# Propagating the uncertainty of the market value by the use of a Bayesian regression approach

## XXV FIG Congress

Sebastian Zaddach and Hamza Alkhatib

Geodetic Institute
Leibniz Universität Hannover

Kuala Lumpur, Malaysia, June 19$^{th}$ 2014

# Contents

# Motivation

- Aim of real estate valuation: determination of market value
- Sales comparison approach: method with the highest marketability
- Hedonic model of the classical linear regression analysis
  - Most frequently applied model in valuation practice
  - Uncertainty is not taken into account explicitly


**Aim of this research**

Introduction of a Bayesian approach:

- **Quantifying the uncertainty** of single variables
- **Reduction of uncertainty** concerning the estimation of parameters

# Motivation

- Aim of real estate valuation: determination of market value
- Sales comparison approach: method with the highest marketability
- Hedonic model of the classical linear regression analysis
  - Most frequently applied model in valuation practice
  - Uncertainty is not taken into account explicitly


**Aim of this research**

Introduction of a Bayesian approach:

- **Quantifying the uncertainty** of single variables
- **Reduction of uncertainty** concerning the estimation of parameters

# Theory of Multiple Linear Regression Analysis

- Standard procedure for analysing the real estate market
- Approximation method:
  - Explanation of variations of a dependent variable $\boldsymbol{y}$ by the variabilty of independent variables $\boldsymbol{X}$
  - Aim: Estimation of trend parameters (coefficients) $\beta$

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}, \quad \boldsymbol{P} = \sigma_0^2 \cdot \boldsymbol{I}$$

$$\hat{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}, \quad \boldsymbol{V} = (\boldsymbol{X}'\boldsymbol{X})^{-1}$$

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1m} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nm} \end{bmatrix}, \quad \beta = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix}, \quad \boldsymbol{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

$\boldsymbol{y}$: Independent variables (comparative values)
$\beta$: Regression coefficients (parameters)
$m$: Number of independent variables

$\boldsymbol{X}$: Dependent variables
$\boldsymbol{e}$: Residuals
$i = 1, \ldots, n$ : number of data sets

# Theory of Multiple Linear Regression Analysis

- Standard procedure for analysing the real estate market
- Approximation method:
  - Explanation of variations of a dependent variable $\boldsymbol{y}$ by the variabilty of independent variables $\boldsymbol{X}$
  - Aim: Estimation of trend parameters (coefficients) $\beta$

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{e}, \quad \boldsymbol{P} = \sigma_0^2 \cdot \boldsymbol{I}$$

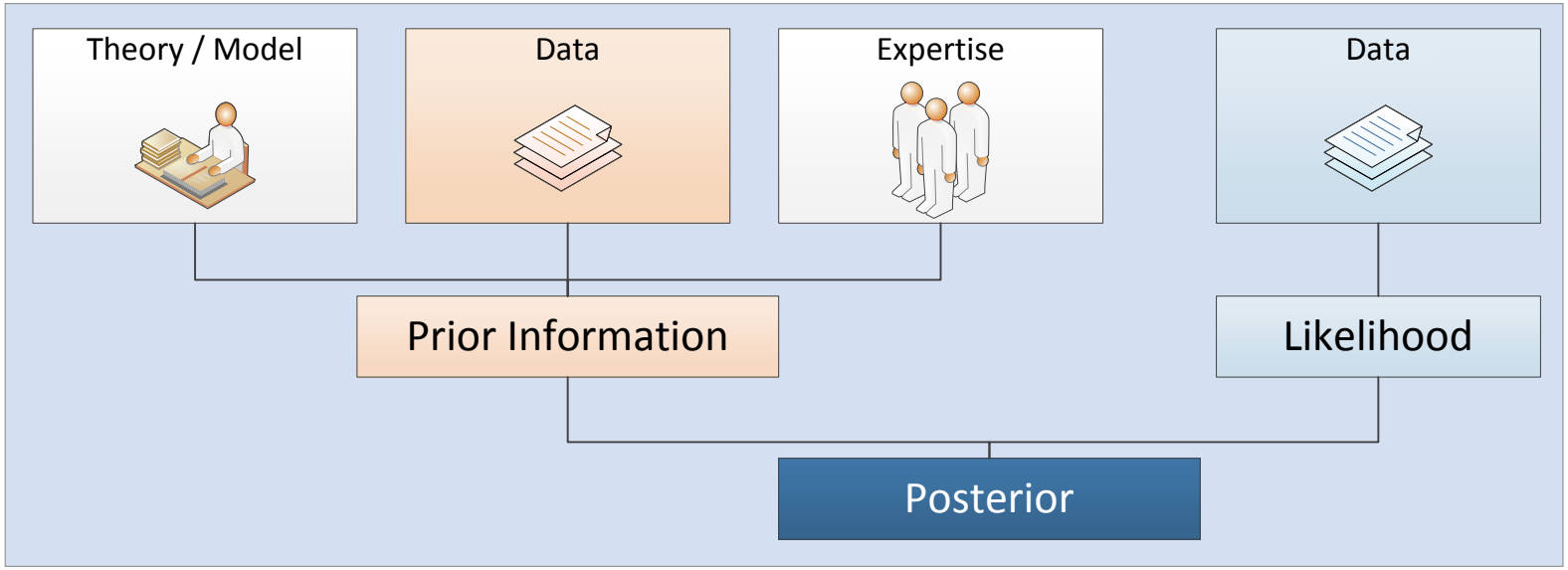
$$\hat{\beta} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{y}, \quad \boldsymbol{V} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}$$

