Universiteit Utrecht

A Gentle Introduction to Bayesian Statistics

Rens van de Schoot
Sara van Erp
Duco Veen
Beth Grandfield

Why do researchers use Bayes

- Bayes is not based on large samples (i.e., the central limit theorem) and hence large samples are not required to obtain accurate results.

Universiteit Utrecht

How large should the sample size be at the highest level in multilevel analyses

????

With ML-estimation:

-> Boomsma (1983): 200 OK, at least 100

-> Hox, Maas Brinkhuis (2010): at least 100 groups

With ML-estimation:

-> Boomsma (1983): 200 OK, at least 100
-> Hox, Maas Brinkhuis (2010): at least 100 groups

With Bayesian estimation:

-> Hox et al (2012): 20-25 OK!

Hox, J., van de Schoot. R., & Matthijsse, S. (2012). How few countries will do? Comparative survey analysis from a Bayesian perspective. Survey Research Methods, 6, 87-93.

EHP
BULLETIN OF THE EUROPEAN HEALTH PSYCHOLOGY SOCIETY
The European Health Psychologist

original article

Bayesian analyses: where to start and what to report

Rens van de Schoot
Utrecht University & North-West University
Sarah Depaoli
University of California

Most researchers in the social and behavioral sciences will probably have heard of Bayesian statistics in which probability is defined differently compared to classical statistics (probability as the long-run frequency versus probability as the subjective experience of uncertainty). At the same time, many may be unsure of whether they should or would like to use Bayesian methods to answer their research questions (note: all types of conventional questions can also be addressed with Bayesian statistics). As an attempt to track how popular the methods are, we searched all papers published in 2013 in the field of Psychology (source: Scopus), and we identified 79 empirical papers that used Bayesian methods (see e.g.

Bayesian Statistics: A brief introduction

Before providing arguments why one would use Bayesian statistics, we first provide a brief introduction. Within conventional statistical techniques, the null hypothesis is always set up to assume no relation between the variables of interest. This null hypothesis makes sense when you have absolutely no idea of the relationship between the variables. However, it is often the case that researchers do have *a priori* knowledge about likely relationships between variables, which may be based on earlier research. With Bayesian

Small samples – example 1

- Two groups with M1=0, M2=0.45, SD=2

- With n=100 the t-test produced a just significant effect of p=.045.

- Also, when using objective Bayes with a very low prior precision the p-value appeared to be p=.046.

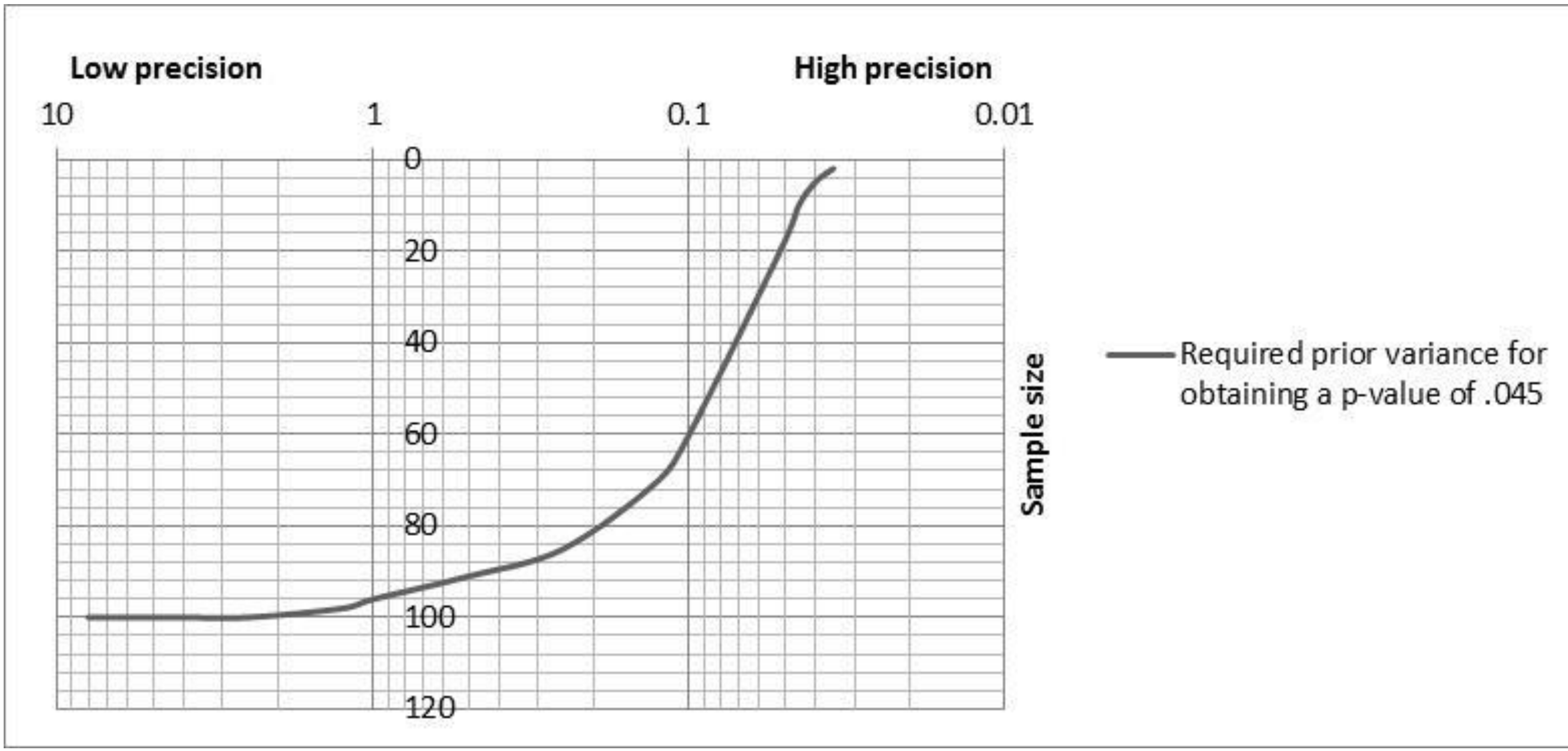
van de Schoot, R. & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. EHP, 16(2), 75-84.

Small samples – example 1

- Two groups with M1=0, M2=0.45, SD=2
- With n=100 the t-test produced a just significant effect of p=.045.
- Also, when using objective Bayes with a very low prior precision the p-value appeared to be p=.046.
- Then I generated new data sets with smaller sample sizes but with same mean difference
- Goal -> to obtain for every data set the same p-value

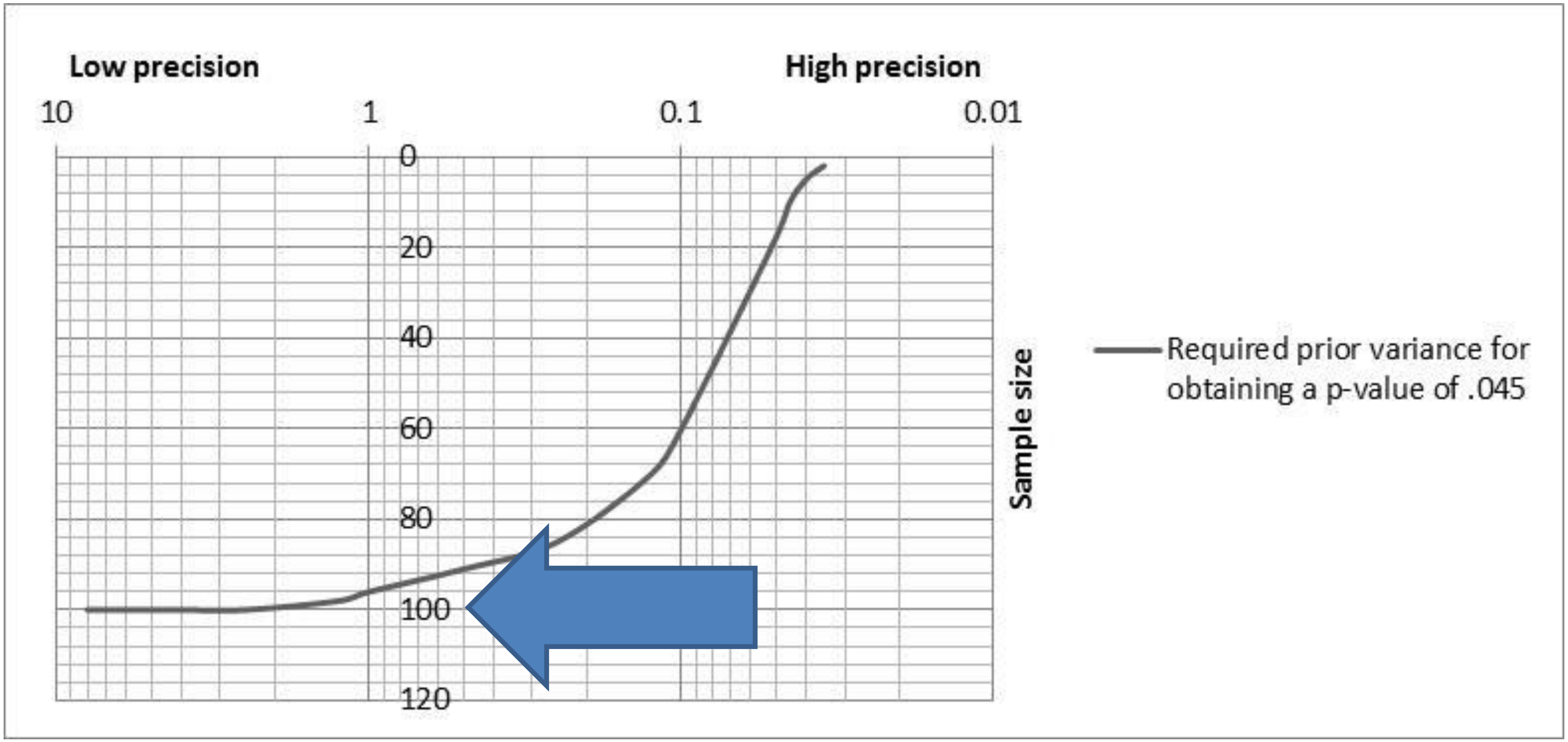
van de Schoot, R. & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. EHP, 16(2), 75-84.

