Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Daniel Gotthardt[1], Christoph Naefgen[2] & Anne Reinarz[3]

Thinking beyond the mean – how to escape a paradigm of averages

[1]Universität Hamburg [2]FernUniversität in Hagen [3]Durham University

Outline

- Sources of differences
- Mechanisms of conditional dispersion
- Testing regression models for dispersion
- Secularization, religiosity and moral attitudes
- Recommendations and open directions

Stuck in the first moment

- long tradition (OLS, ANOVA)
- various extensions
  - generalized linear models (Nelder & Wedderburn 1972)
  - multilevel models (Lindley & Smith 1972)
- focus always on **conditional expectations** (e.g. mean)

> *I suppose it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail.*
>
> Abraham Maslow (1966)

Heteroscedasticity just a statistical problem?

- Heteroscedasticity as **assumption violation**
  - to be detected and resolved (White 1980)
  - but caused by **misspecification** (King & Roberts 2015)
- **Robustifying**
  - leads to „white-washing“ (Leamer 2010: 43)

➢**conceals causal relationships**

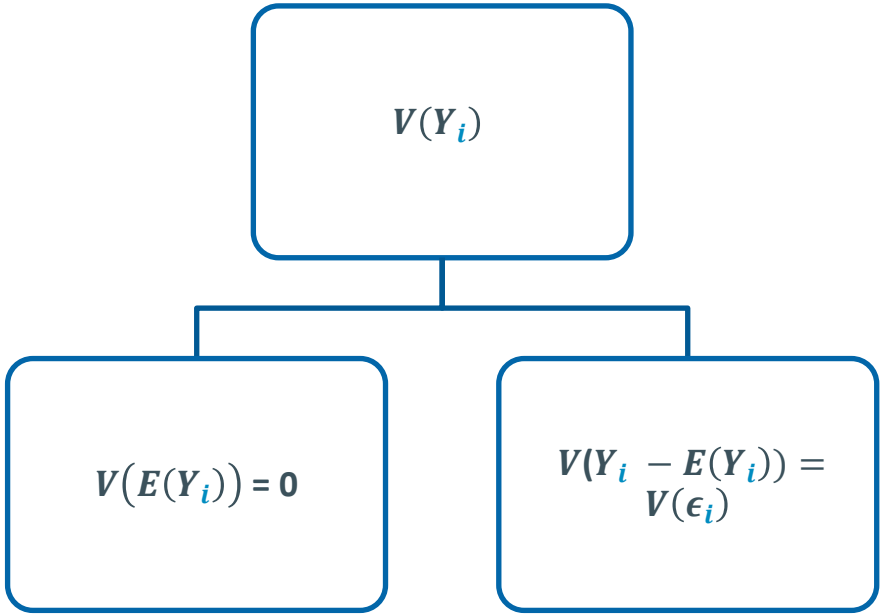
- *Heterogeneous variability*
  - heteroscedasticity as **social facts**
    - e.g., inequality (Western & Bloome 2009), consensus (Lang et al. 2018) and study heterogeneity (Viechtbauer & López-López 2022)
- *(Un-)predictability*
  - conditional dispersion as **non-reducible** problem for **measurement**
    - e.g., test instruments (Ghiselli 1960), survey responses (Hardt 2011) and conditional non-invariance (Martin et al. 2021)

