall on symmetric distributions with more probability mass on the tails or on non-symmetric distributions. Bell and Huang (2006) for example use a Bayesian approach and choose a t-distribution for the random effects. Datta and Lahiri (1995) suggest the use of a Cauchy distribution or a mixture distribution. Outside the SAE field examples for robust modeling can be found in Lange et al. (1989) and Peel and McLachlan (2000) who formulate a likelihood based on a multivariate t-distribution.

Although most of the references focus on robust small area prediction under unit level models, several possibilities have been explored for area level models. Bell and Huang (2006) and Huang and Bell (2006) both use the t-distribution to model the random effect in the FH model within the context of poverty mapping. Xie et al. (2007)

also use a hierarchical Bayes approach using a t-distribution to predict the proportion of overweight individuals in small areas. Fabrizi and Trivisano (2010) explore the possibility to use exponential power distributions to model the random effect; put simply, this is a normal distribution where the skewness and kurtosis are parameterised such that this distribution can be used to model heavy tails or non-symmetry. They also use a Baysian approach. Ghosh, Maiti et al. (2008) modify the Hirarchical Bayes and Empirical Bayes estimators by using an influence function. This is a similar approach taken as proposed in Chapter 3 however in the setting of a Baysian framework – whereas the estimation methodology underpinning the methods of this Thesis is based on Maximum Likelihood. Gershunskaya (2010) and Chakraborty et al. (2015) propose to use the FH model but replace the distribution of the random effects with a mixture distribution. This essentially means that they have to detect which observations are outliers and fit a separate model. All these studies can claim some advantages – in terms of MSPE – for their respective method when outliers are present.

3

ROBUST AREA LEVEL MODELS

[Outliers:] observations obtained
under seemingly normal circumstances,
but that turn out to be extremely deviant
from the main body of observations.

— Abelson (1995, p. 69)

3.1 THE PROBLEM OF OUTLYING OBSERVATIONS FOR AREA LEVEL MODELS

Outliers can have severe and unfortunate impact on predictions. This is true both for the case of linear mixed models as well as for hierarchical Bayes; both these methodologies are being used for area level models. Section 2.5.4 above contains a review of literature on robust area level models. Here it is noteworthy that those studies focused on how outliers can be dealt with in the context of hierarchical Bayes. In what follows the robust estimation methodology of Sinha and Rao (2009) is adapted to area level models which leads to outlier robust predictions under mixed linear models. Before the robust methods are discussed the different scenarios in which we may be interested in using robust methods are first described; i.e. what exactly are outliers at the area level?

Chambers (1986) has introduced the terms *representative* and *non-representative* outliers to distinguish between two common types:

- Representative outliers are observations with large absolute values that are correctly recorded and cannot be considered unique, i.e. they represent a substantial part of the target population.
- Non-representative outliers may be described as incorrectly collected observations which should not be part of the sample, or observations which are unique in the population.

From a practical point of view substantial knowledge of the data generating process and of the target population is needed to be able to distinguish between these two types. A factor which further complicates matters is that area level models are often used in situations where we have no access to unit level data. With this in view the discussion in this Thesis focuses on the case where we assume to have representative outliers or are in the situation where it is not possible to argue for either of these alternatives.

Given that we consider representative outliers it is important to understand the source and effect of such observations. For the FH model in (2.8) the regression residuals can be derived as:

$$\tilde{y}_i - \mathbf{x}_i^\top \boldsymbol{\beta} = u_i + e_i.$$

Bell and Huang (2006) argue that two types of outliers need to be considered:

- Outliers in u_i are area level outliers. These values indicate that for outlying areas the fixed part, i.e. the regression line, is not a good fit. In this case neither the regression estimator nor the direct estimator is a good predictor for this area; unless it happens to be an area with sufficient sample size to support the direct estimator.
- Outliers in e_i are unit level outliers. The source of such values are single units in the sample and in this case the synthetic estimator, i.e. the regression line, is a better choice.

Given this distinction Bell and Huang (2006) argue that in practice it is not possible to indicate which type of outlying observations are observed and that the existing literature mainly focuses on dealing with area level outliers since they lead to a simple extension of the FH model – see for example Datta and Lahiri (1995) and Xie et al. (2007).

In contrast to this way of classification I want to distinguish between *three* types of outlying observations and furthermore to illustrate that it may be possible to distinguish between unit and area level outliers. In the following sections these types are characterised as a preliminary to the study of robust area level models.

3.1.1 *Unit Level Outliers*

Unit level outliers can be representative or non-representative but in most cases it is beyond our reach to judge which kind they are since we may only have access to aggregates. These observations will influence the direct estimator in that this quantity may have unexpectedly large absolute values. They also influence the estimation of the standard error of the direct estimator, which in turn may be used as the true or rather the given variance parameters in the Fay-Herriot model.

In the situation where the direct variance estimators are treated as having the known variance structure two conflicting effects need to be considered. First, we use an obviously unreliable estimator for the sampling error and assume that such values can be used as the *true* variance parameters. Such estimators will have poor properties when the target is the *true* sampling error especially in the context of outliers. This problem has stimulated the discussion around smoothing

the variance estimates prior to using them in a Fay-Herriot model – see Section 2.3.1.2 for a review of this debate. However, unit level outliers may lead to *outlying estimates of the standard errors*, which in turn may have an unwanted effect on the smoothing of these parameters. This situation is further discussed in Section 3.1.3.

Second, we may consider the direct variance estimators not as having an informative quantity for the *true* sampling error as it is assumed under the FH model, but as informative about the unit level sample instead. This is an important aspect because these estimates give an index of reliability. For a given sample size the estimated sampling variance is large if outliers are present or, in general, if the sample is heterogeneous; and it is low for areas in which we have a reliable direct estimator. In this case the Fay-Herriot model weights down unreliable direct estimators, which is why this model may be able to adjust itself for unit level outliers. This property is explored in Section 6.3 using model based simulations.

In addition to these two scenarios it is also relevant to consider the use of robust direct estimators. From a practical point of view we may only have access to non-robust estimators. On the other hand it may be crucial feedback for a data provider to additionally report robust estimates, e.g. a robust mean or a median. Also it is unclear how this relates to the *self-adjusting* effect of the Fay-Herriot model as it was characterized above. This aspect is also addressed in the a model based simulation study in Section 6.3.

3.1.2 *Area Level Outliers*

Area level outliers are domains or areas which are far away from the main body of the observations. The source of such observations are not single units influencing the direct estimator, but the fact that there exist domains which are substantially different.

Under a linear mixed model this means that there are outliers in the random effect, u_i. From a model perspective, where we treat the random effect as an i.i.d. random variable following a normal distribution, such values may arise by chance alone. This would mean that in truth the distribution is indeed normal but we are in one of those rare cases in which we observe abnormal behaviour of single domains.

This line of argumentation seems to be unrealistic for applications in SAE since under repeated sampling we would expect the same domains to be abnormal. This opens up a more fundamental discussion about linear mixed models in Small Area Estimation; and it gives reason to frame a random effect as fixed but unknown quantity. In this setting the existence of outliers means that the normal distribution may well be a good approximation for the majority of areas but not for them all.

In an application in which such values are observed a robust estimation technique can be beneficial. The main effect of area level outliers is that the variance parameter of the random effect will be overestimated. Also the parameter estimates of the fixed effects part can be influenced by such observations, although it is unclear whether they are under or over estimated. Looking at the EBLUP under the FH model in (2.9):

$$\hat{\theta}_i^{FH} = \gamma_i \tilde{y}_i + (1 - \gamma_i)\mathbf{x}_i^\top \hat{\boldsymbol{\beta}}$$

where $\hat{\gamma}_i = \hat{\sigma}_u^2/(\hat{\sigma}_u^2 + \sigma_{e_i}^2)$ we see that the larger the estimated variance parameter $\hat{\sigma}_u^2$ – for fixed values of $\sigma_{e_i}^2$ – the more weight is given to the direct estimator. Thus an overestimation of $\hat{\sigma}_u^2$ will reduce the beneficial effect for the predictions in terms of MSPE since we fall back to the direct estimator.

This type of outlying observations can be seen as one of the main reasons for studying a robust extension to the FH model. Estimating the model parameters using robust estimation equations will reduce the impact of outlying areas and preserve the improvement in MSPE for the non-outlying areas. However as was pointed out by Bell and Huang (2006) this also means that we will have a poor fit for the outlying observations. This will be partially addressed by the bias correction presented in Section 3.4 which means that we fall back on the direct estimator for outlying areas which is only appropriate when the direct estimator has sufficient precision.

3.1.3 *The Creation of Overly Influencial Observations*

The third kind of outlying observations can best be understood when framed as *overly influencial observations*. They are not necessarily far away from the centre of the data, i.e. large in absolute values. Furthermore they cannot be understood by looking at the direct estimator alone but must be seen as pairs of direct estimator and sampling variance. Together they determine the impact of a single observation on the overall predictions under a FH model.

To be able to distinguish them from area and unit level outliers let us consider two hypothetical scenarios from which these values can arise:

- There are no outliers present and the unit level population model is correctly specified as a linear mixed model. However the sample sizes for most areas are very small and only for very few domains do we have sufficient sample size to support reliable direct domain predictions. This may result in heterogeneous area level information with respect to estimated sampling variances. And the effect can be that direct estimates of domains with sufficient sample size will dominate the global mean and thus effectively become the value we are shrinking against.

- We have unit level outliers influencing domain specific direct estimators. However instead of using the estimated sampling variance we smooth the variances against a global parameter – essentially shrinking both means and variances for example as suggested by Maiti et al. (2014). This will result in outlying domains for which, after smoothing, the direct estimators appear to be more reliable than they in reality are. In substance this reduces the self adjusting effect described in Section 3.1.1.

Both these scenarios are situations in which we have heterogeneous variance estimates and where the FH model may not yield stable and reliable results. Especially the second scenario should not be interpreted as an argument against smoothing or that we can ignore unstable variance estimation. Although the variances are not smoothed the influence of observations is nevertheless bounded within a robust estimation procedure. Here the robust model may present an alternative to smoothing strategies – although this may not have been the initial intention – without loosing the ability to weight down areas with unit level outliers. This claim is in fact not easy to support by empirical evidence. It will however be further discussed in a model based simulation in Section 6.3.

3.1.4 *Outlook*

In the following sections the robust estimation methodology of Section 2.5.1 is adapted for area level models. Results for this approach can also be found in Rao and Molina (2015, 146 ff) who derive robust estimation methodology for the FH model based on the results of Sinha and Rao (2009). The approach of Sinha and Rao (2009) is directly applicable to the FH model but is here extended, firstly, by proposing robust estimation equations for area level models with correlated random effects; and secondly, by proposing algorithms for the model parameters and predictions which do not rely on the Taylor series expansion used by Sinha and Rao (2009).

The reason for these extensions are that the original approach focused on the case where the variance structure of the random effects is a diagonal matrix; this is not so in the case of the spatial and temporal extensions. Furthermore the proposed Newton-Raphson algorithm for the random effects was found to be surprisingly unstable for area level models in that it manifested itself in thousands of iterations until numerical convergence was reached. A possible reason may be the ratio between observations and the random effects, e.g. where we have D areas and make D predictions. However a conclusive analysis of this problem is not at hand.

In addition simulation studies are conducted in order to complete the link between unit and area level outliers and their effect on the

area level predictions. This approach is discussed in more detail in Chapter 6.

3.2 EXTENSIONS TO THE FAY-HERRIOT MODEL

In the following sections the models under consideration are reviewed. This includes the methods reviewed in Section 2.3.2 as well as the FH model. These models will be framed in the context of linear mixed models since the robust estimation equations are – where possible – stated for the general case of linear mixed models. For the particular optimisation strategies of the model specific variance parameters it will be necessary to give some results for each model. In these situations the following sections will act as reference points in order to maintain the flow of the arguments.

3.2.1 *Fay-Herriot Model*

For a review of the FH model see Section 2.3.1. In what follows the model is framed as a linear mixed model in order to achieve a consistent notation for the robust parameter estimation. The model can be stated as:

$$\begin{aligned}\mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}\\ \mathbf{u} &\sim \mathcal{N}(\mathbf{o}_D, \mathbf{V}_u)\\ \mathbf{e} &\sim \mathcal{N}(\mathbf{o}_D, \mathbf{V}_e)\end{aligned}$$

where $\mathbf{y}$ denotes the $(D \times 1)$ vector of observed direct estimates; $\mathbf{X}$ the $(D \times P)$ design matrix; $\boldsymbol{\beta}$ the $(P \times 1)$ vector of regression coefficients; $\mathbf{Z} = \mathbf{I}_D$ is an identity matrix; $\mathbf{u}$ is the $(D \times 1)$ vector of random effects; and $\mathbf{e}$ is the $(D \times 1)$ vector of sampling errors. Furthermore $\mathbf{o}_D$ denotes a $(D \times 1)$ vector containing zeros; $\mathbf{V}_u = \mathbf{V}_u(\sigma_u^2) = \sigma_u^2 \mathbf{I}_{D\times D}$ is the variance covariance matrix of $\mathbf{u}$; and $\mathbf{V}_e = diag\left(\{\sigma_{ei}^2\}_{i=1}^D\right)$ with known sampling variances σ_{ei}^2. Furthermore $\mathbf{u}$ and $\mathbf{e}$ are independent. The unknown model parameters are $\boldsymbol{\beta}$ and $\boldsymbol{\delta} = (\sigma_u^2)$.

3.2.2 *Spatial Fay-Herriot Model*

For a review of the spatial extension to the FH model see Section 2.3.2.1. Here the model by Petrucci and Salvati (2006) in (2.10) is represented as linear mixed model and as such it can be represented as:

$$\begin{aligned}\mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}\\ \mathbf{u} &\sim \mathcal{N}(\mathbf{o}_D, \mathbf{V}_u)\\ \mathbf{e} &\sim \mathcal{N}(\mathbf{o}_D, \mathbf{V}_e)\end{aligned}$$

where the notation is the same as for the FH model with the following differences: $\mathbf{u} = \mathbf{u}_1$ with $\mathbf{u}_1$ being the $(D \times 1)$ vector of spatially correlated random effects following a SAR(1):

$$u_{1i} = \rho_1 \sum_{l \neq i} w_{il} u_{1l} + \epsilon_{1i}$$
$$\epsilon_{1i} \sim \mathcal{N}(0, \sigma_1^2) \text{ i.i.d.}$$

with $i = 1, \ldots, D$, where w_{il} are the elements of the row standardised proximity matrix $\mathbf{W}$. Let $\mathbf{W}^0$ denote the proximity matrix then one possible definition of $\mathbf{W}^0$ is that the elements are equal to one if area i and l are neighboured and zero otherwise – see Section 2.3.2.1 for a discussion of such definitions. This leads to the following distribution of $\mathbf{u}_1$:

$$\mathbf{u}_1 \sim \mathcal{N}\left(\mathbf{o}_D, \sigma_1^2 \boldsymbol{\Omega}_1(\rho_1)\right)$$

where

$$\boldsymbol{\Omega}_1 = \boldsymbol{\Omega}_1(\rho_1) = \left((\mathbf{I}_D - \rho_1 \mathbf{W})^\top (\mathbf{I}_D - \rho_1 \mathbf{W})\right)^{-1}$$

thus $\mathbf{V}_u = \sigma_1^2 \boldsymbol{\Omega}_1(\rho_1)$. Furthermore $\mathbf{u}_1$ and $\mathbf{e}$ are independent. The unknown model parameters are $\boldsymbol{\beta}$ and $\boldsymbol{\delta} = (\rho_1, \sigma_1^2)$. Because the results will be needed in later sections the partial derivatives of $\mathbf{V}(\boldsymbol{\delta})$ with respect to each element in $\boldsymbol{\delta}$ are given here: $\partial \mathbf{V} / \partial \sigma_1^2 = \boldsymbol{\Omega}_1(\rho_1)$ and

$$\frac{\partial \mathbf{V}}{\partial \rho_1} = -\sigma_1^2 \boldsymbol{\Omega}_1 \frac{\partial \boldsymbol{\Omega}_1^{-1}}{\partial \rho_1} \boldsymbol{\Omega}_1$$

with

$$\frac{\partial \boldsymbol{\Omega}_1^{-1}}{\partial \rho_1} = -\mathbf{W} - \mathbf{W}^\top + 2\rho_1 \mathbf{W}^\top \mathbf{W}.$$

3.2.3 *Temporal Fay-Herriot Model*

A review of temporal extensions can be found in Section 2.3.2.2. Here the representation of the temporal model by Rao and Yu (1994) in (2.12) is reviewed as a linear mixed model of the form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$
$$\mathbf{u} \sim \mathcal{N}(\mathbf{o}_{D+DT}, \mathbf{V}_u)$$
$$\mathbf{e} \sim \mathcal{N}(\mathbf{o}_{DT}, \mathbf{V}_e)$$

where $\mathbf{y}$ denotes the $(DT \times 1)$ vector of observed direct estimates; $\mathbf{X}$ the $(DT \times P)$ design matrix; $\boldsymbol{\beta}$ the $(P \times 1)$ vector of regression coefficients; $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2)$ where $\mathbf{Z}_1 = \mathbf{I}_D \otimes (\mathbf{1}_T)$, $\mathbf{1}_T$ is a $T \times 1$ vector of ones,

and $\mathbf{Z}_2 = \mathbf{I}_{DT}$; $\mathbf{u} = (\mathbf{u}_0, \mathbf{u}_2)$ is the $(D + DT \times 1)$ vector of random effects; and $\mathbf{e}$ is the $(DT \times 1)$ vector of sampling errors. Note that $\mathbf{u}_0$ denotes the vector of random effects similar to the FH model in that $\mathbf{u}_0 \sim \mathcal{N}(\mathbf{o}_D, \sigma_u^2 \mathbf{I}_D)$. $\mathbf{u}_2$ is a $(DT \times 1)$ vector of autocorrelated random effects following an AR(1):

$$u_{2it} = \rho_2 u_{2i,t-1} + \epsilon_{2it}$$

with $i = 1, \ldots, D$ and $t = 1, \ldots, T$, where ρ_2 is the autocorrelation coefficient with $|\rho_2| < 1$ and $\epsilon_{2it} \sim \mathcal{N}(0, \sigma_2^2)$. Furthermore $\mathbf{u}_0$, $\mathbf{u}_2$ and $\mathbf{e}$ are pairwise independent. The distribution of $\mathbf{u}_2$ is given by:

$$\mathbf{u}_2 \sim \mathcal{N}\left(\mathbf{o}_{DT}, \sigma_2^2 \mathbf{\Omega}_2(\rho_2)\right)$$

where $\mathbf{\Omega}_2 = \mathbf{\Omega}_2(\rho_2)$ is a block diagonal matrix in which each block is defined by:

$$\mathbf{\Omega}_{2i}(\rho_2) = \frac{1}{1-\rho_2^2} \begin{pmatrix} 1 & \rho_2 & \cdots & \rho_2^{T-2} & \rho_2^{T-1} \\ \rho_2 & 1 & & & \rho_2^{T-2} \\ \vdots & & \ddots & & \vdots \\ \rho_2^{T-2} & & & 1 & \rho_2 \\ \rho_2^{T-1} & \rho_2^{T-2} & \cdots & \rho_2 & 1 \end{pmatrix}_{T \times T}$$

with $i = 1, \ldots, D$. Thus $\mathbf{V}_u = \mathbf{V}_u(\delta) = diag\left(\sigma_u^2 \mathbf{I}_D, \sigma_2^2 \mathbf{\Omega}(\rho_2)\right)$ is the covariance matrix of $\mathbf{u}$; and $\mathbf{V}_e = diag\left(\{\{\sigma_{eit}^2\}_{t=1}^T\}_{i=1}^D\right)$ is the covariance matrix of $\mathbf{e}$ with known sampling variances for each time period and area. Hence the unknown model parameters are $\boldsymbol{\beta}$ and $\boldsymbol{\delta} = (\sigma_u^2, \rho_2, \sigma_2^2)$. The partial derivatives of $\mathbf{V}(\boldsymbol{\delta})$ with respect to the variance parameters are then given by: $\partial \mathbf{V}/\partial \sigma_u^2 = \mathbf{Z}_1 \mathbf{I}_D \mathbf{Z}_1^\top$, $\partial \mathbf{V}/\partial \sigma_2^2 = \mathbf{\Omega}_2$, and

$$\frac{\partial \mathbf{V}}{\partial \rho_2} = \sigma_2^2 diag\left(\left\{\frac{\partial \mathbf{\Omega}_{2i}}{\partial \rho_2}\right\}_{i=1}^D\right)$$

with

$$\frac{\partial \mathbf{\Omega}_{2i}}{\partial \rho_2} = \frac{1}{1-\rho_2^2} \begin{pmatrix} 0 & 1 & \cdots & \cdots & (T-1)\rho_2^{T-2} \\ 1 & 0 & & & (T-2)\rho_2^{T-3} \\ \vdots & & \ddots & & \vdots \\ (T-2)\rho_2^{T-3} & & & 0 & 1 \\ (T-1)\rho_2^{T-2} & \cdots & \cdots & 1 & 0 \end{pmatrix} + \frac{2\rho_2}{1-\rho_2^2} \mathbf{\Omega}_{2i}.$$

3.2.4 *Spatio-Temporal Fay-Herriot Model*

A review of this model can be found in Section 2.3.2.3. Here the model (2.14) by Marhuenda et al. (2013) is reviewed as a linear mixed model of the from:

$$\begin{aligned}\mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}\\ \mathbf{u} &\sim \mathcal{N}(\mathbf{o}_{D+DT}, \mathbf{V}_u)\\ \mathbf{e} &\sim \mathcal{N}(\mathbf{o}_{DT}, \mathbf{V}_e)\end{aligned}$$

where the notation is the same as for the temporal model with the following differences: here $\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2)$ where $\mathbf{u}_1$ denotes the spatially correlated random effect as also defined for the spatial model; and $\mathbf{u}_2$ the autocorrelated random effect as defined for the temporal model. Thus the variance structure of $\mathbf{u}$ is given by $\mathbf{V}_u = \mathbf{V}_u(\boldsymbol{\delta}) = \mathtt{diag}\left(\sigma_1^2\boldsymbol{\Omega}_1(\rho_1), \sigma_2^2\boldsymbol{\Omega}_2(\rho_2)\right)$ and hence the unknown model parameters are $\boldsymbol{\beta}$ and $\boldsymbol{\delta} = (\rho_1, \sigma_1^2, \rho_2, \sigma_2^2)$. The partial derivatives of $\mathbf{V}(\boldsymbol{\delta})$ with respect to each element in $\boldsymbol{\delta}$ are given by: $\partial\mathbf{V}/\partial\sigma_2^2$ and $\partial\mathbf{V}/\partial\rho_2$ are defined as for the temporal model; $\partial\mathbf{V}/\partial\sigma_1^2 = \mathbf{Z}_1\boldsymbol{\Omega}_1\mathbf{Z}_1^\top$ and

$$\frac{\partial\mathbf{V}}{\partial\rho_1} = -\sigma_1^2\mathbf{Z}_1\boldsymbol{\Omega}_1\frac{\partial\boldsymbol{\Omega}_1^{-1}}{\partial\rho_1}\boldsymbol{\Omega}_1\mathbf{Z}_1^\top$$

with

$$\frac{\partial\boldsymbol{\Omega}_1^{-1}}{\partial\rho_1} = -\mathbf{W} - \mathbf{W}^\top + 2\rho_1\mathbf{W}^\top\mathbf{W}.$$

3.3 ROBUST PREDICTIONS UNDER AREA LEVEL MODELS

The concrete approach for the outlier robust predictions under area level models is set out hereunder. The approach is kept in conformity with the general notation for linear mixed models where the spatio and temporal extensions can be framed as special cases within this class of models.

3.3.1 *Robust Estimation Equations*

In principle I follow the procedure given by Sinha and Rao (2009) which is to start with the derivatives with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ of the marginal log-likelihood of $\mathbf{y}$:

$$\mathbf{X}^\top\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = 0$$

and

$$(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\delta_l}\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) - \mathrm{tr}\left(\mathbf{V}^{-1}\frac{\partial\mathbf{V}}{\partial\delta_l}\right) = 0$$

with $l = 1, \ldots, Q$ where Q denotes the total number of variance parameters in δ. Sinha and Rao (2009) propose to *robustify* these equations by replacing them with:

$$\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \psi(\mathbf{r}) = 0 \tag{3.1}$$

and

$$\psi(\mathbf{r})^\top \mathbf{U}^{\frac{1}{2}} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta_l} \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \psi(\mathbf{r}) - \operatorname{tr}\left(\mathbf{K} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta_l} \right) = 0 \tag{3.2}$$

where $\mathbf{K} = \mathbb{E}(\psi_b^2(z))\mathbf{I}_n$ is a diagonal matrix of the same order as $\mathbf{V}$ with z following a standard normal distribution; and where $\mathbf{r} = \mathbf{r}(\boldsymbol{\beta}) = \mathbf{U}^{-\frac{1}{2}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ denotes the vector of the standardised residuals and $\mathbf{U} = \mathbf{U}(\delta)$ is the matrix containing the diagonal elements of $\mathbf{V}$. Furthermore $\psi_b(x) = x \min\left(1, \frac{b}{|x|}\right)$ is the influence function proposed by Huber (1964) for a given tuning constant b where a common choice for b is 1.345 – see Sinha and Rao (2009). These robust estimation equations have been utilised also for the area level models under consideration. A difference will be to find solutions to these equations as outlined later in this Section.

For the random effect, $\mathbf{u}$, Sinha and Rao (2009) propose to use the estimation equations of Fellner (1986) given in (2.21). This estimation equation is based on the derivative of the log-likelihood function of the joint probability distribution of $\mathbf{u}$ and $\mathbf{y}$ with respect to $\mathbf{u}$:

$$\mathbf{Z}^\top \mathbf{V}_e^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) - \mathbf{V}_u^{-1} \mathbf{u} = 0.$$

The idea of the robust version is now to restrict the influence of $\mathbf{u}$ and $\mathbf{e}$ in the estimation equation using the same influence function as before. However, instead of using the robust estimation equation of Fellner (1986):

$$\mathbf{Z}^\top \mathbf{V}_e^{-\frac{1}{2}} \psi\left(\mathbf{V}_e^{-\frac{1}{2}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) \right) - \mathbf{V}_u^{-\frac{1}{2}} \psi\left(\mathbf{V}_u^{-\frac{1}{2}} \mathbf{u} \right) = 0$$

I propose the following form:

$$\mathbf{Z}^\top \mathbf{V}_e^{-1} \mathbf{U}_e^{\frac{1}{2}} \psi\left(\mathbf{U}_e^{-\frac{1}{2}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}) \right) - \mathbf{V}_u^{-1} \mathbf{U}_u^{\frac{1}{2}} \psi\left(\mathbf{U}_u^{-\frac{1}{2}} \mathbf{u} \right) = 0 \tag{3.3}$$

where $\mathbf{U}_e$ and $\mathbf{U}_u$ denote diagonal matrices with the diagonal elements of $\mathbf{V}_e$ and $\mathbf{V}_u$, respectively. This modification is very similar to the idea in Sinha and Rao (2009) in that only the diagonal elements are used to standardise the residuals. In fact this modification is not necessary for the robust BHF model or for a robust FH model because for these models $\mathbf{V}_u$ is a diagonal matrix. This is not the case however in the setting of correlated random effects – e.g. in the case of a SAR(1) or AR(1) model – which seek to rely on this extension.

Let $\tilde{\boldsymbol{\beta}}^{\Psi}$ and $\tilde{\mathbf{u}}^{\Psi}$ define the solutions for known variance parameters, $\boldsymbol{\delta}$, to the robust estimation equations (3.1) and (3.3) respectively. Then the area level RBLUP can be defined as:

$$\tilde{\theta}_i^{\text{RBLUP}} = \mathbf{x}_i^{\top}\tilde{\boldsymbol{\beta}}^{\Psi} + \mathbf{z}_i^{\top}\tilde{\mathbf{u}}^{\Psi} \tag{3.4}$$

with $i = 1, \ldots, n$ where n can be the number of domains, D, or for the temporal and spatio-temporal model DT. Here $\mathbf{x}_i$ denotes the $(P \times 1)$ vector of auxiliary information for the ith observation and $\mathbf{z}_i$ similarly the column vector containing the ith row in $\mathbf{Z}$ where the form of $\mathbf{Z}$ depends on the specific model. Here (3.4) acts as a place-holder for the area level spatial RBLUP (SRBLUP), the temporal RBLUP (TRBLUP), and the spatio-temporal RBLUP (STRBLUP). The RBLUP depends on the unknown variance parameters, $\boldsymbol{\delta}$, which need to be estimated. If we replace $\boldsymbol{\delta}$ with the robust parameter estimates $\hat{\boldsymbol{\delta}}^{\Psi}$ we denote $\hat{\boldsymbol{\beta}}^{\Psi}$ and $\hat{\mathbf{u}}^{\Psi}$ as robust solutions and hence the REBLUP can be derived as:

$$\hat{\theta}_i^{\text{REBLUP}} = \mathbf{x}_i^{\top}\hat{\boldsymbol{\beta}}^{\Psi} + \mathbf{z}_i^{\top}\hat{\mathbf{u}}^{\Psi}. \tag{3.5}$$

Here the area level spatial REBLUP (SREBLUP), the temporal REBLUP (TREBLUP), and the spatio-temporal REBLUP (STREBLUP) constitute special cases of (3.5).

3.3.2 *Pseudolinear Representation*

The solutions to the estimation equations (3.1) and (3.3) are derived from the pseudolinear form of the RBLUP which follows a similar form as for the RBLUP under the unit level model by Chambers, Chandra, Salvati et al. (2014) – the latter was reviewed in Section 2.5.3.2. Following this idea the area level RBLUP (3.4) can be rewritten into:

$$\begin{aligned}
\tilde{\theta}_i^{\text{RBLUP}} &= \mathbf{x}_i^{\top}\tilde{\boldsymbol{\beta}}^{\Psi} + \mathbf{z}_i^{\top}\tilde{\mathbf{u}}^{\Psi} \\
&= \left(\mathbf{x}_i^{\top}\mathbf{A} + \mathbf{z}_i^{\top}\mathbf{B}\left(\mathbf{I}_n - \mathbf{X}\mathbf{A}\right)\right)\mathbf{y} \\
&= \tilde{\mathbf{w}}_i^{\top}\mathbf{y}
\end{aligned}$$

where $\tilde{\mathbf{w}}_i$ denotes the vector of weights. In the following I motivate the specific form of $\mathbf{A}$ and $\mathbf{B}$. Note that $\mathbf{A}(\tilde{\boldsymbol{\beta}}^{\Psi})\mathbf{y} = \tilde{\boldsymbol{\beta}}^{\Psi}$ and that

the specific form of $\mathbf{A}$ can be derived starting from the estimation equation in (3.1):

$$\begin{aligned}
\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \psi\left(\mathbf{U}^{-\frac{1}{2}}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right) &= 0 \\
\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \mathbf{W}_1 \mathbf{U}^{-\frac{1}{2}}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) &= 0 \\
\underbrace{\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{W}_1}_{\mathbf{W}_1^*}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) &= 0 \\
\mathbf{W}_1^* \mathbf{y} &= \mathbf{W}_1^* \mathbf{X}\boldsymbol{\beta} \\
\underbrace{\left(\mathbf{W}_1^* \mathbf{X}\right)^{-1} \mathbf{W}_1^*}_{\mathbf{A}} \mathbf{y} &= \boldsymbol{\beta}
\end{aligned}$$

where $\mathbf{W}_1 = \mathbf{W}_1(\boldsymbol{\beta}) = \mathrm{diag}\left(\{w_{1i}\}_{i=1}^n\right)$ is a diagonal matrix with the elements defined by:

$$w_{1i} = \psi\left(u_i^{-\frac{1}{2}}(y_i - \mathbf{x}_i^\top \boldsymbol{\beta})\right)\left(u_i^{-\frac{1}{2}}(y_i - \mathbf{x}_i^\top \boldsymbol{\beta})\right)^{-1}$$

where $u_i^{-\frac{1}{2}}$ is the ith diagonal element of $\mathbf{U}^{-\frac{1}{2}}$. The extension is possible because $\mathbf{U}$ is a diagonal matrix. Furthermore the form can be simplified because $\mathbf{U}^{\frac{1}{2}}$, $\mathbf{W}_1$, and $\mathbf{U}^{-\frac{1}{2}}$ are also diagonal matrices. Hence

$$\mathbf{A} = \mathbf{A}(\boldsymbol{\beta}) = \left(\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{W}_1 \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{V}^{-1} \mathbf{W}_1$$

which is similar to the form for the robust BHF model derived by Chambers, Chandra, Salvati et al. (2014) – see Section 2.5.3.2 for details.