Small samples – example 1



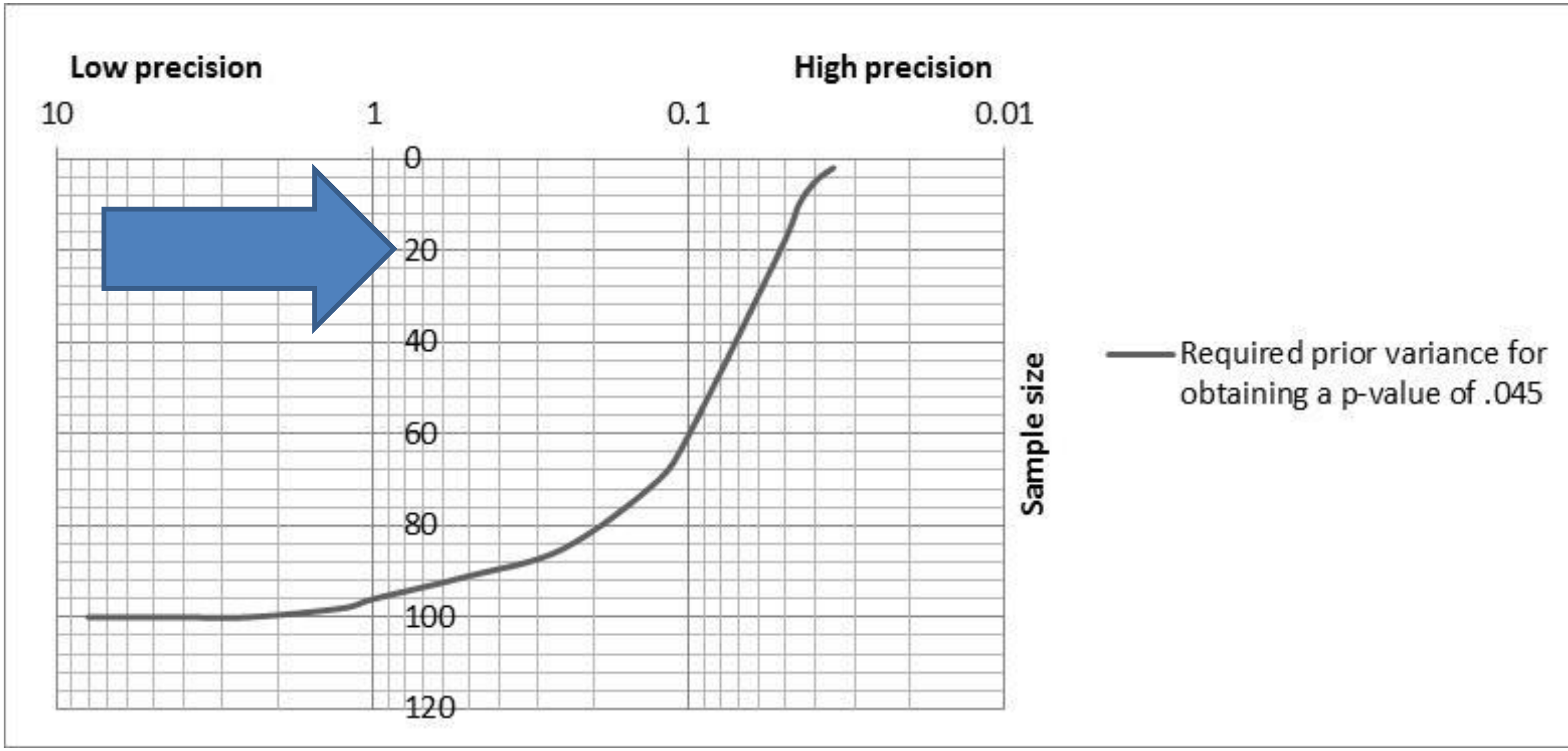
van de Schoot, R. & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. EHP, 16(2), 75-84.

Small samples – example 1



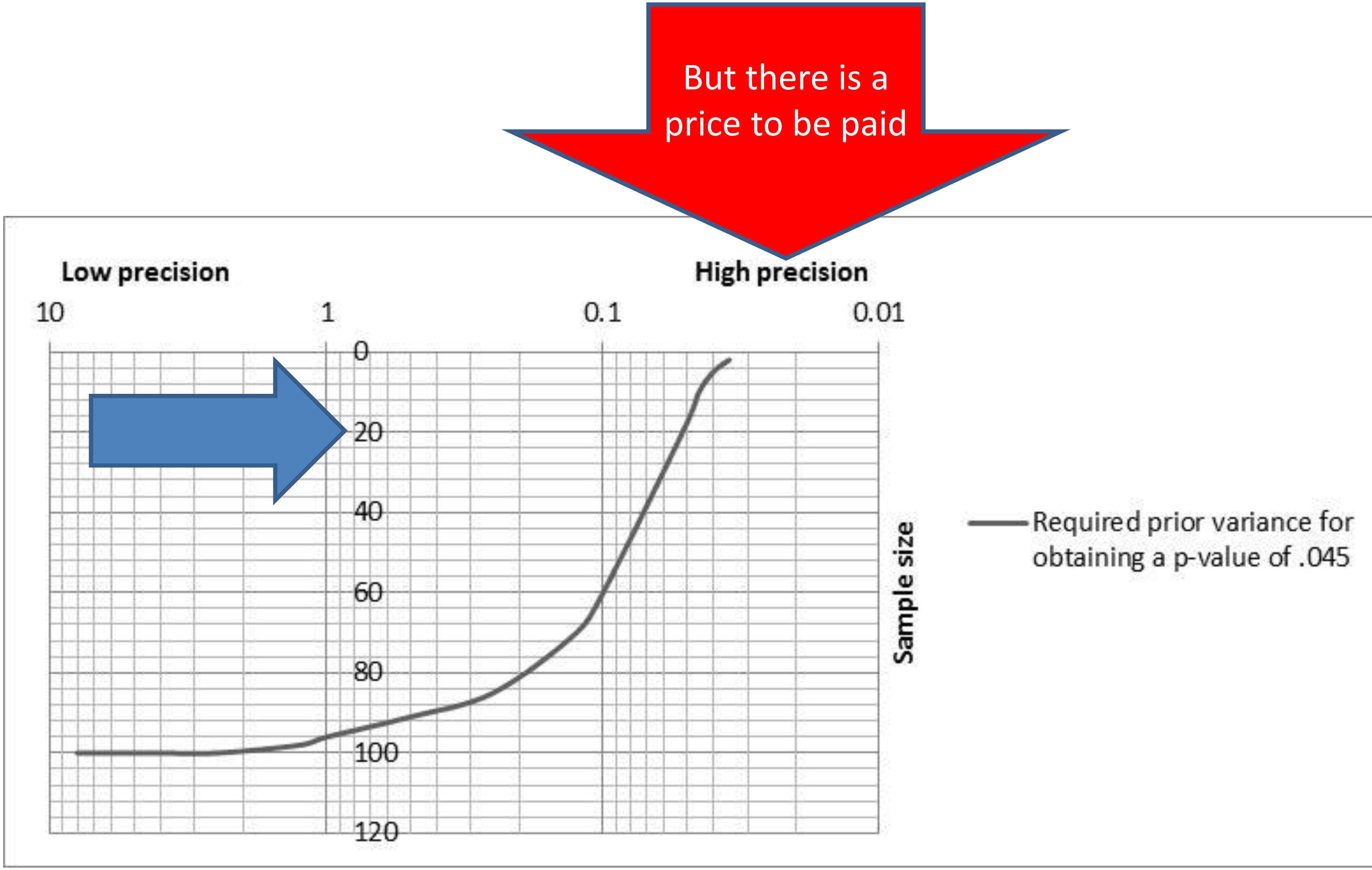
van de Schoot, R. & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. EHP, 16(2), 75-84.

Small samples – example 1



van de Schoot, R. & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. EHP, 16(2), 75-84.

Small samples – example 1



van de Schoot, R. & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. EHP, 16(2), 75-84.

Small samples – example 2

Universiteit Utrecht

EUROPEAN JOURNAL OF
PSYCHO-TRAUMATOLOGY
THE OFFICIAL JOURNAL OF THE EUROPEAN SOCIETY FOR TRAUMATIC STRESS STUDIES

Analyzing small data sets using Bayesian estimation: the case of posttraumatic stress symptoms following mechanical ventilation in burn survivors

Rens van de Schoot[1,2*], *Joris J. Broere*[1], *Koen H. Perryck*[1], *Mariëlle Zondervan-Zwijnenburg*[1] *and Nancy E. van Loey*[3,4]

[1]Department of Methods and Statistics, Utrecht University, Utrecht, The Netherlands; [2]Optentia Research Program, Faculty of Humanities, North-West University, Vanderbijlpark, South Africa; [3]Department of Clinical & Health Psychology, Utrecht University, Utrecht, The Netherlands; [4]Department Behavioural Research, Association of Dutch Burn Centres, Beverwijk, The Netherlands

› Article

- Abstract
- Empirical example
- Study 1: sensitivity analysis
- Study 2: simulation study
- Results empirical example
- Conclusion
- Conflict of interest and funding
- References
- FOOTNOTES

› About the Authors

› Article Metrics

ABSTRACT

Background: The analysis of small data sets in longitudinal studies can lead to power issues and often suffers from biased parameter values. These issues can be solved by using Bayesian estimation in conjunction with informative prior distributions. By means of a simulation study and an empirical example concerning posttraumatic stress symptoms (PTSS) following mechanical ventilation in burn survivors, we demonstrate the advantages and potential pitfalls of using Bayesian estimation.