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1m} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nm} \end{bmatrix}, \quad \beta = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix}, \quad \boldsymbol{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

$\boldsymbol{y}$: Independent variables (comparative values)
$\beta$: Regression coefficients (parameters)
$m$: Number of independent variables

$\boldsymbol{X}$: Dependent variables
$\boldsymbol{e}$: Residuals
$i = 1, \ldots, n$ : number of data sets

## *Basic idea of the Bayesian approach*



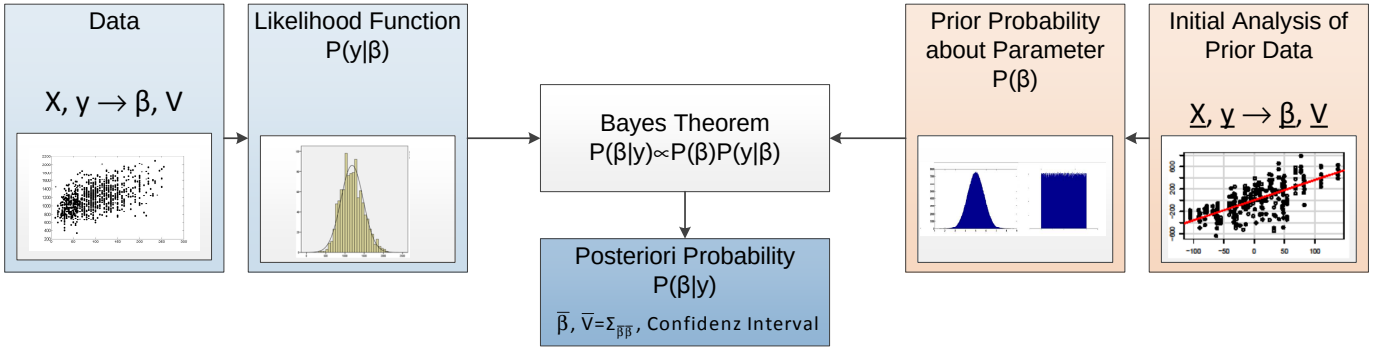

cf. Weitkamp/Alkhatib 2012: The Bayesian approach in the valuation. Proceedings of FIG Working Week 2012, Rome, Italy,