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Sources of differences

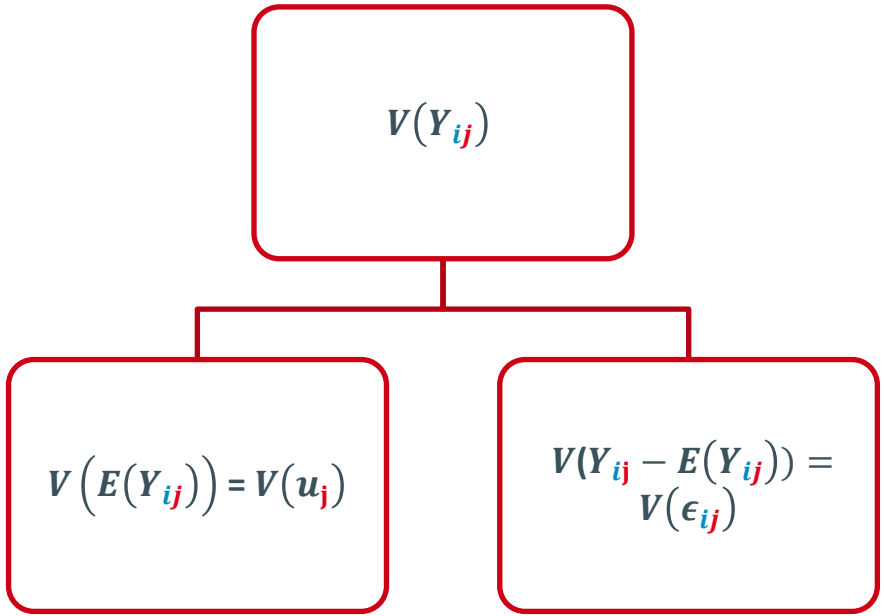
Potential sources of differences

- **Observational variability**
  - dispersion of Y in i



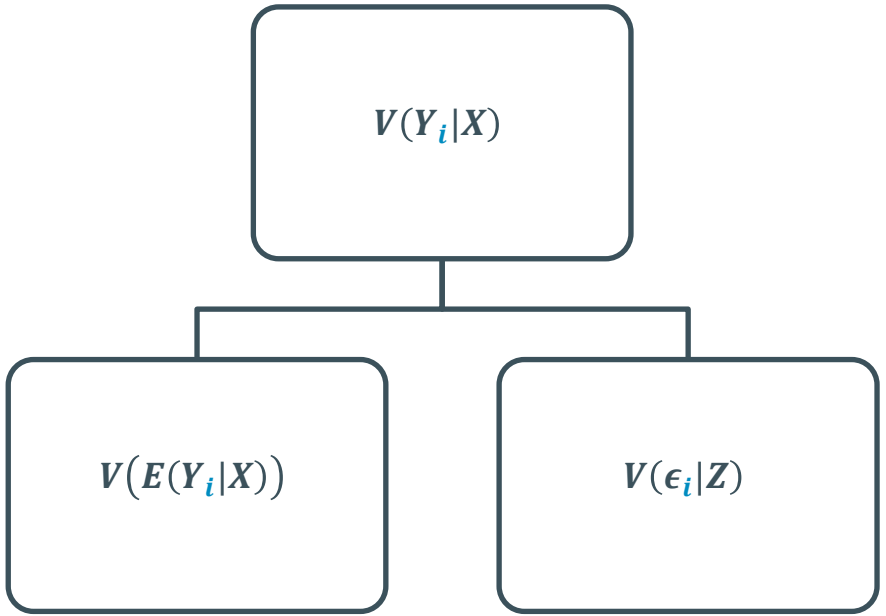
Potential sources of differences

- **Observational variability**
  - dispersion of Y in i
- **Unit variability**
  - $E(Y_{ij}) = b_0 + u_j$ becomes random in j
  - **within-unit** vs. **between-unit**



Potential sources of differences

- **Observational variability**
  - dispersion of Y in i
- **Unit variability**
  - $E(Y_{ij}) = b_0 + u_j$ becomes random in j
  - **within-unit** vs. **between-unit**
- **Treatments**
  - Conditional on X and Z, $E(Y_i|X)$ and $V(\epsilon_i|Z)$ can be random
  - **between-treatment** vs. **within-treatment**