Methods: First, we show how to specify prior distributions and by means of a sensitivity analysis we demonstrate how to check the exact influence of the prior (mis-) specification. Thereafter, we show by means of a simulation the situations in which the Bayesian approach outperforms the default, maximum likelihood and approach. Finally, we re-analyze empirical data on burn survivors which provided preliminary evidence of an aversive influence of a period of mechanical ventilation on the course of PTSS following burns.

Results: Not suprisingly, maximum likelihood estimation showed insufficient coverage as well as power with very small samples. Only when Bayesian analysis, in conjunction with informative priors, was used power increased to acceptable levels. As expected, we showed that the smaller the sample size the more the results rely on the prior specification.

Conclusion: We show that two issues often encountered during analysis of small samples, power and biased parameters, can be solved by including prior information into Bayesian analysis. We argue that the use of informative priors should always be reported together with a sensitivity analysis.

Keywords: *Bayesian estimation; maximum likelihood; prior specification; power; repeated measures analyses; small samples; burn survivors; mechanical ventilation; PTSS*

Download Article

Cite Article | Share

Keywords

adolescents anxiety assessment children combat complex ptsd depression dissociation intimate partner violence mental health military posttraumatic stress posttraumatic stress disorder psychotherapy psychotraumatology ptsd refugees resilience social support trauma treatment

Data on 15 persons who were mechanically ventilated and 63 who did not have mechanical ventilation after a burn event.

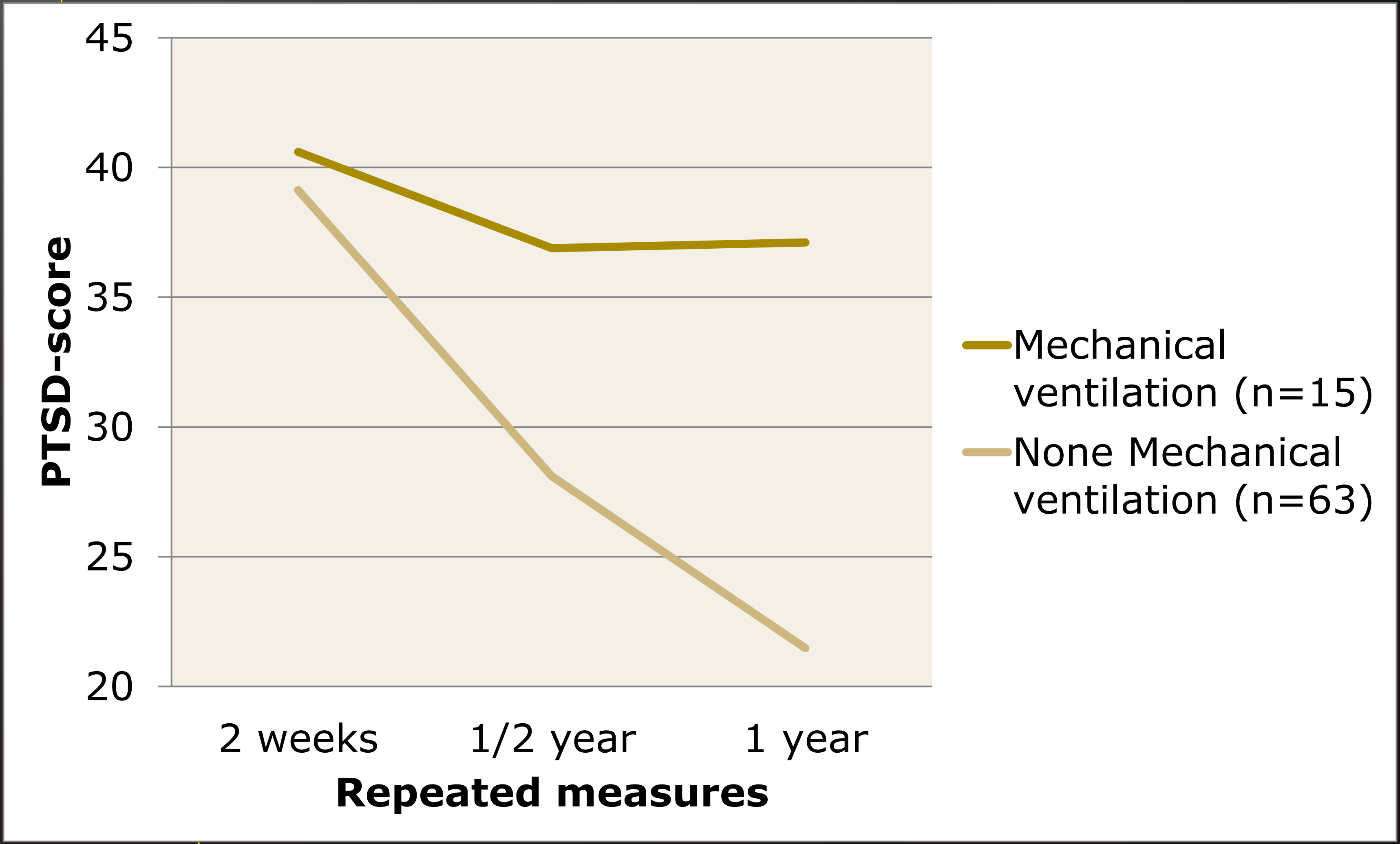
The average length of hospitalization was 27.51 days

The patients burned surface area ranged from 1% till 60%, M = 15.99%.

Dependent variable: post-traumatic stress symptoms (Dutch version of the IES) after 2 weeks, 6 months and 1 year.

Example taken from: Van Loey, et al (2003). Predictors of Chronic Posttraumatic Stress Symptons Following Burn Injury: Results of a Longitudinal Study. *Journal of Traumatic Stress*, *16*(4), 361 - 369.

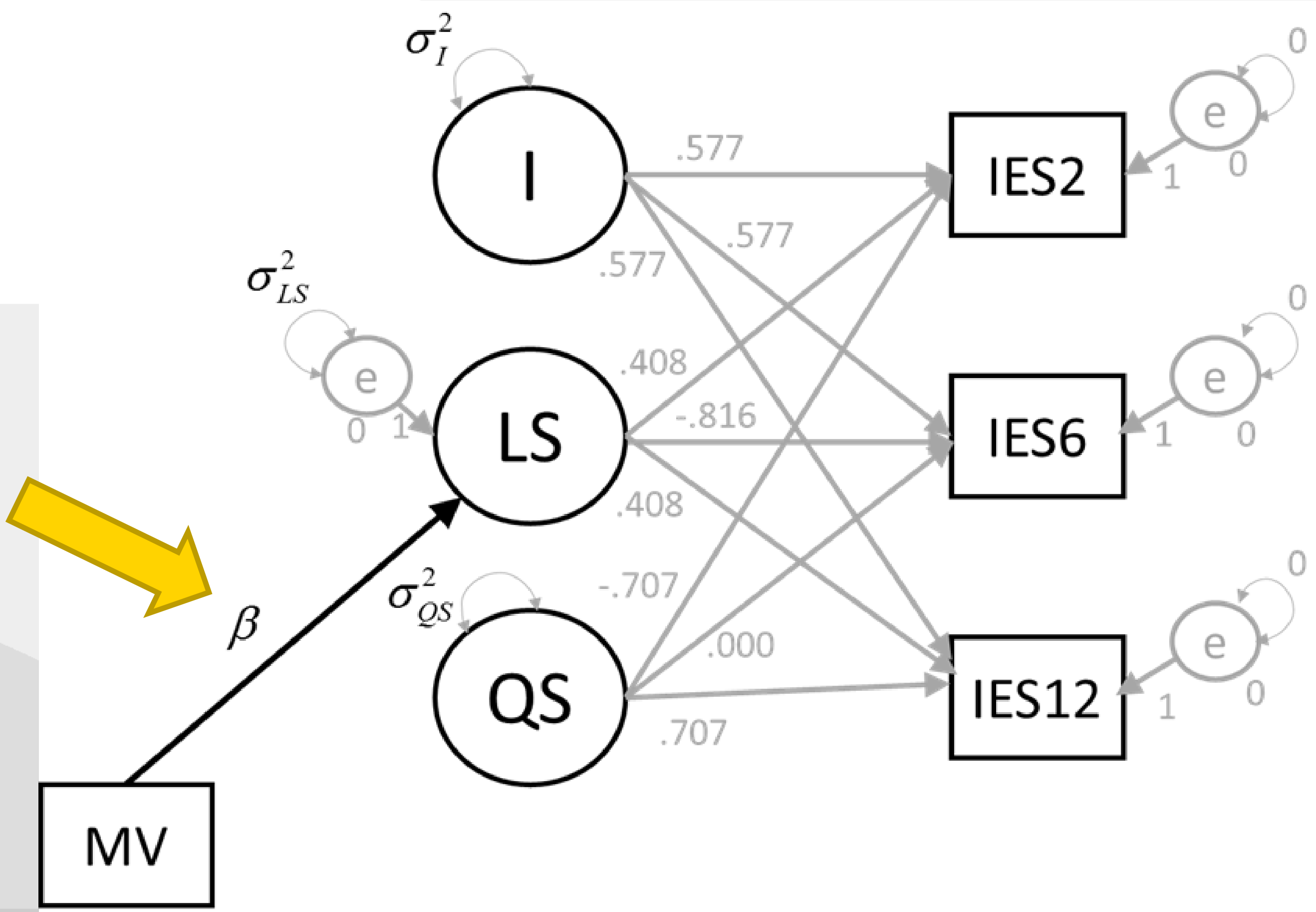
Data



45
40
35
30
25
20
PTSD-score
2 weeks
1/2 year
1 year
Repeated measures
Mechanical ventilation (n=15)
None Mechanical ventilation (n=63)