URL: www.fig.net

$$P(\beta|\mathbf{y}) \propto P(\beta)\, P(\mathbf{y}|\beta)$$

Posterior density $\propto$ Prior density $\cdot$ Likelihood

## *Bayesian Regression Approach*



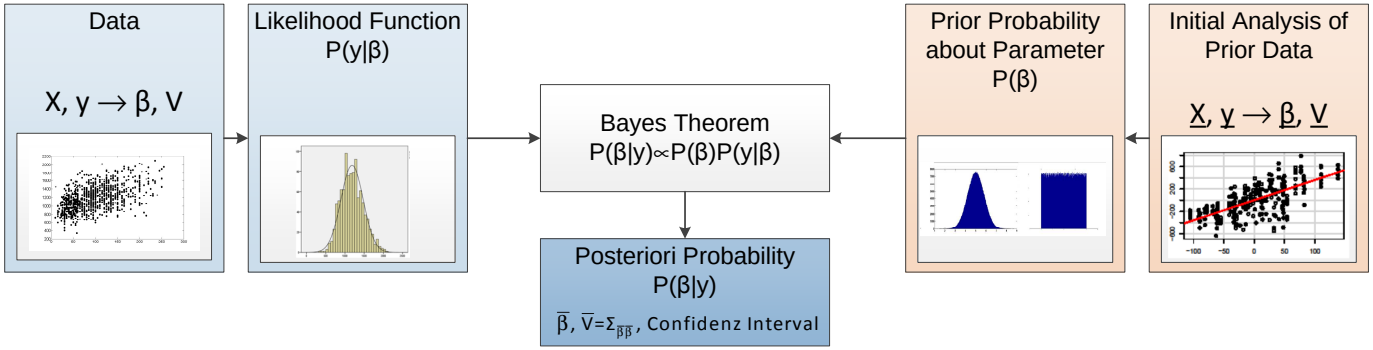

| Classical Approach | Bayesian Approach |
|---|---|
| $\beta = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ | $\overline{\beta} = (\mathbf{X}'\mathbf{X} + \underline{\mathbf{V}}^{-1})^{-1}(\mathbf{X}'\mathbf{y} + \underline{\mathbf{V}}^{-1}\underline{\beta})$ |
| $\mathbf{V} = (\mathbf{X}'\mathbf{X})^{-1}$ | $\overline{\mathbf{V}} = (\mathbf{X}'\mathbf{X} + \underline{\mathbf{V}}^{-1})^{-1}$ |

Posterior parameters: overlined / Prior parts: underlined und red

$\beta$: Regression Coefficients (Parameters)
$\mathbf{X}$: Design Matrix (Independent Variables)
$\mathbf{V}$: Cofaktor Matrix
$\mathbf{y}$: Dependent Variable

cf. Weitkamp/Alkhatib 2012: The Bayesian approach in the valuation. Proceedings of FIG Working Week 2012, Rome, Italy

