DISSERTATION

SEBASTIAN WARNHOLZ

Freie Universität Berlin

Robust Area-Level Models in Small Area Estimation:
Theory, Software and Simulation Studies

May 29, 2015 – version 0.1

CONTENTS

i THEORY 1

1 REVIEW OF ROBUST METHODS IN SMALL AREA ESTIMATION 3

1.1 new section 4

2 THE ROBUST FAY-HERRIOT MODEL 5

2.1 Mean squared error 6

2.1.1 pseudo linear FH 6

ii SOFTWARE 7

3 SIMULATION TOOLS FOR SMALL AREA ESTIMATION 9

4 IMPLEMENTATION OF ALGORITHMS 11

5 ROBUST AREA-LEVEL MODELS IN SMALL AREA ESTIMATION 13

iii SIMULATION STUDIES 15

6 MODEL-BASED SIMULATIONS 17

7 DESIGN-BASED SIMULATIONS 19

iv APPENDIX 21

BIBLIOGRAPHY 23

Part I

THEORY

This is the chapter where I want to present the theoretical concepts underpinning the development of software and application. Most notably is the robust version of a Fay-Herriot Type model with different variance-covariance structures.

1

REVIEW OF ROBUST METHODS IN SMALL AREA ESTIMATION

1.1 NEW SECTION

- item
  - item 1.1
  - item 1.2
  - new item

$$x_i = y_i \tag{1}$$
$$x_i + y_i = y_i \tag{2}$$

cite me: Abberger (1997)
Reference math: (1)
Math without numbering:

$$x_i = y_i$$
$$x_i + y_i = y_i$$

```
x <- 1
```

```
ggplot(data.frame(x = 1:10, y = 11:20)) +
  geom_point(aes(x, y)) +
  theme_thesis()
```

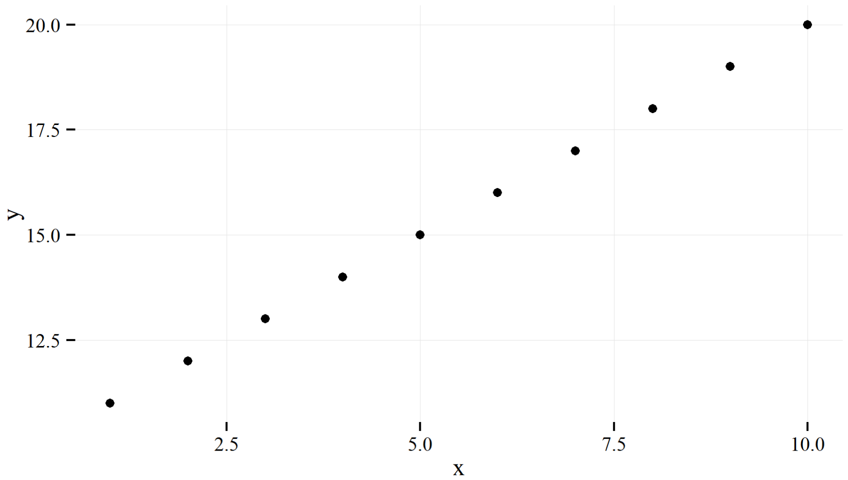

Figure 1: plot of chunk unnamed-chunk-2

2 THE ROBUST FAY-HERRIOT MODEL

2.1 MEAN SQUARED ERROR

2.1.1 *pseudo linear FH*

This is the representation of the pseudo linear representation of the FH model. As it is introduced in Chambers, J. Chandra, and Tzavidis (2011) and Chambers, H. Chandra, et al. (2014).

Presenting the FH in pseudo linear form means to present the area means as a weighted sum of the response vector y. The FH model is given by

$$\theta_i = \gamma_i y_i + (1 - \gamma_i) x_i^\top \beta \tag{3}$$

where $\gamma_i = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$, so it can be represented as

$$\theta_i = w_i^\top y$$

where

$$w_i^\top = \gamma_i \mathbf{I}_i^\top + (1 - \gamma_i) x_i^\top \mathbf{A}$$

and

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

with

$$\mathbf{W} = \mathrm{Diag}(w_j), \text{ with } j = 1, \dots, n$$

and

$$w_j = \frac{\psi \left(u_j^{-\frac{1}{2}} (y_j - x_j^\top \beta) \right)}{u_j^{-\frac{1}{2}} (y_j - x_j^\top \beta}$$

Note that if ψ is the identity or equally a huber influence function with a large smoothing constant, i. e.inf:

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

Part II

SOFTWARE

This is the part where I want to introduce the software where the theoretical concepts find implementation.

3 SIMULATION TOOLS FOR SMALL AREA ESTIMATION

4 IMPLEMENTATION OF ALGORITHMS

5 ROBUST AREA-LEVEL MODELS IN SMALL AREA ESTIMATION

Part III

SIMULATION STUDIES

This is the part where I will present all results. Most certainly they will contain a lot of model- and design-based simulation studies for various settings. Maybe there will be more data available and I can present some applications.

6 MODEL-BASED SIMULATIONS

7 DESIGN-BASED SIMULATIONS

Part IV

APPENDIX

BIBLIOGRAPHY

Abberger, K. (1997). "Quantile smoothing in financial time series." In: *Statistical Papers* 38 (2), pp. 125–148.

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

Chambers, R., J. 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.