Universiteit Utrecht

Prior Precision



Prior Precision

$$p(\theta) * p(\mathrm{D} \mid \theta) \propto p(\theta \mid \mathrm{D})$$

Universiteit Utrecht

Prior Precision

$$p(\theta) * p(D \mid \theta) \propto p(\theta \mid D)$$

$$p(\theta) = p(I, LS, QS, \beta, \sigma_I^2, \sigma_{LS}^2, \sigma_{QS}^2)$$

$$= p(I), p(LS), p(QS), p(\beta), p(\sigma_I^2), p(\sigma_{LS}^2), p(\sigma_{QS}^2)$$

Universiteit Utrecht

Prior Precision

$$p(\theta) * p(D \mid \theta) \propto p(\theta \mid D)$$

$$p(\theta) = p(I, LS, QS, \beta, \sigma^2_I, \sigma^2_{LS}, \sigma^2_{QS})$$

$$= p(I), p(LS), p(QS), p(\beta), p(\sigma^2_I), p(\sigma^2_{LS}), p(\sigma^2_{QS})$$

$$\theta^1 = p(I), p(LS), p(QS), p(\beta)$$

$$\theta^2 = p(\sigma^2_I), p(\sigma^2_{LS}), p(\sigma^2_{QS})$$

Prior Precision

$$p(\theta) * p(D \mid \theta) \propto p(\theta \mid D)$$

$$p(\theta) = p(I, LS, QS, \beta, \sigma^2_I, \sigma^2_{LS}, \sigma^2_{QS})$$

$$= p(I), p(LS), p(QS), p(\beta), p(\sigma^2_I), p(\sigma^2_{LS}), p(\sigma^2_{QS})$$

$$\theta^1 = p(I), p(LS), p(QS), p(\beta) \Rightarrow p(\theta^1) \sim N(\mu_0, \sigma^2_0)$$

$$\theta^2 = p(\sigma^2_I), p(\sigma^2_{LS}), p(\sigma^2_{QS}) \Rightarrow p(\theta^2) \sim IG(\alpha, \nu)$$

Prior Precision

$$p(\theta) * p(D \mid \theta) \propto p(\theta \mid D)$$

$$p(\theta) = p(I, LS, QS, \beta, \sigma^2_I, \sigma^2_{LS}, \sigma^2_{QS})$$

$$= p(I), p(LS), p(QS), p(\beta), p(\sigma^2_I), p(\sigma^2_{LS}), p(\sigma^2_{QS})$$

$$\theta^1 = p(I), p(LS), p(QS), p(\beta) \Rightarrow p(\theta^1) \sim N(0, 10^{10})$$

$$\theta^2 = p(\sigma^2_I), p(\sigma^2_{LS}), p(\sigma^2_{QS}) \Rightarrow p(\theta^2) \sim IG(-1, 0)$$

Default settings ~ Objective Bayes

Prior Precision

Changing the default settings for beta:

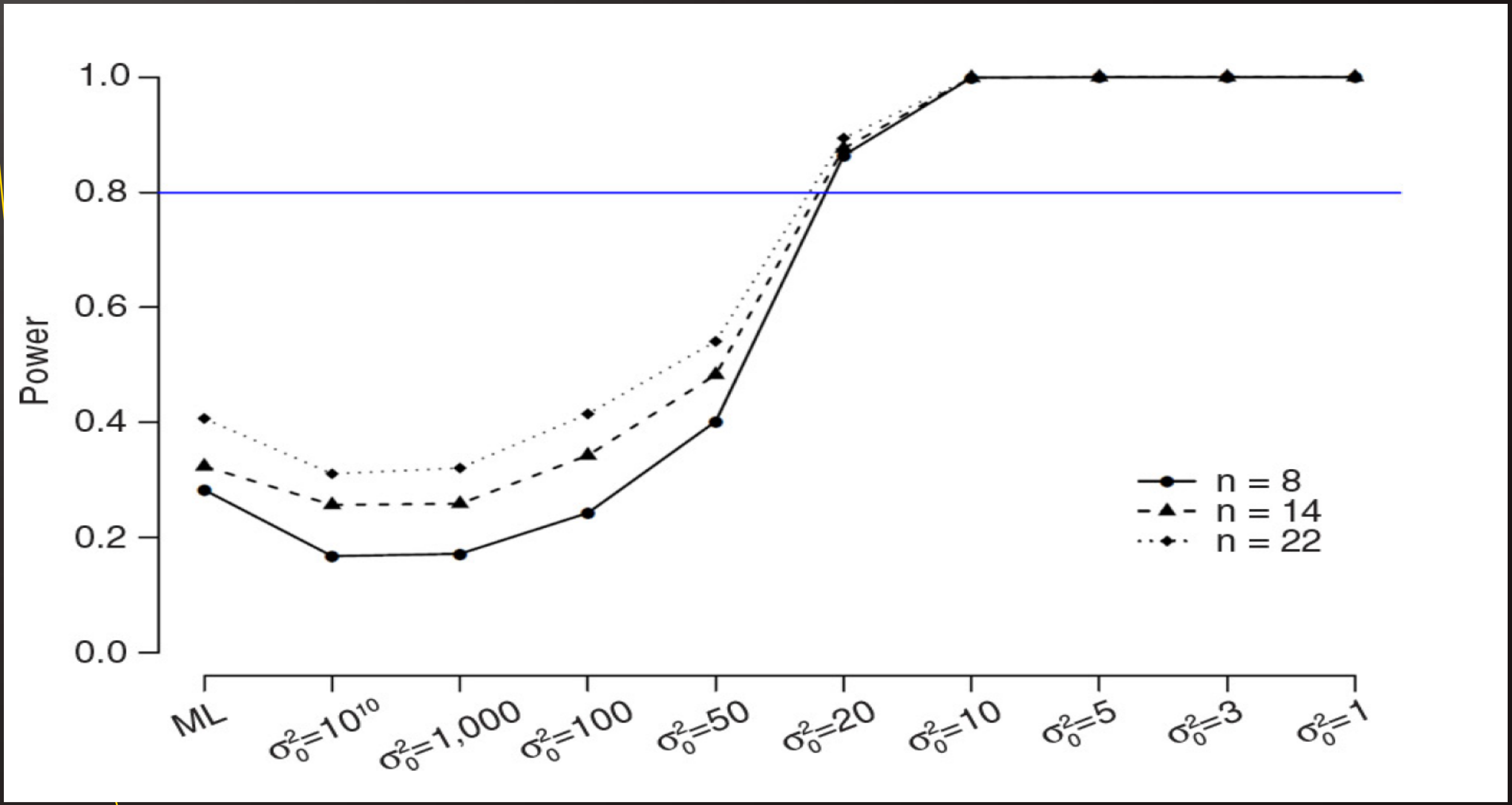
$$p(\theta) = p(\mathrm{I}, \mathrm{LS}, \mathrm{QS}, \beta, \sigma^2_{\mathrm{I}}, \sigma^2_{\mathrm{LS}}, \sigma^2_{\mathrm{QS}})$$

$$= p(\mathrm{I})\, p(\mathrm{LS})\, p(\mathrm{QS})\, p(\beta)\, p(\sigma^2_{\mathrm{I}})\, p(\sigma^2_{\mathrm{LS}})\, p(\sigma^2_{\mathrm{QS}})$$

$$\beta \sim \mathrm{N}\,(10.007,\ 1)$$

$$\vdots$$

$$\beta \sim \mathrm{N}\,(10.007,\ 1000)$$



Power
1.0
0.8
0.6
0.4
0.2
0.0
ML
σ_0^2=10^10
σ_0^2=1,000
σ_0^2=100
σ_0^2=50
σ_0^2=20
σ_0^2=10
σ_0^2=5
σ_0^2=3
σ_0^2=1
n = 8
n = 14
n = 22