## *Bayesian Regression Approach*



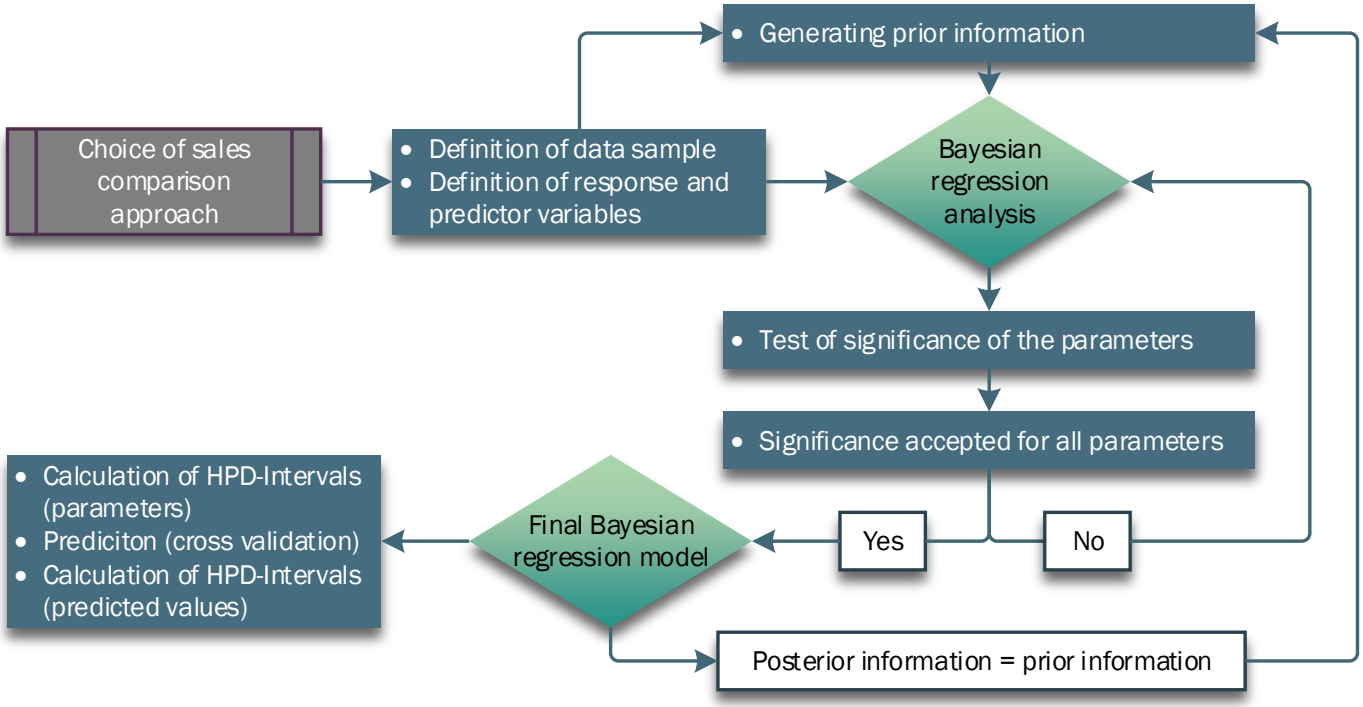

| Classical Approach | Bayesian Approach |
|---|---|
| $\beta = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ | $\overline{\beta} = (\mathbf{X}'\mathbf{X} + \underline{\mathbf{V}}^{-1})^{-1}(\mathbf{X}'\mathbf{y} + \underline{\mathbf{V}}^{-1}\underline{\beta})$ |
| $\mathbf{V} = (\mathbf{X}'\mathbf{X})^{-1}$ | $\overline{\mathbf{V}} = (\mathbf{X}'\mathbf{X} + \underline{\mathbf{V}}^{-1})^{-1}$ |

Posterior parameters: overlined / Prior parts: underlined und red

$\beta$: Regression Coefficients (Parameters)
**X**: Design Matrix (Independent Variables)
**V**: Cofaktor Matrix
**y**: Dependent Variable

cf. Weitkamp/Alkhatib 2012: The Bayesian approach in the valuation. Proceedings of FIG Working Week 2012, Rome, Italy