To derive $\mathbf{B}$ note that:

$$\begin{aligned}
\tilde{\mathbf{u}}^\psi &= \mathbf{B}(\tilde{\mathbf{u}}^\psi)\,(\mathbf{I}-\mathbf{X}\mathbf{A})\,\mathbf{y} \\
&= \mathbf{B}(\tilde{\mathbf{u}}^\psi)\left(\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\beta}}^\psi\right)
\end{aligned}$$

and that the approach taken is similar as for $\mathbf{A}$. However the form of $\mathbf{B}$ is different from the approach by Chambers, Chandra, Salvati et al. (2014) because of the different estimation equations for the random effects (3.3). This different form simplifies the form of the two weighting matrices – $\mathbf{W}_2$ and $\mathbf{W}_3$ in the equations below – as they are also diagonal matrices. Thus starting from (3.3) the specific form of $\mathbf{B}$ can be derived as:

$$\begin{aligned}
\mathbf{Z}^\top \mathbf{V}_e^{-1} \mathbf{U}_e^{\frac{1}{2}} \psi\left(\mathbf{U}_e^{-\frac{1}{2}}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{Z}\mathbf{u})\right) - \mathbf{V}_u^{-1} \mathbf{U}_u^{\frac{1}{2}} \psi\left(\mathbf{U}_u^{-\frac{1}{2}}\mathbf{u}\right) &= 0 \\
\mathbf{Z}^\top \mathbf{V}_e^{-1} \mathbf{U}_e^{\frac{1}{2}} \mathbf{W}_2 \mathbf{U}_e^{-\frac{1}{2}}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{Z}\mathbf{u}) - \mathbf{V}_u^{-1} \mathbf{U}_u^{\frac{1}{2}} \mathbf{W}_3 \mathbf{U}_u^{-\frac{1}{2}}\mathbf{u} &= 0 \\
\underbrace{\mathbf{Z}^\top \mathbf{V}_e^{-1} \mathbf{W}_2}_{\mathbf{W}_2^*}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}-\mathbf{Z}\mathbf{u}) - \underbrace{\mathbf{V}_u^{-1} \mathbf{W}_3}_{\mathbf{W}_3^*}\mathbf{u} &= 0
\end{aligned}$$

where I continue with $\mathbf{W}_2^*$ and $\mathbf{W}_3^*$ to simplify the notation. As before $\mathbf{U}_e^{\frac{1}{2}}$, $\mathbf{U}_e^{-\frac{1}{2}}$, $\mathbf{U}_u^{\frac{1}{2}}$, and $\mathbf{U}_u^{-\frac{1}{2}}$ cancel out because they are diagonal matrices as are $\mathbf{W}_2$ and $\mathbf{W}_3$. Continuing:

$$\begin{aligned}\mathbf{W}_2^*(\mathbf{y}-\mathbf{X}\beta-\mathbf{Z}\mathbf{u}) &= \mathbf{W}_3^*\mathbf{u}\\ \mathbf{W}_2^*(\mathbf{y}-\mathbf{X}\beta) &= (\mathbf{W}_2^*\mathbf{Z}+\mathbf{W}_3^*)\mathbf{u}\\ \underbrace{(\mathbf{W}_2^*\mathbf{Z}+\mathbf{W}_3^*)^{-1}\mathbf{W}_2^*}_{\mathbf{B}}(\mathbf{y}-\mathbf{X}\beta) &= \mathbf{u}\end{aligned}$$

and thus

$$\mathbf{B}(\mathbf{u},\beta) = \mathbf{B} = \left(\mathbf{Z}^\top\mathbf{V}_e^{-1}\mathbf{W}_2\mathbf{Z}+\mathbf{V}_u^{-1}\mathbf{W}_3\right)^{-1}\mathbf{Z}^\top\mathbf{V}_e^{-1}\mathbf{W}_2$$

where $\mathbf{W}_2 = \mathbf{W}_2(\beta)$ is a $(n\times n)$ diagonal matrix with elements:

$$w_{2i} = \psi\left(u_{ei}^{-\frac{1}{2}}(y_i-\mathbf{x}_i^\top\beta-\mathbf{z}_i^\top\mathbf{u})\right)\left(u_{ei}^{-\frac{1}{2}}(y_i-\mathbf{x}_i^\top\beta-\mathbf{z}_i^\top\mathbf{u})\right)^{-1}$$

where $u_{ei}^{-\frac{1}{2}}$ denote the diagonal elements of $\mathbf{U}_e^{-\frac{1}{2}}$; and $\mathbf{W}_3 = \mathbf{W}_3(\mathbf{u})$ is a $(m\times m)$ diagonal matrix with elements:

$$w_{3j} = \psi\left(u_{uj}^{-\frac{1}{2}}u_j\right)\left(u_{uj}^{-\frac{1}{2}}u_j\right)^{-1}$$

where $u_{uj}^{-\frac{1}{2}}$ denote the diagonal elements of $\mathbf{U}_u^{-\frac{1}{2}}$ and m the number of elements in $\mathbf{u}$. Now we can state more precisely:

$$\mathbf{B}(\tilde{\mathbf{u}}^\psi,\tilde{\beta}^\psi)\left(\mathbf{y}-\mathbf{X}\tilde{\beta}^\psi\right) = \tilde{\mathbf{u}}^\psi.$$

This approach can be adapted for the REBLUP in (3.5) when we replace the model parameters by their estimators $\hat{\beta}^\psi$ and $\hat{\delta}^\psi$ and the prediction for $\mathbf{u}$ with $\hat{\mathbf{u}}^\psi$. The area level REBLUP is then given by:

$$\hat{\theta}_i^{\text{REBLUP}} = \hat{\mathbf{w}}_i^\top\mathbf{y}.$$

where $\hat{\mathbf{w}}_i$ denotes the vector of weights depending on the estimated model parameters. This form of representation leads to two important results. First it leads to the derivation of an iterative solution for the estimation equations (3.1) and (3.3) presented in the following Section; and second to an analytical MSPE estimator similar to that proposed by Chambers, Chandra and Tzavidis (2011) – this is presented in Section 3.5.2.

3.3.3 *Solutions for the Robust Estimation Equations*

3.3.3.1 *Solution for* β

For the solutions of the robust estimation equations I begin with the solution for the regression coefficients. Recall from the pseudolinear representation that:

$$\beta = \mathbf{A}(\beta)\mathbf{y}$$

and that $\hat{\boldsymbol{\beta}}^{\psi} = \mathbf{A}\left(\hat{\boldsymbol{\beta}}^{\psi}\right)\mathbf{y}$. This representation indicates that we have a fixed point function for $\boldsymbol{\beta}$ which, in this special case, leads to an iterative re-weighted least squares (IRWLS) algorithm to solve (3.1) for $\boldsymbol{\beta}$:

$$\begin{aligned}\boldsymbol{\beta}^{(m+1)} &= \mathbf{A}\left(\boldsymbol{\beta}^{(m)}\right)\mathbf{y} \\ &= \left(\mathbf{X}^{\top}\mathbf{V}^{-1}\mathbf{W}_1(\boldsymbol{\beta}^{(m)})\mathbf{X}\right)^{-1}\mathbf{X}^{\top}\mathbf{V}^{-1}\mathbf{W}_1(\boldsymbol{\beta}^{(m)})\mathbf{y} \end{aligned} \tag{3.6}$$

with adequate starting values $\boldsymbol{\beta}^{(0)}$ and known variance parameters $\boldsymbol{\delta}$. This proposition is very similar to the IRWLS algorithm suggested by Schoch (2012); however it depends on a different choice for the robust estimation equation. Details of IRWLS algorithms for robust M-estimation can also be found in Maronna et al. (2006, pp. 104-105). A natural choice as a starting value for the weighting matrix is to set the diagonal elements to one. This will lead to the non-robust generalised least squares estimator in the first iteration. This solution is acceptable since – as Maronna et al. (2006, pp. 104-105) note – for a continuous choice of ψ the algorithm is non-sensitive to the starting values with respect to convergence. Bad starting values can however increase the number of iterations. This property is investigated in Section 4.2. Note also that this iterative algorithm has the desirable property that it reduces to the GLS estimator if we set the tuning constant of ψ_b to infinity because of the special form of $\mathbf{W}_1$.

3.3.3.2 *Solution for* $\mathbf{u}$

Similarly a fixed point function for $\mathbf{u}$ can be defined by:

$$\mathbf{u} = \mathbf{B}(\mathbf{u})\left(\mathbf{y} - \mathbf{X}\beta\right)$$

where we can define an algorithm for known regression coefficients, $\boldsymbol{\beta}$, and variance parameters, $\boldsymbol{\delta}$:

$$\begin{aligned}\mathbf{u}^{(m+1)} =& \mathbf{B}(\mathbf{u}^{(m)})\left(\mathbf{y} - \mathbf{X}\beta\right) \\ =& \left(\mathbf{Z}^{\top}\mathbf{V}_e^{-1}\mathbf{W}_2(\mathbf{u}^{(m)})\mathbf{Z} + \mathbf{V}_u^{-1}\mathbf{W}_3(\mathbf{u}^{(m)})\right)^{-1} \\ &\times \mathbf{Z}^{\top}\mathbf{V}_e^{-1}\mathbf{W}_2(\mathbf{u}^{(m)})\left(\mathbf{y} - \mathbf{X}\beta\right)\end{aligned} \tag{3.7}$$

with suitable starting values $\mathbf{u}^{(0)}$. The starting values can be chosen similarly to the IRWLS algorithm in that we set the diagonal elements in $\mathbf{W}_2$ and $\mathbf{W}_3$ to one. However what I propose here is to use the non-iterative but outlier robust estimator of (2.24):

$$\hat{\mathbf{u}}^{\psi} = \mathbf{V}_u\mathbf{Z}^{\top}\mathbf{V}^{-1}\mathbf{U}^{\frac{1}{2}}\psi(\mathbf{r})$$

to compute starting values. This estimator has been reviewed in Section 2.5.1 and can be found in Rao and Molina (2015, p. 196). Like the iterative solution in (3.7) it also depends on the known parameter

vectors $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$. Also non-robust starting values showed acceptable results in the simulation presented in Section 4.2; however robust starting values have a significant impact on the number of iterations needed when outliers are present in the data. Similar to the IRWLS algorithm for $\boldsymbol{\beta}$, if the tuning constant of the influence function is chosen to be infinity the algorithm in (3.7) has the desirable property that it reduces to the form of the non-robust estimation equation of (3.3) because of the special form of $\mathbf{W}_2$ and $\mathbf{W}_3$, i.e. they reduce to identity matrices. And – as reviewed in Section 2.5.1 – the solution to the non-robust equation (3.3) is equivalent to the BLUP for $\mathbf{u}$ given by (2.4).

3.3.3.3 *Solutions for* $\boldsymbol{\delta}$

An algorithm to solve the robust estimation equation (3.2) for $\boldsymbol{\delta}$ was derived by Chatrchi (2012) for the case of the robust BHF model in (2.25). This approach is adapted for the robust area level models under consideration to derive fixed point equations for the variance parameters based on (3.2). In order to stay as general as possible and yet present simple solutions it is necessary to distinguish between two cases. First when $\boldsymbol{\delta}$ is a scalar under a given model. Second when $\boldsymbol{\delta}$ has more than one element. This will lead to a form in which the solutions for the spatial and temporal extensions are just special cases. When we set $\boldsymbol{\delta} = \delta$, i.e. the first case, (3.2) can be written as:

$$
\begin{aligned}
0 &= \psi(\mathbf{r})^\top \mathbf{U}^{\frac{1}{2}} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta} \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \psi(\mathbf{r}) - \operatorname{tr}\left(\mathbf{K}\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta} \delta^{-1} \delta \right) \\
\implies \operatorname{tr}\left(\mathbf{K}\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta} \delta^{-1} \delta \right) &= \psi(\mathbf{r})^\top \mathbf{U}^{\frac{1}{2}} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta} \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \psi(\mathbf{r}) \\
\implies \delta &= \underbrace{\operatorname{tr}\left(\mathbf{K}\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta} \delta^{-1} \right)^{-1}}_{c_1^{-1}(\delta)} \underbrace{\psi(\mathbf{r})^\top \mathbf{U}^{\frac{1}{2}} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta} \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \psi(\mathbf{r})}_{c_2(\delta)}
\end{aligned}
$$

where $c_1(\delta)$ and $c_2(\delta)$ are both scalar depending on δ. An iterative equation to solve for δ can now be formulated as:

$$
\delta^{(m+1)} = c_1^{-1}(\delta^{(m)}) c_2(\delta^{(m)}) \tag{3.8}
$$

for an adequate starting value $\delta^{(0)}$ and known regression coefficients $\boldsymbol{\beta}$. If (3.2) is a convex and continuous function the fixed point function converges against the true parameter regardless of the starting value – as long as they are chosen as values in the parameter space (Maronna et al., 2006, p. 328). Section 4.2 provides some empirical results on the matter.

The solution in 3.8 proved to be more robust with respect to starting values especially for the variance parameters. However for the correlation parameters in the spatial and temporal extensions the method often needs many iterations. On the other hand starting values are

not as critical since the correlation parameters are bounded between minus and plus one. Hence for the correlation parameters a Newton-Raphson algorithm is utilised instead:

$$\delta^{(m+1)} = \delta^{(m)} - \frac{d(\delta^{(m)})}{d'(\delta^{(m)})} \tag{3.9}$$

where $d(\cdot)$ denotes the robust estimation equation of (3.2) and $d'(\cdot)$ its first derivative with respect to the parameter δ. The simulation of Section 4.2 shows that a numerical approximation to the derivative is sufficient and already yields satisfying results. Thus the first derivative, $d'(\delta^{(m)})$, is approximated by $d(\delta^{(m)} - h)$ where h is some small number. For a review of such methods see for example Maronna et al. (2006, 46 ff) and more generally McCullough (2004).

Now turning to the more general case where δ has more than one element; the solution of Chatrchi (2012) is based on the idea to add $\mathbf{V}^{-1}\mathbf{V}$ to the term inside the trace in (3.2). This was necessary because under the unit level model, $\mathbf{V}_e$ contains an unknown parameter. In the case of the FH model and its extensions, $\mathbf{V}_e$ is assumed to be known and thus can be omitted. The following approach is specific if the parameters in $\mathbf{V}_e$ are known but can be extended easily to the more general case. Starting from (3.2) we can write:

$$\begin{aligned} 0 =& \psi(\mathbf{r})^\top \mathbf{U}^{\frac{1}{2}} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta_l} \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \psi(\mathbf{r}) \\ & - \operatorname{tr}\left(\mathbf{K}\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta_l} (\mathbf{Z}\mathbf{V}_u\mathbf{Z}^\top)^{-1} (\mathbf{Z}\mathbf{V}_u\mathbf{Z}^\top) \right) \end{aligned}$$

with $l = 1, \ldots, Q$ where we now need to find matrices $\bar{\mathbf{V}}_{ul}$ such that:

$$\begin{aligned} 0 =& \underbrace{\psi(\mathbf{r})^\top \mathbf{U}^{\frac{1}{2}} \mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta_l} \mathbf{V}^{-1} \mathbf{U}^{\frac{1}{2}} \psi(\mathbf{r})}_{c_{2l}(\delta)} \\ & - \operatorname{tr}\left(\underbrace{\mathbf{K}\mathbf{V}^{-1} \frac{\partial \mathbf{V}}{\partial \delta_l} (\mathbf{Z}\mathbf{V}_u\mathbf{Z}^\top)^{-1}}_{\mathbf{C}_l^*(\delta)} (\mathbf{Z}\bar{\mathbf{V}}_{u1}\mathbf{Z}^\top \quad \ldots \quad \mathbf{Z}\bar{\mathbf{V}}_{uQ}\mathbf{Z}^\top)\delta \right) \end{aligned}$$

with $l = 1, \ldots, Q$. Because the sum of matrices in the trace are square matrices we can now solve for δ. To simplify the solution we can set $Q = 2$ because this is the case relevant for the specific area level models as will be shown later:

$$\delta = \mathbf{C}_1^{-1}(\delta)\mathbf{c}_2(\delta) \tag{3.10}$$

where

$$\mathbf{C}_1(\delta) = \begin{pmatrix} \operatorname{tr}\left(\mathbf{C}_1^*(\delta)\mathbf{Z}\bar{\mathbf{V}}_{u1}\mathbf{Z}^\top\right) & \operatorname{tr}\left(\mathbf{C}_1^*(\delta)\mathbf{Z}\bar{\mathbf{V}}_{u2}\mathbf{Z}^\top\right) \\ \operatorname{tr}\left(\mathbf{C}_2^*(\delta)\mathbf{Z}\bar{\mathbf{V}}_{u1}\mathbf{Z}^\top\right) & \operatorname{tr}\left(\mathbf{C}_2^*(\delta)\mathbf{Z}\bar{\mathbf{V}}_{u2}\mathbf{Z}^\top\right) \end{pmatrix}$$

and

$$\mathbf{c}_2(\boldsymbol{\delta}) = \begin{pmatrix} c_{21}(\boldsymbol{\delta}) \\ c_{22}(\boldsymbol{\delta}) \end{pmatrix}.$$

Similarly to the case discussed earlier we can now state an iterative equation as:

$$\boldsymbol{\delta}^{(m+1)} = \mathbf{C}_1^{-1}(\boldsymbol{\delta}^{(m)})\mathbf{c}_2(\boldsymbol{\delta}^{(m)})$$

for adequate starting values $\boldsymbol{\delta}^{(0)}$ and known regression coefficients $\boldsymbol{\beta}$. Some adjustments to this solution need to be made to take the different definitions of $\mathbf{V}_u$ into account. For the specific form of $\partial\mathbf{V}/\partial\delta_l$ see the results in the respective section for each model, i.e. Sections 3.2.1, 3.2.2, 3.2.3, and 3.2.4. In the following sections the necessary adjustments are highlighted and specifically the special form of $\bar{\mathbf{V}}_{ul}$ are derived where necessary. The FH model is omitted since (3.8) is directly applicable when we set $\delta = \sigma_u^2$ and use the notation of Section 3.2.1.

SOLUTIONS FOR THE ROBUST SPATIAL FH MODEL

In principle $\boldsymbol{\delta} = (\rho_1, \sigma_1^2)$ under the spatial FH model of Section 3.2.2. Hence the specific form of (3.10) needs to be derived. In this case each $\bar{V}_{ul}$ with $l = 1, 2$ needs to be defined. Although it is possible to find such matrices, an implementation has lead to unstable results. This difficulty can be met by splitting the problem into finding solutions for $\delta_\sigma = (\sigma_1^2)$ and $\delta_\rho = (\rho_1)$ separately. This means that each iterative equation depends on the other *known* parameter. Admittedly this is numerically not optimal since it will lead to an increase in overall iterations. However, as will be shown in Section 4.2, it leads to acceptable and stable results. For the specific adaptation of the iterative equations for $\boldsymbol{\delta}$ this means that the iterative fixed point function (3.8) is used to find solutions for δ_σ and the Newton-Raphson algorithm in (3.9) for δ_ρ separately; all other specifications under the spatial model can be found in Section 3.2.2.

SOLUTIONS FOR THE ROBUST TEMPORAL FH MODEL

Under the temporal model of Section 3.2.3 the variance parameters are defined as $\boldsymbol{\delta} = (\sigma_u^2, \rho_2, \sigma_2^2)$. Although not all combinations have been empirically validated the same numerical problems may be suspected to arise when we try to optimise the variance and the correlation parameter jointly. Hence in this case the problem is split into finding solutions for $\boldsymbol{\delta}_\sigma = (\sigma_u^2, \sigma_2^2)$ and $\delta_\rho = (\rho_2)$. Both iterative equations assume known parameters of the left out elements.

For δ_ρ we can simply use the Newton-Raphson algorithm in (3.9) which is directly compatible with the results of Section 3.2.3. For $\boldsymbol{\delta}_\sigma$ we need to find $\bar{\mathbf{V}}_{u1}$ and $\bar{\mathbf{V}}_{u2}$ where ρ_2 is treated as

known parameter. Using the results of Section 3.2.3 we can find these solutions:

$$\begin{aligned}\mathbf{V}_u &= \begin{pmatrix} \sigma_u^2 \mathbf{I}_D & \mathbf{o}_{D\times DT} \\ \mathbf{o}_{DT\times D} & \sigma_2^2 \mathbf{\Omega}_2 \end{pmatrix} \\ &= \sigma_u^2 \begin{pmatrix} \mathbf{I}_D & \mathbf{o}_{D\times DT} \\ \mathbf{o}_{DT\times D} & \mathbf{o}_{DT\times DT} \end{pmatrix} + \sigma_2^2 \begin{pmatrix} \mathbf{o}_{D\times D} & \mathbf{o}_{D\times DT} \\ \mathbf{o}_{DT\times D} & \mathbf{\Omega}_2 \end{pmatrix} \\ &= \begin{pmatrix} \bar{\mathbf{V}}_{u1} & \bar{\mathbf{V}}_{u2} \end{pmatrix} \begin{pmatrix} \sigma_u^2 \\ \sigma_2^2 \end{pmatrix}\end{aligned}$$

and with the definition of $\bar{\mathbf{V}}_{u1}$ and $\bar{\mathbf{V}}_{u2}$ read in conjunction with the results in Section 3.2.3 the iterative equation (3.10) can be used.

SOLUTIONS FOR THE ROBUST SPATIO-TEMPORAL FH MODEL

For the solutions under the spatio-temporal model of Section 3.2.4 we recall that $\boldsymbol{\delta} = (\rho_1, \sigma_1^2, \rho_2, \sigma_2^2)$. By the same reasoning as before we can split the problem into finding solutions for $\boldsymbol{\delta}_\sigma = (\sigma_1^2, \sigma_2^2)$ and $\boldsymbol{\delta}_\rho = (\rho_1, \rho_2)$. Beginning with deriving the matrices for $\boldsymbol{\delta}_\sigma$ with known parameters $\boldsymbol{\delta}_\rho$ we can take a similar approach as for the temporal model:

$$\begin{aligned}\mathbf{V}_u &= \begin{pmatrix} \sigma_1^2 \mathbf{\Omega}_1 & \mathbf{o}_{D\times DT} \\ \mathbf{o}_{DT\times D} & \sigma_2^2 \mathbf{\Omega}_2 \end{pmatrix} \\ &= \sigma_1^2 \begin{pmatrix} \mathbf{\Omega}_1 & \mathbf{o}_{D\times DT} \\ \mathbf{o}_{DT\times D} & \mathbf{o}_{DT\times DT} \end{pmatrix} + \sigma_2^2 \begin{pmatrix} \mathbf{o}_{D\times D} & \mathbf{o}_{D\times DT} \\ \mathbf{o}_{DT\times D} & \mathbf{\Omega}_2 \end{pmatrix} \\ &= \begin{pmatrix} \bar{\mathbf{\Omega}}_1 & \bar{\mathbf{\Omega}}_2 \end{pmatrix} \begin{pmatrix} \sigma_1^2 \\ \sigma_2^2 \end{pmatrix}.\end{aligned}$$

With this result we can now adapt the iterative equation (3.10) when we set $\bar{\mathbf{V}}_{u1} = \bar{\mathbf{\Omega}}_1$ and $\bar{\mathbf{V}}_{u2} = \bar{\mathbf{\Omega}}_2$ and use the results in Section 3.2.4.

Similarly we can derive solutions for $\boldsymbol{\delta}_\rho$ with known parameters $\boldsymbol{\delta}_\sigma$. In this case however the presented solution has not lead to a stable implementation. To enable and simplify future research the derivation is stated here:

$$\begin{aligned}\mathbf{V}_u &= \begin{pmatrix} \sigma_1^2 \rho_1 \mathbf{\Omega}_1^* & \mathbf{o}_{D\times DT} \\ \mathbf{o}_{DT\times D} & \sigma_2^2 \rho_2 \mathbf{\Omega}_2^* \end{pmatrix} \\ &= \rho_1 \begin{pmatrix} \sigma_1^2 \mathbf{\Omega}_1^* & \mathbf{o}_{D\times DT} \\ \mathbf{o}_{DT\times D} & \mathbf{o}_{DT\times DT} \end{pmatrix} + \rho_2 \begin{pmatrix} \mathbf{o}_{D\times D} & \mathbf{o}_{D\times DT} \\ \mathbf{o}_{DT\times D} & \sigma_2^2 \mathbf{\Omega}_2^* \end{pmatrix} \\ &= \begin{pmatrix} \bar{\mathbf{\Omega}}_1^* & \bar{\mathbf{\Omega}}_2^* \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}\end{aligned}$$

where $\boldsymbol{\Omega}_1^*$ can be derived by:

$$\begin{aligned}\boldsymbol{\Omega}_1(\rho_1) &= \left((\mathbf{I}_D - \rho_1\mathbf{W})^\top(\mathbf{I}_D - \rho_1\mathbf{W})\right)^{-1} \\ &= \boldsymbol{\Omega}_1\boldsymbol{\Omega}_1(\mathbf{I}_D - \rho_1\mathbf{W})^\top(\mathbf{I}_D - \rho_1\mathbf{W}) \\ &= \rho_1 \underbrace{\boldsymbol{\Omega}_1\boldsymbol{\Omega}_1(\mathbf{I}_D - \rho_1\mathbf{W})^\top\left(\rho_1^{-1}\mathbf{I}_D - \mathbf{W}\right)}_{\boldsymbol{\Omega}_1^*}\end{aligned}$$

and $\boldsymbol{\Omega}_2^*$ can be derived by:

$$\boldsymbol{\Omega}_2(\rho_2) = \rho_2 \underbrace{diag\left(\{\boldsymbol{\Omega}_{2i}^*\}_{i=1}^D\right)}_{\boldsymbol{\Omega}_2^*}$$

with

$$\boldsymbol{\Omega}_{2i}^* = \frac{1}{\rho_2 - \rho_2^3}\begin{pmatrix} 1 & \rho_2 & \cdots & \rho_2^{T-2} & \rho_2^{T-1} \\ \rho_2 & 1 & & & \rho_2^{T-2} \\ \vdots & & \ddots & & \vdots \\ \rho_2^{T-2} & & & 1 & \rho_2 \\ \rho_2^{T-1} & \rho_2^{T-2} & \cdots & \rho_2 & 1 \end{pmatrix}_{T\times T}.$$

As said before this solution can formally be used to adapt the iterative equation (3.10) when we set $\bar{\mathbf{V}}_{u1} = \bar{\boldsymbol{\Omega}}_1^*$ and $\bar{\mathbf{V}}_{u2} = \bar{\boldsymbol{\Omega}}_2^*$ in conjunction with the results in Section 3.2.4. At present it is unknown as to why this method does not yield stable results. This is why instead a stepwise Newton-Raphson algorithm is used: when we split the problem of finding solutions for $\boldsymbol{\delta}_\rho$ into $\delta_{\rho_1} = (\rho_1)$ and $\delta_{\rho_2} = (\rho_2)$ solutions can be found using the Newton-Raphson algorithm in (3.9) when the left out parameter is known. Although this result is unsatisfying the results in Section 4.2 show that this method produces stable and reliable results.

3.3.3.4 *Algorithm*

Given the previously adapted and proposed iterative equations we can now define the overall optimisation strategy. In fact the overall algorithm is similar to the approach suggested by Sinha and Rao (2009) and the literature reviewed in Section 2.5.1. Of course the concrete implementations to find the solutions for $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, and $\mathbf{u}$ are different.

1. Choose initial values $\boldsymbol{\beta}^{(0)}$ and $\boldsymbol{\delta}^{(0)}$ and set $m = 0$. Initial values for $\boldsymbol{\beta}$ are chosen by setting the diagonal elements in the weighting matrix to one. For $\boldsymbol{\delta}$ an arbitrary value from the respective parameter space can be chosen – see Section 4.2 for results.
2. Compute $\boldsymbol{\beta}^{(m+1)}$ using equation (3.6) with given parameters $\boldsymbol{\delta}$ until convergence is reached.

3. With the solution of 2. for $\boldsymbol{\beta}$ compute $\boldsymbol{\delta}^{(m+1)}$ using the model specific solutions for (3.8), (3.9), and (3.10). This step itself can involve several algorithms for the respective variance and correlation components of the variance structure as was enunciated in Section 3.3.3.3.
4. Set $m = m + 1$.
5. Repeat steps 2. to 4. until the set of parameters jointly reaches convergence. Declare these solutions to be the robust parameter estimates $\hat{\boldsymbol{\beta}}^{\Psi}$ and $\hat{\boldsymbol{\delta}}^{\Psi}$.
6. With the solutions $\hat{\boldsymbol{\beta}}^{\Psi}$ and $\hat{\boldsymbol{\delta}}^{\Psi}$ compute $\mathbf{u}^{(m+1)}$ using equation (3.7) until convergence is reached. Declare these solutions to be the robust predictors, $\hat{\mathbf{u}}^{\Psi}$. The starting values, $\mathbf{u}^{(0)}$, are computed using the robust predictor in (2.24).

The stopping rule utilised to determine whether convergence is reached is defined in a form which proved to be useful when we optimise many parameters at the same time. Remember that in the case of the random effects D or DT parameters are sought for at the same time. To account for this fact, let $\mathbf{p}$ denote a parameter vector of length P and p_l an element with $l = 1, \ldots, P$ then the stopping rule is defined as:

$$\log\left(1 + \frac{1}{P}\sum_{l=1}^{P}\frac{|\, p_l^{(m+1)} - p_l^{(m)}\,|}{\max(1, |\, p_l^{(m)}\,|)}\right) < \epsilon$$

where $p_l^{(m)}$ denotes a parameter value in the iteration m and ϵ is some small constant. This stopping rule may appear to be artificial but it is motivated for the following reasons: it shows good properties when many, say hundreds or thousands, parameters are optimised jointly. This is so because first it is scaled down by taking the mean over all parameters and second by taking the log. The one is added to avoid negative values of the logarithm. Inside the sum the relative absolute differences of the parameter estimates at iteration m and $m + 1$ are computed. In fact only the relative difference is used when the denominator is larger than 1; otherwise the absolute difference is being used. Here the advantage of the relative difference is that the numerical precision is independent of the magnitude of the true parameters – this is also pointed out by Weihs et al. (2014, pp. 143-144). Weihs et al. (2014) and McCullough (2004) also provide a discussion of the choice of such stopping rules. The properties and results of this algorithm is the subject matter of Section 4.2.

3.4 BIAS CORRECTION

The discussion in Section 2.5.2 already highlighted the importance of addressing the potential bias of a robust prediction when we suspect

non-symmetric outliers. Chambers, Chandra, Salvati et al. (2014) and Jiongo, Haziza et al. (2013) have suggested different corrections to the unit level REBLUP to take this prediction bias into account. However the methods reviewed are based on the estimation of the within domain prediction error; this approach is not directly applicable to area level models since we may only have a single *observation* per domain.

However the general problem remains. To find a solution we need to review the different kinds of outliers observable at the area level. In Section 3.1 we distiguished between two types: area level and unit level outliers. Following Bell and Huang (2006) area level outliers are values with large absolute values in the random effect and thus are values for which the synthetic regression estimator is a poor fit. Independently of this question Fay and Herriot (1979) refer to Efron and Morris (1971) and Efron and Morris (1972) who argue that the Bayes Estimator and the Emperical Bayes Estimator may improve the overall prediction performance but may be ill suited for specific domains. In essence this also describes the situation in which area level outliers are present in the data and we are driven to use a robust estimation procedure since we shrink the prediction towards a global conditional mean. Although this may improve the overall quality of the predictions it will not improve the prediction for the outlying domains.