Misspecification

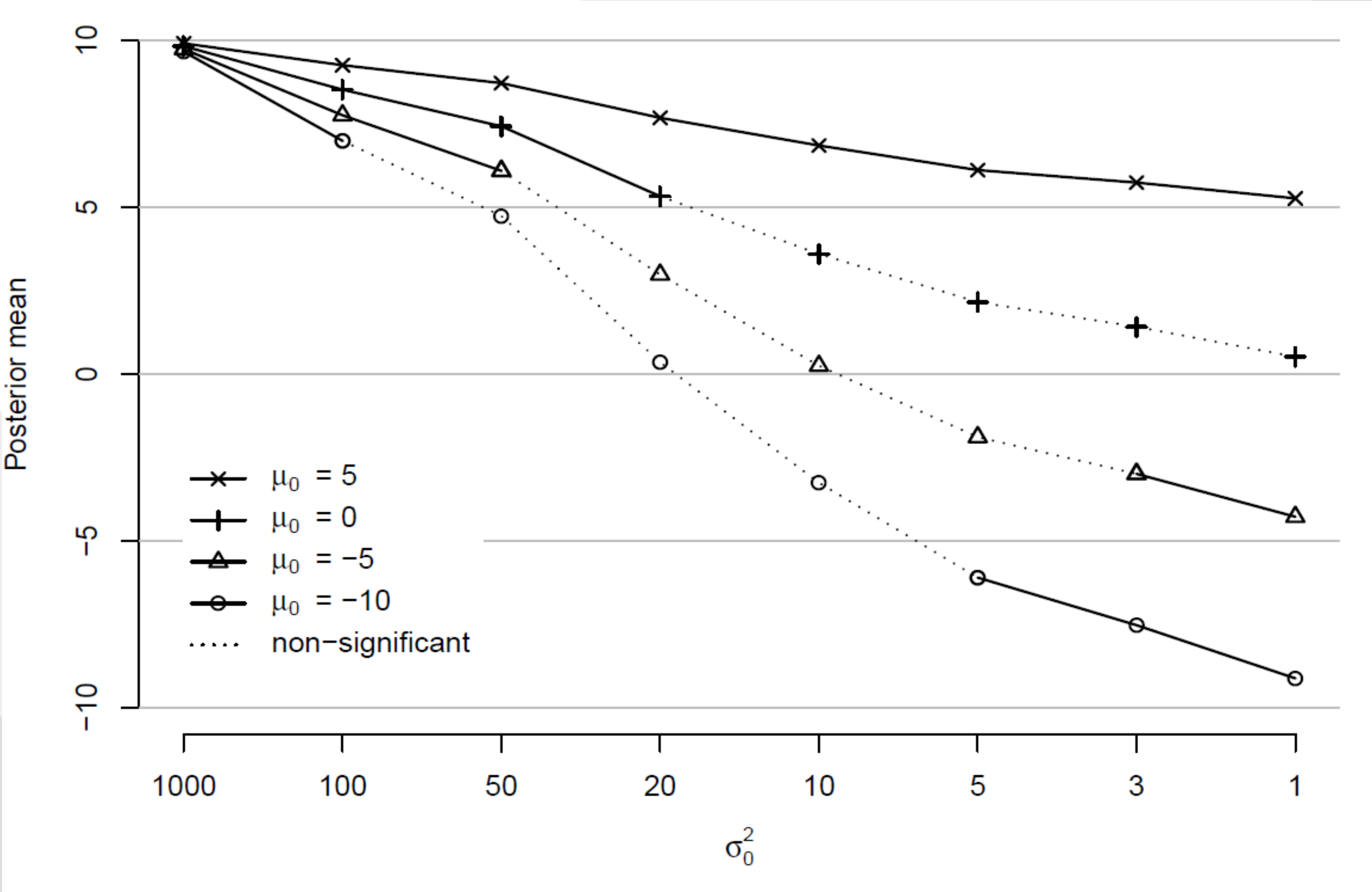
The previous results were dependent on a correctly specified prior mean.

Actually, we cheated... ☹

What if we specified the prior mean incorrectly?

van de Schoot, R., Broere, J., Perryck, K., Zondervan-Zwijnenburg, M., & van Loey, N. (2015). Analyzing small data sets using Bayesian estimation: the case of posttraumatic stress symptoms following mechanical ventilation in burn survivors. *European Journal Of Psychotraumatology, 6*. doi:http://dx.doi.org/10.3402/ejpt.v6.25216

Misspecification



Influence of misspecification of prior means of β with different prior variances ~correct β = 10.007

Simulation

Results:

- the relative mean bias defined as

$$((M - \beta^{pop}) / \beta^{pop}) * 100$$

where M is the average beta obtained from the simulation study.

Cut-off value of <10% as a criterion, as suggested by Hoogland and Boomsma (1998) for ‘reasonable’ accuracy.

Universiteit Utrecht

Simulation results - ML

	n = 8					
	Pop	**M**	**MSE**	**% Bias**	**95% cover**	**Power**
I	54.0900	54.1551	74.0206	0.12%	.907	1.000
LS	-12.4810	-12.5464	51.6079	0.52%	.853	.557
QS	1.7590	1.6870	17.2733	4.09%	.880	.137
β	10.0080	9.9196	100.8752	-0.88%	.864	.283
Var I	593.6330	542.7769	78909.335	-8.57%	.800	1.000
Var S	199.3820	152.9436	9769.9697	-23.29%	.695	1.000
Var Q	134.6140	114.1505	4169.3696	-15.20%	.756	1.000

Universiteit Utrecht

Simulation results - ML

	n = 8					
	Pop	**M**	**MSE**	**% Bias**	**95% cover**	**Power**
I	54.0900	54.1551	74.0206	0.12%	.907	1.000
LS	-12.4810	-12.5464	51.6079	0.52%	.853	.557
QS	1.7590	1.6870	17.2733	4.09%	.880	.137
β	10.0080	9.9196	100.8752	-0.88%	.864	.283
Var I	593.6330	542.7769	78909.335	-8.57%	.800	1.000
Var S	199.3820	152.9436	9769.9697	-23.29%	.695	1.000
Var Q	134.6140	114.1505	4169.3696	-15.20%	.756	1.000

Universiteit Utrecht

Simulation results - ML

	n = 8					
	Pop	**M**	**MSE**	**% Bias**	**95% cover**	**Power**
I	54.0900	54.1551	74.0206	0.12%	.907	1.000
LS	-12.4810	-12.5464	51.6079	0.52%	.853	.557
QS	1.7590	1.6870	17.2733	4.09%	.880	.137
β	10.0080	9.9196	100.8752	-0.88%	.864	.283
Var I	593.6330	542.7769	78909.335	-8.57%	.800	1.000
Var S	199.3820	152.9436	9769.9697	-23.29%	.695	1.000
Var Q	134.6140	114.1505	4169.3696	-15.20%	.756	1.000

Universiteit Utrecht

Simulation results - ML

			n = 8			
	Pop	M	MSE	% Bias	95% cover	Power
I	54.0900	54.1551	74.0206	0.12%	.907	1.000
LS	-12.4810	-12.5464	51.6079	0.52%	.853	.557
QS	1.7590	1.6870	17.2733	4.09%	.880	.137
β	10.0080	9.9196	100.8752	-0.88%	.864	.283
Var I	593.6330	542.7769	78909.335	-8.57%	.800	1.000
Var S	199.3820	152.9436	9769.9697	-23.29%	.695	1.000
Var Q	134.6140	114.1505	4169.3696	-15.20%	.756	1.000

Simulation results - ML

	n = 8					
	Pop	**M**	**MSE**	**% Bias**	**95% cover**	**Power**
I	54.0900	54.1551	74.0206	0.12%	.907	1.000
LS	-12.4810	-12.5464	51.6079	0.52%	.853	.557
QS	1.7590	1.6870	17.2733	4.09%	.880	.137
β	10.0080	9.9196	100.8752	-0.88%	.864	.283
Var I	593.6330	542.7769	78909.335	-8.57%	.800	1.000
Var S	199.3820	152.9436	9769.9697	-23.29%	.695	1.000
Var Q	134.6140	114.1505	4169.3696	-15.20%	.756	1.000

Universiteit Utrecht

Simulation results – Bayes obj

	n = 8					
	Pop	M	MSE	% Bias	95% cover	Power
I	54.0900	53.8515	74.0908	-0.44%	.981	.993
LS	-12.4810	-12.2739	51.6691	-1.66%	.982	.161
QS	1.7590	1.8140	17.2806	3.13%	.972	.033
β	10.0080	9.9278	100.8716	-0.80%	.987	.049
Var I	593.6330	1004.8042	*********	69.26%	.918	1.000
Var S	199.3820	367.7034	72300.7656	84.42%	.906	1.000
Var Q	134.6140	208.4106	17948.2188	54.82%	.923	1.000

Universiteit Utrecht

Simulation results – Bayes obj

			n = 8			
	Pop	**M**	**MSE**	**% Bias**	**95% cover**	**Power**
I	54.0900	53.8515	74.0908	-0.44%	.981	.993
LS	-12.4810	-12.2739	51.6691	-1.66%	.982	.161
QS	1.7590	1.8140	17.2806	3.13%	.972	.033
β	10.0080	9.9278	100.8716	-0.80%	.987	.049
Var I	593.6330	1004.8042	*********	69.26%	.918	1.000
Var S	199.3820	367.7034	72300.7656	84.42%	.906	1.000
Var Q	134.6140	208.4106	17948.2188	54.82%	.923	1.000

Universiteit Utrecht

Simulation results – Bayes obj

What happened?