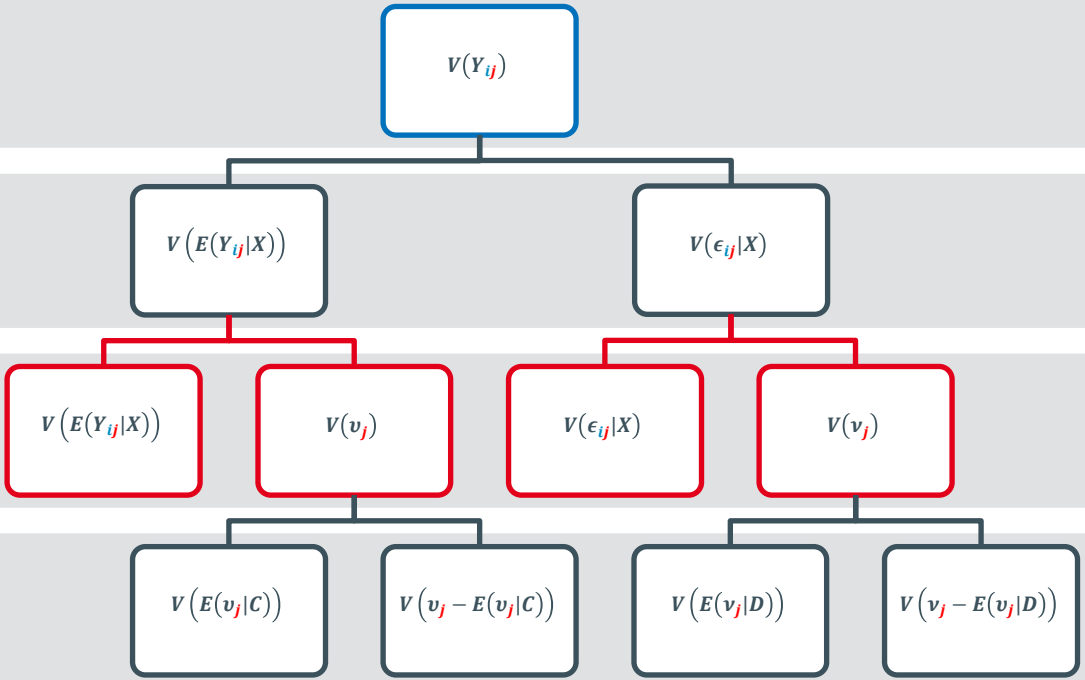


Observations

Between and within treatments

Within and between units

Between and within treatments



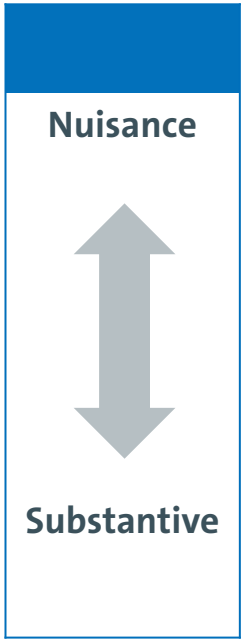
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Mechanisms of conditional dispersion

A glimpse beyond – reasons for conditional dispersion

	Probabilistic		Deterministic
Nuisance	*Noise by chance*	*Truncation* (Holm et al. 2015)	*Unmodelled non-linearities and moderators* (King & Roberts 2015)
	Measurement error (Ghiselli 1960)	*Asymmetric limitations of change* (Rosenberg et al. 2017)	*Heterogeneous aggregation of conflicting interests* (Braumoeller 2006)
	P(dY>0|X) > P(dY>0|!X) (Dzhafarov & Kujala 2012)	*Homogeneity pressure* (Ostroff & Fulmer 2014)	*Constrained necessary but insufficienct conditions and v.v.* (Downs & Rocke 1979)
Substantive	*Mixture of distributions* (Haaf & Rouder 2019)	*Hardening and polarisation* (Baumgaertner et al. 2016)	*Diffusion with relative adaptation* (Braumoeller 2006)

A glimpse beyond – from nuisance to substantive

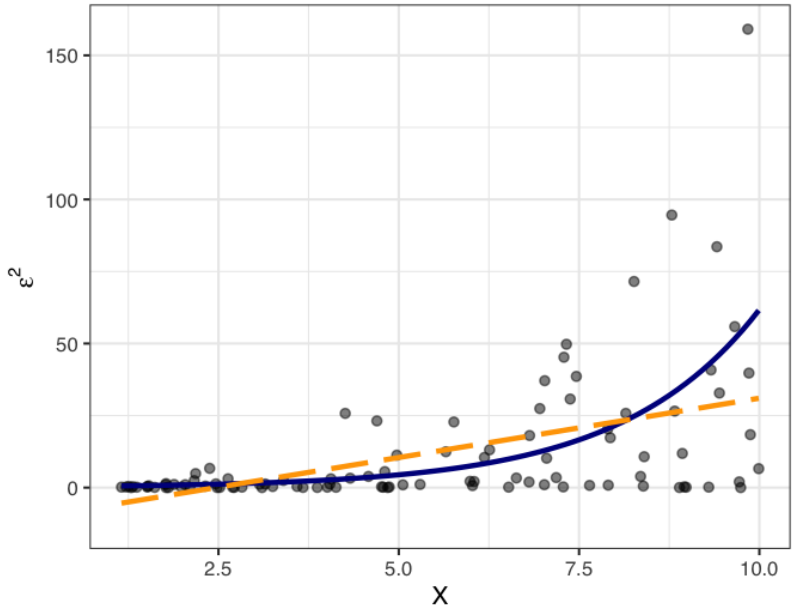


- **Nuiscance** can be ignored or **eliminated** by **improvements in the location model**
- Variability modeling as **precursor to complex location** models with **additional data** (e.g. network, diffusion)
- **Irreducible** mechanisms make **full variability modeling** necessary