In this setting we can address area level outliers in the same way as was already discussed by Fay and Herriot (1979). Following their results the idea is to bound the area level predictions using the known sampling variances – Efron and Morris (1971) and Efron and Morris (1972) refer to this strategy as a *limited translation estimator*. Hence we can modify the REBLUP of equation (3.5) into a bias corrected version:

$$\hat{\theta}_i^{\text{REBLUP}-\text{BC}} = \begin{cases} y_i - c & \text{if } \hat{\theta}_i^{\text{REBLUP}} < y_i - c \\ \hat{\theta}_i^{\text{REBLUP}} & \text{if } y_i - c \leqslant \hat{\theta}_i^{\text{REBLUP}} \leqslant y_i + c \\ y_i + c & \text{if } \hat{\theta}_i^{\text{REBLUP}} > y_i + c \end{cases} \tag{3.11}$$

with $i = 1, \ldots, n$ where c is a constant. Fay and Herriot (1979) suggest to set $c = \sigma_{ei}$ to balance the tradeoff between loss in efficiency and gain in reduced bias. This approach can be directly applied to the spatial and temporal extensions and thus to the area level REBLUP; the area level REBLUP-BC in (3.11) acts as a placeholder for all extensions.

Since we need the pseudolinear representation below, we can state (3.11) as:

$$\hat{\theta}_i^{\text{REBLUP}-\text{BC}} = \hat{\mathbf{w}}_i^{*\top} \mathbf{y}$$

where $\hat{\mathbf{w}}_i^{*\top}$ denotes the vector of weights given the parameter estimates for $\boldsymbol{\beta}$, $\boldsymbol{\delta}$, and the prediction of $\mathbf{u}$:

$$\hat{\mathbf{w}}_i^* = \begin{cases} (1 - c/y_i)\mathbf{1}(j = i) & \text{if } \hat{\theta}_i^{\text{REBLUP}} < y_i - c \\ \hat{\mathbf{w}}_i & \text{if } y_i - c \leqslant \hat{\theta}_i^{\text{REBLUP}} \leqslant y_i + c \\ (1 + c/y_i)\mathbf{1}(j = i) & \text{if } \hat{\theta}_i^{\text{REBLUP}} > y_i + c \end{cases}$$

where $\mathbf{1}(j = i)$ denotes a vector where the element $j = i$ is equal to one and zero otherwise.

This solution may be beneficial when we observe area level outliers. However it is unclear how this modification relates to the scenario in which we have asymmetric unit level outliers. In fact this is a situation which can hardly be targeted by an area level correction without additional information. In Chapter 6 this issue is addressed again in model based simulation studies.

3.5 MEAN SQUARED PREDICTION ERROR

In this Section the parametric bootstrap by Sinha and Rao (2009) is adapted for the methods under consideration. Also a pseudolinearisation-based MSPE estimator is derived based on the results by Chambers, Chandra and Tzavidis (2011).

3.5.1 *Parametric Bootstrap Methods*

To estimate the MSPE of the robust prediction under the BHF model in (2.25) Sinha and Rao (2009) have proposed a parametric bootstrap method – see Section 2.5.3.1 for a review. This method is here adapted for the case of robust area level models and – as before – kept in general notation to include all models under consideration.

With the robust parameter estimates, $\hat{\boldsymbol{\beta}}^{\Psi}$ and $\hat{\boldsymbol{\delta}}^{\Psi}$, the bootstrap samples are generated from:

$$\begin{aligned} \mathbf{y}^* &= \mathbf{X}\hat{\boldsymbol{\beta}}^{\Psi} + \mathbf{Z}\mathbf{u}^* + \mathbf{e}^* \\ \mathbf{u}^* &\sim \mathcal{N}(\mathbf{o}_n, \mathbf{V}_u(\hat{\boldsymbol{\delta}}^{\Psi})) \\ \mathbf{e}^* &\sim \mathcal{N}(\mathbf{o}_n, \mathbf{V}_e). \end{aligned}$$

Note that here n can either be set to the number of domains, D, or for the temporal extensions to DT. For each bootstrap sample the target statistic is computed as:

$$\theta^{*(b)} = \mathbf{X}\hat{\boldsymbol{\beta}}^{\Psi} + \mathbf{Z}\mathbf{u}^{*(b)}.$$

Using one of the robust area level models – including its respective bias corrected version – the target statistic is predicted using the bth bootstrap sample. Hence the MSPE estimator is given by:

$$\widehat{\text{MSPE}}\left(\hat{\theta}_i^{\Psi}\right) = \sum_{b=1}^{B}\left(\hat{\theta}_i^{\Psi(b)} - \theta_i^{*(b)}\right)^2$$

where $\hat{\theta}_i^{\Psi(b)}$ denotes the robust prediction for area i using the bth bootstrap sample and $\theta_i^{*(b)}$ denotes the true value of the respective bootstrap sample for area i. The extension by Jiongo, Haziza et al. (2013) to the bootstrap method of Sinha and Rao (2009) can be applied directly when we use non-robust parameter estimates to generate the bootstrap samples. An empirical investigation of this method can be found in Section 6.2 in a model based simulation and then again in Chapter 7 in a design based simulation.

3.5.2 *Pseudolinearisation-based MSPE*

In the following I propose to adapt the results of Chambers, Chandra and Tzavidis (2011) – the pseudolinearisation-based approach to MSPE estimation – for the models under consideration. This idea was adapted for the robust prediction under the BHF model in Chambers, Chandra, Salvati et al. (2014) and extended to account for the extra variability associated with the estimation of the variance parameters δ. In principle the extension would be desirable but is not being adapted for the models under consideration; this can be an avenue for further research but at this point it is beyond the scope of the present Thesis. A review of the pseudolinear MSPE estimator for the unit level REBLUP can be found in Section 2.5.3.2 above.

The MSPE estimator can be formulated as:

$$\widehat{\text{MSPE}}\left(\hat{\theta}_i^{\Psi}\right) = \widehat{\mathbb{V}}\left(\hat{\theta}_i^{\Psi} - \theta_i\right) + \widehat{\mathbb{B}}\left(\hat{\theta}_i^{\Psi}\right)^2 \tag{3.12}$$

where we are now interested in an estimator for the prediction variance and bias. Following Chambers, Chandra and Tzavidis (2011) we can formulate a robust predictor as the weighted sum of the direct estimators:

$$\hat{\theta}_i^{\Psi} = \mathbf{w}_i^{\top}\mathbf{y}$$

where $\mathbf{w}_i$ denotes the vector of fixed weights. *Fixed* here refers to the requirement that they do not depend on $\mathbf{y}$ and furthermore $\mathbf{w}_i^{\top}\mathbf{1}_n = 1$. The prediction bias can be derived as

$$\begin{aligned}\mathbb{E}(\hat{\theta}_i^{\Psi} - \theta_i) &= \sum_{k=1}^{n} w_{ik}\mathbb{E}(y_k) - \theta_i \\ &= \sum_{k=1}^{n} w_{ik}\theta_k - \theta_i \end{aligned} \tag{3.13}$$

where w_{ik} denotes the kth element in $\mathbf{w}_i$. And the prediction variance can be obtained by:

$$\begin{aligned}
\mathbb{V}\left(\hat{\theta}_i - \theta_i\right) =& \mathbb{V}\left(\sum_{k=1}^{n} w_{ik} y_k - (\mathbf{x}_i^\top \boldsymbol{\beta} + \mathbf{z}_i^\top \mathbf{u})\right) \\
=& \mathbb{V}\left(\sum_{k=1}^{n} w_{ik}(\mathbf{z}_k^\top \mathbf{u} + e_k) - \mathbf{z}_i^\top \mathbf{u}\right) \\
=& \mathbb{V}\left(\sum_{k=1}^{n} w_{ik}\mathbf{z}_k^\top \mathbf{u} + \sum_{k=1}^{n} w_{ik} e_k - \mathbf{z}_i^\top \mathbf{u}\right) \\
=& \mathbb{V}\left(\sum_{k=1}^{n} a_{ik}\mathbf{z}_k^\top \mathbf{u} + \sum_{k=1}^{n} w_{ik} e_k\right) \\
=& \sum_{k=1}^{n} a_{ik}^2 \mathbf{z}_k^\top \mathbf{V}_u \mathbf{z}_k + 2\sum_{k<l}\sum a_{ik} a_{il} \mathbf{z}_k^\top \mathbf{V}_u \mathbf{z}_l \\
&+ \sum_{k=1}^{n} w_{ik}^2 \sigma_{ek}^2
\end{aligned} \tag{3.14}$$

where $a_{ik} = (w_{ik} - I(k = i))$ and $I(k = i)$ denotes an indicator function which is one if $k = i$ and zero otherwise. Note that we need the term $\mathbf{z}_k^\top \mathbf{V}_u \mathbf{z}_k$ for the case of correlated random effects; and in fact it reduces to σ_u^2 under the FH model. The variance of the sum can be separated as above because $\mathbf{u}$ and $\mathbf{e}$ are independent under all models under consideration. The first two terms are the variance and covariance of the weighted sum of random effects – see for example Bamberg et al. (2012, p. 118) – and the remaining third term is the variance due to sampling which is here assumed to be independent between all areas. The difference to the approach of Chambers, Chandra and Tzavidis (2011) is that the leading two terms here are stated for the more general case of correlated random effects; and the third term is stated for a more specific case, i.e. when the sampling variances are known.

Following the results by Chambers, Chandra, Salvati et al. (2014) we can now find an estimator for the conditional MSPE of the REBLUP by using (3.14) with an estimated variance covariance structure $\widehat{\mathbf{V}}_u$ and the weights defined in Section 3.3.2 for the robust area level EBLUP. The bias can be estimated accordingly when we use an estimator $\hat{\theta}_i$ for θ_i in (3.13). Chambers, Chandra, Salvati et al. (2014) note that this estimator should be an unbiased estimator and they have suggested to use the unshrunken version of the respective method. For area level models it may be tempting to use the direct estimator since it is unbiased. However this strategy has lead to very unstable results. Instead the regression estimator is used.

This MSPE estimator is based on the assumption that the weights are *fixed*. This is of course not the case under the respective models and hence it is based on a *pseudo*-linear form. This approach directly

extends to the bias correction when we use the derived weights of Section 3.4 instead. In this case Chambers, Chandra, Salvati et al. (2014) recommend that the squared bias term in (3.12) can be omitted since the unit level correction leads to an approximately unbiased estimator. This may not be the case for the area level models since we cannot correct for a bias due to unit level asymmetric outliers. The estimator is evaluated in Section 6.2 where its properties are compared to the bootstrap methods presented earlier.

3.6 DISCUSSION

This Section gives a short overview of the main findings of this Chapter. Furthermore I will outline how the remaining parts of this Thesis connect to these findings. As a summary I want to highlight the following items:

- The approach by Sinha and Rao (2009) for the estimation of robust mixed linear models in the SAE field has been modified to account for the case correlated random effects.
- These findings have been combined with the results by Chambers, Chandra and Tzavidis (2011) and Chambers, Chandra, Salvati et al. (2014) to derive the area level REBLUPs – the area level SREBLUP, TREBLUP, and STREBLUP – in a pseudo linear form. The results are presented in a general form to include the general class of mixed linear models.
- This pseudolinear representation has lead to algorithms to find solutions for the robust estimation equations. An IRWLS algorithm for $\boldsymbol{\beta}$ has been derived as well as a fixed point algorithm for the random effects, $\mathbf{u}$.
- Furthermore a simple bias correction for the robust prediction has been adapted based on the results of Efron and Morris (1972).
- And finally the approach by Chambers, Chandra and Tzavidis (2011) to the estimation of the MSPE of domain predictions has been extended to be used with the robust area level models with correlated random effects. In this regard also the parameteric bootstrap method for the MSPE estimation by Sinha and Rao (2009) has been adapted although hardly any changes have been necessary in this case.

The various methods and extensions are now further investigated in two ways. First, it will be described how the robust predictors are implemented inside a software package in the R-language (R Core Team, 2015): `saeRobust` (Warnholz, 2016). In this case especially the stability of the algorithms are of interest – see Chapter 4. Second, these methods are used in model based – Chapter 6 – and design

based – Chapter 7 – simulation studies. These simulations aim at exploring the statistical properties of the robust predictions compared to their non-robust counterparts. These simulation studies are conducted using the R-language and a simulation framework developed alongside this Thesis. This framework is implemented in the R-package `saeSim` (Warnholz and Schmid, 2016) and some discussion regarding this implementation is provided in Chapter 5.

Part II

IMPLEMENTATION

4 ROBUST SMALL AREA ESTIMATION IN R

It is important to distinguish between an algorithm and its implementation. The former is a theoretical approach to a problem and leaves many practical details unanswered. The latter is how the approach is applied practically.

— McCullough (2004)

4.1 OUTLINE

In what follows the numerical properties of the software implementation are investigated. The robust estimators are implemented in the package `saeRobust` (Warnholz, 2016) for the R-language (R Core Team, 2015) which is available as supplementary material to this Thesis. The main reason for providing this package is to simplify the process of reproducing the results presented in Part iii. A second reason is to provide an initial version of a software package which is ready to be used in practice – although further extensions and research may be needed in order to provide a comprehensive suit of tools for conducting *real world* analyses. This Section provides a general discussion of the current version of the software but it does not aim to be a manual of instructions. The package itself provides documentation which can be used for that purpose.

At present several software packages are available for the R-language. Molina and Marhuenda (2015) have introduced the package `sae` which is a comprehensive collection of common unit level models, e.g. the BHF model reviewed in Section 2.4, and the non-robust spatial- and temporal extensions to the FH model reviewed in Section 2.3.2. Other packages focus on the implementation of single estimators: e.g. `saery` (Lefler et al., 2014) implements the EBLUP in (2.12) of Rao and Yu (1994). Robust methods in the SAE field are implemented in the package `rsae` (Schoch, 2014) which implements the methods introduced in Schoch (2012) reviewed in Section 2.5.1.

An implementation concerning robust methods for area level models is not available at present. Also the methods and advances reviewed in Section 2.5.1 have not been published in terms of software – except for the results by Schoch (2012). In this respect `saeRobust` aims to provide a first – but stable – version implementing the models under consideration. This may also prove useful as a vessel for robust unit level models since most of the software components involved are designed to be reused. E.g. the robust score functions, the proposed algorithms for the regression coefficients and random effects, as well

as more general features such as the pseudolinear representation of a robust estimator are directly reusable as functions.

In this Chapter some numerical problems and in general the stability of the software implementation is investigated. This analysis is strongly influenced by the ideas in Weihs et al. (2014, 87 ff) and Zielke (1974) which have been invaluable in contributing towards a stable implementation. The idea utilised here is to create a numerically challenging scenario by imposing a high condition number in the testing data. Zielke (1974) has proposed simple testing matrices to this extend, and Weihs et al. (2014) review a general procedure to test the solution in linear least squares problems. By contrast to these approaches we are interested in the creation of a scenario in which spatial and temporal structures need to be identified. Here I assume the correctness of the underlying subroutines in the `Matrix` package by Bates and Maechler (2016) and the R-language and focus on the unknown model parameters of these respective models.

Two scenarios are created and investigated in a Monte Carlo simulation study. Both suffer from bad starting values as the solutions should not be influenced by the choice of these values. In one of the scenarios we can then see if and to what extent results vary when we impose very extreme values in the data. In this respect the current implementation does find solutions under both scenarios without software failure; however this is a result of conducting this study by tuning the implementation. *Finding solutions* refers to the implemented algorithm being able to find solutions to the estimation equations in (3.1), (3.2), and (3.3). Hence some references are given to the value at the last iteration of these estimation equations. The statistical properties of the estimators are investigated separately in Chapter 6.

Some code examples are provided in Section 4.3 to illustrate the current state of the implementation. As stated earlier this software version may lack some software features which should be available during a data analysis. In this regard some open research questions and remarks are mentioned in Section 4.4.

4.2 STABILITY

To find a stable implementation various configurations of the algorithm have been investigated. This includes the order of the nested algorithms as well as starting values and boundaries. A main concern has been the stability with respect to software failure in combination with the ability to find solutions to the estimation equations. This leads to some choices which increase the number of iterations and hence the computational demand. The latter can be dramatically reduced at the cost of stability: some remarks in this respect are given in Section 4.4.

4.2.1 *Algorithm*

To avoid confusion in discussing the results some details on the implementation which are not part of the algorithms of Section 3.3.3 themselves are given here.

The implementations of algorithms where the parameter space is known to be restricted have modified return values. Thus the correlation parameter of the spatial and temporal extensions are bounded between $10^{-5} - 1$ and $1 - 10^{-5}$; and all variance parameters have a lower bound of 10^{-5} to avoid zero or negative variance parameters – see also the discussion in Rao and Molina (2015, 151 ff). These restrictions are mostly relevant in optimisations with bad starting values; and they mainly prevent software failures. In these situations they also enable the algorithm to find solutions at all.

The algorithm for the model parameters, i.e. the parameter in the fixed effects part and the variance components, is nested. This means that solutions for the regression coefficients are found in a separate algorithm; also solutions for the variance parameter are sought for in one or more individual algorithms. These algorithms are then iterated over until all model parameters jointly reach the stopping rule given in Section 3.3.3.4. This means that we observe an overall number of iterations as well as a number of iterations for each nested algorithm. Hence we have two maximum numbers of iterations as additional stopping rules: one for each nested algorithm and one for the overall optimisation. The value for ϵ in the convergence criterion given in Section 3.3.3.4 is set to 10^{-6} to guarantee sufficient numerical accuracy.

4.2.2 *Testing Scenarios*

To test the stability of the algorithms two scenarios may be compared. In both scenarios non optimal starting values are used. A main focus lies on the ability of the algorithms to find a solution in a given number of iterations. The overall number of iterations as well as the number of iterations in each nested algorithm is restricted to 100. The maximum number of iterations for the random effects is restricted to 1000 and some problems with this optimisation strategy are discussed below. 500 Monte Carlo repetitions are conducted for each scenario. The following two scenarios are compared:

- *Base*: The basic scenario is an area level scenario in which we draw random numbers from:

$$
\begin{aligned}
y_i &= 100 + 10x_i + \mathbf{z}_i^\top \mathbf{u} + e_i \\
x_i &\sim \mathcal{N}(0, 16) \text{ i.i.d.} \\
\mathbf{u} &\sim \mathcal{N}(0, \mathbf{V}_u) \\
e_i &\sim \mathcal{N}\left(0, \sigma_{ei}^2\right)
\end{aligned}
$$

  where $\mathbf{V}_u$ and $\mathbf{z}_i$ for each model are chosen correctly. This means that each model is specified with the correct variance structure during testing. The variance parameters in all the models are set to 100 and correlation parameters are set to 0.5. The sampling variances, σ_{ei}^2, are generated as an equidistant sequence of real numbers between 25 and 225. Furthermore they are time invariant in scenarios with a time dimension and are used as the known model parameters. The number of domains is set to $D = 40$ and the number of time periods – in scenarios with a time dimension – is set to $T = 10$. In scenarios with spatial correlation the proximity matrix is defined as type *rook* (Bivand et al., 2008, p. 250). This means that each domain has two neighbouring units; with an exception being the first and last domain which only have one neighbouring unit. This definition is used for data generation as well as during the parameter estimation.
- *Outlier*: The outlier scenario imposes deterministic outliers: i.e. they are not generated randomly but are fixed at a value of 10000 in e_i for 10 per cent of the domains. The outlier domains are chosen randomly in each Monte Carlo repetition to avoid an artificial scenario in combination with the values of σ_{ei}^2. The choice of the value 10000 is arbitrary and may give rise to a numerically challenging situation. However it is more extreme than scenarios typically considered in the literature when studying robust methods – see the corresponding Chapter 6 and referenced literature.

The starting values for the regression coefficients are computed by setting the values of the diagonal weighting matrix in the IRWLS algorithm to one,
yielding non-robust starting values. All variance parameters are set to one and correlation parameters to zero. Starting values for the random effects are computed using the non iterative but robust estimator of equation (2.24). The tuning constant for the influence function is fixed at 1.345.

4.2.3 *Results*

In order to avoid unnecessary repetition for each separate model, the solutions for the regression coefficients and random effects for the RFH models are discussed in more detail. The solutions for the spatial and temporal extensions then focus on the specific solutions for the variance components. Furthermore it should be noted that figures with kernel density estimates omit the labels of the estimated density, i.e. the y-axis. Because of the scale of the different entities investigated the concrete realisations of these values are non-informative; the figures should only be used as a descriptive tool to evaluate the solutions.

To begin with, consider Figures 4.1 and 4.2 which present the estimates of the regression coefficients of the robust estimation under the FH model. The distribution on the right side presents the corresponding values of the estimation equation (3.1) at the solution for the respective parameter. Since the algorithm aims to find the root of the estimation equation we should expect values close to zero. In this regard the IRWLS algorithm shows acceptable performance regardless of the scenario.

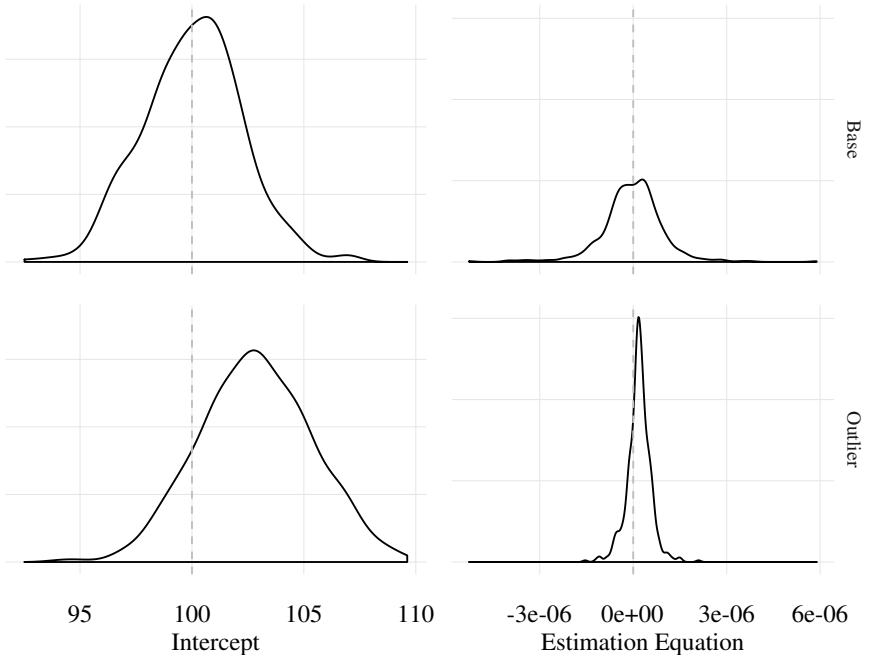

Figure 4.1: RFH – Parameter Estimates: Intercept – Robust parameter estimation under the FH model.

A distinction between the two scenarios manifests itself in the different estimations of the intercept parameter. Although a robust estimation technique is applied on average a higher intercept is estimated in the outlier scenario. The reason we observe this effect is that outlying domains are not removed but weighted down, hence on av-

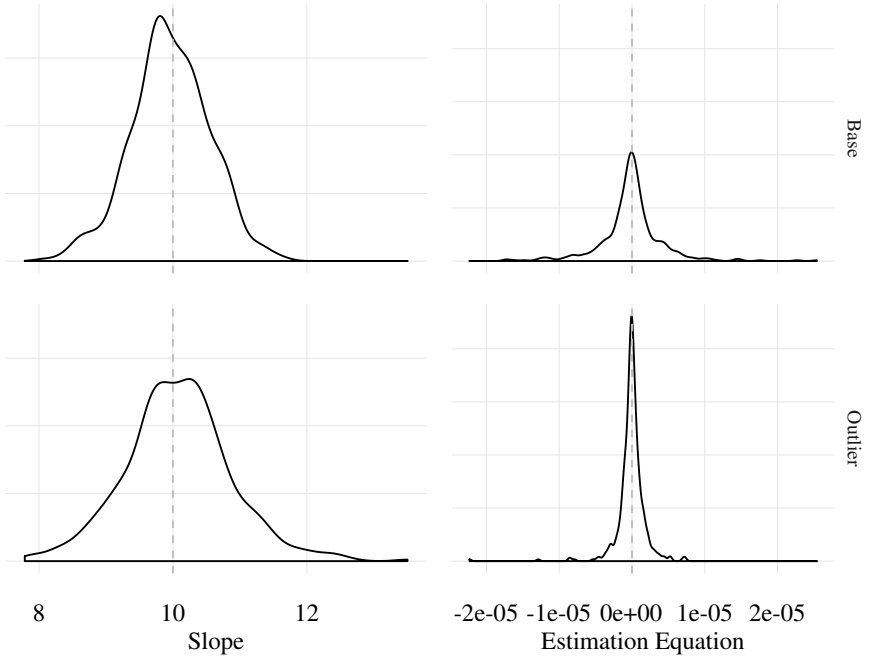

Figure 4.2: RFH – Parameter Estimates: Slope – Robust parameter estimation under the FH model.

erage more observations with higher values are present in the data which explains this effect.

The estimation of the variance parameter σ_u^2 yields stable results overall in the sense that the algorithms converge. A known issue also present here is that we can estimate values close to zero as can be seen in Figure 4.3. This effect is more visible under the *Base* scenario in combination with a situation where we have a non optimal ratio between observations and variability in the data resulting in a wide range of solutions – between 0 and 200. The right hand side of Figure 4.3 also indicates that the algorithm may reach the stopping rule; however the value of the estimation equation is not satisfactory – again we would expect values close to zero – when the solution for the parameter is close to zero. In this regard the value of the estimation equation may be a good indicator for the quality of the solution.

Similar to the case of the intercept we can observe on average higher estimates for the outlier scenario – see Figure 4.3. This may suggest an overestimation of the parameter despite the fact that we are using a robust method; but this can be explained by the increased variability under the outlier scenario. To put this result in perspective: when we estimate the variance with a non-robust method we would tend to reach values close to 10^5.

In all the results presented so far the stopping rule to indicate convergence has been reached. Neither the optimisation in the base scenario nor in the outlier scenario has reached the maximum of the

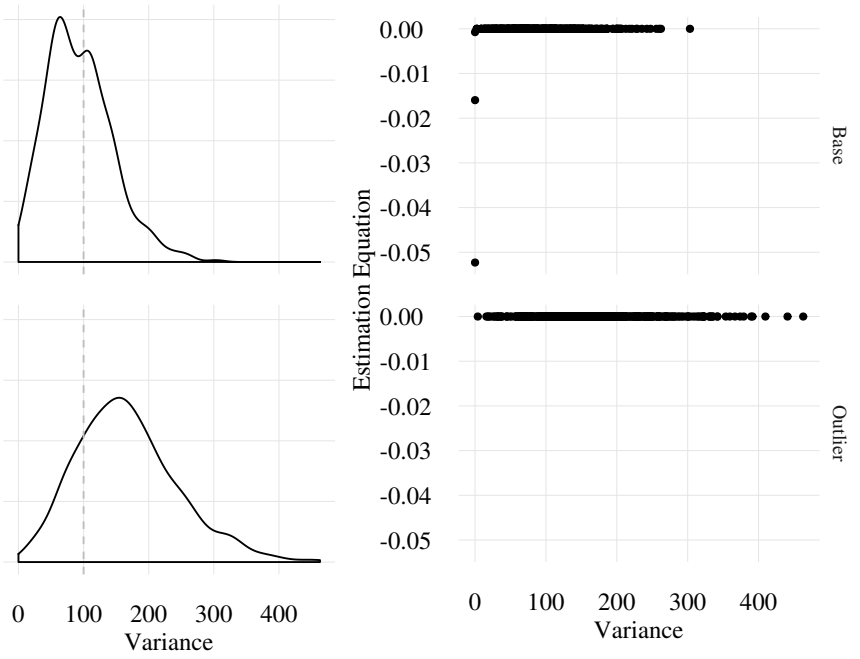

Figure 4.3: RFH – Parameter Estimates: Variance – Robust parameter estimation under the FH model.

	Scenario	First	Second	Remaining	Max	Converged
Overall	Base			3	28	1.000
	Outlier			3	14	1.000
- Coefficients	Base	9.0	6	2	26	
	Outlier	13.0	5	2	26	
- Variance	Base	53.5	28	7	100	
	Outlier	50.0	26	8	100	
Random Effect	Base			9	989	1.000
	Outlier			101	1000	0.992

Table 4.1: Median Number of Iterations in Optimisation until Convergence Was Reached. The columns *converged* contains the relative frequency of runs in which the stopping rule was reached before the maximum number of iterations.

allowed iterations – see Table 4.1. In most cases few iterations are needed.

These overall iterations represent only part of the solution. We can observe that the optimisation is more involved in the case of the outlier scenario to optimise the regression coefficients, i.e. more iterations are needed in the first overall iteration. This can be explained by starting from non-robust starting values which may result in large absolute values for the intercept. We can also see that the algorithm of the variance parameter takes approximately 50 iterations in the first overall iteration: in some cases even up to 100; however iterating between finding solutions for the regression coefficients and variance parameter adjusts itself only after a few overall iterations. There is no notable difference between the two scenarios; many iterations are needed in both cases where bad starting values are involved.

The solutions for the random effects on the other hand present more concern. Although in most cases 1000 iterations are sufficient we still observe very high median values, especially for the outlier scenario. With this effect in mind consider Figure 4.4 which shows the median of the predicted random effect in each solution. Given their distributional assumption we should expect values around zero which can be confirmed for the base scenario. Also there seems to be no relationship between the solution for the variance, σ_u^2, and the median random effect. Though this is appears to be different under the outlier scenario. One possible explanation is that the prediction of the random effect compensates the over estimation of the intercept – thus we observe negative median values of the random effects.

With respect to the relatively high median number of iterations for the random effects presented in Table 4.1 this solution can be unacceptable. In fact we not only observe a large number of iterations for the outlier scenario but in some cases also for the base scenario. To give some details regarding the underlying process consider Figure 4.5 which presents the first 10 iterations of 1000 of one Monte Carlo repetition – this is one selected repetition for the sake of illustration. Here we can compare the evolution of the fitted values and the corresponding values of the estimation equation (3.3) which, after all, is what should be close to zero at its solution. What we observe is that for most domain predictions the value of the estimation equation is close to zero only after a few iterations. Only for some domains are better solutions searched for. This effect becomes stronger as the number of domains increase. This suggests to investigate the solution for the random effects, i.e. the values of the estimation equation, after just a few iterations – especially when computation time is relevant.

Most of these effects are also present in the case of the spatial and temporal extensions to the FH model. Hence the main interest in the following discussion centres around the extensions with respect to the robust estimates of the variance parameters. When we look

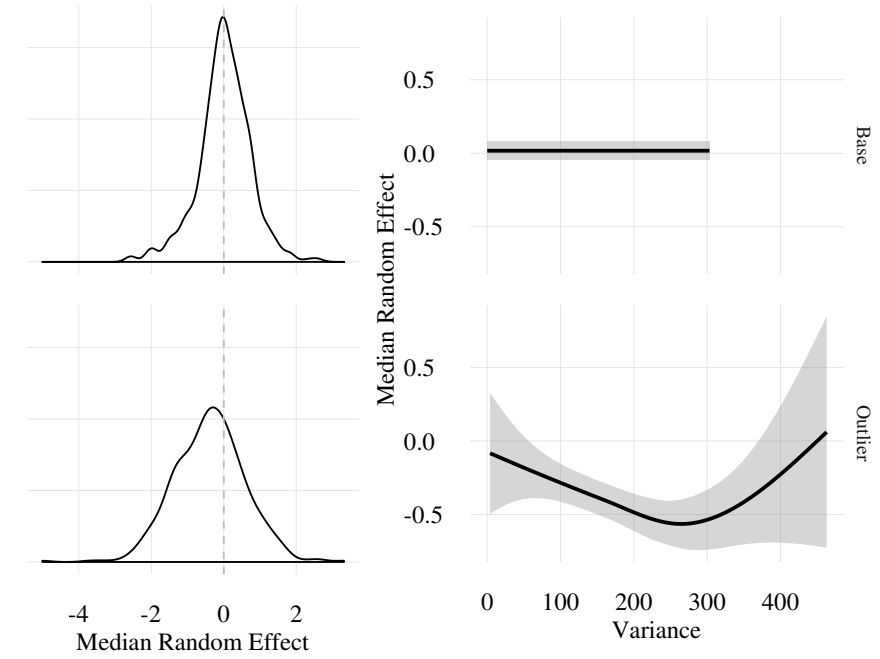

Figure 4.4: RFH – Median of Predicted Random Effects – Robust prediction under the FH model.