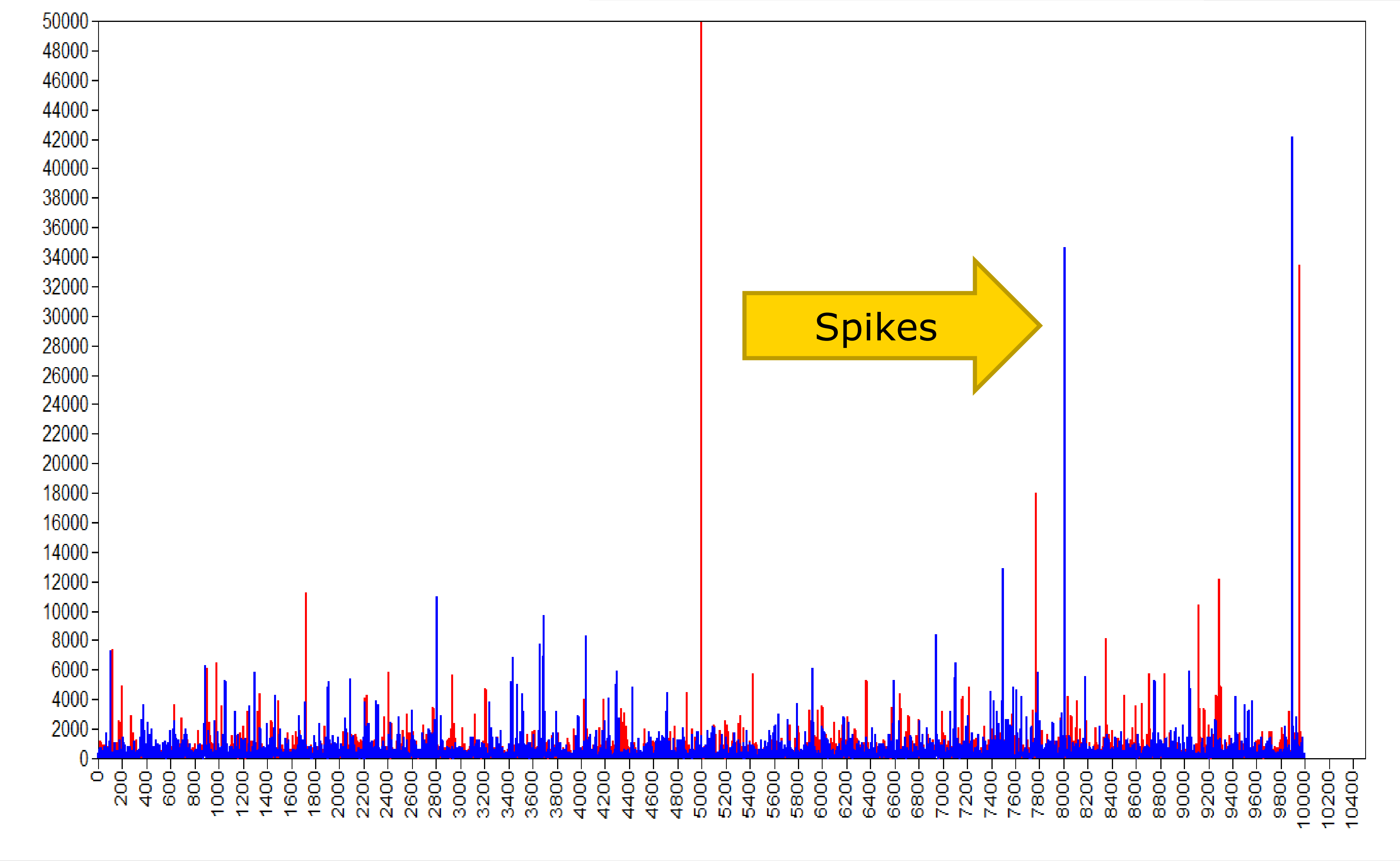
Bayes with default prior settings resulted in huge variances...

So, the default settings should not be trusted with small sample sizes

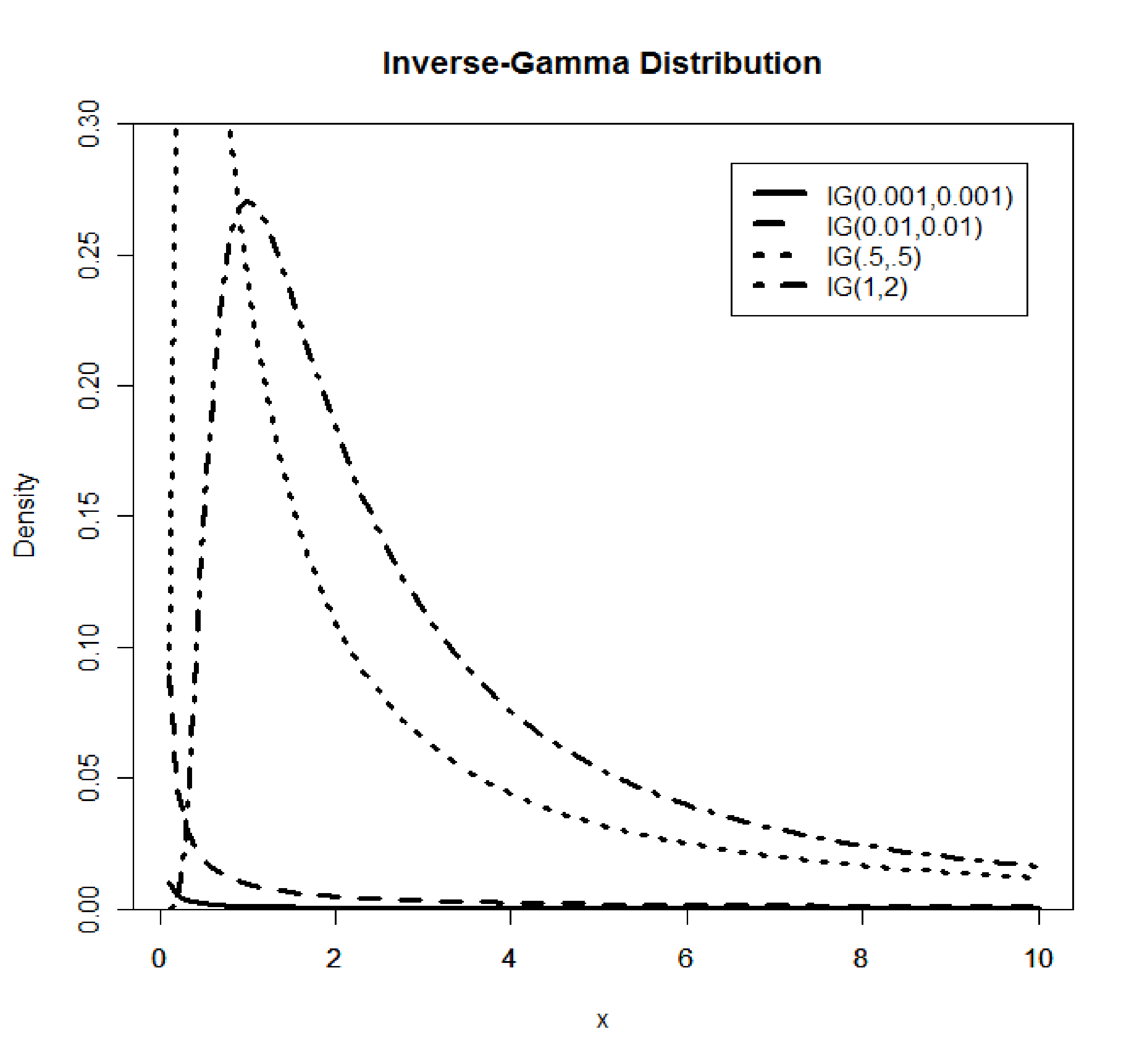
Let’s inspect the trace plots of the variance terms and adjust the prior settings a bit

Universiteit Utrecht

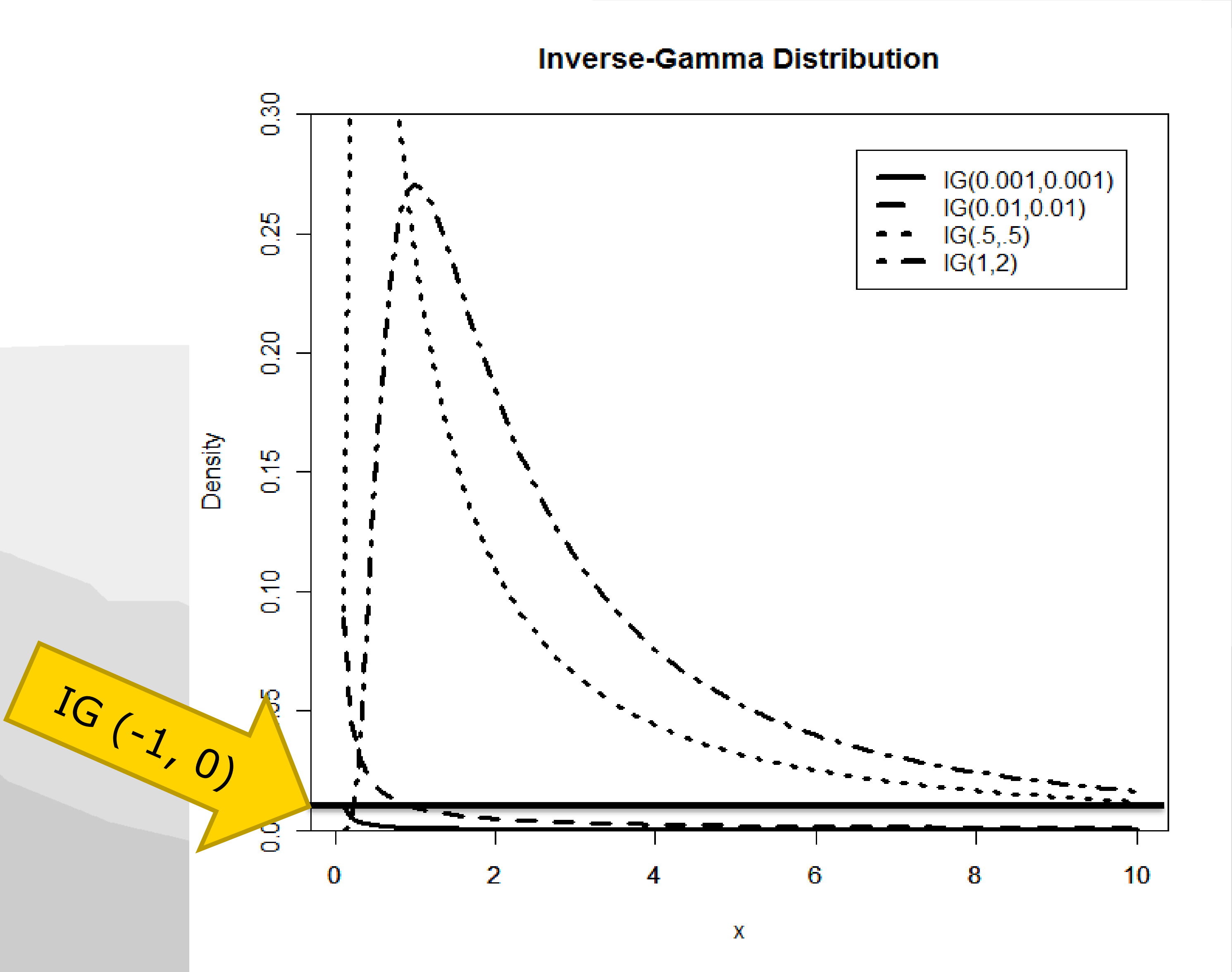
Trace plot for the variance of the Slope
Default prior setting IG(-1,0)