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

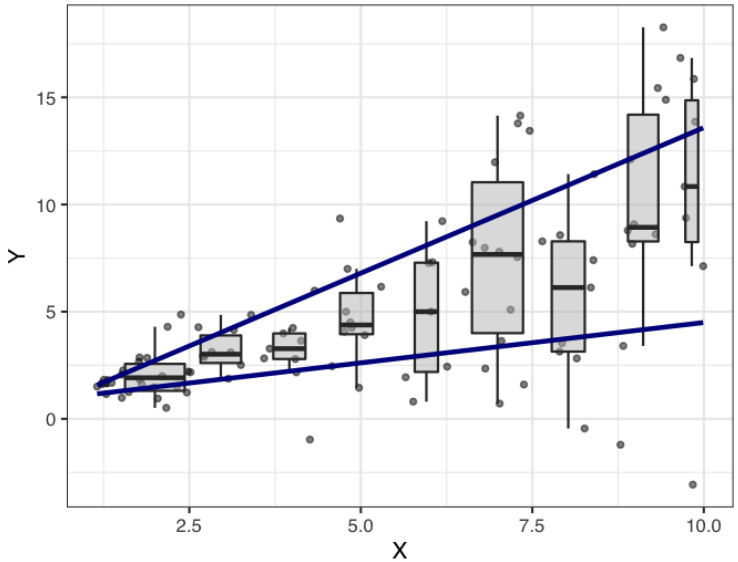
Regression models for dispersion

Variance function regression



- Two-step glm of gamma distributed **residuals** or **joint model** (Smyth 1989)
- VFR with ML- but REML-estimation for low N (Smyth 2002, Western and Bloome 2009)
- Assumes **linearity** on **link scale** and constant residual variance (after model) and **no outliers**
- Multilevel extension mixed-effects location-scale models (Hedeker et al. 2008)

Quantile regression

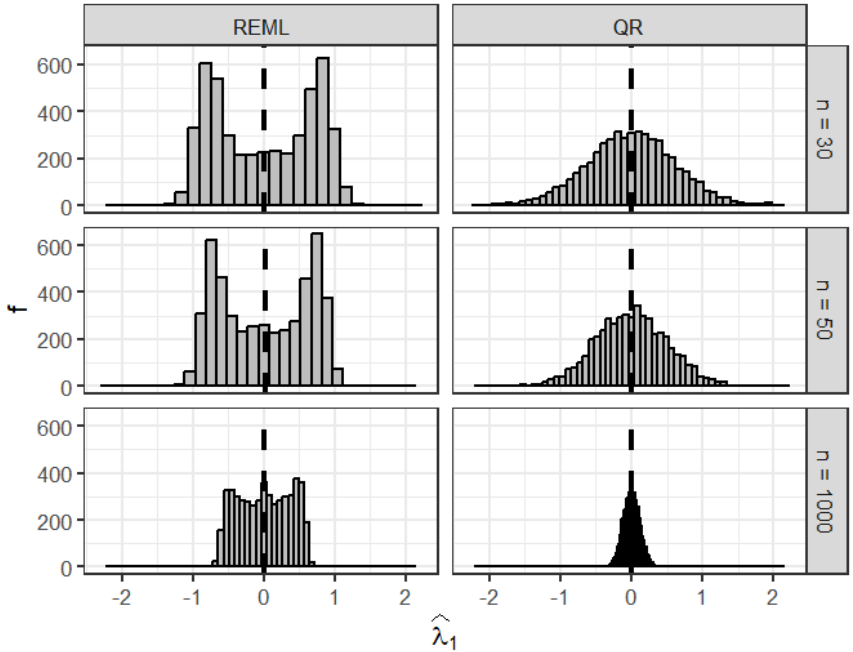


- minimization of **weighted absolute deviations** (Bassett Jr and Koenker 1978)
- **few distributional assumptions** but quantile crossing possible, and dependent variable needs to be **non-discrete** (Wenz 2019)
- **robust** against outliers
- multilevel extensions with asymmetric laplace approximation (Geraci & Bottai 2014)