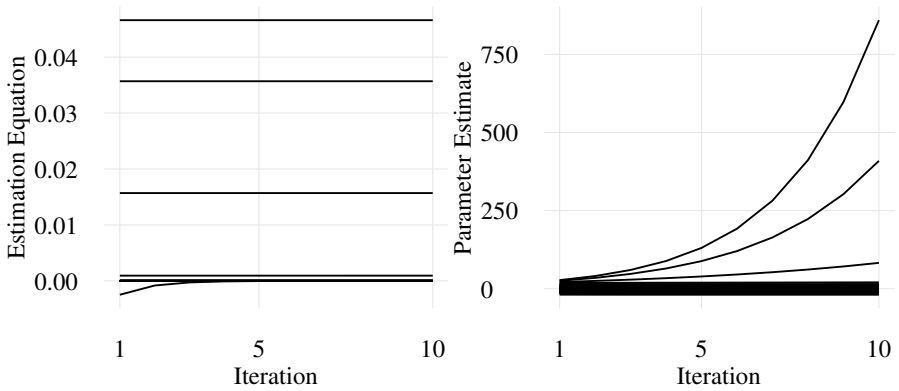

Figure 4.5: RFH – Convergence of Random Effects – Robust prediction under the FH model. First 10 iterations of 1000 total (no convergence according to stopping rule).

at Figure 4.6 we are directly comparing the estimates of the correlation parameter and the variance. Remember that these results are for a scenario based on a model with spatial correlation and a true correlation parameter of 0.5 with a variance of 100. Two effects can be seen similar to what was discussed earlier. First in the base scenario we estimate variance parameters close to zero which coincides with larger values for the estimation equation as before. Furthermore we observe a higher value for the variance parameter and a lower value for the correlation parameter in the outlier scenario. What happens is that the additional variation due to the outliers is captured by the variance component and the spatial correlation structure is somewhat shadowed.

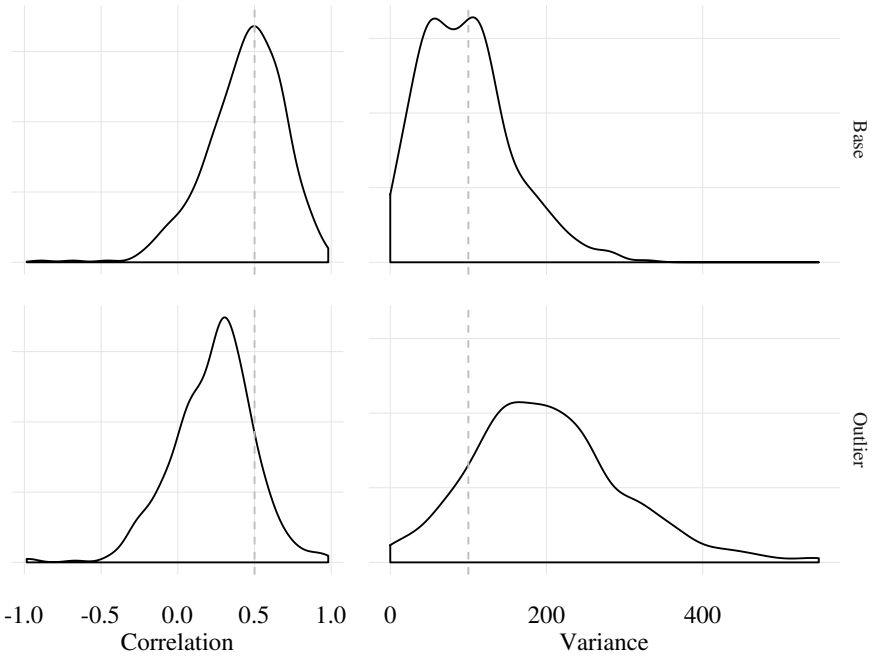

Figure 4.6: SFH – Parameter Estimates: Variance Components – Robust parameter estimation under the spatial FH model.

As stated earlier the overall algorithm has a maximum number of 100 iterations as well as has each nested algorithm – referred to as the *100 Iter* strategy. This is a setting suboptimal with respect to the number of iterations; however it is one which has proofed to be without a single software failure during the simulation. Figure 4.7 illustrates the choice when we set the maximum number of iterations of each nested algorithm to one – referred to as the *1 Iter* strategy. It is counter intuitive that the *100 Iter* strategy needs a lot more iterations but this is because the figure compares the results of each overall iteration. In fact it needs 1160 iterations to find a solution for the variance and 147 for the correlation parameter compared to 232 for the *1 Iter* strategy.

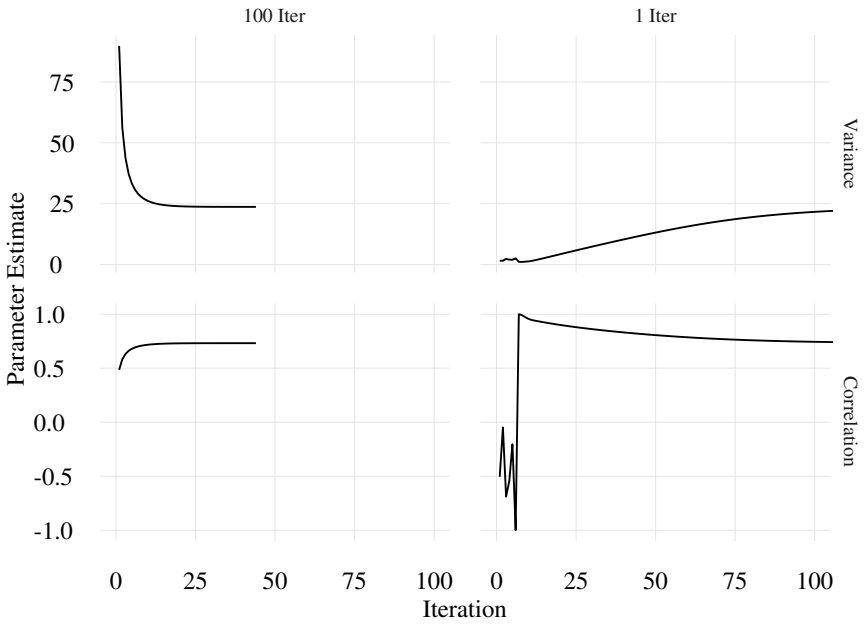

Figure 4.7: SFH – Parameter Estimates: Variance Components – Different optimisation strategies: *100 Iter* allows for a maximum of 100 nested iterations in each overall repetition; *1 Iter* allows only for 1 iteration in nested algorithms. Cumulative number of iterations for *100 Iter* is 1160 for the variance parameter and 147 for the correlation parameter.

Studying Figure 4.7 further we can observe how the choice of the number of iterations changes the overall behaviour. In the left panels we see that we start from a large initial estimate for the variance and a lower for the correlation parameter. This happens because in the first iteration all the variation is captured by the variance parameter since the correlation parameter is fixed at its initial value of zero. In each further iteration the relationship between correlation parameter and variance is balanced out. The evolution in the right hand side panels presents a different path since here each algorithm only has one step. In this setting the correlation parameter has a higher initial estimate and the variance a lower one. Both strategies yield approximately the same solutions as indicated by the figure. The panel on the right hand side for the correlation parameter also reveals that the algorithm can be unstable when we have bad starting values.

The *1 Iter* strategy often needs fewer iterations – which is a trade-off with a software failure rate of approximately 10 per cent in these testing scenarios. In preliminary tests it has often proofed useful to set the maximum number of iterations at some small number, say 10, to achieve a balance between the two settings. For applications this suggests that even if the algorithm fails, a different choice of the number of iterations may still provide useful results.

Finally still open for discussion are the temporal and spatio-temporal extensions. Figure 4.8 and 4.9 show respectively the parameter estimates of the extensions. A main overall effect is that the estimation of the variance components of the AR(1) process are influenced by outliers under both models. In contrast to the results seen earlier the estimation of the variance parameter of the random intercept and SAR(1) process is much more stable. One explanation for this effect may be that the random effect components, i.e. the random intercept and SAR(1) process, are constant over time during this simulation. This may result in a stronger signal of the structure opposed to the AR(1) process. In this setting the AR(1) process captures the higher variability in the data due to the outlying observations which masks the correlation structure in this case. What may also have an effect here is that we do not consider *area level* outliers; under area level outliers all observations belonging to one domain are outlying observations. However the testing framework only sets single observations to 10000. Hence these results may be artificial in their statistical properties.

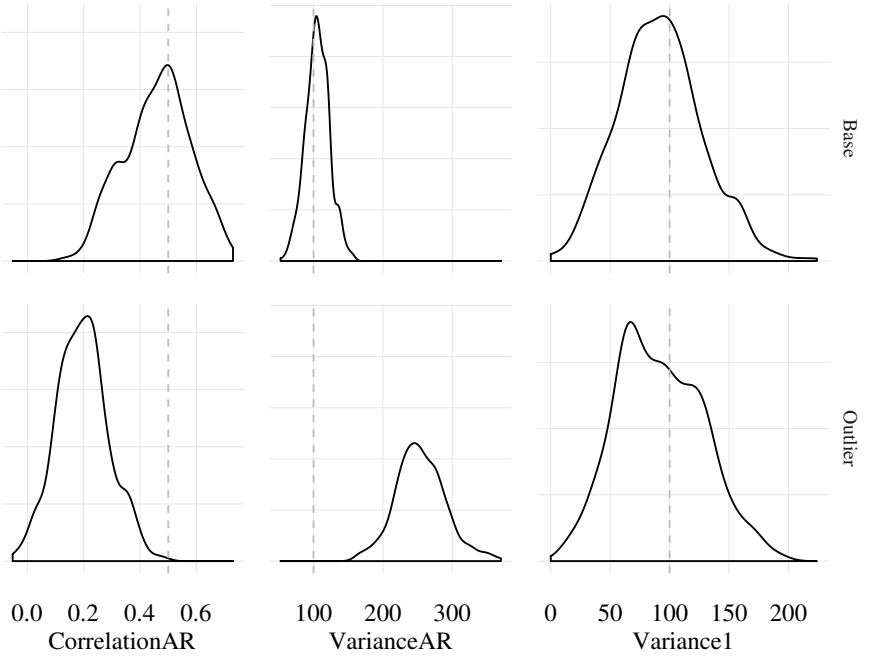

Figure 4.8: TFH – Parameter Estimates: Variance Components – Robust parameter estimation under the temporal FH model.

In conclusion we move to the number of iterations needed to find solutions for the spatial and temporal extensions which can be found in Table 4.2. Here we see that the stopping criterion has not been reached in all cases. The main reason for this result is the choice of the starting values. Allowing for more iterations leads to better results in all cases. One positive aspect is that the overall number of iterations is kept relatively low in most scenarios. Furthermore the

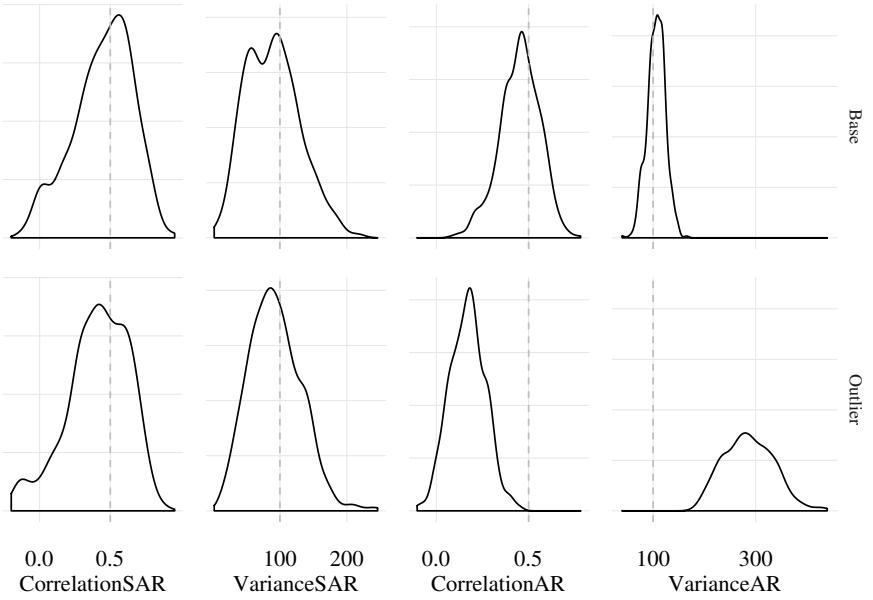

Figure 4.9: STFH – Parameter Estimates: Variance Components – Robust parameter estimation under the spatio-temporal FH model.

algorithm for the correlation parameter in the AR(1) as well as the SAR(1) shows very promising behaviour. Although they are based on a numeric approximation of the derivative of the estimation equation convergence is reached rapidly. The main reason may be the restriction of the parameter space to $\rho < |1|$ which makes it more robust against the choice of starting values.

4.3 CODE EXAMPLES

In this Section I present some code examples using the R-package `saeRobust` to illustrate the current state of its development. An important software package implementing the SFH and STFH as well as the basic FH model in the R-language is the `sae` package of Molina and Marhuenda (2015). They also provide several examples on how to use their implementation with various data sets. For this purpose I adapt these examples.

As an introductory example we can make use of the synthetic data set `grapes` (Molina and Marhuenda, 2015) which contains information on the geographical area for grape production in Tuscany, Italy. The data set comprises the information of the 274 municipalities in Tuscany. The target variable is the number of hectare used for grape production: `grapehect`. The sampling variance of this direct estimator is stored in the variable `var`; additional covariates are the area used for agriculture within each municipality: `area`; and the average number of working days in the reference year (2000): `workdays`. Additionally we have the proximity matrix of these municipalities stored in `grapesprox`.

	Scenario	First	Second	Remaining	Max	Converged
SFH - Overall	Base			7	100	0.996
	Outlier			5	100	0.996
- CorSAR	Base	5	4	3	100	
	Outlier	5	4	2	100	
- VarSAR	Base	44	25	15	100	
	Outlier	45	26	13	100	
TFH - Overall	Base			5	40	1.000
	Outlier			4	12	1.000
- CorAR	Base	5	4	2	8	
	Outlier	4	3	2	5	
- Variances	Base	36	23	8	52	
	Outlier	37	20	4	50	
STFH - Overall	Base			8	100	0.994
	Outlier			7	57	1.000
- CorSAR	Base	5	4	3	7	
	Outlier	5	4	3	7	
- CorAR	Base	4	4	2	7	
	Outlier	4	3	2	5	
- Variances	Base	35	25	9	63	
	Outlier	37	23	7	49	

Table 4.2: Median Number of Iterations in Optimisation until Convergence Was Reached. The columns *converged* contains the relative frequency of runs in which the stopping rule was reached before the maximum number of iterations.

Before I illustrate the incorporation of the spatial correlation structure, we can see how the basic, but robust, FH model can be estimated:

```
library("saeRobust")
data("grapes", package = "sae")
data("grapesprox", package = "sae")

fitRFH <- rfh(
  grapehect ~ area + workdays,
  data = grapes,
  samplingVar = "var"
)

fitRFH
```

```
##
## Call:
## rfh(formula = grapehect ~ area + workdays, data =
## grapes, samplingVar = "var")
##
```

```
##
## Coefficients:
## (Intercept)         area     workdays
##    -6.33547     -0.01069      0.52660
##
## Variance Components:
## variance
##    91.03
```

The function `rfh` is the central interface to the robust prediction under all models as will be illustrated below. We specify the fixed effects part of the model using a `formula` object in `R` and the sampling variance is identified by the variable name in the data set. The simple output we see here is in direct correspondence to a linear model in `R` with the additional information about the estimated variance parameters.

As we saw in the previous section it is relevant to check the number of iterations and the values of the estimation equations at their respective solutions:

```
summary(fitRFH)
```

```
##
## Call:
## rfh(formula = grapehect ~ area + workdays, data = grapes, samplingVar
## = "var")
##
##
## Coefficients:
## (Intercept)         area     workdays
##    -6.33547     -0.01069      0.52660
##
## Variance Components:
## variance
##    91.03
##
##                 Min.        1st Qu.     Median      Mean
## Random Effects  -2.835e+01  -2.264e+00   3.926e-02  -3.022e-02
## Residuals       -1.046e+02  -4.979e+00   6.298e-05   3.926e+00
##                 3rd Qu.     Max.
## Random Effects   2.643e+00   2.226e+01
## Residuals        5.066e+00   2.084e+02
##
##
## ## Solutions to the Robust Estimation Equations:
##
## (Intercept)         area     workdays     variance
##   2.924e-07    4.050e-04    7.079e-05    2.338e-07
##
## Random Effects:
##       Min.     1st Qu.      Median        Mean     3rd Qu.
## -3.234e-05   0.000e+00   0.000e+00  -1.160e-07   0.000e+00
##       Max.
##  2.888e-07
##
```

```
##
## ## Iterations:
## Allowed: 100 (10) -- 100
## model parameter   random effects
##              31                9
```

When we call the `summary` generic function we get detailed information on the solutions and the needed iterations. Here the values of the respective robust estimation equations are sufficiently close to zero. Overall the algorithm needed 31 iterations. Each nested algorithm has been restricted to 10 iterations. The optimiser for the random effects needed 9 iterations. With regard to the algorithm the user has the option to specify starting values, the number of iterations, and numerical tolerance used in the stopping criterion.

Some diagnostic plots can be obtained for the model residuals, the predictions, and the MSPE estimation. They are triggered using the `plot` generic function on the respective data type; that is the respective object containing the fitted model, the prediction, or the MSPE estimation. In the following presentation the returned plots are modified to fit into the general theme of this Thesis. Beginning with the residual plots we can use:

```
plot(fitRFH)
```

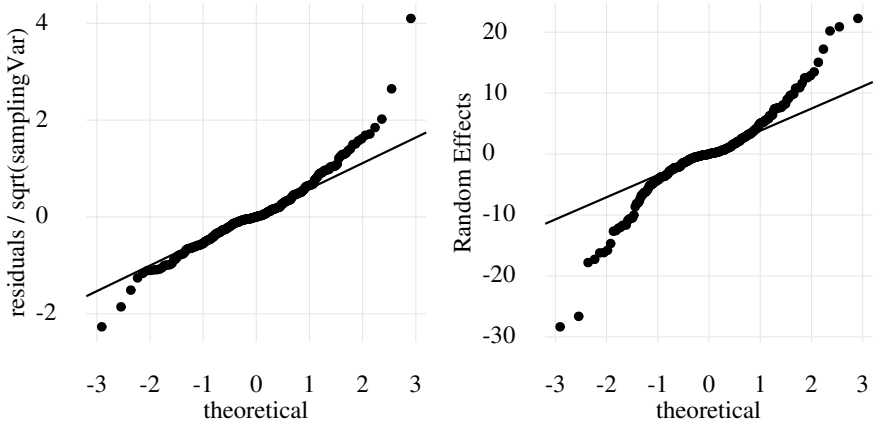

Figure 4.10: Quantile-Quantile Plots for Standardised Residuals (left) and Random Effects (right)

and obtain the quantile-quantile plots in Figure 4.10. Since the FH model and all extensions under consideration assume a heteroscedastic sampling distribution it is not meaningful to plot the estimated sampling errors – here denoted as residuals. Thus the standardised residuals are used instead:

$$\hat{e}_i^* = \hat{e}_i / \sigma_{ei}$$

where $\hat{e}_i$ is the realisation under the fitted model of the sampling error for area i. Note that this extends to all models when we set $i =$

$1, \ldots, n$ and n can be equal to D or DT for a given model. See Section 3.2 for the notation and models underpinning the above formulation.

Figure 4.10 shows here deviation from the normal distribution for residuals and random effects. In both plots we see more extreme values than we should expect; hence this is a situation in which the robust estimation technique may be beneficial. We can estimate the non-robust version by setting the tuning constant to some large value:

```
rfh(
  grapehect ~ area + workdays,
  data = grapes,
  samplingVar = "var",
  k = 10000
)
```

```
##
## Call:
## rfh(formula = grapehect ~ area + workdays, data = grapes, samplingVar
## = "var", k = 10000)
##
##
## Coefficients:
## (Intercept)         area     workdays
##    -5.75112     -0.01049      0.52206
##
## Variance Components:
## variance
##    97.43
```

What we can observe is that the intercept and variance parameter are estimated at a higher value than before. This is an indicates that outlying observations may have an effect on the parameter parameter estimation.

To obtain the predictions we can make use of the `predict` generic function. Here a `data.frame` is returned containing the realised random effects, the direct estimator, the REBLUP, and the bias corrected REBLUP:

```
predsRFH <- predict(fitRFH, c("reblup", "reblupbc"))
head(predsRFH)
```

```
##           re   direct   reblup reblupbc
## 1  0.4184722 30.94776 30.84878 30.84878
## 2 -3.8894206 57.21614 65.82273 65.82274
## 3 -3.7106460 73.75407 73.86895 73.86895
## 4 17.2199779 66.24203 63.22792 63.22792
## 5 -0.3898053 36.93180 37.20691 37.20691
## 6 -5.1509866 78.53393 78.54315 78.54315
```

Now we can plot these predictions using, again, the `plot` generic function:

```
plot(predsRFH)
```

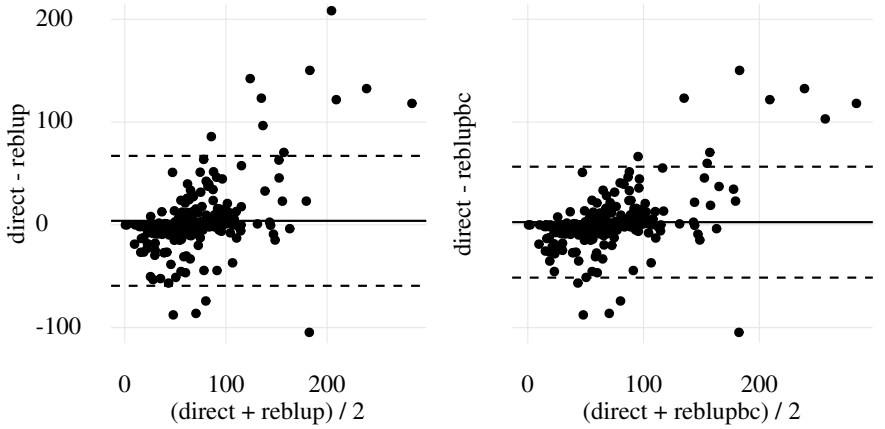

Figure 4.11: Bland-Altman Plots for Predictions – Solid line is the mean of the differences; dashed lines are the limits-of-agreements: mean plus / minus the 97.5% quantile of the standard normal distribution times the standard deviation of the differences.

and obtain the plots in Figure 4.11 where for each predictor a mean-difference plot is utilised to compare the direct predictor with the respective robust estimation technique. These plots are inspired by Bland and Altman (1986) who use a mean-difference plot to judge whether two methods of measurement are equivalent. The goal here is to compare a new method to a *gold standard* in the field of Medicine. Interpreting this more freely in the SAE field we want to achieve a similar feat: to understand the difference between the direct estimator and an alternative modelling strategy. As additional visual reference the solid line is the mean of the differences which may be expected to be close to zero when the REBLUP is unbiased. The *limits-of-agreement*, the dashed lines, indicate larger divergence between the two techniques. They are here constructed as the mean plus and minus the 97.5% quantile of the standard normal distribution times the standard deviation of the differences. We may suspect here that for larger values of the direct estimator we predict lower values using the REBLUP; although there are only few observations supporting this observation. Furthermore we see that the bias correction has little impact on the overall presentation.

Now turning to the estimation of the MSPE we make use of the function `mse` which returns a `data.frame` containing the direct estimator with the sampling variances as well as the REBLUP and the MSPE estimation using the pseudolinear-based approach of Section 3.5.2. The parametric bootstrap of Section 3.5.1 can be triggered when we set the argument `type` to `"boot"`:

```
mseRFH <- mse(fitRFH, type = "pseudo", predType = "reblup")
head(mseRFH)
```

```
##     direct   reblup samplingVar     pseudo
## 1 30.94776 30.84878  21.5305443 17.4731401
## 2 57.21614 65.82273 201.4299002 63.3683787
## 3 73.75407 73.86895   2.8182181  2.7346395
## 4 66.24203 63.22792  21.3808940 17.5369665
## 5 36.93180 37.20691  64.2439319 37.8925166
## 6 78.53393 78.54315   0.1629171  0.1626368
```

Again we can trigger some diagnostic plots using the `plot` generic function:

```
plot(mseRFH, ylim = c(0, 400))
```

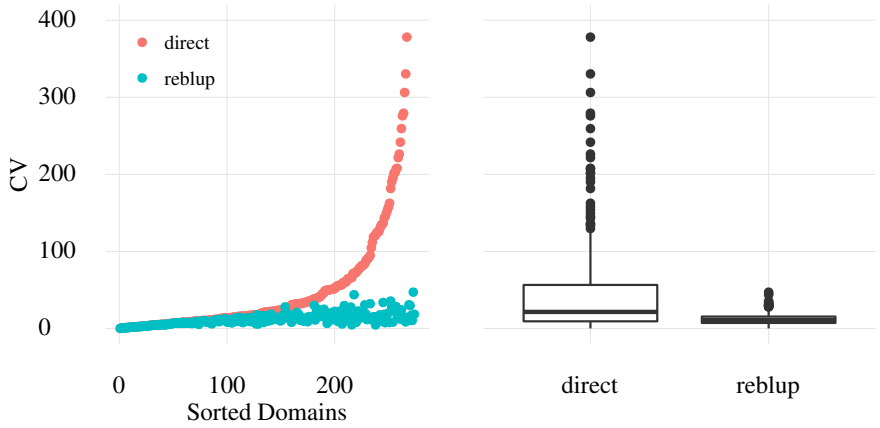

Figure 4.12: Plots for Coefficient of Variation – direct vs. REBLUP. The domains in the left panel are sorted for the CV of the direct estimator.

to obtain the plots in Figure 4.12. The plot on the left hand side is similar to the one provided by Molina and Marhuenda (2015). Here we plot the coefficient of variation (CV) against a sorted sequence of domain identifiers. As in Molina and Marhuenda (2015) the presentation is restricted to values which have a smaller CV than 400. What we can observe is a clear benefit of using the REBLUP for domain predictions; here the *very unreliable* domains – measured in terms of CV – benefit the most.

The above presented methods, that are the `mse`, `predict`, and `plot` methods, are implemeted for the robust estimation of the standard FH model – as presented above – and furthermore the spatial and temporal extensions. The same function names can be used. From the perspective taken in `saeRobust` we estimate a robust FH model with varying variance structures. Hence the specification of the variance is simply an argument to the function `rfh`. This design choice is close

to the functionality implemented for example in the package `nlme` (Pinheiro et al., 2016) for nonlinear mixed effects models.

As an example I present the same analysis as above for the robust estimation of the spatial FH model. From the previous analysis we have an initial estimate for the variance and furthermore we may suspect a positive spatial correlation between the municipalities hence we can supply some initial values:

```
fitRSFH <- rfh(
  grapehect ~ area + workdays,
  data = grapes,
  samplingVar = "var",
  correlation = corSAR1(as.matrix(grapesprox)),
  x0Var = c(0.5, 90)
)
fitRSFH
```

```
##
## Call:
## rfh(formula = grapehect ~ area + workdays, data = grapes, samplingVar
## = "var", correlation = corSAR1(as.matrix(grapesprox)), x0Var = c(0.5,
## 90))
##
##
## Coefficients:
## (Intercept)         area     workdays
##    -3.93102     -0.01176      0.51633
##
## Variance Components:
## correlation     variance
##      0.5456      65.5172
```

Remember that `grapeprox` is the row standardised proximity matrix also supplied by the package `sae`. In this case an additional parameter, `correlation`, is added to the output. The analysis can now be continued using the previously used methods but is omitted as there are no new features of `saeRobust` to be discovered. To trigger the estimation of the temporal model we can make use of `corAR1` and for the spatio-temporal model we can use `corSAR1AR1` to specify the respective correlation structure in the random effects.

Additional features which have been ommitted from the presentation so far are for example the possibility to compute the MSPE using the parameteric bootstrap. This can be done using the function `mse` and by setting the argument `type` to `"boot"`. As an example consider the following code:

```
mse(fitRSFH, type = "boot", B = 50)
```

As before, this method is implemented for all models under consideration; the *correct* bootstrap method is selected depending on the first argument to the method.

4.4 DISCUSSION

In principle the results of the stability tests are promising as in both scenarios acceptable solutions can be found when we set the number of iterations to a higher value. *Acceptable* here refers to a value close to zero at the solution for the respective estimation equation. This however presents a trade off between the number of iterations, stability, and computational demand, which can become relevant with temporal data. The key to computationally less demanding solutions is the choice of starting values and the restriction of the maximum number of iterations of the nested algorithms. To address this trade-off the implementation in `saeRobust` allows to set both parameters as well as the number of iterations for the optimisation of the random effects.

Starting values for the regression coefficients, the variance parameters, and the random effects can be supplied by the user. And related to computational demand users can run a model with only a subset of the data at hand to produce better starting values and then update the analysis and continue with an updated data set. This can be accomplished using the `update` function. This strategy was originally implemented to implement the parametric bootstrap method however it may also be valuable in situations with large data sets. Also it may be important to first optimise the model parameters and then find solutions for the random effects which is possible by continuing a model fitting process with updated parameters.

Furthermore it is important to investigate the estimation equations at their solutions. A common return value of such fitting procedures is to provide the reason of *convergence* – see for example the function `sae::eblupFH` in the R package `sae`. Such a value may indicate that the maximum number of iterations has been exceeded or that the convergence criterion has been reached. However as McCullough (2004) notes the fact that we reach the stopping rule can be very misleading – it is in fact the estimation equation we should evaluate and in addition the second derivative of the log-likelihood to ensure that we found indeed a maximum. This is supported and illustrated by the results above when the numerical solver reaches the stopping criterion but the value of the estimation equation is not approximately zero. E.g. this happens with variance estimates close to zero. For this reason a design choice in the package's output is to report the value of the estimation equation as well as each step during optimisation. However what is not currently provided is the possibility to evaluate the second derivative of the log-likelihood which may present a possible future extension.

Two practical issues which have not been discussed in the SAE literature are model selection and inference on parameter estimates for robust methods. The implementation in Schoch (2014) relies for

example on the asymptotic normality of the regression parameters. However there is little empirical evidence on using such results. Specifically for area level models where we may have only a few observations – e.g. 40 – it is not clear whether it is advisable to place reliance on such results. For this reason the current version of `saeRobust` only supports a link to the parametric bootstrap with the function `bootstrap`. With respect to model selection the actual log-likelihood function for robust methods is generally unknown since robust versions of its partial derivatives are used. Hence it is not clear how to provide information criteria for the robust methods under consideration – see also the discussion in Koller (2016).

The discussion in this Chapter focused on the properties of the implementation. The statistical properties of the robust methods have not been investigated so far and are the subject matter of the Chapters 6 and 7 in model and design based simulations. An important outcome of this Thesis are the accompanying software implementations. With this regard also the package `saeSim` (Warnholz and Schmid, 2016) can be seen as one result of this Thesis; it aims to simplify the process of setting up simulation studies and hence is introduced in the following Chapter 5.

5

SIMULATION TOOLS FOR SMALL AREA ESTIMATION

Instead of imagining that our main task is
to instruct a computer what to do,
let us concentrate rather on explaining
to human beings what we want a computer to do.

— Knuth (1992, p.99)

5.1 OUTLINE

In this Chapter I want to present a framework for simulation studies within the SAE field. The framework is implemented in the R-package `saeSim` (Warnholz and Schmid, 2016). It should be noted that the content here presented has in part been previously published as an Article: "Simulation Tools for Small Area Estimation: Introducing the R-package saeSim" – Warnholz and Schmid (2016). Here a shorter version of that Article is presented. Content specific to a general introduction to the SAE field is not included here; in contrast to the code examples presented in the Article here the SAE methods implemented in `saeRobust` (Warnholz, 2016) are applied instead of of the utilities in `sae` (Molina and Marhuenda, 2015). Furthermore where appropriate text passages have been altered to integrate the content into the general context of this Thesis.