Default Prior settings



Default Prior settings



Improper prior

- Probability distribution does not sum or integrate to one
- Shape and scale parameter need to be larger than zero

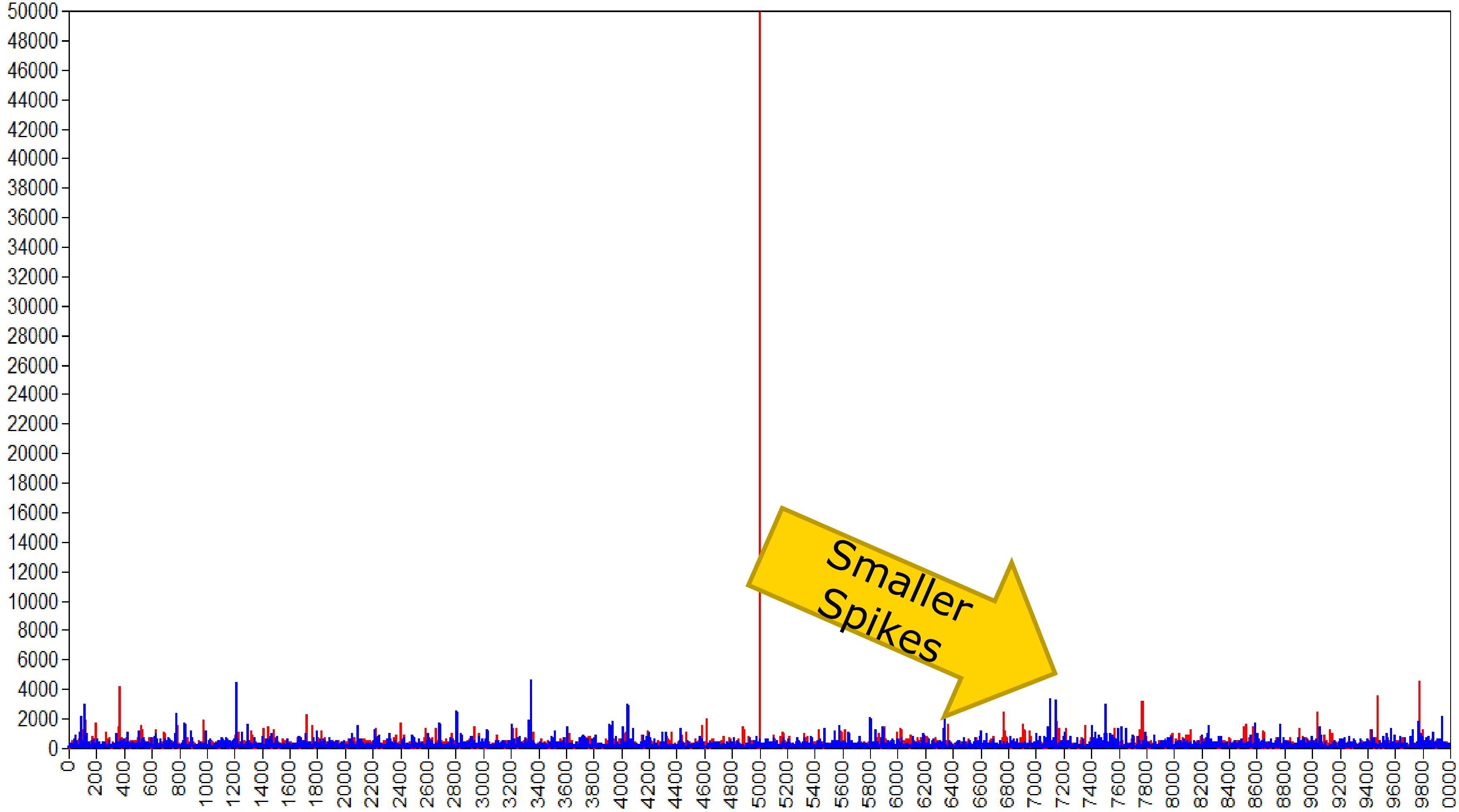
Improper prior:

$p(\theta_l) \sim \mathrm{IG}(-1,0)$
$p(\theta_l) \sim \mathrm{IG}(0,0)$

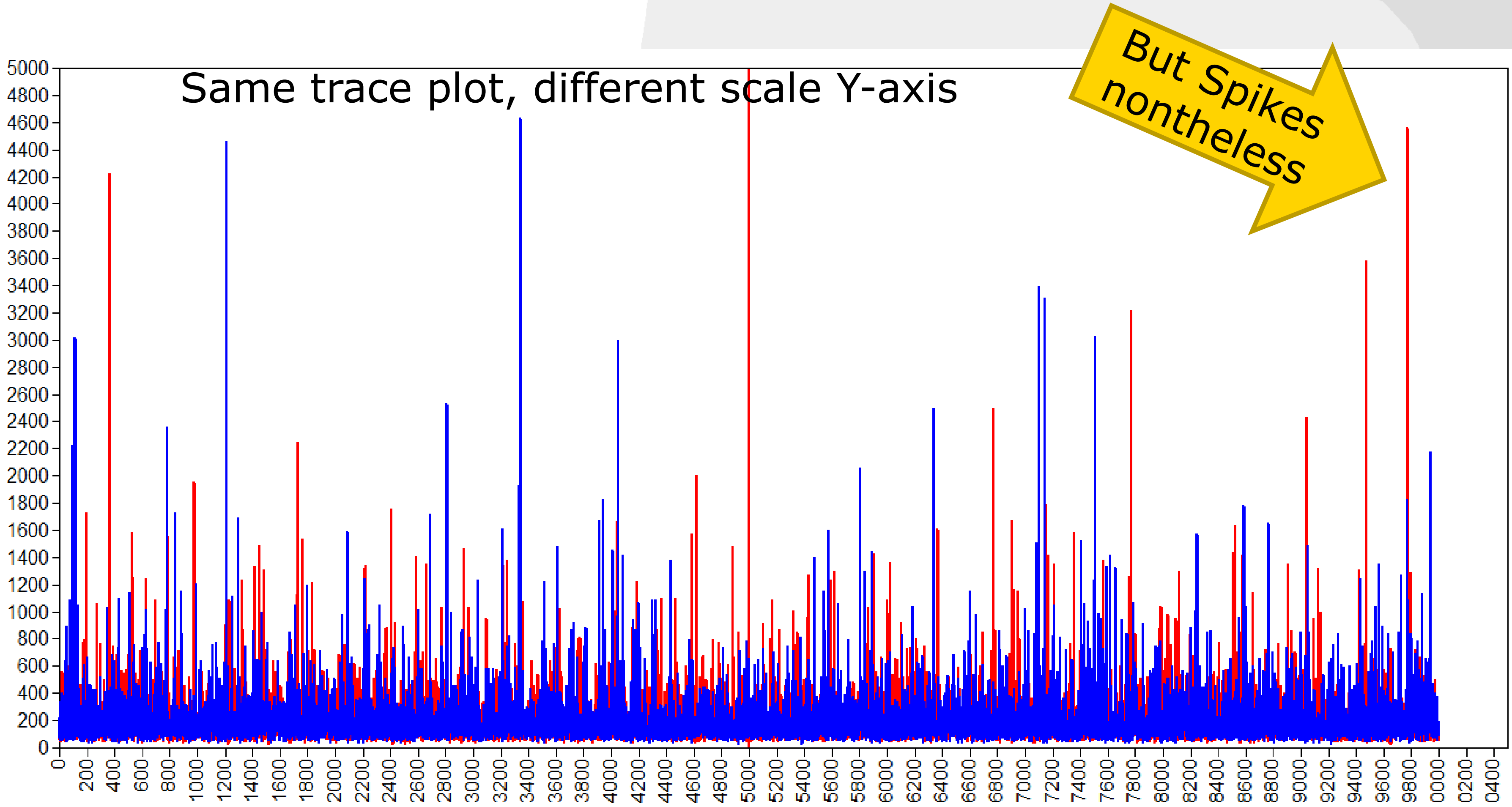
Proper prior:

$p(\theta_l) \sim \mathrm{IG}(.001,.001)$
$p(\theta_l) \sim \mathrm{IG}(.5,.5)$

Trace plot for the variance of the Slope Prior setting IG(0,0)