## *Recursive Bayesian Estimation*



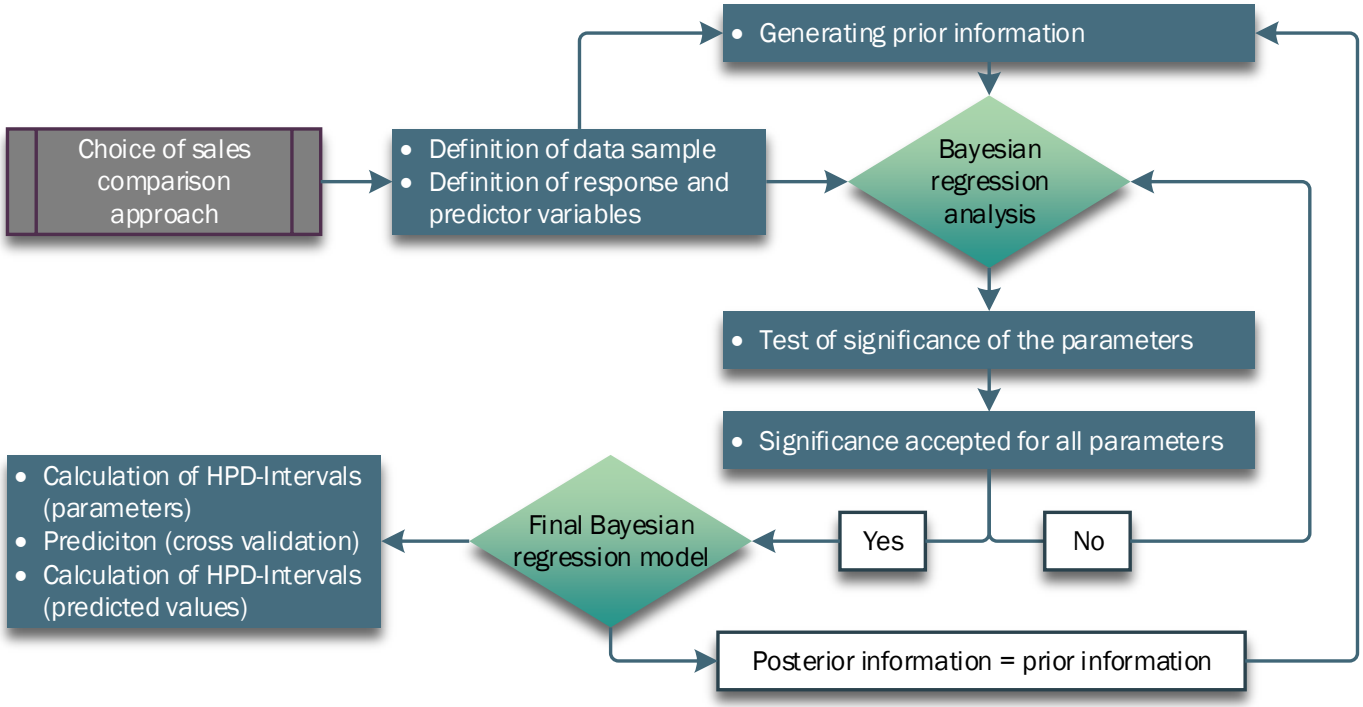

## *Pracital research approach: Description of submarket*

- Data sets based on the Automated Purchase Records (APR)
- Spatial submarket: Hanover; Objective submarket: condominiums
- Number of data sets: 489 (2008), 132 (2009), 184 (2010)
- Value affecting charasteristics: standard land value (€/qm), age (years), area of living space (qm), distance green areas (m), quality of location (-)

# The Bayesian Approach in Real Estate Valuation



## *Recursive Bayesian Estimation*




## *Pracital research approach: Description of submarket*

- Data sets based on the Automated Purchase Records (APR)
- Spatial submarket: Hanover; Objective submarket: condominiums
- Number of data sets: 489 (2008), 132 (2009), 184 (2010)
- Value affecting charasteristics: standard land value (€/qm), age (years), area of living space (qm), distance green areas (m), quality of location (-)

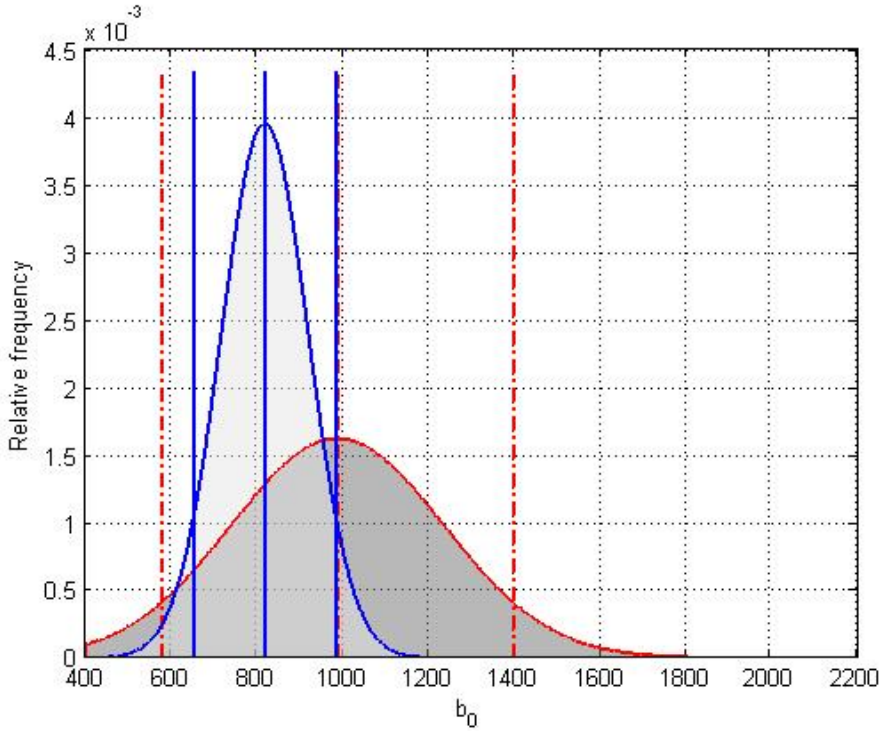
## *Practical results: Estimation*

| Coefficient | | Prior (2008) | Posterior based on non-informative prior (2009) | Posterior based on informative prior (2009) |
|---|---|---|---|---|
| Intercept | $\beta_0$ | 776.97 | 926.18 | 820.56 |
| Standard land value | $\beta_1$ | 9.65 | 2.37 | 8.67 |
| Age | $\beta_2$ | -3.05 | -2.40 | -3.36 |
| Area of living space | $\beta_3$ | 5.43 | 6.00 | 5.49 |
| Dist. recreation area | $\beta_4$ | -13.15 | -16.99 | -14.69 |
| Quality of location | $\beta_5$ | -70.91 | -51.33 | -69.86 |