The set of tools available in the package `saeSim` have been designed to provide an infrastructure which makes it easier to reproduce the results of model and design based simulation studies. Reproducibility here comprises the availability of the full academic research, including data and the source code. Open source tools like the R-language and LaTeX can be used in tight integration to combine the statistical analysis with the written words in an article. This can be achieved by using tools like `knitr` (Xie, 2013), `sweave` (Leisch, 2002), and, more recently, `rmarkdown` (Allaire et al., 2014). Such tools can assist in making research more reproducible. The tools provided by `saeSim` aim at simplifying the process of writing the source code for a simulation study; in the context of reproducible research this may prove to be useful in the development of script files as well as in combination with the tools discussed above.

Real data is often very sensitive and can be subject to strict confidentiality restrictions. Synthetic data generation mechanisms can be used to provide safe data which can be made publicly available – for a more detailed discussion see Rubin (1993), Alfons, Kraft et al.

(2011), and Kolb (2013). Burgard et al. (2014) describe this as an open research philosophy. Such synthetic data sets can be used to test newly proposed statistical methods in a close-to-reality framework. In general, simulation studies in statistics can be divided into two concepts:

- Design based: Here the simulation study is based on true or synthetic data of a fixed population. Samples are then selected repeatedly from the underlying finite population and different estimation methods are applied in each replication. The estimates so obtained are compared to the true population values in terms of, for instance, the relative bias (RBIAS) and relative root mean squared prediction error (RRMSPE).
- Model based: Here the simulation study uses data drawn directly from a model. In each iteration the population is generated from a model and a sample is selected according to a specific sampling scheme. The sample is used to estimate the target statistic for which quality measures (like the RBIAS and RRMSPE) are derived.

Further discussion partaining to model and design based simulations can be found in Münnich et al. (2003), Salvati, Chandra et al. (2010), and Alfons, Templ et al. (2010).

A closely related software packages in the R-language is `simFrame` (Alfons, Templ et al., 2010) which helps to configure simulation studies in a reproducible environment. It includes a wide range of features – like data generation, sampling schemes, outlier contamination mechanisms, and missing values – and was originally developed for simulations in the context of survey statistics but is now designed to be as general as possible (Alfons, Templ et al., 2010). The package `simPop` (Meindl et al., 2014) supports the generation of synthetic population data. This can be a suitable environment in scenarios where the reproducibility of results and confidentiality issues play an important role.

Compared to `simFrame` the package `saeSim` is different in design and its main focus is to assist applications in the SAE field. Most importantly it is based on a framework which is mapped into the software package. This framework defines the overall structure and aims for unifying the shared elements between simulation studies. A simulation is here defined as a stream of data to be manipulated in a sequence of steps. Furthermore it provides the definition of the interface between these steps. The package `saeSim` maps this framework into the R-language and combines it with commonly used facilities in this context; e.g. tools for data generation, sampling, and a link to the parallel computing capabilities in R.

The framework is presented more concretely in the following Section 5.2. This is followed by code examples in Section 5.3 implement-

ing a simple model based simulation study designed to illustrate the capabilities of the package. Section 5.4 concludes this Chapter with a brief discussion.

5.2 A SIMULATION FRAMEWORK

The framework strongly relies on the idea of describing a simulation as a process of data manipulation. Independent of simulation studies, Wickham and Francois (2015) and Wickham (2016) promote this idea by providing tools for cleaning and transforming data. In those frameworks every defined function takes a `data.frame` as input and returns it modified. This leads to a natural connection between all defined functions, since the result of one function can be directly passed to the next as an argument. The symbioses of these packages with the pipe operator, `%>%`, from the package `magrittr` (Bache and Wickham, 2014) only emphasizes this process.

In `saeSim` this approach is extended to simulation studies in the SAE field. The main focus here is the description of a simulation as a process of data manipulation. Each step in this process can be defined as a self contained component – a function in `R` – and thus can be easily replaced, extended, and reused.

Simulation studies address three different levels: these are the population, the sample, and the data on aggregated level. Figure 5.1 illustrates these levels. The left column describes the steps in the data manipulation, the right column presents the function names used in `saeSim` to define the corresponding steps. The *population-level* defines the data on which the study is conducted and may be based on real population data, a synthetic population, or randomly generated variates from a model. In the context of a design based simulation the simulation study would be based on true or synthetic data of *one* population. In model based simulations the population can be randomly drawn from a population model in each repetition.

The scope of the framework is not to opt for viewpoints. The aim is to incorporate the different simulation concepts in a common framework. The *base* (first component in Figure 5.1) of a simulation study is a data table; here the question is whether this data is *fixed* or *random* over repetitions: or from a more technical point of view, is the data generation (the second step in Figure 5.1) repeated in each repetition or omitted in the study. Depending on the choice of a fixed or random population it is necessary to re-compute the population target statistics like domain means and variances, and other statistics of interest (third component in Figure 5.1).

The *sample-level* is necessary when domain predictions are conducted for unit-level models. Independently of how the population is treated – whether as fixed or random – this phase consists of two steps: firstly, of drawing a sample according to a specific sampling

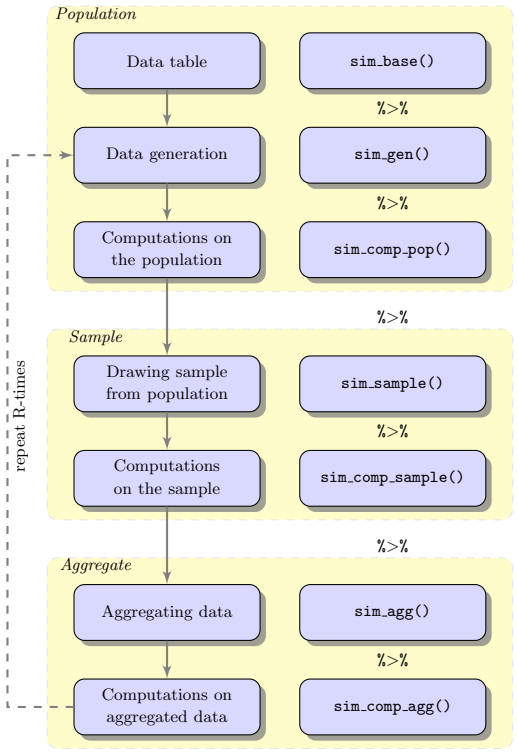

Figure 5.1: Process of Simulation – Left column contains the steps in a simulation. Right column contains the corresponding function names to represent those steps in R.

scheme and secondly, of conducting computations on the samples (fourth and fifth component in Figure 5.1). Given the sample, small area methods are applied. Of interest here are, for instance, estimated model parameters, domain predictions, or measures of uncertainty for the predictions.

Since the sample-level is necessary when unit level models are applied, the *aggregate-level* is relevant when area level models are applied (the seventh and last component in Figure 5.1). Area level models in SAE typically only use information available for domains – in contrast to units. Thus the question for simulation studies for area level methods is whether the data is generated on unit level and is used after the aggregation (sixth component in Figure 5.1) or whether the data is generated directly on area level, i.e. drawn from an area level model. Depending on whether or not unit-level data and sampling are part of the simulation process, the aggregate-level follows the generation of the population or is based on the aggregated sample.

Depending on the scope of the research, some steps in this simulation framework can be more relevant than others. The framework defines a complete list of steps which may be relevant. Single components may be omitted if they are not relevant in specific applications. For example *data generation* is not relevant if we have popula-

tion data; or the *sample-level* is not used when the sample is directly drawn from the model.

Seen in this way, `saeSim` maps the different steps into `R`. Two layers with separate responsibilities need to be discussed. The first is *how* different simulation components can be combined; and the second is *when* they are applied. Regarding the first, in `saeSim` the emphasis is on the interface of each component. To be precise, functions are used which take a `data.frame` as argument and have a `data.frame` as return value. The return value of one component is the input of the next. This definition of interfaces is used for all existing tools in `saeSim`. The second column in Figure 5.1 shows how the different steps in a simulation can be accessed. It is important to note that the functions in Figure 5.1 control the process, the second layer, i.e. *when* components are applied. Each of these functions take a simulation setup object to be modified and a function with the interface discussed above as arguments. To illustrate these implementation details the following Section gives some code examples to implement a model based simulation.

5.3 CODE EXAMPLES

In this Section I want to illustrate some of the features of `saeSim` using code examples. To begin with I introduce some basic functionalities since the package relies on the so called *pipe operator* (`%>%`) which may need some explanation. The pipe operator is designed to make otherwise nested expressions more readable as a line can be read from left to right, instead from inside out (Bache and Wickham, 2014). As a simple example consider the following lines which are equivalent in terms of their functionality:

```
library("magrittr")
colMeans(matrix(rnorm(10), ncol = 2))
rnorm(10) %>% matrix(ncol = 2) %>% colMeans
```

Code written using `saeSim` is based on the idea of passing data forward through a sequence of data manipulation steps. This idea is emphasised by using the pipe operator. The following example illustrates some design aspects of the package as well as the use of the pipe operator:

```
library("saeSim")
setup1 <- sim_base_lm() %>% sim_sample(sample_number(5))
setup2 <- sim_base_lm() %>% sim_sample(sample_fraction(0.05))
```

Without knowing anything about the setup defined in `sim_base_lm` we notice that `setup1` and `setup2` only differ in the sampling scheme applied. `sim_sample` operates as a control when a function is applied (after the population-level) and `sample_number(5)` and `sample_-`

`fraction(0.05)` define the explicit ways of drawing samples. Separating the responsibility of each component into what is applied and when it is applied makes it possible to add new components to any step in the process. The composition of a simulation in this manner will focus on the definition of components and hide control structures. Any function can be passed to `sim_sample` which has a `data.frame` both as an input and as return value. The only responsibility of that function is to draw a sample – which makes it easy to find, understand, and reuse when published. The pipe operator is used to add new components to the setup.

In what follows one way of constructing a simulation in a model based setting is reviewed. The aim is to make domain predictions using the FH model and the robust extension introduced. In this example we start with a unit level population model. From this population samples are drawn in each iteration with SRSWOR where the domain specific sample sizes are: $n_i \in \{5, \ldots, 15\}$. Considered here is the case of 40 domains with 1000 units each. The first step is to generate the data under the following model:

$$y_{ij} = 100 + 5 \cdot x_i + u_i + e_{ij}$$

where $x_i \sim N(0, 16)$, $u_i \sim N(0, 4)$ and $e_{ij} \sim N(0, 32)$. The variance parameter of the unit level error, e_{ij} is chosen such as to lead to sampling errors of the sample mean between 2 and 6 – see Section 6.3 below for a more detailed presentation of a similar scenario.

In this case the *base-component* is a data frame with an id variable named `idD` and constructed with the function `base_id`. Any random number generator in `R` can be used. For the reproducibility of the following results the seed is set to one. The seed is not part of a simulation setup in `saeSim` but needs to be defined by the user:

```
set.seed(1)
setup <-
  base_id(nDomains = 40, nUnits = 1000) %>%
  sim_gen_generic(rnorm, sd=4, name = "x", groupVars = "idD") %>%
  sim_gen_generic(rnorm, sd=2, name = "u", groupVars = "idD") %>%
  sim_gen_generic(rnorm, sd=sqrt(32), name = "e")
setup
```

```
## data.frame [40,000 x 5]
##
##      idD   idU         x          u          e
##    (int) (int)     (dbl)      (dbl)      (dbl)
## 1      1     1 -2.505815 -0.3290472 -3.2168761
## 2      1     2 -2.505815 -0.3290472 -0.7646857
## 3      1     3 -2.505815 -0.3290472  6.6642664
## 4      1     4 -2.505815 -0.3290472 -8.6185953
## 5      1     5 -2.505815 -0.3290472  3.3598670
## 6      1     6 -2.505815 -0.3290472  1.8834517
## ..   ...   ...       ...        ...        ...
```

Note that if you print a simulation setup to the console, as in the above example, one simulation run is performed and only the first rows (the head) of the resulting data are printed. This enables interactivity with the object itself; however, it hides the fact that the setup object is a collection of functions to be called. We can also see that the variables `x` and `u` are generated to be constant for all units within a domain. This is an operation triggered by the usage of the argument `groupVars`.

Using this simulation setup we can now compute the response variable and also the domain mean in the population which is going to be our target variable:

```
setup <- setup %>%
  sim_comp_popMean %>%
  sim_resp_eq(y = 100 + 5 * x + u + e)
setup
```

```
## data.frame [40,000 x 7]
##
##      idD   idU         x        u           e        y  popMean
##    (int) (int)     (dbl)    (dbl)       (dbl)    (dbl)    (dbl)
## 1      1     1 -5.988335 2.248824  -0.2295507 72.07760 72.35218
## 2      1     2 -5.988335 2.248824   4.2040069 76.51116 72.35218
## 3      1     3 -5.988335 2.248824  -2.0089078 70.29824 72.35218
## 4      1     4 -5.988335 2.248824   0.5039513 72.81110 72.35218
## 5      1     5 -5.988335 2.248824   2.3808408 74.68799 72.35218
## 6      1     6 -5.988335 2.248824 -12.8642923 59.44286 72.35218
## ..   ...   ...       ...      ...         ...      ...      ...
```

Here we make use of a set of preconfigured convenience functions to compute the population mean in each domain. Also we can see how to add arbitrary variable definitions to the setup using `sim_resp_eq` for creating the response variable. Now we can draw samples from this population model as follows:

```
sampleSizes <- round(seq(5, 15, length.out = 40))
setup <- setup %>%
  sim_sample(sample_numbers(sampleSizes, groupVars = "idD")) %>%
  sim_comp_n()
setup
```

```
## data.frame [400 x 8]
##
##      idD   idU         x         u         e        y  popMean     n
##    (int) (int)     (dbl)     (dbl)     (dbl)    (dbl)    (dbl) (int)
## 1      1   508 -4.497571 -1.906189 -7.511993 68.09396 75.42649     5
## 2      1   346 -4.497571 -1.906189 -4.632248 70.97371 75.42649     5
## 3      1   582 -4.497571 -1.906189  3.313183 78.91914 75.42649     5
## 4      1   573 -4.497571 -1.906189  5.666321 81.27228 75.42649     5
## 5      1   174 -4.497571 -1.906189 -6.090321 69.51563 75.42649     5
## 6      2   206 -1.353329 -2.315472 -3.750388 87.16750 90.94664     5
## ..   ...   ...       ...       ...       ...      ...      ...   ...
```

In this step `sampleSizes` is used to store the sample sizes to be drawn from each domain. The function `sample_numbers` is responsible for drawing samples and the additional argument `groupVars` triggers the operation to be made in each domain. The default is to draw samples without replacement. Furthermore we store the sample size as variable in the data. Since in this example we are interested in applying area level models we need to add an aggregation step. In this setting we compute the sample mean for the response variable and the single regressor, `x`. Furthermore we need to add an estimator for the sampling variance within the domains which is here implemented using a direct variance estimator:

```
setup <- setup %>%
  sim_agg() %>%
  sim_comp_sample(
    comp_var(samplingVariance = var(y) / n),
    by = "idD"
  )
setup
```

```
## data.frame [40 x 8]
##
##      idD          x          u           e         y   popMean     n
##    (dbl)      (dbl)      (dbl)       (dbl)     (dbl)     (dbl) (dbl)
## 1      1  1.3794532 -0.3774379  0.40906571 106.92889 106.63586     5
## 2      2  2.0420706  1.1389641  0.09954804 111.44887 111.17720     5
## 3      3 -1.1771188 -2.8497507 -3.76998827  87.49467  91.11176     6
## 4      4  5.0512936  3.4489851 -1.19986561 127.50559 128.60690     6
## 5      5 -1.3484770  0.4596911 -3.60295546  90.11435  93.94134     6
## 6      6  0.5125505  1.9468516 -2.08368623 102.42592 104.46338     6
## ..   ...        ...        ...         ...       ...       ...   ...
## Variables not shown: samplingVariance (dbl)
```

Notice that we use the function `sim_comp_agg` with the additional argument `by` to repeat this computation within each domain. In this example `sim_agg` will compute the means in the sample for each numeric variable. Hence we arrive at a data set containing 40 rows: one for each domain.

Thus far it has been possible to utilise preconfigured features from the package. In this sense it is useful to have a tested set of tools which can assist in configuring with relative ease the repetitive elements across simulation studies. However an important aspect of the package is the definition of the interface between components: each component – defined by a function – takes a `data.frame` as input and returns the modified version. Making predictions using the FH model and its robust extension is something which is not covered by the package. Here the non-robust predictions are made by setting the tuning constant to a large numeric value. Hence we need to define these steps:

```
library("saeRobust")
comp_fh <- function(dat) {
  modelFit <- rfh(y ~ x, dat, "samplingVariance", k = 1000)
  dat$FH <- modelFit$reblup
  dat
}

comp_rfh <- function(dat) {
  modelFit <- rfh(y ~ x, dat, "samplingVariance")
  dat$RFH <- modelFit$reblup
  dat
}

setup <- setup %>%
  sim_comp_agg(comp_fh) %>%
  sim_comp_agg(comp_rfh)
```

The above code illustrates how new components can be defined. Notice that the definition of these components only need a few lines of code and have a single purpose. This can contribute to the reproducibility of simulation studies since these components are easily understood. A main aspect contributing to the readability is that these definitions are decoupled from any control structures which are often present in script files for simulation studies.

The object `setup` stores all necessary information to run one iteration of the simulation. In what follows $R = 10$ repetitions are performed. The result is a `list` of `data.frames`. These results are then combined into a single data set:

```
simResults <- sim(setup, R = 10) %>% do.call(what = rbind)
simResults[c("idD", "idR", "popMean", "y", "FH", "RFH")] %>%
  head
```

```
##   idD idR   popMean         y        FH       RFH
## 1   1   1  79.11658  80.36806  80.84265  80.94357
## 2   2   1  97.87564  97.49684  97.36191  97.23226
## 3   3   1 124.94693 123.25169 124.46248 124.64608
## 4   4   1  86.26846  90.63665  88.70489  88.27641
## 5   5   1  78.91773  77.32984  77.53927  77.63569
## 6   6   1 114.04013 117.34472 116.61518 116.73686
```

An additional variable `idR` is automatically added as an ID-variable to distinguish between iterations. In `saeSim` no further tools for processing the resulting data are implemented. There are many tools readily available in R for that purpose.

5.4 DISCUSSION

In the previous example I illustrated how a model based simulation can be configured using the R-package `saeSim`. A design based configuration differs simply in that it starts with sampling instead of with

data generation; otherwise the same tools can be utilised. Code examples for a design based simulation can be found in Warnholz and Schmid (2016).

The main *design* aspect making `saeSim` a useful tool is that simulations can be composed by combining different components. Furthermore it may contribute towards a reasonable way for defining such components within the R-language: as short single argument functions which take a `data.frame` as input and return it modified.

The presentation of this package here has not been exhaustive but has focused on communicating the main idea in the composition of simulation studies. Other features which are available and worthy of note are the generation of outlying data points. Such values are always generated as part of the population and hence focus on the presence of representative outliers as defined by Chambers (1986). Furthermore some effort has gone into building a connection to parallel and high performance computing facilities. In this regard the package `parallelMap` (Bischl and Lang, 2015) is utilised as an interface to Rs parallel computing capabilities. This also includes a link to the package `BatchJobs` (Bischl, Lang et al., 2015) which can be used in conjunction with many high performance infrastructures.

Part III

RESULTS

6 MODEL BASED SIMULATION STUDIES

With the advent of computers it became possible to carry out simulations of models which were intractable using 'classical' theoretical techniques.

— Landau and Binder (2014, p. xv)

6.1 PERFORMANCE OF ROBUST AREA LEVEL PREDICTIONS

In this Section I investigate the statistical properties of the proposed predictors. Using model based simulation studies the empirical bias and the mean squared prediction error are compared over a variety of simulation scenarios as well as between the spatial and temporal extensions of the FH model. The proposed methods for estimating the MSPE of predictions will be presented separately in Section 6.2.

The simulation scenarios are developed under an area level model; thus we consider the situation in which we know the true sampling variances and we generate the Monte Carlo sample directly from the area level model. This is in conceptual contrast to the setting in Section 6.3 where we consider the situation in which the sampling variances are estimated as part of the study and thus need to generate the data on the unit level. However the area level perspective is considered in the literature for evaluating area level models – see for example Fabrizi and Trivisano (2010) and Marhuenda et al. (2013) – and hence this will be considered first.

Although the area level perspective on the data generating process may be incomplete – in that we assume the sampling variances to be known – it is the model under consideration. It presents the possibility of studying the performance of the estimation procedures where the model assumptions appear to be violated due to outlying observations. However in this setting we are limited to the study of area level outliers since it is not obvious how we can simulate unit level outliers into an area level data generating process – see also Section 3.1 for a discussion of unit and area level outliers.

Since several models with various correlation structures have been considered it is tempting to investigate a multitude of specifications. In this respect some choices have been made to restrict the presentation of results to a concise set to support the understanding of different model specifications in the context of outliers. Although we will see some benefits in utilising spatial and temporal correlation structures, the study does not aim at showing the superiority of such

methods but focuses on the effect of outliers instead. We will only consider non-symmetric outliers since they represent the realistic use-case scenario. The choices of the simulation settings are presented in Section 6.1.1; the results are then presented in Section 6.1.4; this is followed by a discussion taking other literature into account in Section 6.1.4.

6.1.1 *Simulation Scenarios*

To avoid the specification of too many scenarios only a temporal setting is considered. Similar to the approach of Marhuenda et al. (2013) the FH model and spatial FH model are treated as special cases when we use only the data from the current time period. The temporal and spatio-temporal extensions then use all available information; however predictions are only made for the current time period. In such a setting it is possible to compare the different model specifications in terms of a unified strategy.

The area level model from which the data is generated in each Monte Carlo repetition is defined as:

$$y_{it} = 100 + 5x_i + u_{1i} + u_{2it} + e_{it}$$

with $i = 1, \dots, D$ and $t = 1, \dots, T$ where $D = 40$ and $T = 10$. The *current* time period is defined to be $t = T$.

- The single regressor, x_i, is a deterministic sequence defined by $x_i = \frac{i}{2D} + 1$. This is a similar setting to the choices made by Marhuenda et al. (2013); however the specification here is constant over time.
- The sampling errors, e_{it}, are drawn from $e_{it} \sim \mathcal{N}(0, \sigma^2_{ei})$. The sampling variances, $\sigma^2_{eit} = \sigma^2_{ei}$, are known during the estimation; they are defined as an ascending and equidistant sequence between 2 and 6: $\sigma^2_{ei} = \frac{4(i-1)}{D-1} + 2$.
- The random effect components are generally generated from the spatio-temporal model, i.e. $u_{1i} \sim \text{SAR}(1)$ and $u_{2it} \sim \text{AR}(1)$ – see also the more concrete presentation in Section 3.2.4. The proximity matrix to generate the spatial correlation structure is of type *rook* (Bivand et al., 2008, p. 250) and correctly specified in the estimation. Models with uncorrelated random effects are deduced where the respective correlation parameter is set to zero. Since the impact of outliers is our main concern only those scenarios are considered where the spatial and temporal correlation is set to $\rho = \rho_1 = \rho_2$ and the respective variance parameters to $\sigma^2_u = \sigma^2_1 = \sigma^2_2$. The concrete choices for ρ and σ^2_u depend on the respective scenarios as defined below.

The scenarios which are considered are defined now by the combination of spatial and temporal correlation in combination with the presence of outliers:

- *(o, o)* denotes the scenario in which we set $\rho = 0$ and $\sigma_u^2 = 2$, i.e. we have uncorrelated random effects and no outliers.
- *(0.5, o)* denotes the scenario in which we set $\rho = 0.5$ and $\sigma_u^2 = 2$. Here we have no outliers but have both spatial and temporal correlation in the random effects.
- *(o, u)* denotes the outlier scenario when we set $\rho = 0$. Area outliers are the domains for which $i \in \{5, 15, 25, 35\}$ to avoid an artificial setting in combination with σ_{ei}^2. For regular observations $\sigma_u^2 = 2$. Outliers are drawn from $u_i \sim \mathcal{N}(9, 25)$ and they ignore any correlation structure, i.e. $u_{1i} = u_{2it} = 0$ for these observations and are replaced by u_i.
- *(0.5, u)* denotes the scenario in which outliers are generated in the same way as for *(o, u)* with the difference that $\rho = 0.5$ for the generation of the regular observations.

6.1.2 *Quality Measures*

The methods to be compared are the robust spatial and temporal extensions of the FH model. The non-robust methods are denoted by FH, SFH, TFH, and STFH; and their robust counterparts by RFH, RSFH, RTFH, and RSTFH. In fact these methods were introduced as area level REBLUP, SREBLUP, TREBLUP, and STREBLUP in Section 3.3 and are here abbreviated for notational simplicity. Furthermore we have the bias corrected versions of the predictors using the correction proposed in Section 3.4. This correction is only applied to the robust predictions and the respective models are referred to by RFH.BC, RSFH.BC, RTFH.BC, and RSTFH.BC. The tuning constant for the robust predictions is fixed at 1.345. Furthermore the direct estimator is denoted by *Direct*; this is the generated value for y_{iT}.

To assess the quality of the predictions under the various methods two measures are utilised: the relative bias (RBIAS) and the relative root mean squared prediction error (RRMSPE). These measures are computed over all realisations of the Monte Carlo repetitions; in total $R = 500$ repetitions have been conducted. Let $\hat{\theta}_{ir}^M$ denote the prediction for the ith area in the rth repetition with $r = 1, \ldots, R$ under the model M where M is one of the considered models, e.g. the RFH model. The Monte Carlo RRMSPE is here defined as:

$$\text{RRMSPE}_i^M = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(\frac{\hat{\theta}_{ir}^M - \theta_{ir}}{\theta_{ir}}\right)^2}$$

where θ_{ir} denotes the true target statistic in the Monte Carlo repetition r and is defined by:

$$\theta_{ir} = \theta_{iTr} = 100 + 5x_i + u_{1ir} + u_{2iTr}$$

where the target statistic is defined as the true value in time period $t = T$. The relative bias for the ith area can accordingly be defined as:

$$\text{RBIAS}_i^M = \frac{1}{R} \sum_{r=1}^{R} \frac{\hat{\theta}_{ir}^M - \theta_{ir}}{\theta_{ir}}.$$

6.1.3 *Results*

The main results are summarised in Figures 6.1 and 6.2. They present the RBIAS and RRMSPE in *per cent* over 500 Monte Carlo repetitions. Most of the findings conform to expectations in that the robust methods have a beneficial effect in terms of MSPE when outliers are present and are comparable but never superior otherwise. However a number of findings are more surprising and need some explanation. To that extent consider three observations we can make when studying the figures:

1. In terms of bias the robust temporal and spatio-temporal extensions are unbiased in the presence of outliers, whereas all other methods are positively biased.
2. The non-robust temporal and spatio-temporal models are less efficient in terms of MSPE even in scenarios without outliers.
3. The bias correction shows good overall results. However it seems to add to the MSPE in the case of the temporal and spatio-temporal models.

Firstly it may be observed that the models that take the temporal structure into account are unbiased. Several properties of the models as well as the simulation setting contribute to this effect. Identifying the spatial correlation structure is problematic in this simulation scenario. One emerging point is the amount of variance which is due to the spatial correlation structure; this variation has a variance parameter set to $\sigma_1^2 = 2$ which is small compared to the overall variance in the data. Also the small sample size of $D = 40$ may have an effect – the more data we have the easier it is to identify small effects or in this case a correlation structure. This is also present when comparing the FH and SFH models where no difference in terms of MSPE is visible – even in the scenarios *(0.5, 0)* and *(0.5, u)*.

The inability to identify this spatial effect leads to a similar result as was evident in Figure 4.9. There we observed that under contamination the temporal correlation structure is influenced by outliers and at the same time the spatial structure has been identified correctly.