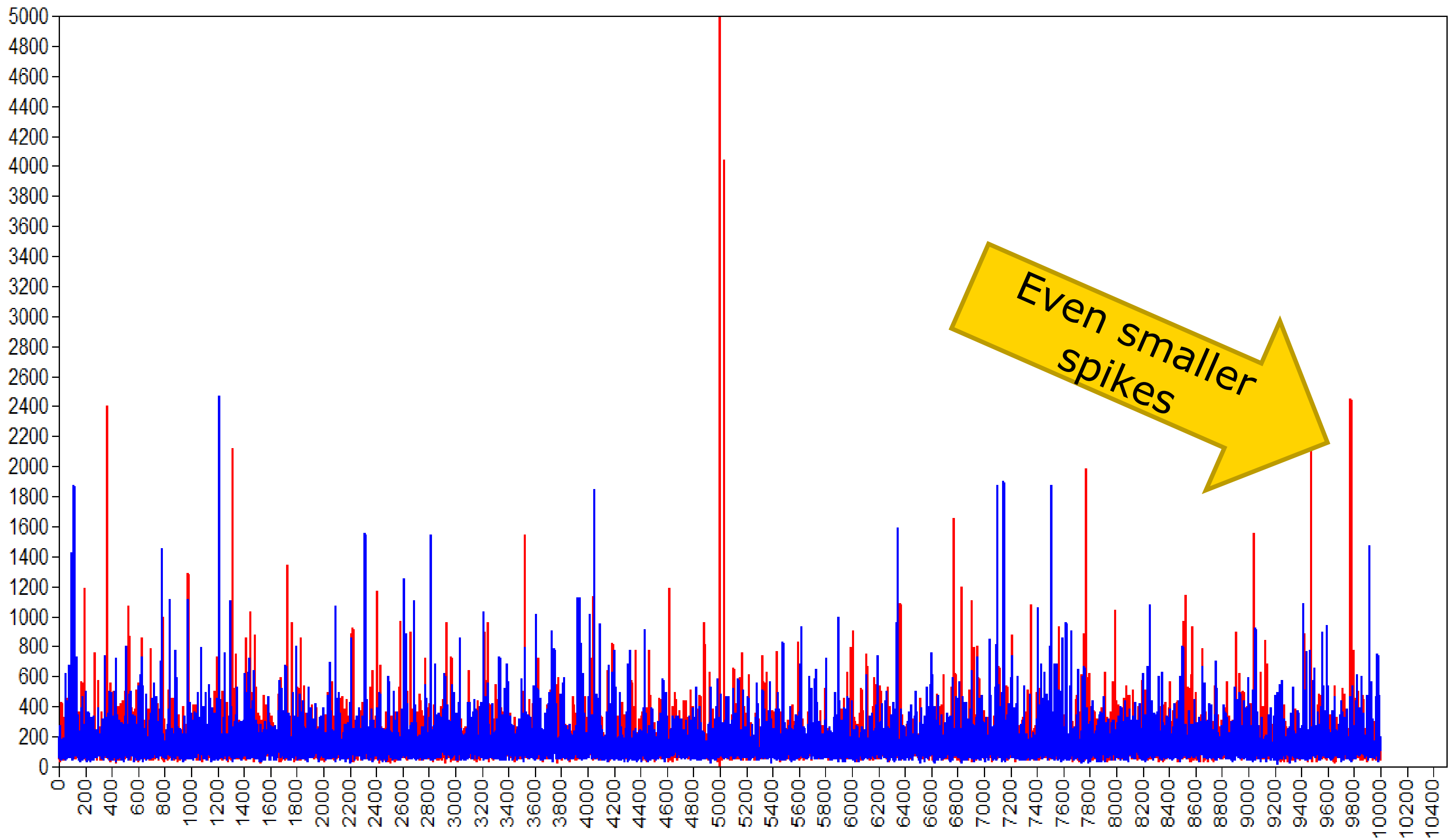


Trace plot for the variance of the Slope Prior setting IG(0,0)

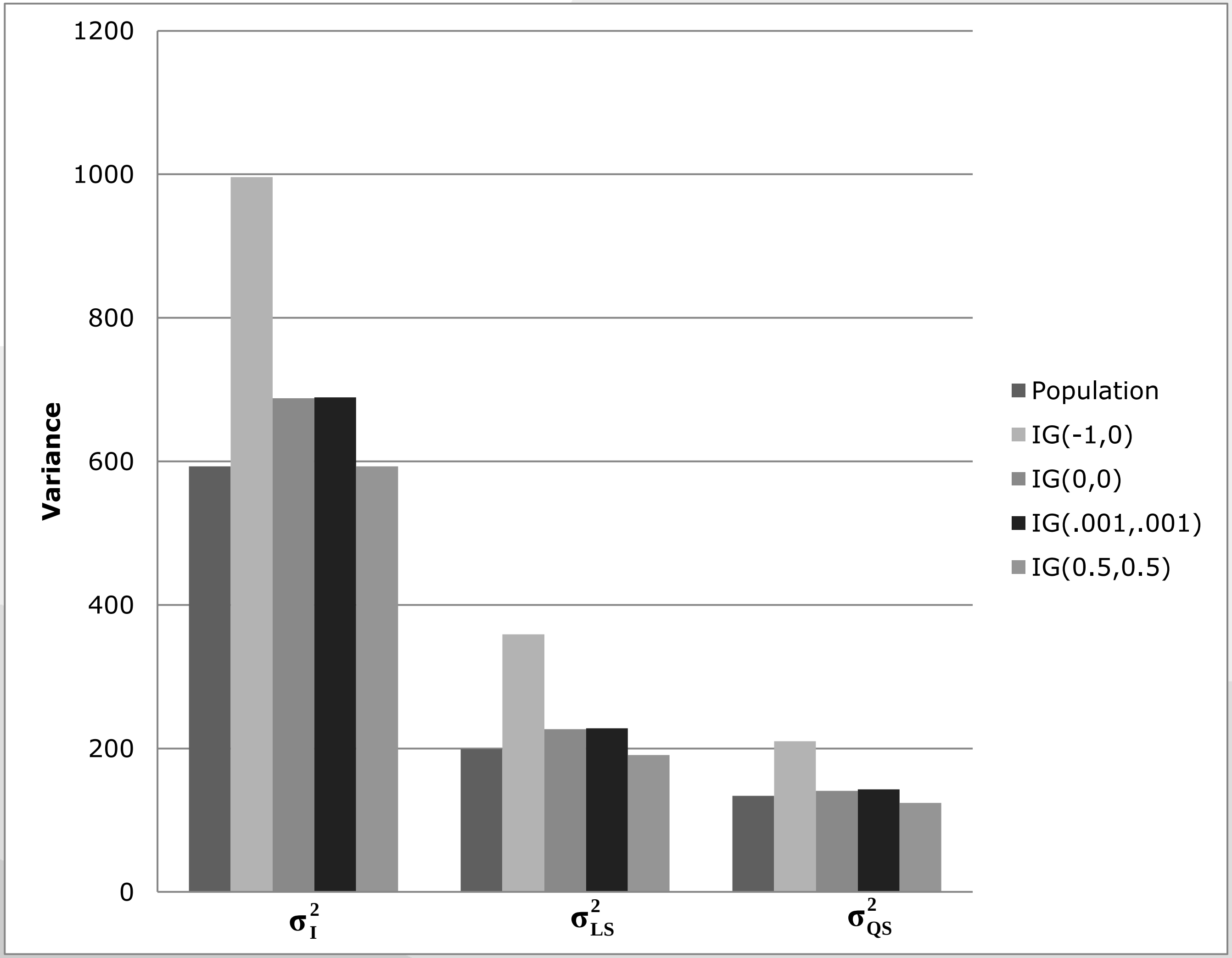


Universiteit Utrecht

Trace plot for the variance of the Slope
Prior setting IG(.5,.5)



Estimated posterior variances for different priors on the variance terms



Simulation results – Bayes obj

			n = 8			
	Pop	**M**	**MSE**	**% Bias**	**95% cover**	**Power**
I	54.0900	53.8515	74.0908	-0.44%	.981	.993
LS	-12.4810	-12.2739	51.6691	-1.66%	.982	.161
QS	1.7590	1.814	17.2806	3.13%	.972	.033
β	10.0080	9.9278	[illegible]	-0.80%	.987	.049
Var I	593.6330	1004.8042	*******	69.26%	.918	1.000
Var S	199.3820	367.7034	72300.7656	84.42%	.906	1.000
Var Q	134.6140	208.4106	17948.2188	54.82%	.923	1.000

Remember...

Universiteit Utrecht

Simulation results – with IG(.5,.5)

	n = 8					
	Pop	M	MSE	% Bias	95% cover	Power
I	54.0900	54.0606	74.0217	-0.05%	.940	1.000
LS	-12.4810	-12.6530	51.6272	1.38%	.927	.408
QS	1.7590	1.693[illegible]	[illegible]7.2728	-3.75%	.922	.089
β	10.0080	9.8741	1[illegible]	-1.34%	.937	.169
Var I	593.6330	593.1871	91110.703[illegible]	-0.08%	.938	1.000
Var S	199.3820	191.7761	12008.5742	-3.81%	.932	1.000
Var Q	134.6140	124.9042	4575.5659	-6.86%	.932	1.000

Much better!

Simulation results – with IG(.5,.5)

			n = 8			
	Pop	**M**	**MSE**	**% Bias**	**95% cover**	**Power**
I	54.0900	54.0606	74.0217	-0.05%	.940	1.000
LS	-12.4810	-12.6530	51.6272	1.38%	.927	.408
QS	1.7590	1.6930	17.2728	-3.75%	[illegible]	.089
β	10.0080	9.8741	100.8888	-1.34%	.937	.169
Var I	593.6330	593.1871	91110.7031	-0.08%	.938	1.000
Var S	199.3820	191.7761	12008.5742	-3.81%	.932	1.000
Var Q	134.6140	124.9042	4575.5659	-6.86%	.932	1.000

Low power!