| Coefficient | | Prior (2009) | Posterior based on non-informative prior (2010) | Posterior based on informative prior (2010) |
|---|---|---|---|---|
| Intercept | $\beta_0$ | 820.56 | 740.13 | 836.59 |
| Standard land value | $\beta_1$ | 8.67 | 15.12 | 9.13 |
| Age | $\beta_2$ | -3.36 | -3.29 | -3.33 |
| Area of living space | $\beta_3$ | 5.49 | 5.40 | 5.50 |
| Dist. recreation area | $\beta_4$ | -14.69 | -11.59 | -13.17 |
| Quality of location | $\beta_5$ | -69.86 | -74.49 | -80.71 |

## *Practical results: Confidence regions - HPDI (coefficients)*

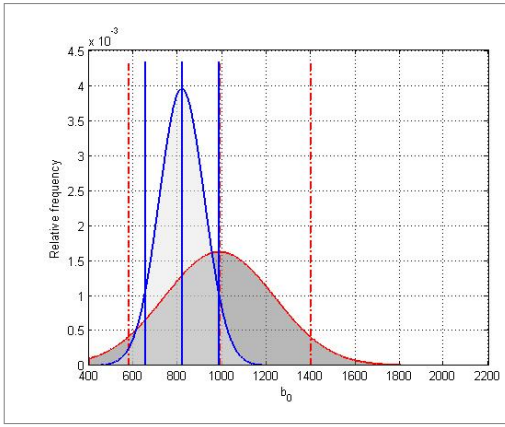
- Results for the posterior solution 2009



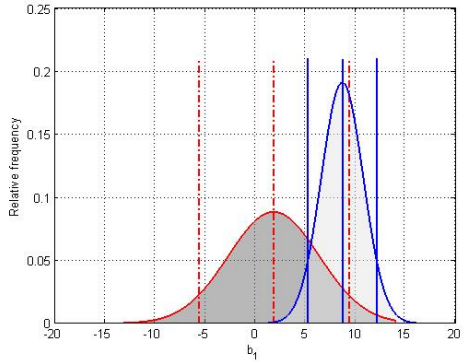

Blue color: Result of informative solution —— : 95 % HPDI
Red color : Result of non-informative solution - - - : 95 % HPDI

## *Practical results: Confidence regions - HPDI (2009)*

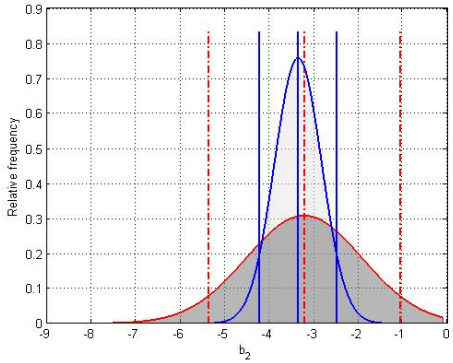
Offset $\overline{\beta}_0$:



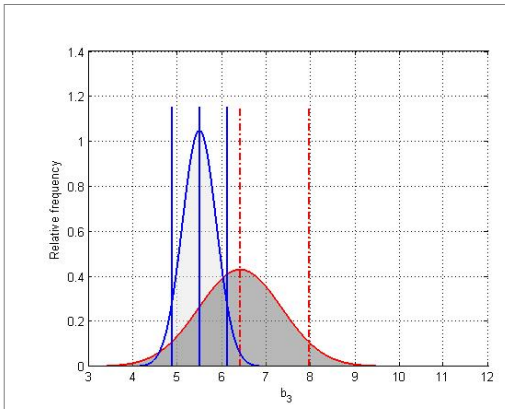

Standard land value $\overline{\beta}_1$:



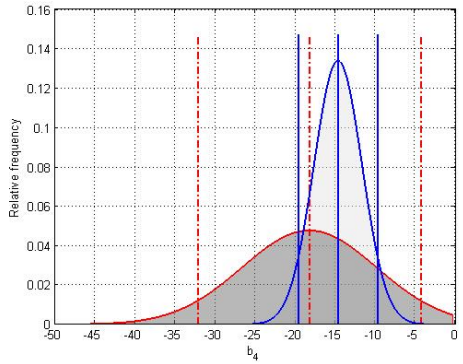

Age $\overline{\beta}_2$:



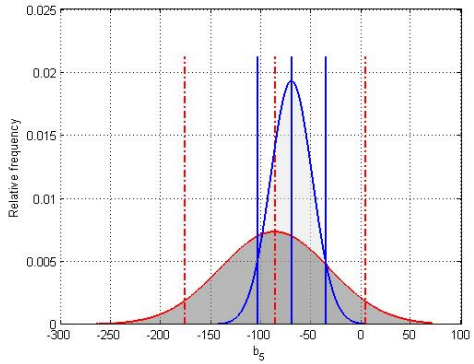

Area of living space $\overline{\beta}_3$:




Dist. to recreation area $\overline{\beta}_4$:

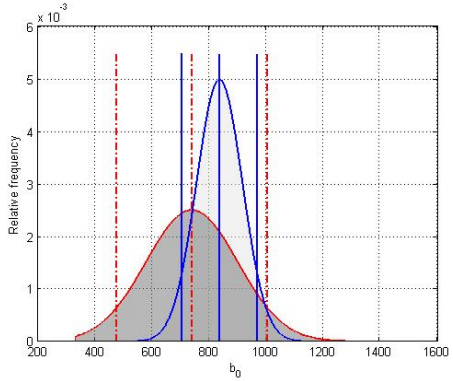



Quality of Location $\overline{\beta}_5$:

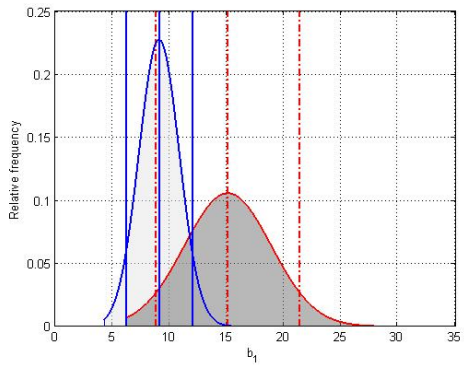


## *Practical results: Confidence regions - HPDI (2010)*

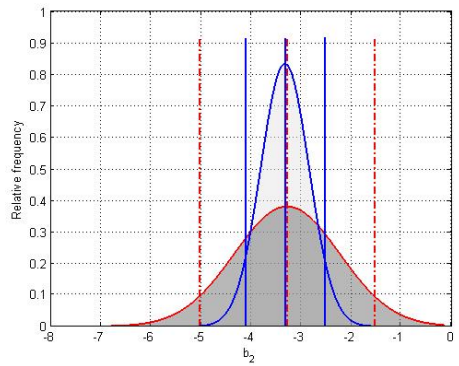
Offset $\overline{\beta}_0$:



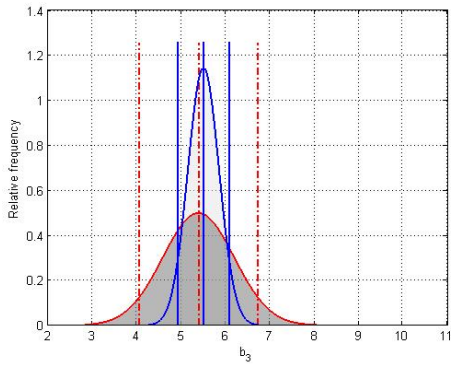

Standard land value $\overline{\beta}_1$:



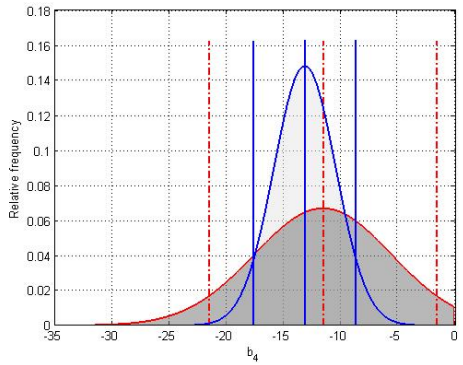

Age $\overline{\beta}_2$:



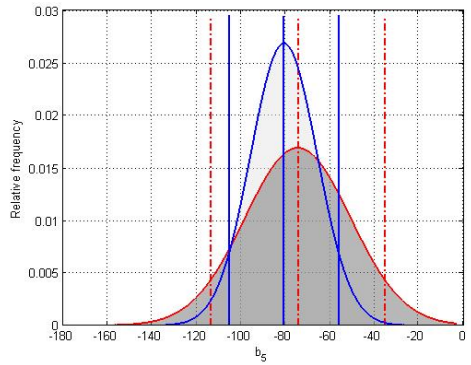

Area of living space $\overline{\beta}_3$:




Dist. to recreation area $\overline{\beta}_4$:

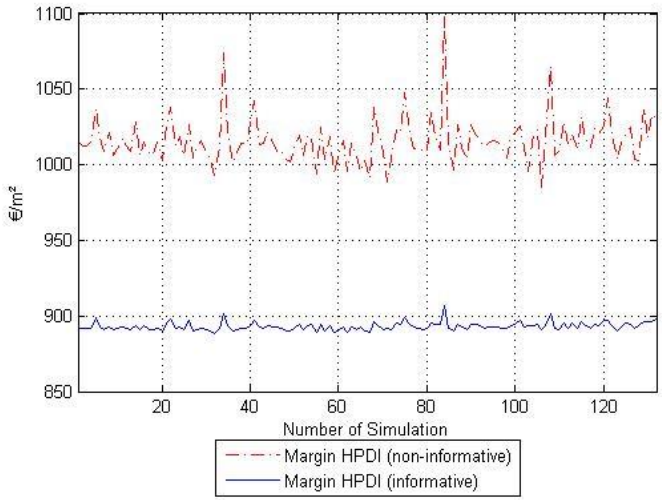


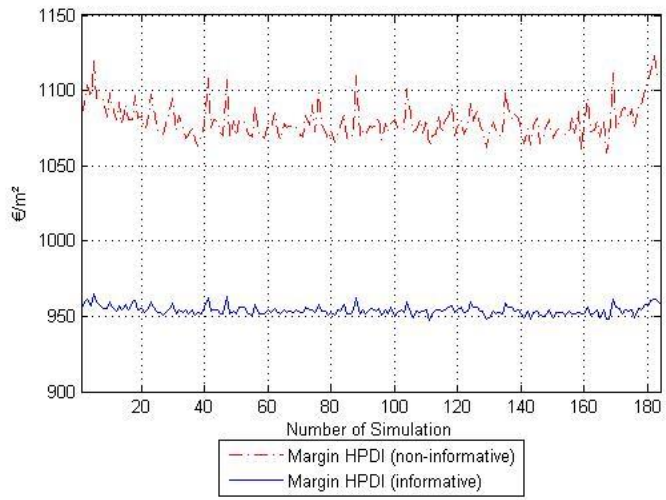

Quality of Location $\overline{\beta}_5$:



## *Practical results - Prediction*

- Leave-one-out cross validation
- Margin: difference of upper and lower boundary of confidence regions
- Margin of confidence regions 2009 (left) and 2010 (right)




- 2009: 132 predicted values; 2010: 184 predicted values
- 2009: Mean increase of 12.1 %, corresponding to about 100 €/qm
- 2010: Mean increase of 11.6 %, corresponding to about 150 €/qm

# Conclusion

- Recursive Bayesian approach implemented successfully to issues of real estate valuation
- Integration of **empirically derived** prior knowledge enables

  ... the possibility of **quantifying the uncertainty** of single variables
  ... the **reduction of uncertainty** concerning the estimation of parameters and the prediction of comparative values

- Future goals of research:
  - Introduction of weights for the influence of passed periods
  - Improvement of functional model

# Conclusion

- Recursive Bayesian approach implemented successfully to issues of real estate valuation
- Integration of **empirically derived** prior knowledge enables

  ... the possibility of **quantifying the uncertainty** of single variables
  ... the **reduction of uncertainty** concerning the estimation of parameters and the prediction of comparative values

- Future goals of research:
  - Introduction of weights for the influence of passed periods
  - Improvement of functional model

# Contact

**Thank you for your attention!**

**Dipl.-Ing. Sebastian Zaddach**

Leibniz Universitaet Hannover
Geodetic Insitute
Land and Real Estate Management

Nienburger Str. 1
Hannover D - 30167
GERMANY

zaddach@gih.uni-hannover.de
www.gih.uni-hannover.de

**Dr.-Ing. Hamza Alkhatib**

Leibniz Universitaet Hannover
Geodetic Insitute
Evaluation methods

Nienburger Str. 1
Hannover D - 30167
GERMANY

alkhatib@gih.uni-hannover.de
www.gih.uni-hannover.de