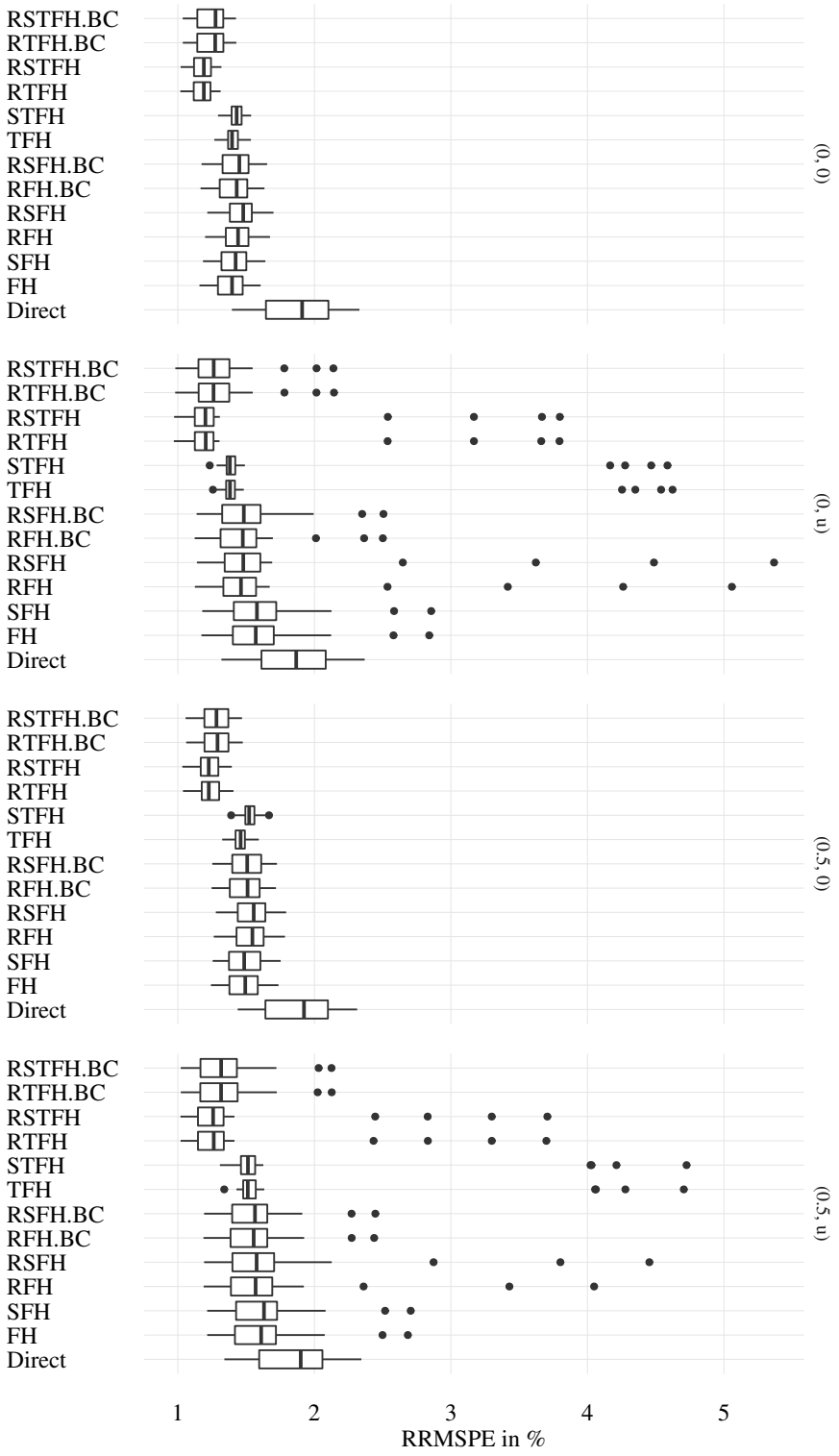

Figure 6.1: Relative Root Mean Squared Prediction Error of Spatial and Temporal REBLUPs

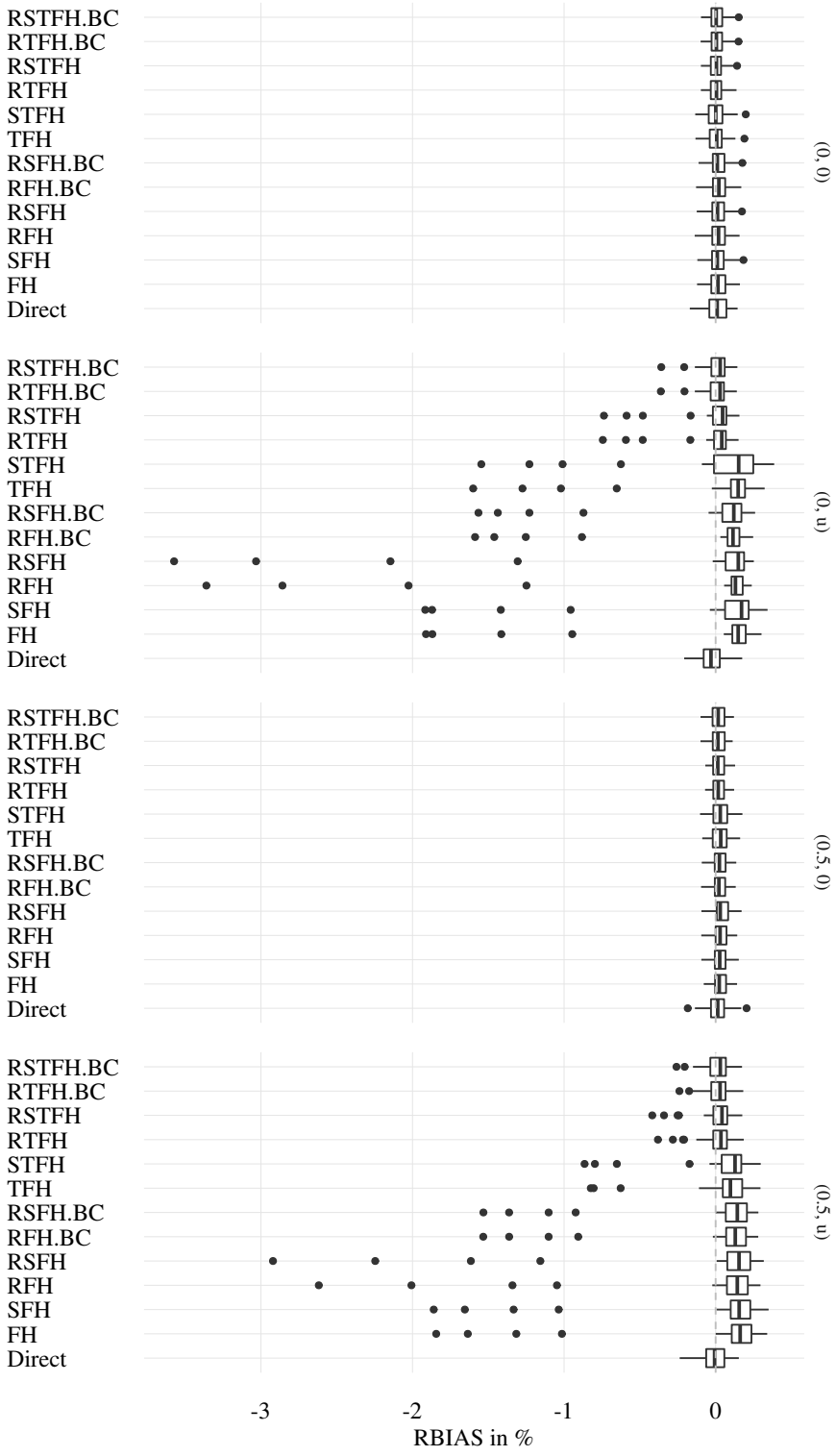

Figure 6.2: Relative BIAS of Spatial and Temporal REBLUPs

Here we observe a similar effect: yet in contrast the spatial structure captures the outlier contamination and the temporal autocorrelation can be identified correctly. So in contrast to the stability tests the effect has changed due to the different nature of outlying areas; in Section 4.2 outliers are single observations; here outlying areas include all observations over time. Thus in the results we see the ability of the temporal models to borrow strength from the correctly identified temporal autocorrelation. The reduced bias for outlying areas is due to the spatial correlation structure used for the prediction of such domains. In fact this effect would not only be present for the robust methods but also for the non-robust methods. Basically we have an over parametrised model – since the signal of the spatial correlation is too weak – which captures the mixture distribution used to induce the outlier contamination.

Now the question is why do non-robust methods seem to be less beneficial in terms of MSPE even when there are no outliers; this is the second observation. As this result is surprising a comparison has been made using the implementation from the R-package `sae` (Molina and Marhuenda, 2015) to rule out errors in the software. The results remain the same. What we really observe here is that in approximately 30 per cent of the simulation runs, the variance parameters for the temporal and spatio-temporal models are estimated to be close to zero. Using the implementation in `sae` such results are sometimes denoted as *not converged*; using `saeRobust` the evaluation of the estimation equations at their respective solutions reveal unsatisfactory results, i.e. values which cannot be considered to be close to zero. In terms of predictions the estimation of zero variance parameters results in random effects close to zero: hence only the regression estimates are used as a synthetic estimator. This also explains the small variation between area predictions which is revealed by the *small* boxes in Figure 6.1 for the TFH and STFH model.

Two settings in the simulation can be tweaked to improve the results for the non-robust methods. First, the ratio of variance due to the random effects can be increased, thus simplifying the identification of these correlation structures. Second, the strength of the outlier contamination can be increased. In a study where the mean of the outlier distribution was set to 100 the estimation under the TFH and STFH models yielded results comparable to their robust counterparts. As was discussed earlier in such a setting these models are over parametrised and can thus model the mixture distribution used for inducing the outlier contamination.

Still open to question in this line of argumentation is the problem of why the predictions under the RTFH and RSTFH models do not suffer from a similar fate. The concrete mechanism explaining the difference is unknown at present. In various simulations the robust methods have revealed to be less sensitive to the misspecification of

the sampling variances. In many cases zero variance parameters for the random effects can be explained by settings in which the specified sampling variances are large compared to the overall variance in the data. Using the *true* sampling variances means that we use the parameters from a super population model; such values may be less optimal in concrete realisations. However this is a broad and more conceptual discussion which is again addressed in Section 6.3 and somewhat beyond the scope of the discussion here.

The third observation is that the bias correction does not improve the results for the temporal models. The bias correction constructs an interval around the inefficient but unbiased direct estimates in which we allow predictions to be made. The choice made in Section 3.4 for the size of this interval may be too conservative in this simulation study. This leads in effect to numerous repetitions in which domain predictions are unnecessarily bias corrected. However for the models which do not take the temporal structure into account we can see that the bias correction has a beneficial effect on the prediction of the outlying domains, both in terms of bias and consequently also in terms of MSPE. For the non-outlying areas no additional gain is discernible from this correction.

6.1.4 *Discussion*

The aim of this study has not been to promote the use of a specific model in terms of a correlation structure but rather to focus on revealing the differences between the robust and non-robust estimation methods. The benefits of utilising temporal autocorrelation have been demonstrated in different studies – see Rao and Yu (1994) for the temporal model and Marhuenda et al. (2013) for the spatio-temporal model. The main findings of this Section are:

- Using an over-parameterised model may have a positive effect in terms of RBIAS and RRMSPE in the presence of outliers. Although not explicitly shown this can be true also for the non-robust methods. This may be due to the ability of the fitting process to approximate the mixture distribution.
- The proposed bias correction may prove to be useful especially for outlying domains. However the choice of the width of the interval in which predictions can be made should be handled with care in practice.
- The robust methods have an expected positive effect in terms of RRMSPE in the presence of outliers. Also they may be more robust against the choice of the sampling variances – however this claim needs further investigation.

6.2 PERFORMANCE OF MEAN SQUARED PREDICTION ERROR ESTIMATORS

In addition to the prediction of a target statistic we are also interested in estimating the measure of uncertainty surrounding this quantity. To this extent two MSPE estimators have been proposed in Section 3.5: an adaptation of the parametric bootstrap proposed by Sinha and Rao (2009) and an MSPE estimator based on the pseudolinearisation approach of Chambers, Chandra and Tzavidis (2011). In what follows these two MSPE estimators are compared in different simulation settings.

In the previously conducted simulation study we saw that the simulation settings revealed both advantages and disadvantages of the robust methods. With respect to the MSPE estimation the scenarios are chosen to conform better to the underlying models. However the computational effort involved in the bootstrap estimator is relatively high; this made it necessary to restrict the number of scenarios to a minimum. Thus for each model – the RFH, RSFH, RTFH, and RSTFH – data is generated under the correct model and only non-outlier are compared to outlier scenarios.

The results for the non-robust methods are omitted in order to reduce the number of necessary comparisons. In principle a comparison with established methods like the MSPE estimator of Prasad and Rao (1990) can be useful. However the MSPE estimators associated with the non-robust predictors are too diverse and have on that account been omitted from the discussion here. The corresponding review of these methods is provided in Section 2.3.2.

In Section 6.2.1 below the simulation settings for each model are described in detail. Section 6.2.2 describes which measures are utilised to assess the performance of the MSPE estimators. A presentation of the results can be found in Section 6.2.3 and this is followed by a discussion of the results in the context of the existing literature in Section 6.2.4.

6.2.1 *Simulation Scenarios*

In the simulation study each model is fitted on data generated using the corresponding model which can be represented in general form as:

$$y_{it} = 100 + 5x_i + \mathbf{z}_{it}^\top \mathbf{u} + e_{it}$$

with $i = 1, \dots, D$ and $t = 1, \dots, T$. The regressor is defined as before: $x_i = \frac{i}{2D} + 1$; and the sampling error structure is the same for all scenarios: $e_{it} \sim \mathcal{N}(0, \sigma^2_{ei})$ with $\sigma^2_{eit} = \sigma^2_{ei}$ and $\sigma^2_{ei} = \frac{4(i-1)}{D-1} + 2$. Furthermore a distinction is made between:

- RFH-(o) – denotes the non-outlier scenario used for evaluating the MSPE estimators for predictions using the RFH method. Here $D = 40$ and $T = 1$; $\mathbf{z}_{it}^\top \mathbf{u} = u_i$ with $u_i \sim \mathcal{N}(0, 9)$.
- RSFH-(o): denotes the non-outlier scenario for the spatial FH model. $D = 40$ and $T = 1$; $\mathbf{z}_{it}^\top \mathbf{u} = u_{1i}$ with $u_{1i} \sim SAR(1)$ where $\rho_1 = 0.5$ and $\sigma_1^2 = 9$. Similar to before a *rook* proximity matrix is used to generate the spatial structure (Bivand et al., 2008, p. 250).
- RTFH-(o): denotes the non-outlier scenario under a TFH model. Here $D = 40$ and $T = 10$; $\mathbf{z}_{it}^\top \mathbf{u} = u_{0i} + u_{2it}$ with $u_{0i} \sim \mathcal{N}(0, 9)$ and $u_{2it} \sim AR(1)$ where $\rho_2 = 0.5$ and $\sigma_2^2 = 9$.
- RSTFH-(o): denotes the non-outlier scenario under the spatio-temporal FH model. Here $D = 40$ and $T = 10$; $\mathbf{z}_{it}^\top \mathbf{u} = u_{1i} + u_{2it}$ with $u_{1i} \sim SAR(1)$ where $\rho_1 = 0.5$ and $\sigma_1^2 = 9$ and $u_{2it} \sim AR(1)$ where $\rho_2 = 0.5$ and $\sigma_2^2 = 9$. Also a proximity matrix of type *rook* is utilised to generate the spatial process.

Here *(o)* is used to denote the non-outlier scenario. For each of these scenarios one outlier scenario is considered which is denoted by *(u)* where we replace the random effect for the outlying domains:

- *-(u): $\mathbf{z}_{it}^\top \mathbf{u} = u_i$ with $u_i \sim \mathcal{N}(9, 25)$ for all $i \in \{5, 15, 25, 35\}$. The set of outlying domains is chosen such as to avoid an artificial scenario in combination with the choice for σ_{ei}^2.

6.2.2 *Quality Measures*

In this study we are interested in the performance of the MSPE estimators. To assess the quality of these estimators the relative root mean squared error (RRMSE) and RBIAS of the estimated root MSPE (RMSPE) are compared with the *true* values. Let $\widehat{RMSPE}_{ir}^{M}$ denote the estimated RMSPE for area i in the rth Monte Carlo repetition using method M. M is either the parametric bootstrap referred to by BOOT or the pseudolinearisation based approach which is referred to by CCT. Similar to the previous study, predictions are made for the *current* time period, which is again defined as $t = T$ for the respective scenario. This also means that we evaluate only the RMSPE for these predictions. In the bootstrap 100 repetitions are conducted; changing the number of repetitions did not change the results significantly. Overall 500 repetitions are conducted. We can then define the RRMSE as:

$$RRMSE_i^M = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(\frac{\widehat{RMSPE}_{ir}^{M} - RMSPE_i^M}{RMSPE_i^M}\right)^2}$$

where we define the *true* RMSPE as the Monte Carlo RMSPE over all repetitions:

$$\text{RMSPE}_i^M = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(\widehat{\text{RMSPE}}_{ir}^M - \text{RMSPE}_i^M\right)^2}$$

Furthermore we have the RBIAS of the MSPE estimator defined as:

$$\text{RBIAS}_i^M = \frac{1}{R}\sum_{r=1}^{R}\frac{\widehat{\text{RMSPE}}_{ir}^M - \text{RMSPE}_i^M}{\text{RMSPE}_i^M}$$

These measures are computed for both MSPE estimators and for all robust predictions; the respective predictors are referred to by RFH, RSFH, RTFH, and RSTFH; and their bias-corrected counterparts by RFH.BC, RSFH.BC, RTFH.BC, and RSTFH.BC. The tuning constant is fixed at 1.345 for all models.

6.2.3 *Results*

The main results of the simulation study are given in Table 6.1. The following observations need to be discussed:

1. The CCT has not a negative bias for all estimators.
2. The CCT has a high MSE for the bias corrected predictors.
3. The advantage of the CCT compared to BOOT is small for the outlying domains.

The first observation is that the CCT has not a negative bias in all cases. However as we disregard the uncertainty due to the estimation of the variance parameters and assume independence between the weights and the response in the pseudolinear form we should expect an underestimation. This is observable for the RFH and RSFH in the non-outlier scenario. Under the same scenario the RTFH and RSTFH can have a positive bias since only the estimated MSPE for the *current* time period is considered. In an analysis using all the time periods the median RBIAS is indeed negative. The estimation for regular domains under the outlier scenario can also have a positive bias. This happens because the variation due to the random effect is higher when outliers are present. Thus the variance parameters used to compute the MSPE for the regular observations are slightly higher resulting in a higher estimate for the MSPE.

The second observation is the higher MSE of the CCT for the predictions with bias correction. The bias correction will bind the domain prediction. The choice of the interval in which predictions can be made leads – in this simulation – to the situation that more predictions than necessary are *bias corrected*. Since the weights for the bias correction – see Section 3.4 – are also used for estimating the MSPE

Predictor	MSPE	(o)-Regular	(u)-Regular	(u)-Outlier
Median RBIAS:				
RFH	CCT	−7.59	−0.93	−36.65
	BOOT	−3.93	1.91	−36.66
RFH.BC	CCT	0.88	1.80	−20.80
	BOOT	−3.73	1.09	−28.52
RSFH	CCT	−7.19	−1.09	−33.81
	BOOT	−3.34	2.10	−33.77
RSFH.BC	CCT	−2.00	0.51	−17.44
	BOOT	−3.14	1.22	−24.86
RTFH	CCT	0.18	0.32	−41.60
	BOOT	1.10	1.17	−41.95
RTFH.BC	CCT	8.16	6.67	−3.84
	BOOT	1.15	1.15	−21.52
RSTFH	CCT	0.25	0.25	−16.43
	BOOT	1.21	1.02	−16.71
RSTFH.BC	CCT	2.04	2.66	−9.22
	BOOT	1.15	0.90	−13.73
Median RRMSE:				
RFH	CCT	11.21	6.78	36.79
	BOOT	11.59	9.42	37.04
RFH.BC	CCT	21.54	12.56	25.09
	BOOT	11.02	9.34	29.23
RSFH	CCT	10.05	6.28	34.01
	BOOT	10.34	9.14	34.25
RSFH.BC	CCT	18.16	10.33	23.21
	BOOT	10.11	9.09	25.67
RTFH	CCT	4.40	3.88	41.69
	BOOT	8.43	8.06	42.18
RTFH.BC	CCT	28.91	26.52	27.57
	BOOT	8.15	8.03	22.30
RSTFH	CCT	2.32	3.32	16.50
	BOOT	7.70	8.11	17.82
RSTFH.BC	CCT	11.79	11.26	17.94
	BOOT	7.64	8.04	15.23

Table 6.1: Performance of RMSPE Estimators in Model-Based Simulation. Results are in %. *Regular* denotes non-outlier observations. (o) is the model specific scenario without contamination; (u) is with outlier contamination.

more weight is given to the sampling variance. This leads to more unstable results over all. This may be less relevant in an application in which we can choose the width of the interval. Also this effect was observed in the model based simulation in Section 6.1.

The third observation we can make is that the CCT only has a small advantage in terms of bias and MSE for the outlying observations. In general the results regarding the performance of the MSPE estimators are very sensitive with respect to the choices made for the simulation setting. Here the bias we can make in a prediction is relatively small since the mean for outlying domains is shifted by 9 units compared to an overall intercept of 100. In the design based study of Chapter 7 we will observe outlying observations which are multiple times larger than the main body of observations. Hence the benefit in terms of bias depends largely on the magnitude of the intercept of the outlying observations. Furthermore it must be noted that the MSPE predictor will gain an advantage in terms of MSE due to the reduced bias in the prediction.

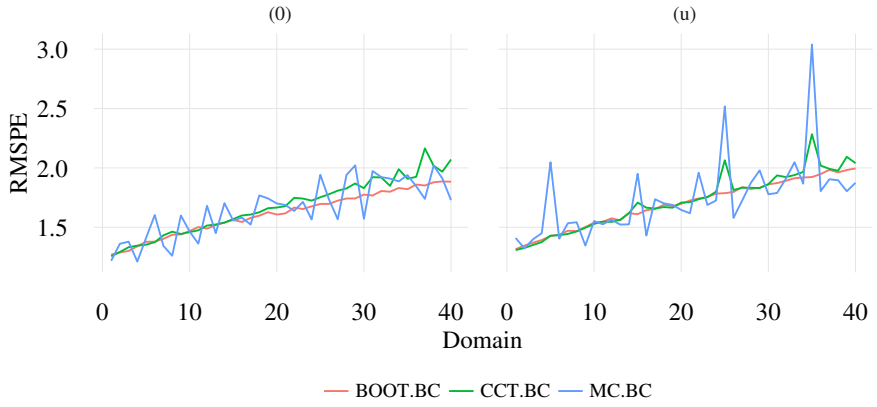

Figure 6.3: Estimated Root Mean Squared Prediction Error for the Bias Corrected Robust Spatial FH Model. Compared are the bootstrap (BOOT.BC) and the pseudolinearisation-based estimator (CCT.BC) with the Monte Carlo MSPE (MC.BC).

Figure 6.3 also illustrates the way the estimated RMSPE relates to outliers. Note that observations 5, 15, 25, and 35 are outlying domains and that the sampling variances increase with i. What we see here is that the CCT for the bias corrected RSFH is able to follow the Monte Carlo MSPE better when the sampling variances are large. Since there is no information on the *true* variation in the outlier distribution we observe a better fit the more relevant the sampling variance is. This shows that these results depend largely on the scenario in which the CCT method is applied. The bootstrap, in contrast, has no means of mirroring the Monte Carlo MSPE for outlying observations regardless of the scenario.

6.2.4 *Discussion*

In general the performance of the MSPE estimators is promising. However the concrete results strongly depend on the scenario settings. In general the CCT for the bias correction suffers under repeated sampling when the prediction interval of the correction is too conservative. For the same reason we observe a correlation between the sampling variances and the bias. The bootstrap, in contrast, shows very stable results – approximately 10 per cent in terms of RRMSE – over a variety of different settings. However this method cannot capture the variation for outlying domains. In scenarios in which the respective models repeatedly yield estimates close to zero variances the bootstrap will underestimate the true variation since more weight is given to the linear predictor; this can be seen in Chapter 7. This setting has been avoided – in contrast to Section 6.1 – by choosing larger values for the variance parameters in this study.

The settings in this study have been chosen to be close to the approach taken by Chambers, Chandra, Salvati et al. (2014). Although a comparison is difficult to make because of the transition between area level and unit level models, some similarities in the results can be found. Most notably Chambers, Chandra, Salvati et al. (2014) report values of the RRMSE for the bootstrap for the unit level REBLUP of approximately 10 per cent in scenarios having area level outliers and scenarios without contamination. This result is very close to what we can observe in this study. In the specific scenario above the results for the CCT show similar performance in terms of RRMSE to the bootstrap except for the bias corrected predictions. In the study conducted by Chambers, Chandra, Salvati et al. (2014) the CCT method does not show as good results as we can see here; however this can possibly be attributed to the differences in the scenarios.

6.3 FROM UNIT TO AREA LEVEL CONTAMINATION

In this Section a different approach is taken in that a model based simulation study is conducted beginning with a unit level population. This is typical for unit level models; see for example Chambers, Chandra, Salvati et al. (2014) and Sinha and Rao (2009). However results for robust area level models are often produced using area level data – see for example Fabrizi and Trivisano (2010) – or studies using survey data – see for example Bell and Huang (2006) and Xie et al. (2007). This area level perspective implicitly rules out any impact of unit level outliers on area level models; and whether such observations can influence predictions is the subject matter of this Section.

This changed perspective enables a more thorough investigation of the impact of unit level contamination in conjunction with area level models. Furthermore it is necessary to illuminate how we es-

timate the sampling variance of the direct estimator as this can also have an impact on domain predictions; so far these quantities have been assumed to be known. In the literature the estimation of the sampling variances is often connected to the application of smoothing techniques to stabilise these parameters – see Section 2.3.1.2 for a review. Hence in this context several points require scrutiny:

- Under the FH model we specify an area level model with a heteroscedastic sampling distribution. The model does not explain the source for this heteroscedasticity. In what follows two different sources are induced: one is the sample size varying across domains, and the other source are unit level outliers.
- When we begin from the unit level we may ask what the true sampling variance is; under the model this quantity is assumed to be known. In practice these values are estimated and are themselves unreliable if the target is the population variance. So it will be relevant to discuss how the FH estimator and the robust extension perform when the direct variance estimator instils an additional source of uncertainty. Also it is unclear how smoothing techniques interact with unit level outliers.
- In the context of unit level outliers it may be intuitive to suggest a robust direct estimator instead of the sample mean. Hence different robust direct estimators are used as an alternative to the domain specific sample mean.

This discussion has practical implications but is conceptually not related to the question if we exploit the spatial and temporal correlation structures present in the data. For this reason the following study focuses on the FH model and excludes the spatial and temporal extensions. In principle these results should apply to these extensions also; however this remains an avenue for further research.

The present study is structured as follows: Section 6.3.1 completes the formal link between unit and area level. Essentially we need to specify what assumptions have been made with respect to the unit level population model in order to establish a coherent link between the population level, the sampled level, and the area level. Based on these results Section 6.3.2 presents the simulation settings, Section 6.3.3 the investigated estimation strategies; then in Section 6.3.4 the results of the Monte Carlo Simulation are presented.

6.3.1 *A Unit Level Population Model*

In the following the unit level population model assumed for this study is presented. Hence we begin with a liner mixed model for the population of the form:

$$y_{ij} = x_i^\top \beta + u_i + e_{ij}$$

with $i = 1, \dots, D$ and $j = 1, \dots, N_i$. Here x_i is a vector of auxiliary information and β the vector of regression coefficients; u_i is an i.i.d. random variable: $u_i \sim \mathcal{N}(0, \sigma_u^2)$; and e_{ij} is also an i.i.d. random variable: $e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$. Given this simple unit level model consider the case in which we draw samples with simple random sampling without replacement (SRSWOR) in each domain. Only the case is considered where units are observed for all domains. The domain specific sample sizes are denoted by n_i. When we use the sample mean as a direct estimator we can derive the area level model as:

$$\bar{y}_i = \underbrace{x_i^\top \beta + u_i}_{\theta_i} + \bar{e}_i$$

with $i = 1, \dots, D$. Here $\bar{y}_i$ denotes the direct estimator and θ_i the true target statistic. The sampling errors, $\bar{e}_i$, are i.i.d. with $\bar{e}_i \sim \mathcal{N}(0, \sigma_{ei}^2)$ where $\sigma_{ei}^2 = \frac{\sigma_e^2}{n_i}$. In this simple setting x_i and u_i are the same for the unit and area level since they are constant within domains. Also we see in the definition of σ_{ei}^2 that the only source for the heteroscedasticity are the sample sizes, n_i. An additional source for heteroscedastic sampling errors may also be induced by unit level outliers.

6.3.2 *Simulation Scenarios*

The model underlying the Monte Carlo study is given by:

$$y_{ij} = 100 + 5x_i + u_i + e_{ij}$$

- The regressor, x_i, is a deterministic sequence defined as in the model based scenarios before: $x_i = \frac{i}{2D} + 1$.
- From this model a population is generated in each Monte Carlo repetition. Each domain is of size $N_i = 1000$ and we consider $D = 40$ as before.
- Samples are drawn with SRSWOR in each domain. The sample sizes are defined as $n_i \in \{5, \dots, 15\}$. They are sorted as ascending sequence in i. The sample sizes are chosen in correspondence to the variance of the unit error in the non-outlier scenario such that the sampling variances of the sample means are an ascending sequence between 2 and 6.

With these settings several different choices for outlier contamination are investigated. Here we can now consider unit and area level outliers:

- *(o, o)* – no contamination. Here $u_i \sim \mathcal{N}(0, 2)$ and $e_{ij} \sim \mathcal{N}(0, 32)$.
- *(u, o)* – area level outliers. Similar to the generation before the non-contaminated domains are generated from $u_i \sim \mathcal{N}(0, 2)$ and $e_{ij} \sim \mathcal{N}(0, 32)$. Outlying domains are domains with $i \in \{5, 15, 25, 35\}$ and their respective random effect is given by $u_i \sim \mathcal{N}(9, 25)$.

- *(o, e)* – unit level outliers. The random effect is drawn from $u_i \sim \mathcal{N}(0, 2)$. Individual outliers are generated from a mixture of two independent normals: $e_{ij} \sim \delta_i \mathcal{N}(0, 32) + (1 - \delta_i)\mathcal{N}(20, 200)$ where δ is a Bernoulli random variable with $\mathbb{P}(\delta = 1) = 0.8$ for domains with $i \in \{4, 14, 24, 34\}$ and set to one otherwise. I.e. 10 per cent of the areas have 20 per cent of unit level observations which are generated from the outlier distribution.
- *(o, e-sym)* – symmetric unit level outliers. This scenario is the same as *(o, e)* with the unit level outlier distribution replaced by $\mathcal{N}(0, 200)$ – hence symmetric unit level outliers are imposed.
- *(u, e)* – unit and area level outliers. The area level contamination is imposed as in *(u, o)* and the unit level contamination as in *(o, e)*.

6.3.3 *Quality Measures and Considered Methods*

Two measures are utilised to assess the quality of predictions: the RRMSPE and RBIAS. These quantities are computed according to Section 6.1.2. The performance is studied for several methods; the non-robust predictions under the FH model and robust prediction referred to as RFH. For the robust predictions the tuning constant is fixed at a value of 1.345. The following modelling strategies are considered:

- SM – denotes the sample mean. Since we sample with SRSWOR an unbiased direct estimator is the sample mean.
- RM – denotes a robust direct estimator. Here the sample median is used.
- FH.SM – is the non-robust prediction under the FH model. This model is based on the sample mean and uses the direct variance estimate as sampling variance for each domain: $\sigma^2_{ei} := \frac{s_i^2}{n_i}$.
- RFH.SM – is the robust prediction based on the sample mean. The same setup as for FH.SM is used.
- FH.SM.GVF – is the non-robust prediction under the FH model using a generalised variance function (GVF). In contrast to FH.SM a generalised variance function is used as an estimator for the sampling variance: $\sigma^2_{ei} := \frac{\tilde{s}^2}{n_i}$ with

$$\tilde{s}^2 = \frac{1}{n - D} \sum_{i=1}^{D} (n_i - 1) s_i^2$$

  where s_i^2 denotes the sample variance for domain i and $n = \sum_{i=1}^{D} n_i$.
- RFH.SM.GVF – is the robust prediction using the generalised variance function. This has the same setup as the FH.SM.GVF using the robust prediction.

- RFH.SM.BC – is the robust prediction with the sample mean as direct estimator and direct variance estimates. The predictions are bias corrected as described in Section 3.4.
- FH.RM – is the non-robust prediction based on the sample median. As sampling variance a robust estimator is used based on the median absolute deviation (MAD): $\sigma^2_{ei} := \frac{MAD_i^2}{n_i}$.
- RFH.RM – is the robust prediction based on the sample median. The sampling variances are estimated as for the FH.RM method.

6.3.4 *Results*

To summarise the results of this study consider Figure 6.4 and 6.5. With these results the following observations can be made:

1. There is only a small advantage in using generalised variance functions as an alternative to direct variance estimates.
2. The robust FH prediction has no additional effect under unit level outliers.
3. A robust direct estimator has no advantage over the sample mean.

The first observation to make is that the particular choice for the generalised variance function does not show a significant gain in performance. Wolter (2007, pp. 272-273) points out that the main benefits in using GVFs may be to reduce the computational effort, to ease the communication of results, and achieve some gain in stability. He argues that we can expect a gain in stability since instead of many parameters – one for each domain – fewer and hence more stable parameters can be estimated. However he also notes that there is no theoretical underpinning for this claim. The discussion in Section 2.3.1.2 showed that despite this lack of theory the use of generalised variance functions is often preferred in practical applications. Hence it may well be that we tend to observe results which are specific to this simulation setting.

Some differences when using GVFs can be observed with respect to the RBIAS. In a scenario with non-symmetric unit level outliers we can observe an increase of the RBIAS. What happens here is that the overall variance is smoothed. Hence smaller sampling variances, σ^2_{ei}, are assigned to the outlying domains making them appear to be more reliable measurements. This results in a stronger impact of the outlying domains on the predictions and thus in an increase in bias. In Section 3.1.3 this effect is referred to as the *creation of overly influential observations*.