Simulation study

- **Monte-Carlo-Simulation** with 5000 runs for each scenario
- **7 scenarios** each for N (30, 50, 1000) including
  - 3 (1 **misspecified**) null models
  - 2 (contaminated) **outlier** models
  - 2 semi-**realistic/complex** models
- Comparing performance measures for Two-Step, ML, REML and QR
  - *Bias*, *precision*, *coverage* and *true error rates*

Misspecification of expectation (S3)



- Missing non-linearity in expectation model:
  - **bimodal** distribution of the **estimates** for **VFR** with **inflated rejection rates**
  - **QR** is on average acceptable but **heavy tails** in the estimates imply potential for **high M-error**

Semi-realistic data generation scenarios

- More realistic but still simplified model

$$y_i = 3 + 0{,}5 \cdot x_i + 0{,}2 \cdot x_i^2 - 1 \cdot z_i - 3 \cdot x_i \cdot z_i + \sigma_i \cdot \epsilon_i, \quad \epsilon_i \sim N(0,1)$$

  - with **exponential variance** model to test **VFR**:

$$\sigma_i = \sqrt{\exp(2 + 0{,}5 \cdot x_i + 0{,}2 \cdot z_i - 0{,}4 \cdot x_i \cdot z_i)} \text{ (Szenario 6)}$$

  - And a **linear variance** model to test **QR**:

$$\sigma_i = 5 + 3 \cdot x_i + 1 \cdot z_i - 2 \cdot x_i \cdot z_i \text{ (Szenario 7)}$$

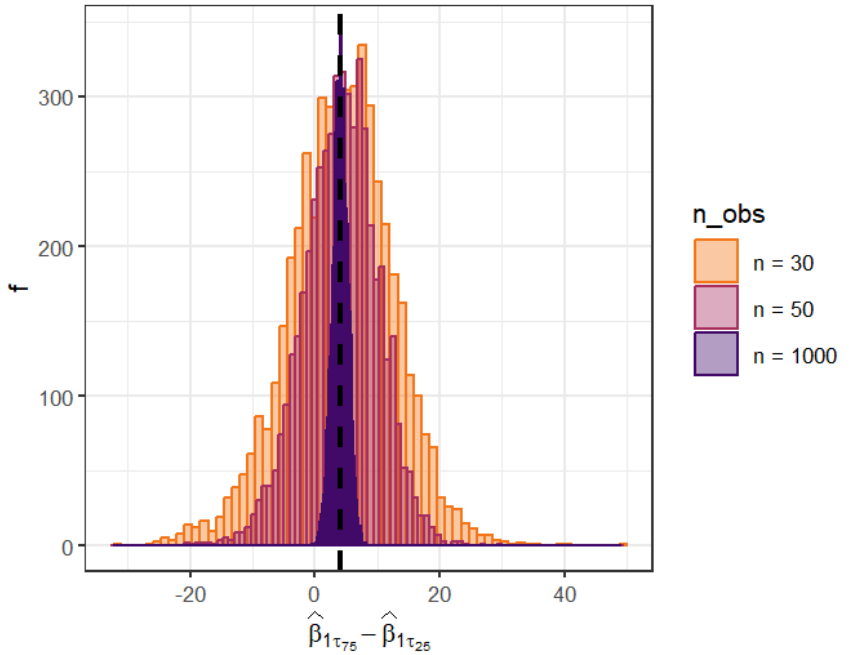
Small sample bias with more complex models for VFR …

		λ_1(X)	λ_2(Z)	λ_3(XZ)
Bias, n = 50				
	ML	**0.022**	**0.055**	**-0.019**
		(0.002)	**(0.018)**	**(0.003)**
	REML	**-0.006**	**-0.027**	0.003
		(0.002)	**(0.017)**	(0.003)
Empirical SE, n = 50				
	ML	0.148	1.279	0.213
		(0.001)	(0.013)	(0.002)
	REML	0.138	1.192	0.198
		(0.001)	(0.012)	(0.002)
Coverage rate of 95%-CI, n = 50				
	ML	**0.886**	**0.881**	**0.871**
		(0.005)	**(0.005)**	**(0.005)**
	REML	0.940	0.944	0.939
		(0.003)	(0.003)	(0.003)