In this simulation scenario the unit level outlier contamination does not influence the estimation of the GVF. In a different setting with a stronger contamination the GVF can be expected to overestimate the

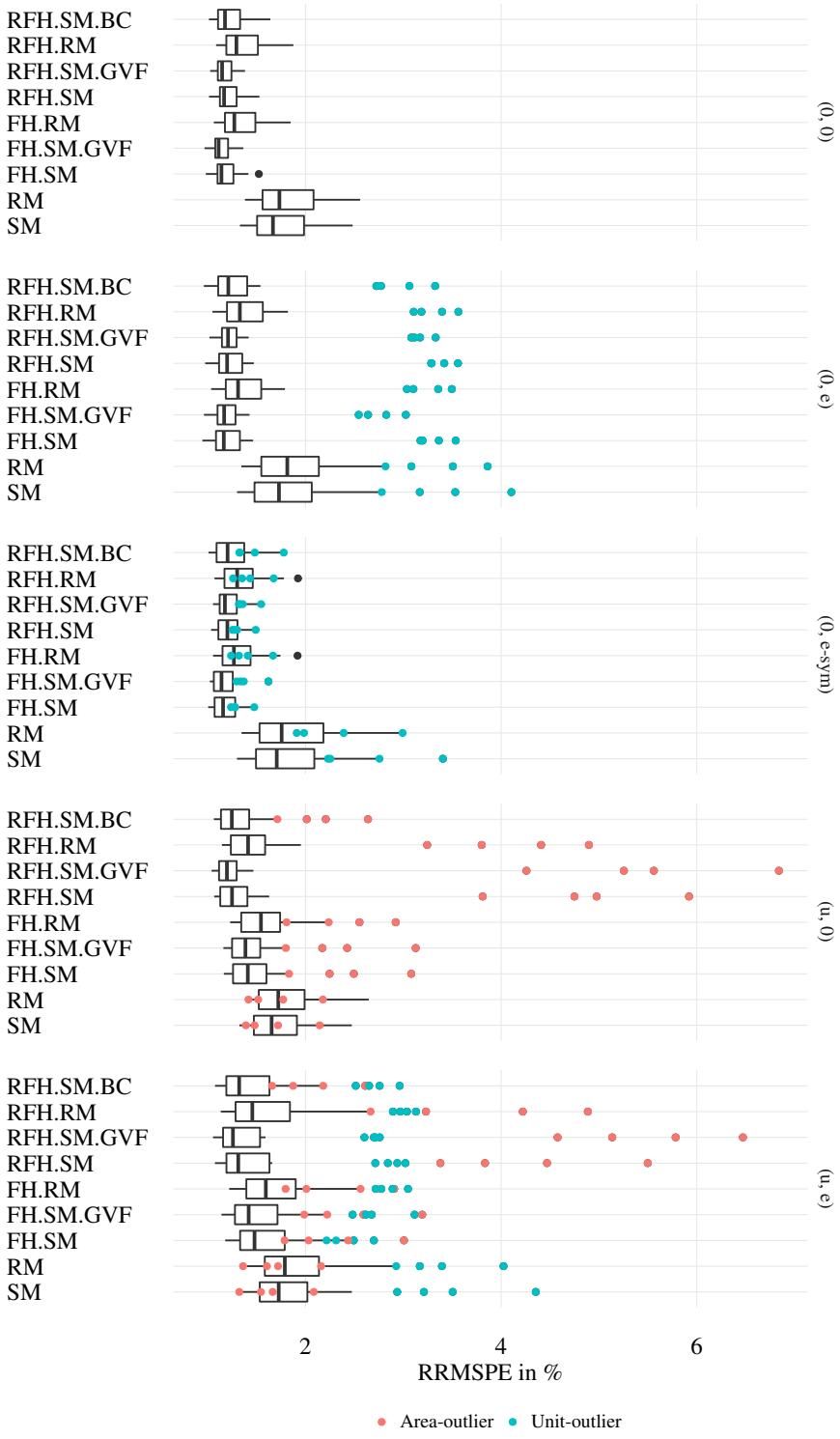

Figure 6.4: Relative Root Mean Squared Prediction Error for Domain Predictions

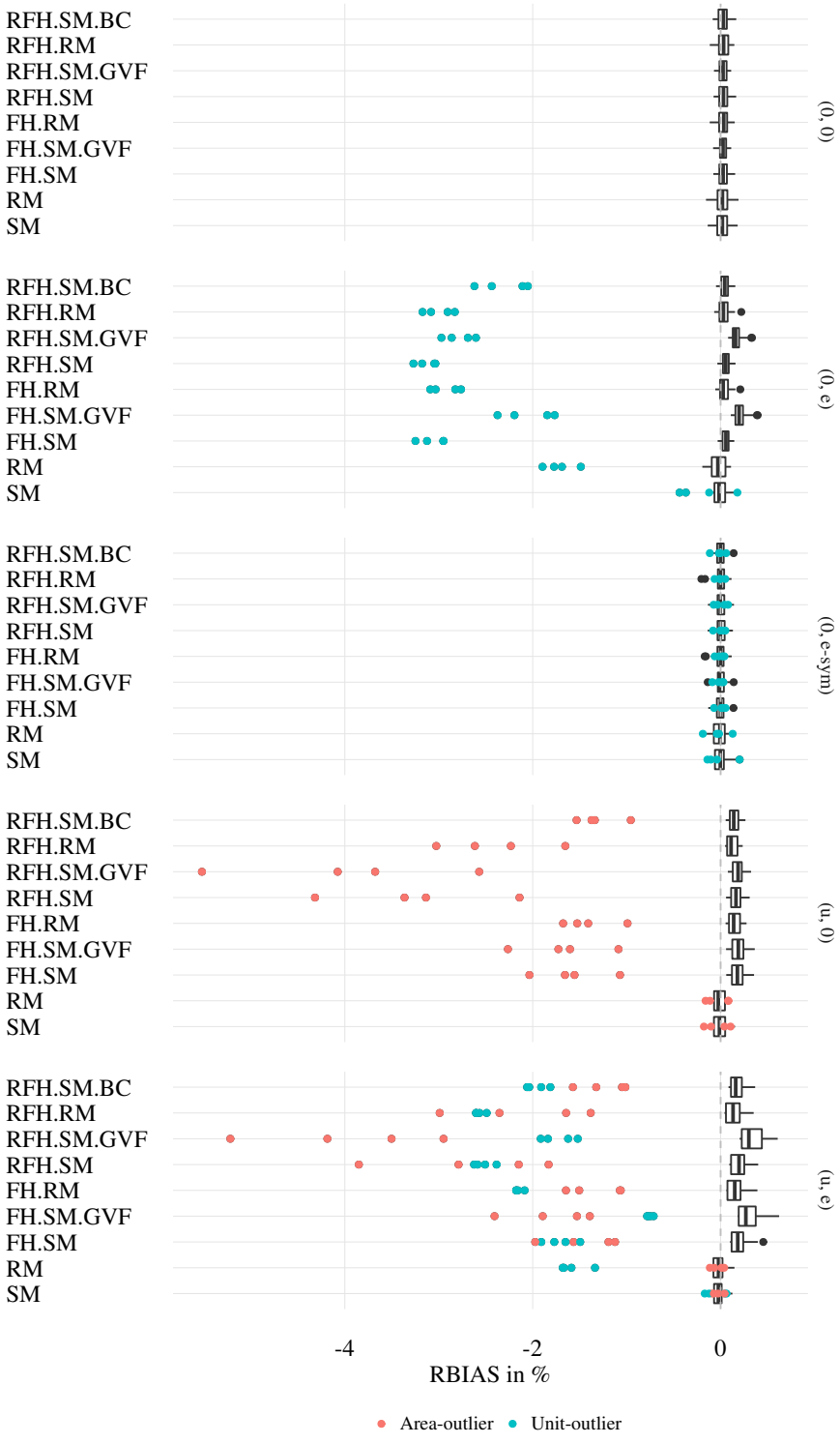

Figure 6.5: Relative Bias for Domain Predictions

true sampling variances. In preliminary experiments this leads to a loss in efficiency in terms of RRMSPE.

The second observation is that the robust prediction under an FH model does not improve the predictions under unit level contamination. The results under symmetric unit level contamination show that the FH model has a self adjusting effect as was described in Section 3.1.1. The prediction mechanism under the FH model gives more weight to the linear predictor for the outlying domains since we observe a larger sampling variance for these type of outliers. This is not an undesirable property as long as the linear prediction is a good fit for these observations. When we compare these results with the non-symmetric outlier scenario – *(o, e)* – we can see that the effect of unit level contamination on the prediction is small overall. This holds for all modelling strategies except when using GVFs for the reasons outlined above. However in these situations we introduce a bias for the outlying domains. In contrast to the bias correction for unit level models – see Section 2.5.2 for a review – such a bias cannot be corrected using area level data.

The use of the RFH model is limited when we consider unit level contamination. The self adjusting effect of the FH model is in general also present in the robust prediction. When we consider the use of the GVF it can happen – particularly with more extreme outlier contamination – that the RFH model becomes preferable. This is so because we create overly influential observations which, from their design, do not differ from area level outliers – at least from the area level perspective.

The third observation is that a robust direct estimator, in this case the median, has no additional advantage over the sample mean. Here several points need to be mentioned. First, using a robust estimator may be problematic when we consider small samples in that the median will itself have an even higher variance than the sample mean. The additional gain with respect to the unit level contamination does not seem to outweigh this shortcoming – at least in the scenarios considered. This is also true when we consider the same symmetric unit level outlier setting in all domains. Second, the variance estimation of a robust direct method is usually also problematic, again, because we are considering small samples. In practice it may be desirable to use a bootstrap for estimating the variance of the sample median. With these small samples the results have been disappointing and have not led to an improvement in the area level predictions. Hence the MAD seems to provide a sensible alternative. Third, it may be argued that a different robust method should be considered. An M-type estimator for the sample mean, in conjunction with a robust variance estimator, has also been considered. Compared to the median very similar results have been obtained in this case; consequently these results have been omitted here.

6.3.5 *Discussion*

There are two main reasons to consider a robust area level model. First this helps to make the prediction robust against area level outliers. Second it facilitates to protect against the overly influential observations which may occur upon using GVFs to smooth the sampling variances. However a more thorough investigation of the use of GVFs may be necessary to come to a plausible recommendation since the simulation setup has not shown a significant improvement in the overall predictions.

With respect to unit level outliers we can see a self adjusting effect in the robust and non-robust prediction under the FH model. This adjustment replaces the predictions for outliers with the linear predictor which constitutes an improvement in those cases in which this predictor is suitable, i.e. unbiased. In principle we face the same problem as for robust methods using unit level data in that again we introduce a bias. However at present it is unclear how a bias correction can be developed for these situations when only area level data is observed. The proposed bias correction can have a beneficial effect when we consider area level contamination; however it does not provide any advantage with respect to unit level outliers.

The use of robust direct estimators has not proved to be encouraging. It may be that trying to solve an efficiency problem of the sample mean by replacing it with an even less efficient estimator is the main reason here. This effect can well be due to the small sample sizes considered in SAE, particularly in this simulation setup. These small sample sizes then lead to a less efficient estimate for the domain specific mean and also its variance. Using an area level model has not shown itself to outweigh this effect.

7 DESIGN BASED SIMULATION STUDY

Because the synthetic microdata looks just like replicated actual microdata, you can use standard complete-data software to draw inferences.

— Rubin (1993)

7.1 OUTLINE

In this final Chapter I want to present a design based simulation study in which the robust and non-robust predictions under the FH model are compared. In contrast to the model based simulations of Chapter 6 the following study is based on one synthetic population. The data is generated and provided by Statistics Netherlands (CBS) and is based on the Structural Business Survey (SBS). The target domains here are industry sectors and the target statistic is the annual tax-turnover within the domains. Using this data we can study several effects on the domain predictions in a more realistic setting. In particular we will see how domain predictions are influenced by outlying units and domains in combination with a sampling design.

More details on the data and aim of the study from a contextual point of view are given in Section 7.2. Also here the sampling design and setup for the simulation as well as details on the methods used can be found. The results are then presented in Section 7.3 and a discussion of the results in a broader context is provided in Section 7.4.

7.2 THE SYNTHETIC POPULATION

The population data is a synthetic population generated using the SBS. The SBS is an annual business survey in the Netherlands. Furthermore the sampling scheme used in this study is similar to that applied in the SBS and it is provided and used without any change. The context of this study is the prediction of the total tax-turnover for 20 industry domains. Possible predictor variables are the tax-turnover in the previous year, the size class in terms of the number of employees, and the actual number of employees.

The unit level population consists of 63981 observations spread across 20 industry sectors. The target variable, tax-turnover in Euro, has a median value in the population of 0.16 million, a mean of 0.42 million, and a maximum of 83 million Euros. Thus the distribution is skewed and contains unit level outliers which are here assumed

to be representative. The domain specific population means range between 0.1 and 2.9 million Euro with a median of 0.42 and a mean of 0.62. Only three domains have values larger than one million Euros and may be suspected to be outlying areas.

A similar sampling design to the one used in the SBS has been used to draw 500 samples from this population. It is stratified for the size class within the domains. Within each stratum samples are drawn with simple random sampling without replacement (SRSWOR). In addition larger firms are selected with a probability of one. The sample always consists of 5074 units. The sample sizes between domains vary between 6 and approximately 1050 observations; however they have some variation over the 500 repetitions.

First results showed that for an area level model it is sufficient to use only one predictor variable – the tax-turnover of the previous year – which has the most predictive power in this setting. Furthermore on the area level we have only 20 observations – the 20 industry sectors – and even in a setting with one predictor we need to estimate three parameters. Hence a simple regression model is estimated:

$$\bar{y}_i = \beta_0 + \beta_1 \tilde{y}_{it-1} + u_i + e_i$$

which is here assumed to follow the FH model. Here $i = 1, \ldots, 20$ and u_i is assumed to follow a normal distribution with zero mean and variance σ_u^2. The sampling error, e_i, is also assumed to follow a normal distribution with zero mean and variance σ_{ei}^2. The direct estimator, $\bar{y}_i$, is a weighted mean using the inverse of the first order inclusion probabilities as weights, i.e. the HT estimator (Horvitz and Thompson, 1952). The standard error is computed using the approximation based on the results of Hansen and Hurwitz (1943) since only the first order inclusion probabilities are known. These standard errors are then squared and used as the true sampling variances, σ_{ei}^2, in the model. The regressor variable, $\tilde{y}_{it-1}$, is the mean value of the previous year. Here the population mean is used; since the target variable is tax-turnover it may be plausible that at some point the true or realised values are known to the analyst. An investigation in which also the lagged dependent variable was estimated using either the HT estimator or the sample mean showed that using an area level method only brings a small advantage over the direct estimator.

The assessment of the quality of the predictions follows along the lines of Section 6.1.2 in that the RBIAS and RRMSPE over all repetitions are computed. The true target statistic is the domain specific population mean of tax-turnover and is constant over all repetitions. Predictions are made under the FH model comparing the standard model referred to by FH, the robust prediction (RFH), and the bias-corrected robust prediction (RFH.BC).

Also the MSPE estimators are compared using the RBIAS and RRMSE as they are defined in Section 6.2.2. The target here is again

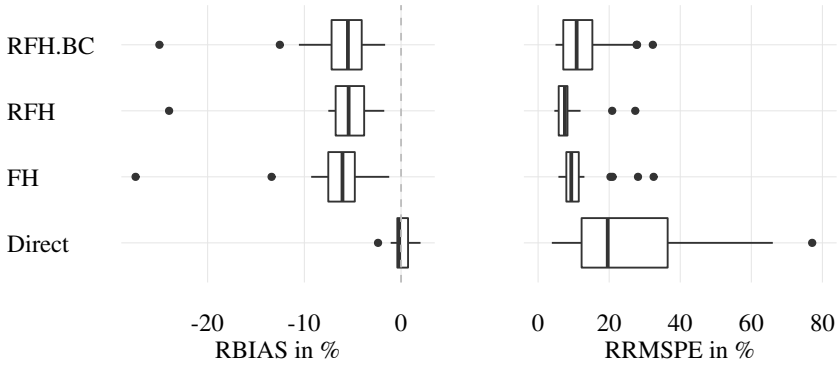

Figure 7.1: Performance of Domain Predictions – Estimated root mean squared prediction error (RRMSPE) and relative bias (RBIAS).

the estimated root MSPE (RMSPE). The methods used are the pseudo-linearisation based approach (CCT) and the parametric bootstrap (BOOT); both are introduced in Section 3.5.

7.3 RESULTS

The performance of the predictions can be evaluated using Figure 7.1. The median values of the RBIAS and RRMSPE for the respective method are also additionally given in Table 7.1. Here a main result is that the prediction can be vastly improved in terms of MSPE using the robust and non-robust methods compared with the direct estimation. This is so since we use a strong predictor variable which is not incorporated into the direct domain estimation. However the improvement in terms of MSPE has the price of an introduced bias which can be observed for the area level models; the direct estimator is design-unbiased as should be expected. We can also see that the robust method can improve the prediction for approximately another two percentage points in terms of RRMSPE. The bias correction cannot add any advantage over the robust method and can even be a draw back compared to the non-robust predictions. A possible explanation is that in this setting we have a strong predictor and also very small estimated variance parameters of the random effects. This means that we favour the linear predictor and hence we are as far away from the direct estimator as is possible using these methods. Thus the bias-correction is applied more frequently.

Turning now towards the performance of the MSPE estimators the main results can be found in Table 7.2 and Figure 7.2. What we see here is that for predictions under the FH and RFH method the CCT and parametric bootstrap methods show similar performance. In both cases we have a strong negative bias. This negative bias stems from the fact that the estimated variance parameter under both methods – FH and RFH – are close to zero. In terms of predictions this

Predictor	RBIAS	RRMSPE
RFH.BC	−5.50	10.85
RFH	−5.42	7.54
FH	−6.06	9.34
Direct	−0.19	19.59

Table 7.1: Performance of the Domain Predictions in Design-Based Simulation – Results are in %. Presented are the median values of 20 industry sectors.

means that we are relying strongly on the linear predictor – and as discussed earlier we can see the overall improvement in the predictions in terms of MSPE. However these small variance parameters mean that the bootstrap samples have too small a variation; and also the CCT estimator relies on this parameter; however neither the bootstrap nor the CCT can account for the uncertainty related to these parameter estimates. In terms of bias the CCT produces acceptable results for the predictions using RFH.BC, i.e. the robust bias corrected predictions. However compared to the bootstrap we observe a higher RRMSE.

Predictor	MSPE	RBIAS	RRMSE
FH	CCT	−31.70	55.50
	BOOT	−29.77	56.03
RFH	CCT	−36.91	48.06
	BOOT	−34.48	49.15
RFH.BC	CCT	1.87	88.19
	BOOT	−20.28	47.11

Table 7.2: Performance of RMSPE Estimators in Design-Based Simulation. Results are in %. Presented are the median values of 20 industry sectors.

7.4 DISCUSSION

The results of this Chapter show the overall beneficial effect of small area methods in terms of MSPE. However the point predictions also reveal an introduced bias; here we observe the typical variance-bias trade-off associated to mixed linear models and Empirical Bayes methods (Efron and Morris, 1972). Interestingly we do not see any improvement using the bias correction technique; especially in terms of bias. However we do see an improvement by applying the robust estimation technique in terms of MSPE.

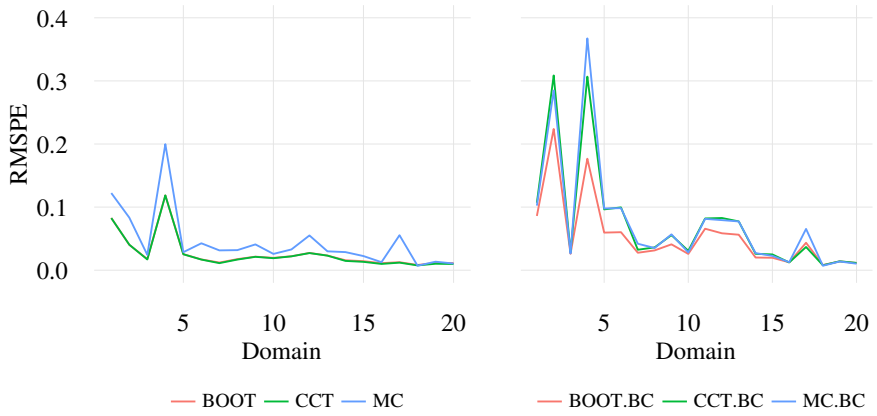

Figure 7.2: Performance of the Estimated Root Mean Squared Prediction Error – Compared are the root MSPE estimates using the methods CCT and BOOT with the MSPE under the Monte Carlo simulation. The left panel shows the RMSPE estimates for the robust predictions; the right panel for the robust bias-corrected predictions.

In this study we have both, unit and area level outliers. It remains somewhat in-transparent if unit level outliers have an effect at all. In this example we have a sampling design which is highly correlated with the target variable. Hence we already address the heterogeneity in the distribution of tax-turnover on the unit level since we incorporate the inclusion probabilities. One property of the direct estimator – also in the presence of outliers – is that we expect it to be design-unbiased. If this property is fulfilled then we should not expect to have an impact of asymmetric unit level outliers. This may not be plausible in scenarios in which the sampling design was chosen for a different target statistic.

The results of the MSPE estimators – CCT and BOOT – reveal some problems in a scenario in which we estimate the variance parameter to be close to zero. Essentially we are here at the boundary of the parameter space of the variance. At present it is not clear how to address this issue. It may be helpful to incorporate the comparison with the more established method by Prasad and Rao (1990) which at this point remains an avenue for further research.

8 CONCLUSION

8.1 THEORY

In Chapter 2 the relevant advances in the SAE field have been reviewed. Two developments are central for this Thesis: First the extension of the FH model by the incorporation of spatial and temporal correlation structures into the random effects part of the model. Second the robust predictions under linear mixed models.

In particular the spatial extension of the FH model by Pratesi and Salvati (2008) is of interest. Here the random effects part of the model is assumed to follow a simultaneous autoregressive process (SAR) of order one. Under this model Pratesi and Salvati (2008) derived a spatial EBLUP (SEBLUP) which they have applied to the Survey on Life Conditions in Tuscany, Italy. They aimed at predicting the mean per capita income within municipalities. Furthermore they found that incorporating the unobserved spatial correlation between domains – the SAR process – into the prediction leads to an improvement in terms of MSPE – here however measured in terms of the domain specific coefficient of variation.

The reviewed temporal extension by Rao and Yu (1994) involves the splitting of the random effects into a random intercept and an autocorrelated process (AR) of order one. The underlying data now has repeated measurements for each domain. The variation within domains and hence between the repeated measurements is addressed using the AR(1) process; the variation between domains is captured by inserting an additional random intercept. In this context they introduced a temporal EBLUP (TEBLUP). Empirical results are here based on a model based simulation study. Later Singh, Shukla et al. (2005) showed the advantage of incorporating temporal autocorrelation into the analysis where they predict the per capita consumption expenditure using an annual consumer survey in India.

Marhuenda et al. (2013) then introduced a combination of these methods: they replaced the random intercept in the temporal model of Rao and Yu (1994) with a SAR(1) process. Using this model they derived a spatio-temporal EBLUP (STEBLUP). They evaluate their method using the Survey on Income and Living Conditions for Spain and predict two poverty indicators. They show results for the spatial, temporal, and their spatio-temporal extension to the FH model. Their results suggest that especially the incorporation of the correlation over time has a beneficial effect.

Next to the different possibilities of incorporating correlated random effects in Section 2.5 I reviewed the recent advances in outlier robust predictions in the SAE field. The robust methodology used in this Thesis are based on the results of Sinha and Rao (2009) and Richardson and Welsh (1995). Here the idea is to *robustify* the score functions of the maximum likelihood function for a given mixed linear model. Based on these robust estimation equations model parameters and domain predictions can be derived. To this extent the residuals in the score function are bounded using an appropriate influence function. In various model and design based simulation studies the advantages of such methods have been demonstrated where outliers are present in the data; see for example Sinha and Rao (2009), Chambers, Chandra, Salvati et al. (2014), Schmid and Münnich (2014), and Schmid, Tzavidis et al. (2016).

The main contribution of this Thesis is now the combination of the robust estimation methodology and the spatial and temporal extensions to the FH model. This connection is established in Chapter 3. There I distinguished between three types of outlying observations: unit level outliers, area level outliers, and overly influential observations. The first two types are common in the robust literature and describe outlying units and outlying areas respectively. The third type may be artificial in the context of area level models: they may present a problem when we smooth direct variance estimators. In doing this we may create observations which appear to be reliable but are in fact heavily influenced by unit level outliers. Especially with respect to area level outliers and overly influential observations the robust extensions to the spatial and temporal FH models may present themselves as being beneficial.

With regard to robust area level models four such models have been under consideration: the *standard* FH model, the spatial extension by Pratesi and Salvati (2008), the temporal extension by Rao and Yu (1994), and the spatio-temporal model by Marhuenda et al. (2013). These methods are extended to obtain (i) an area level robust EBLUP (REBLUP) which is the direct result of applying the results of Sinha and Rao (2009); (ii) an area level spatial REBLUP (SREBLUP) which – to contrast it with Schmid (2011) – is based on the FH model instead of the BHF model; (iii) a temporal REBLUP (TREBLUP); and (iv) a spatio-temporal REBLUP (STREBLUP). Essentially I present these models in a comprehensive framework of robust area level models. This framework includes contributions to the existing scheme of robust methodology as listed below:

- The robust score functions used by Sinha and Rao (2009) have been extended such that they can be used in the context of correlated random effects.
- The robust score functions have been combined with the results of Chambers, Chandra and Tzavidis (2011) and Chambers,

Chandra, Salvati et al. (2014) to derive the area level REBLUPs – the area level SREBLUP, TREBLUP, and STREBLUP – in a pseudolinear form. The results are presented in general notation to include the general class of mixed linear models.

- This pseudolinear representation has led to an IRWLS algorithm for the regression coefficients. Furthermore an algorithm for the optimisation of the random effects has been proposed which is also derived from the pseudolinear form. Both solutions can have the advantage that they reduce to the non robust BLUE estimator and EBLUP for the random effects when the influence function is the identity function.
- Solutions for the variance parameters of the different models are sought for using a fixed point algorithm based on the results of Chatrchi (2012). The solutions for these parameters need to be tuned for each model.
- Furthermore a simple bias correction for the robust prediction has been adapted based on the results of Efron and Morris (1972). This bias correction is simply the *limited translation estimator* proposed by Efron and Morris (1972) which may be useful when we suspect asymmetric area level outliers.
- And finally the approach of Chambers, Chandra and Tzavidis (2011) to the estimation of the MSPE of domain predictions has been extended for area level models with correlated random effects. The robust extension here is simply a special case defined by its pseudolinear representation. In this regard also the parametric bootstrap method for the MSPE estimation of Sinha and Rao (2009) has been adapted: these results have been directly applicable without any significant change as they are formulated for the general case of linear mixed model.

These results may not only be useful for the specific models under consideration but they also present a generic framework for robust area level models which are based on the class of linear mixed models.

8.2 IMPLEMENTATION

The outcome of this Thesis includes also software packages for the R-language (R Core Team, 2015). `saeRobust` (Warnholz, 2016) provides an implementation of the area level REBLUP, SREBLUP, TREBLUP, and STREBLUP. In addition to the predictions for each model, three types of diagnostic plots are provided: normal quantile-quantile plots for the random effects and model residuals; mean-difference plots for the predictions; and diagnostic plots for the estimated MSPE. Additionally predictions can be made using the proposed bias correction.

For both versions of predictions the pseudolinearisation based MSPE estimator and the parametric bootstrap are implemented.

In Chapter 4 the main features of this software implementation have been reviewed where also the stability of the algorithms is investigated. To this end a model based simulation study is conducted where the data is generated to represent a numerically challenging situation. This is accomplished by inducing deterministic outliers of relatively large magnitude and by using bad starting values.

In principle the results of the stability tests are promising since in all scenarios acceptable solutions can be found when we set the number of iterations to a higher value. *Acceptable* here refers to a value close to zero at the solution for the respective estimation equation. This however presents a trade-off between the number of iterations, stability, and computational demand, which can easily become relevant with temporal data. The key to computationally less demanding solutions is the choice of starting values or generally an algorithm which proves to have a higher convergence rate.

In addition to the package `saeRobust` a software package for implementing model and design based simulation studies in the context of the SAE field is also provided: `saeSim` (Warnholz and Schmid, 2016). This package has already been introduced in Warnholz and Schmid (2016); and Chapter 5 is an adaptation of this article to present this package within the context of this Thesis.

The main *design* aspect making `saeSim` a useful tool is that simulations can be composed by combining different components. Furthermore it may provide a coherent framework for the definition of such components within the `R`-language: as short single argument functions which take a `data.frame` as input and return it modified. In this regard one goal of this implementation is to propose a reasonable way of conducting simulation studies in the `R`-language. This may contribute towards the reproducibility of such results within the field since source code is easier to share and tools are tested to be correct.

8.3 RESULTS

Empirical results can be found in model based simulations in Chapter 6 and a design based simulation study in Chapter 7. The aim of the model based simulation studies has not been to investigate the beneficial effect of correlated random effects on domain predictions. Instead these studies show how the robust and non-robust methods behave differently under varying settings. However especially in connection with temporal models clear advantages can be observed in the model based simulations when we make use of the additional information presented by repeated measurements.

Overall four simulation studies have been conducted. The study in Section 6.1 aimed at explaining differences between robust and non-robust domain predictions in scenarios with correlated random effects and area level outliers. The effect of area level outliers on the non-robust methods is that variance parameters of the random effects are estimated too high. This sacrifices efficiency as the non-robust predictions are closer to the direct estimator. In this context the robust counterparts show good properties overall. The bias correction in this setting appears to have a positive effect on the MSPE in that this quantity is smaller for outlying domains. This is due to the fact that the bias for outlying domains is indeed reduced. However the results also indicate that its own tuning parameter needs to be chosen with care in practical applications. In some simulation scenarios the setting has been too conservative and showed a negative effect on the overall domain predictions.

The study in Section 6.2 provides insights into the performance of the MSPE estimators – the pseudolinear based approach and the parametric bootstrap. Both methods provide good results for the robust predictions with two limitations: First, the parametric bootstrap fails at approximating the MSPE for outlying domains. Second, the pseudolinearisation based approach can capture this behaviour; however this depends on the relation between sampling variance and actual MSPE of a domain. This is especially problematic with respect to the bias correction.

Then in a next step the model of simulation is changed in Section 6.3: instead of generating data on the area level we begin with a unit level population model. This has the advantage that we can take different direct estimators into account as well as the variance estimation for these direct estimators. However only the case of the *standard* FH model has been considered. Here a robust direct method and a generalised variance function (GVF) have been utilised. We can see that it is important to use the direct estimator with the smallest variance. The use of robust direct estimators did not show any improvement when such estimates were used in an area level model – even in the presence of unit level outliers. Furthermore the use of a GVF may prove to be problematic in the presence of unit level outliers. This topic is commented on again below as the results cannot be regarded as conclusive but are a first assessment of the issue.

In a design based simulation study presented in Chapter 7 the performance of the non-robust and robust FH model are compared in a more realistic setting. The target statistic in that study is the tax-turnover in 20 industry sectors in the Netherlands. The basis of this analysis is a synthetic population based on the Structural Business Survey provided by Statistics Netherlands. Here we have the combination of unit and area level outliers in combination with an informative sampling design. The study shows an overall improve-

ment achieved by using area level models. The robust domain predictions yield an additional advantage in terms of MSPE; however the bias corrected robust prediction appears to be less effective. Also the MSPE estimators have been investigated: both methods – the parametric bootstrap as well as the pseudolinearisation based approach – are problematic. A possible reason is that we are in an application in which the variance parameter of the FH model is estimated to be close to zero; neither method can handle this.

8.4 REMARKS AND OPEN RESEARCH QUESTIONS

In a final step of discussing the results of this Thesis I want to present three claims. These in themselves may be ambitious but they show the benefits and current limitations of the results overall:

1. We should favour a robust method for small area predictions.
2. Unit level outliers may or may not be present in a sample; however this issue cannot be tackled with area level information .
3. One should prefer an established design-unbiased direct estimator when there is a choice.

Taking the results of the various model based simulation studies we may be tempted to conclude that we should always use robust estimation methodology. In this context results often suggest a clear advantage of robust methods when outliers are present. On the other hand they are never superior but may be considered equivalent when there are no outliers. This can be seen in Section 6.1 and 6.3 as well as in Chapter 7. A certain degree of deviation from the normal distribution may be assumed to be the rule and not an exception in *real world* applications and here robust methods may present themselves as useful.

However an essential point is that practical issues like measuring the overall Goodness of Fit and variable selection have not been investigated yet. Hence in order to make the use of such methods reliable further research is needed. A further limiting factor is the estimation of the MSPE. The parametric bootstrap and pseudolinearisation based approach have proved to be useful in the model based simulation; however in the design based study we saw the limitation of these methods when the variance parameter of the FH model is estimated to be close to zero. At this point it remains unclear how an established method like the MSPE estimator of Prasad and Rao (1990) performs in this setting; a more thorough investigation may be needed. In addition it may be useful to carefully investigate the possibility of utilising recent bootstrap methods of Jiongo and Nguimkeu (2014) and Mokhtarian and Chambers (2013). The main problem here may be the incorporation of correlated random effects.

With respect to unit level outliers the discussion on the area level is complex and the following is mostly based on the results of Section 6.3. When it is plausible to assume unbiased direct estimates as it is under the FH model it may be irrelevant if unit level outliers are considered to be asymmetric or not; on the area level this effect does not exist anymore. In such a case the FH model is already *robust* against unit level outliers since such observations are simply an additional source of heterogeneity and drive the sampling variances. Domains with large sampling variances are already weighted down by the standard FH model. Considerable improvement for such domains can be expected independently of robust estimation techniques.

The effect of unit level outliers may be amplified when we consider the use of generalised variance functions. The estimated sampling variances of the domains containing outlying units may be framed as outlying with respect to the GVF. This topic is discussed in Section 6.3 but this claim needs more investigation and support than offered in this Thesis. The main problem is that the simulation conducted in Section 6.3 may not present a case in which we make a *fair* comparison since a significant benefit from the considered GVF cannot be observed.

While conducting the model based simulation study of Section 6.3 and the design based study in Chapter 7 several robust direct estimators have been applied. The main problems in using such methods are – especially in the context of *small* domains – that these estimators introduce a higher variance for domains which do not contain outliers. Furthermore we not only need a good direct estimator but also an estimator for its variance. In this regard the sample mean and the HT estimator yield more reliable results. Even the problem of unit level outliers in the context of an informative sampling design – informative with respect to the target quantity – may be solved since the inclusion probabilities provide sufficient information to account for the variability in the data. The use of a robust method combined with a GVF may still be open to discussion; however in this respect the issues raised above should be addressed first.

> *It is perfectly proper to use both classical and robust/resistant methods routinely, and only worry when they differ enough to matter. But when they differ, you should think hard.*
>
> — Tukey (1979)

Part IV

APPENDIX

ABSTRACT

The demand for reliable small area statistics from sample surveys has grown substantially over the past decades due to their growing use in public and private sectors. The field of Small Area Estimation aims at producing such statistics. In this Thesis I consider several spatial and temporal extensions to the Fay-Herriot (FH) Model to improve the mean squared prediction error (MSPE) of predictions for small domains. Such predictions can be influenced by single observations in the data; hence the estimation of the model parameters and predictions is based on the estimation methodology around robust empirical best linear unbiased prdictions (REBLUPs).

With regard to robust area level models four such models are under consideration: the FH model (Fay and Herriot, 1979), the spatial extension by Pratesi and Salvati (2008), the temporal extension by Rao and Yu (1994), and the spatio-temporal model by Marhuenda et al. (2013). These methods are extended to obtain (i) an area level robust EBLUP (REBLUP) which is the direct result of applying the results of Sinha and Rao (2009); (ii) an area level spatial REBLUP (SREBLUP); (iii) a temporal REBLUP (TREBLUP); and (iv) a spatio-temporal REBLUP (STREBLUP). I present these methods in a comprehensive framework of robust area level models. For the estimation of the MSPE I adapt the parametric bootstrap method from Sinha and Rao (2009) as well as the analytical solution based on a pseudolinear form of Chambers, Chandra and Tzavidis (2011). In this context also a bias correction based on the *limited translation estimator* of Efron and Morris (1972) is adapted to account for a potential bias associated to robust methods.

In addition to the development of these robust methods their implementation in the R-package `saeRobust` (Warnholz, 2016) is investigated. The package provides an initial version for the application of the developed methodology. In this regard some numerical stability tests are performed and also basic features like diagnostic plots for model residuals are reviewed. Also an outcome of this Thesis is the package `saeSim` (Warnholz and Schmid, 2016) which provides a framework for simulation studies within the R-language. It aims at simplifying the configuration of such studies by providing tools for data generation, sampling, and a link to the parallel computing fascilities in the R-language.

The methods under consideration are then further investigated in model and design based simulation studies. Here the performance of the predictions and the MSPE estimators are studied when area

level outliers are present. Furthermore I discuss how unit level outliers as well as area level outliers may effect domain predictions based on area level models. This discussion is continued in the context of a design based simulation. In this study a synthetic population is utilised based on the Structural Business Survey in the Netherlands. Here the tax-turnover in 20 industry sectors is targeted. This data includes unit level outliers, area level outliers, as well as an informative sampling design.

B

ZUSAMMENFASSUNG

Die Nachfrage nach zuverlässigen Statistiken für kleine Gebiete hat in den vergangenen Jahrzehnten stark zugenommen. Ein Grund dafür ist der Bedarf solcher Statistiken in der amtlichen Statistik. Oft basieren Ergebnisse auf Umfragedaten und die Hauptaufgabe des Forschungsfeldes der *Small Area Estimation* ist es die Vorhersagewerte mit Bezug auf den mittleren quadratischen Vorhersagefehler (MSPE) zu optimieren – selbst wenn nur wenige Beobachtungen innerhalb einzelner Gebiete zur Verfügung stehen. In dieser Arbeit werden dazu räumliche und zeitliche Erweiterungen des Fay-Herriot (FH) Models genutzt. Diese Modelle basieren stark auf Verteilungsannahmen und in diesem Zusammenhang können die Ergebnisse leicht durch einzelne Beobachtungen beeinflusst werden. Hierzu werden deshalb räumliche und zeitliche Erweiterungen des FH-Modells mit Ausreißer-robuster Schätzmethodik kombiniert.

Die Modelle die hierbei genutzt werden sind das FH Modell (Fay und Herriot, 1979), die räumliche Erweiterung von Pratesi und Salvati (2008), die zeitliche Erweiterung von Rao und Yu (1994) und die räumlich-zeitliche Erweiterung von Marhuenda u. a. (2013). Basierend auf diesen Modellen werden robuste Alternativen vorgestellt. Diese sind eingebettet in die Klasse der gemischten linearen Modelle. In der Kombination ergeben sich damit robuste empirische beste lineare unverzerrte Prediktoren (REBLUPs). Das sind (1) der REBLUP basierend auf den Ergebnissen von Sinha und Rao (2009), (2) ein räumlicher REBLUP (SREBLUP), (3) ein zeitlicher REBLUP (TREBLUP), und (4) ein räumlich-zeitlicher REBLUP (STREBLUP). Diese verschiedenen Erweiterungen werden in einem gemeinsamen Rahmen von robusten Area-Level-Modellen dargestellt, wobei dieser auf gemischten linearen Modellen aufsetzt. Zur Schätzung des MSPE werden ein parametrisches Bootstrap-Verfahren nach Sinha und Rao (2009) verwendet und die Ergebnisse von Chambers, Chandra und Tzavidis (2011) für Area-Level-Modelle erweitert. Zudem wird eine einfache Korrektur des Vorhersagefehlers – der bei der Verwendung robuster Verfahren auftreten kann – vorgeschlagen. Diese Korrektur basiert auf dem *limited translation estimator* von Efron und Morris (1972).

Zusätzlich zu der methodischen Erweiterung der robusten gemischten linearen Modelle werden auch die entsprechenden Implementierungen in Form von Software Paketen vorgestellt. Das R-Paket `saeRobust` (Warnholz, 2016) beinhaltet Funktionen, um alle eingeführten Verfahren in der Praxis anzuwenden. Zudem werden einige numerische Eigenschaften der Implementierung dargelegt sowie

typische Grafiken, um die Güte von Ergebnissen beurteilen zu können. Außerdem wird das Paket `saeSim` (Warnholz und Schmid, 2016) vorgestellt. Hierbei handelt es sich um ein `R`-Paket welches das Aufsetzen von Simulationsstudien vereinfacht. Dabei werden Werkzeuge zur Datengenerierung, Stichprobenziehung und die Möglichkeit des parallelen Rechnens gegeben.

Die eingeführten Methoden werden außerdem in model- und designbasierten Simulationsstudien untersucht. Dabei werden die statistischen Eigenschaften der Punktschätzungen und auch die der MSPE Schätzung dargelegt. Diese Untersuchungen basieren auf Daten, die direkt auf der Gebietsebene generiert werden. Um sowohl einzelne Beobachtungen als auch gesamte Gebiete als Ausreißer untersuchen zu können, werden in einer weiteren Simulation Beobachtungen einer Population simuliert. Der Vorteil dabei ist, dass der gesamte Prozess der Datengenerierung bis hin zu der Anwendung von Area-Level-Methoden abgebildet werden kann. In einer designbasierten Studie werden dann die Steuereinnahmen in 20 Industriesektoren vorhergesagt. Diese Studie nutzt eine synthetischen Population, welche auf dem *Structural Business Survey* in den Niederlanden basiert. Diese Untersuchung kombiniert eine informative Stichprobenziehung mit dem Vorkommen von Ausreißern auf Individual- und Gebietsebene.

C

PUBLICATIONS

Some ideas and figures have appeared previously in the following publication:

Warnholz, S. and T. Schmid (2016). "Simulation Tools for Small Area Estimation: Introducing the R-package saeSim". In: *Austrian Journal of Statistics* 45.1, pp. 55–69.

The results of this Article are the subject matter of Chapter 5. There I state, again, explicitly that these results have been previously published and how I intend to use them in the context of this Thesis.

BIBLIOGRAPHY

Abelson, R. P. (1995). *Statistics as Principled Argument*. New York: Taylor & Francis Group.

Alfons, A., S. Kraft, M. Templ and P. Filzmoser (2011). "Simulation of close-to-reality population data for household surveys with application to EU-SILC". In: *Statistical Methods & Applications* 20.3, pp. 383–407.

Alfons, A., M. Templ and P. Filzmoser (2010). "An Object-Oriented Framework for Statistical Simulation: The R Package simFrame". In: *Journal of Statistical Software* 37.3, pp. 1–36.

Allaire, J.J., J. McPherson, Y. Xie, H. Wickham, J. Cheng and J. Allen (2014). *rmarkdown: Dynamic Documents for R*. R package version 0.3.3. URL: http://CRAN.R-project.org/package=rmarkdown.

Auerbach, K. (2015). "Numerische Verfahren zur Lösung von Small-Area-Problemen". MA thesis. Department of Economics, Freie Universität Berlin, Berlin, Germany.

Bache, S. M. and H. Wickham (2014). *magrittr: A Forward-Pipe Operator for R*. R package version 1.5. URL: https://CRAN.R-project.org/package=magrittr.

Bamberg, G., F. Baur and M. Krapp (2012). *Statistik*. 17th ed. Oldenbourg Verlag.

Bates, D. and M. Maechler (2016). *Matrix: Sparse and Dense Matrix Classes and Methods*. R package version 1.2-4. URL: https://CRAN.R-project.org/package=Matrix.

Battese, G. E., R. M. Harter and W. A. Fuller (1988). "An error component model for prediction of county crop areas using survey and satellite data". In: *Journal of the American Statistical Association* 83 (401), pp. 28–36.

Beaumont, J.-F. (2004). "Robust generalized regression estimation". In: *Survey Methodology* 30.2, pp. 195–208.

Beaumont, J.-F. and L.-P. Rivest (2009). "Sample Surveys: Design, Methods and Applications". In: ed. by D. Pfeffermann and C. R. Rao. Vol. 29 A. Handbook of Statistics. Elsevier. Chap. Dealing with Outliers in Survey Data, pp. 247–280.

Bell, W. R. and E. T. Huang (2006). "Using the t-distribution to Deal with Outliers in Small Area Estimation". In: *Proceedings of Statistics Canada Symposium 2006: Methodological Issues in Measuring Population Health*.

Benavent, R. and D. Morales (2016). "Multivariate Fay-Herriot models for small area estimation". In: *Computational Statistics & Data Analysis* 94, pp. 372–390.

Bischl, B. and M. Lang (2015). *parallelMap: Unified Interface to Parallelization Back-Ends*. R package version 1.3. URL: https://CRAN.R-project.org/package=parallelMap.

Bischl, B., M. Lang, O. Mersmann, J. Rahnenführer and C. Weihs (2015). "BatchJobs and BatchExperiments: Abstraction Mechanisms for Using R in Batch Environments". In: *Journal of Statistical Software* 64.11, pp. 1–25.

Bivand, R. S., E. J. Pebesma and V. Gomez-Rubio (2008). *Applied Spatial Data Analysis with R*. Springer.

Bland, J. M. and D. G. Altman (1986). "Statistical methods for assessing agreement between two methods of clinical measurement". In: *The lancet* 327.8476, pp. 307–310.

Burgard, J.-P., J.-P. Kolb and R. Münnich (2014). "Generation of Synthetic Universes for Micro-Simulations in Survey statistics". In: *Working Paper*.

Chakraborty, A., G. Sankar Datta and A. Mandal (2015). "A two-component normal mixture alternative to the Fay-Herriot model". In: *ArXiv e-prints*. arXiv: 1510.04482 [stat.ME].

Chambers, R. (1986). "Outlier robust finite population estimation". In: *Journal of the American Statistical Association* 81, pp. 1063–1069.

Chambers, R. and H. Chandra (2013). "A Random Effect Block Bootstrap for Clustered Data". In: *Journal of Computational and Graphical Statistics* 22.2, pp. 452–470.

Chambers, R., H. Chandra, N. Salvati and N. Tzavidis (2014). "Outlier Robust Small Area Estimation". In: *Journal of the Royal Statistical Society: Series B* 76 (1), pp. 47–69.

Chambers, R., H. Chandra and N. Tzavidis (2011). "On bias-robust mean squared error estimation for pseudo-linear small area estimators". In: *Survey Methodology* 37 (2), pp. 153–170.

Chambers, R. and N. Tzavidis (2006). "M-Quantile Models for Small Area Estimation". In: *Biometrika* 93 (2), pp. 255–268.

Chandra, H., N. Salvati and R. Chambers (2015). "A Spatially Nonstationary Fay-Herriot Model for Small Area Estimation". In: *Journal of Survey Statistics and Methodology* 3.2, pp. 109–135.

Chandra, H., N. Salvati, R. Chambers and N. Tzavidis (2012). "Small Area Estimation under Spatial Nonstationarity". In: *Computational Statistics & Data Analysis* 56.10, pp. 2875–2888.

Chatrchi, G. (2012). "Robust Estimation of Variance Components in Small Area Estimation". MA thesis. School of Mathematics and Statistics, Carleton University, Ottawa, Canada.

Chen, S. and P. Lahiri (2003). "A comparison of different MSPE estimators of EBLUP for the Fay-Herriot model". In: *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 905–911.

— (2008). "On Mean Squared Prediction Error Estimation in Small Area Estimation Problems". In: *Communications in Statistics-Theory and Methods* 37.11, pp. 1792–1798.

Choudhry, G. H. and J. N. K. Rao (1989). "Small area estimation using models that combine time series and cross-sectional data". In: *Proceeding of the Statistics Canada Symposium on Analysis of Data in Time*, pp. 67–74.

Clement, E. P. (2014). "Small Area Estimation with Application to Disease Mapping". In: *International Journal of Probability and Statistics* 3.1, pp. 15–22.

Cressie, N. (1993). *Statistics for spatial data*. New York: Wiley.

Datta, G. S. (2009). "Handbook of Statistics: Inference and Analysis". In: ed. by D. Pfeffermann and C.R. Rao. Vol. 29 B. Elsevier. Chap. Model-Based Approach to Small Area Estimation, pp. 251–288.

Datta, G. S. and P. Lahiri (1995). "Robust Hierarchical Bayes Estimation of Small Area Characteristics in the Presence of Covariates and Outliers". In: *Journal of Multivariate Analysis* 54, pp. 310–328.

— (2000). "A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems". In: *Statistica Sinica* 10, pp. 613–627.

Datta, G. S., P. Lahiri and T. Maiti (2002). "Empirical Bayes estimation of median income of four-person families by state using time series and cross-sectional data". In: *Journal of Statistical Planning and Inference* 102.1, pp. 83–97.

Datta, G. S., J. N. K. Rao and D. D. Smith (2005). "On measuring the variability of small area estimators under a basic area level model". In: *Biometrika* 92 (1), pp. 183–196.

Efron, B. and C. Morris (1971). "Limiting the Risk of Bayes and Empirical Bayes Estimators - Part I: The Bayes Case". In: *Journal of the American Statistical Association* 66.336, pp. 807–815.

— (1972). "Limiting the Risk of Bayes and Empirical Bayes Estimators - Part II: The Empirical Bayes Case". In: *Journal of the American Statistical Association* 67.337, pp. 130–139.

Eurostat (2015). *Regions in the European Union: Nomenclature of territorial units for statistics (NUTS 2013/EU-28)*. Manuals and guidelines. European Union. DOI: doi:10.2785/53780.

Fabrizi, E. and C. Trivisano (2010). "Robust linear mixed models for Small Area Estimation". In: *Journal of Statistical Planning and Inference* 140.2, pp. 433–443.

Fay, R. E. and R. A. Herriot (1979). "Estimation of Income for Small Places: An Application of James-Stein Procedures to Census Data". In: *Journal of the American Statistical Association* 74 (366), pp. 269–277.

Fellner, W. H. (1986). "Robust Estimation of Variance Components". In: *Technometrics* 28 (1), pp. 51–60.

Gershunskaya, J. (2010). "Robust Small Area Estimation Using a Mixture Model". In: *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 2783–2796.

Ghosh, M., T. Maiti and A. Roy (2008). "Influence functions and robust Bayes and empirical Bayes small area estimation". In: *Biometrika* 95.3, pp. 573–585.

Ghosh, M. and J. N. K. Rao (1994). "Small Area Estimation: An Appraisal". In: *Statistical Science* 9 (1), pp. 55–93.

Guadarrama, M., I. Molina and J. N. K. Rao (2015). *A Comparison of Small Area Estimation Methods for Poverty Mapping*. UC3M Working Papers: Statistics and Econometrics 15–05. Universidad Carlos III de Madrid, Departamento de Estadistica.

Hall, P. and T. Maiti (2006). "On Parametric Bootstrap Methods for Small Area Prediction". In: *Journal of the Royal Statistical Society: Series B* 68 (2), pp. 221–238.

Hansen, M. H. and W. N. Hurwitz (1943). "On the Theory of Sampling from Finite Populations". In: *The Annals of Mathematical Statistics* 14.4, pp. 333–362.

Henderson, C. R. (1950). "Estimation of genetic parameters". In: *Annals of Mathematical Statistics* 21 (2), pp. 309–310.

— (1963). "Selection Index and Expected Genetic Advance". In: *Statistical Genetics and Plant Breeding*. Vol. 982. Washington, DC: National Academy of Sciences – National Research Council, pp. 141–163.

Horvitz, D.G. and D.J. Thompson (1952). "A generalization of sampling without replacement from a finite universe". In: *Journal of the American Statistical Association* 47, pp. 663–685.

Huang, E. T. and W. R. Bell (2006). "Using the t-distribution in Small Area Estimation: An Application to SAIPE State Poverty Models". In: *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 3142–3149.

Huber, P. J. (1964). "Robust Estimation of a Location Parameter". In: *The Annals of Mathematical Statistics* 35.1, pp. 73–101.

Huggins, R. M. (1993). "A robust approach to the analysis of repeated measures". In: *Biometrics* 49, pp. 715–720.

Hulliger, B (1999). "Simple and Robust Estimators for Sampling". In: *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 54–63.

Jiang, J. and P. Lahiri (2006). "Mixed Model Prediction and Small Area Estimation". In: *Test* 15 (1), pp. 1–96.

Jiang, J., P. Lahiri and S.-M. Wan (2002). "A Unified Jackknife Theory for Empirical Best Prediction with M-Estimation". In: *The Annals of Statistics* 30 (6), pp. 1782–1810.

Jiongo, V. D., D. Haziza and P. Duchesne (2013). "Controlling the bias of robust small-area estimators". In: *Biometrika*, pp. 843–858.

Jiongo, V. D. and P. Nguimkeu (2014). *Bootstrapping Mean Squared Errors of Robust Small Area Estimators*. Tech. rep. Georgia State University, Department of Economics. URL: http://www2.gsu.edu/~econpn/Pierre_Nguimkeu/%20RESEARCH_files/bootstrap_mse%_24122014u.pdf.

Kacker, R. N. and D. A. Harville (1984). "Approximations for Standard Errors of Estimators of Fixed and Random Effects in Mixed Linear Models". In: *Journal of the American Statistical Association* 79, pp. 853–862.

Kennedy Jr, W. J. and E. James (1980). *Statistical Computing*. New York: Marcel Decker, Inc.

Knuth, Donald E. (1992). *Literate Programming*. Stanford: CSLI.

Kolb, J.-P. (2013). "Generation of Synthetic Universes". PhD thesis. University of Trier. URL: http://ubt.opus.hbz-nrw.de/volltexte/2013/816/.

Koller, M. (2016). *robustlmm: Robust Linear Mixed Effects Models*. R package version 1.8. URL: https://CRAN.R-project.org/package=robustlmm.

Landau, D. P. and K. Binder (2014). *A guide to Monte Carlo Simulations in Statistical Physics*. Cambridge University Press.

Lange, K. L., R. J. A. Little and J. M. G. Taylor (1989). "Robust Statistical Modeling Using the t Distribution". In: *Journal of the American Statistical Association* 84.408, pp. 881–896.

Leek, J. (2013). "A couple of requests for the @Statistics2013 future of statistics workshop". In: *Simply Statistics* 20 August 2013. URL: http://simplystatistics.org/2013/08/20/a-couple-of-requests-for-the-statistics2013-future-of-statistics-workshop/.

Lefler, M. D. Esteban, D. M. Gonzalez and A. P. Martin (2014). *saery: Small Area Estimation for Rao and Yu Model*. R package version 1.0. URL: https://CRAN.R-project.org/package=saery.

Lehtonen, R. and A. Veijanen (2009). "Sample Surveys: Inference and Analysis". In: ed. by D. Pfeffermann and C.R. Rao. Vol. 29 B. Handbook of Statistics. Elsevier. Chap. Design-based Methods of Estimation for Domains and Small Areas, pp. 219–249.

Leisch, F. (2002). "Sweave, Part I: Mixing R and LaTeX". In: *R News* 2.3, pp. 28–31.

Lohr, S. L. and J. N. K. Rao (2009). "Jackknife estimation of mean squared error of small area predictors in nonlinear mixed models". In: *Biometrika*, pp. 1–12.

López-Vizcaíno, E., M. J. Lombardía and D. Morales (2015). "Small area estimation of labour force indicators under a multinomial model with correlated time and area effects". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 178.3, pp. 535–565.

Maiti, T., H. Ren and S. Sinha (2014). "Prediction Error of Small Area Predictors Shrinking Both Means and Variances". In: *Scandinavian Journal of Statistics* 41.3, pp. 775–790.

Marhuenda, Y., I. Molina and D. Morales (2013). "Small area estimation with spatio-temporal Fay-Herriot models". In: *Computational Statistics and Data Analysis* 58, pp. 308–325.

Maronna, R. A., D. Martin and V. J. Yohai (2006). *Robust Statistics: Theory and Methods*. Chichester.

McCullough, B. D. (2004). "Numerical Issues in Statistical Computing for the Social Scientist". In: ed. by D. J. Balding, N. A. C. Cressie and N. I. Fisher. John Wiley & Sons, Inc. Chap. Some Details of Nonlinear Estimation, pp. 199–218.

Meindl, B., M. Templ, A. Alfons, A. Kowarik and with contributions from M. Ribatet (2014). *simPop: Simulation of Synthetic Populations for Survey Data Considering Auxiliary Information*. R package version 0.2.6. URL: http://CRAN.R-project.org/package=simPop.

Mokhtarian, P. and R. Chambers (2013). "An outlier robust block bootstrap for small area estimation". In: *Proceedings of the 59th World Statistics Congress of the International Statistical Institute*. International Statistical Institute, Hague, Netherlands, pp. 798–803.

Molina, I. and Y. Marhuenda (2015). "sae: An R Package for Small Area Estimation". In: *The R Journal* 7.1, pp. 81–98.

Molina, I., N. Salvati and M. Pratesi (2009). "Bootstrap for estimating the mean squared error of the Spatial EBLUP". In: *Computational Statistics* 24 (3), pp. 441–458.

Münnich, R., J. Schürle, W. Bihler, H. Boonstra, P. Knotterus, N. Nieuwenbroek, A. Haslinger, S. Laaksonen, D. Eckmair, A. Quatember, H. Wagner, J. Renfer, U. Oetliker and R. Wiegert (2003). *Monte Carlo Simulation Study of European Surveys*. University of Tübingen. URL: https://www.uni-trier.de/fileadmin/fb4/projekte/SurveyStatisticsNet/Dacseis%5C_Deliverables/DACSEIS-D3-1-D3-2.pdf.

Namazi-Rad, M.-R. and D. Steel (2015). "What Level of Statistical Model Should We Use in Small Area Estimation?" In: *Australian & New Zealand Journal of Statistics* 57.2, pp. 275–298.

Peel, D. and G. J. McLachlan (2000). "Robust mixture modelling using the t distribution". In: *Statistics and Computing* 10.4, pp. 339–348.

Petrucci, A. and N. Salvati (2006). "Small Area Estimation for Spatial Correlation in Watershed Erosion Assessment". In: *Journal of agricultural, biological, and environmental statistics* 11.2, pp. 169–182.

Pfeffermann, D. (2002). "Small Area Estimation - New Developments and Directions". In: *International Statistical Review* 70 (1), pp. 125–143.

— (2013). "New Important Developments in Small Area Estimation". In: *Statistical Science* 28 (1), pp. 40–68.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar and R Core Team (2016). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-126. URL: http://CRAN.R-project.org/package=nlme.

Porter, A. T., S. H. Holan, C. K. Wikle and N. Cressie (2014). "Spatial Fay-Herriot models for small area estimation with functional covariates". In: *Spatial Statistics* 10, pp. 27–42.

Prasad, N. G. N. and J. N. K. Rao (1990). "The Estimation of the Mean Squared Error of Small Area Estimators". In: *Journal of the American Statistical Association* 85 (409), pp. 163–171.

Pratesi, M. and N. Salvati (2008). "Small area estimation: the EBLUP estimator based on spatially correlated random area effects". In: *Statistical Methods & Applications* 17, pp. 113–141.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Rao, J. N. K. (1999). "Some recent advances in model-based small area estimation". In: *Survey Methodology* 25, pp. 175–186.

— (2003). *Small Area Estimation*. New York: Wiley.

Rao, J. N. K. and I. Molina (2015). *Small Area Estimation*. New York: Wiley.

Rao, J. N. K., S. K. Sinha and L. Dumitrescu (2014). "Robust small area estimation under semi-parametric mixed models". In: *Canadian Journal of Statistics* 42.1, pp. 126–141.

Rao, J. N. K. and M. Yu (1994). "Small-Area Estimation by Combining Time-Series and Cross-Sectional Data". In: *Canadian Journal of Statistics* 22.4, pp. 511–528.

Richardson, A. M. and A. H. Welsh (1995). "Robust Restricted Maximum Likelihood in Mixed Linear Models". In: *Biometrics* 51 (4), pp. 1429–1439.

Rivest, L.-P. and N. Vandal (2003). "Mean Squared Error Estimation for Small Areas when the Small Area Variances are Estimated". In: *Proceedings of the International Conference on Recent Advances in Survey Sampling*. Laboratory for Research in Statistics and Probability, Carleton University Ottawa, ON, Canada.

Royall, R. M. and W. G. Cumberland (1978). "Variance estimation in finite population sampling". In: *Journal of the American Statistical Association* 73, pp. 351–358.

Rubin, D. B. (1993). "Statistical Disclosure Limitation". In: *Journal of Official Statistics* 9.2, pp. 461–468.

Salvati, N., H. Chandra, M. Giovanna-Ranalli and R. Chambers (2010). "Small area estimation using nonparametric model-based direct estimator". In: *Computational Statistics and Data Analysis* 54 (9), pp. 2159–2171.

Salvati, N., N. Tzavidis, M. Pratesi and R. Chambers (2012). "Small area estimation via M-quantile geographically weighted regression". In: *Test* 21.1, pp. 1–28.

Särndal, C.-E., B. Swensson and J.H. Wretman (1992). *Model Assisted Survey Sampling*. New York: Springer.

Schmid, T. (2011). "Spatial Robust Small Area Estimation applied on Business Data". PhD thesis. University of Trier.

Schmid, T. and R. T. Münnich (2014). "Spatial robust small area estimation". In: *Statistical Papers* 55.3, pp. 653–670.

Schmid, T., N. Tzavidis, R. T. Münnich and R. Chambers (2016). "Outlier Robust Small-Area Estimation Under Spatial Correlation". In: *Scandinavian Journal of Statistics*.

Schoch, T. (2012). "Robust Unit-Level Small Area Estimation: A Fast Algorithm for Large Datasets". In: *Austrian Journal of Statistics* 41 (4), pp. 243–265.

— (2014). *rsae: Robust Small Area Estimation*. R package version 0.1-5. URL: https://CRAN.R-project.org/package=rsae.

Searle, S. R. (1971). *Linear Models*. New York: Wiley.

Singh, A. C., H. J. Mantel and B. W. Thomas (1991). *Time Series Generalization of Fay-Herriot Estimation for Small Areas*. Methodology Branch, Business Survey Methods Division, Statistics Canada.

Singh, B. B., G. K. Shukla and D. Kundu (2005). "Spatio-temporal models in small area estimation". In: *Survey Methodology* 31.2, p. 183.

Sinha, S. K. and J. N. K. Rao (2008). "Robust Small Area Estimation under Unit Level Models". In: *Proceedings of the Survey Research Methods Section, American Statistical Association*. Vol. 82.

— (2009). "Robust Small Area Estimation". In: *The Canadian Journal of Statistics* 37 (3), pp. 381–399.

Slud, E. V. and T. Maiti (2006). "Mean-squared error estimation in transformed Fay-Herriot models". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68.2, pp. 239–257.

Sugasawa, S. and T. Kubokawa (2015). "Parametric transformed Fay-Herriot model for small area estimation". In: *Journal of Multivariate Analysis* 139, pp. 295–311.

Thisted, R. A. (1988). *Elements of Statistical Computing: Numerical Computation*. New York: CRC Press.

Tukey, J. W. (1979). "Robust Techniques for the User". In: *Robustness in Statistics*, pp. 103–106.

Tzavidis, N. and R. Chambers (2005). "Bias Adjusted Small Area Estimation". In: *Statistics in Transition* 7 (3), pp. 707–713.

Tzavidis, N., S. Marchetti and R. Chambers (2010). "Robust estimation of small area means and quantiles". In: *Australian and New Zealand Journal of Statistics* 52 (2), pp. 167–186.

Wang, J. and W. A. Fuller (2003). "The Mean Squared Error of Small Area Predictors Constructed With Estimated Area Variances". In: *Journal of the American Statistical Association* 98 (463), pp. 716–723.

Warnholz, S. (2016). *saeRobust: Robust Small Area Estimation*. R package.

Warnholz, S. and T. Schmid (2016). *saeSim: Simulation Tools for Small Area Estimation*. R package version 0.8.0. URL: http://CRAN.R-project.org/package=saeSim.

Warnholz, S. and T. Schmid (2016). "Simulation Tools for Small Area Estimation: Introducing the R-package saeSim". In: *Austrian Journal of Statistics* 45.1, pp. 55–69.

Weihs, C., O. Mersmann and U. Ligges (2014). *Foundations of Statistical Algorithms: With References to R Packages*. CRC Press.

Wickham, H. (2016). *tidyr: Easily Tidy Data with Spread and Gather Functions.* R package version 0.4.1. URL: http://CRAN.R-project.org/package=tidyr.

Wickham, H. and R. Francois (2015). *dplyr: A Grammar of Data Manipulation*. R package version 0.4.3. URL: http://CRAN.R-project.org/package=dplyr.

Wolter, K. (2007). *Introduction to Variance Estimation*. Springer Science & Business Media.

Xie, D., T. E. Raghunathan and J. M. Lepkowski (2007). "Estimation of the proportion of overweight individuals in small areas–a robust extension of the Fay-Herriot model". In: *Statistics in Medicine* 26.13, pp. 2699–2715.

Xie, Y. (2013). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC.

Yau, K. K. W. and A. Y. C. Kuk (2002). "Robust Estimation in Generalized Linear Mixed Models". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64.1, pp. 101–117.

You, Y. and B. Chapman (2006). "Small Area Estimation Using Area Level Models and Estimated Sampling Variances". In: *Survey Methodology* 32.1, pp. 97–103.

Zielke, G. (1974). "Testmatrizen mit maximaler Konditionszahl". In: *Computing* 13.1, pp. 33–54.

DECLARATION

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

Berlin, April, 2016

Sebastian Warnholz,
28th April 2016