		λ_1(X)	λ_2(Z)	λ_3(XZ)
Bias, n = 1000				
	ML	0.001	0.007	-0.001
		(0.000)	(0.003)	(0.001)
	REML	0.000	0.005	-0.000
		(0.000)	(0.003)	(0.001)
Empirical SE, n = 1000				
	ML	0.025	0.216	0.035
		(0.000)	(0.002)	(0.000)
	REML	0.025	0.215	0.035
		(0.000)	(0.002)	(0.000)
Coverage rate of 95%-CI, n = 1000				
	ML	**0.943**	**0.943**	**0.944**
		(0.003)	**(0.003)**	**(0.003)**
	REML	0.946	**0.944**	0.947
		(0.003)	**(0.003)**	(0.003)

… and differences of quantile regression estimates

	β_1 (X)	β_2 (XX)	β_3 (Z)	β_4 (XZ)
N = 50				
Bias	**0.148**	**-0.030**	-0.209	**0.083**
	(0.083)	**(0.008)**	(0.202)	**(0.044)**
EmpSE	5.884	0.570	14.287	3.077
	(0.059)	(0.006)	(0.143)	(0.031)
H0-Rejection	0.069	0.029	0.027	0.079
	(0.011)	(0.011)	(0.011)	(0.011)
N = 1000				
Bias	0.031	-0.004	-0.039	0.008
	(0.017)	(0.002)	(0.040)	(0.009)
EmpSE	1.230	0.121	**2.820**	0.643
	(0.012)	(0.001)	**(0.028)**	(0.006)
H0-Rejection	0.901	0.055	**0.072**	0.985
	(0.003)	(0.011)	**(0.011)**	(0.001)



Summary of preliminary simulation results

- **Two-Step biased** & high **undercoverage**
- **Misspecified expectation** leads to and **Type 1 errors** of VFR estimates
- **Small sample bias** and limited power but **QR** and **REML-VFR** with **acceptable coverage** for moderately complex models and N (30, 50)
- **REML-VFR more efficient** but **QR robust** especially to outliers

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

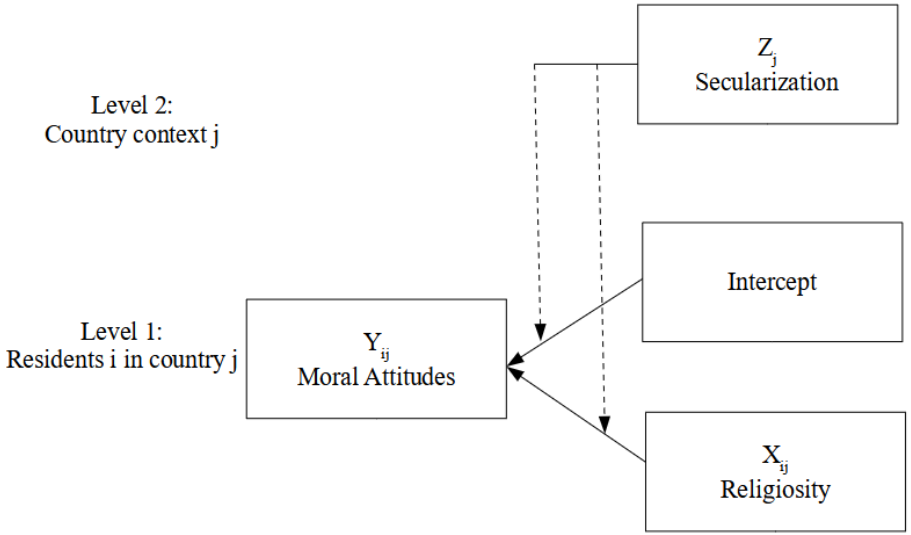
Secularization, religiosity and moral attitudes

Secularization as context of religious influence on moral

- Societal and individual secularization hand in hand (Dobbelaere 2002)
- *Individual secularization*
  - **reduction** of integration in religious communities and beliefs
- *Societal secularization*
  - **decoupling** i.a. of **moral attitudes** and **religiosity**
- With perspective of limited universalism (Nunner-Winkler 1996)
  - reduction and relativization of religion specific norms
  - **less religious foundation** of **universal norms** like civic morality

Multilevel model for moral attitudes

- Civic morality (C_{ij}) for individual ***i*** in country ***j*** by individual level religiosity (R_{ij}) country level secularization (S_j)
- Individual Level
  - $C_{ij} = \beta_{0j} + \beta_{1j} \cdot R_{ij} + Controls + \epsilon_{ij}$
- Country Level
  - $\beta_{0j} = \gamma_{00} + \gamma_{01} \cdot S_j + \upsilon_{0j}$
  - $\beta_{1j} = \gamma_{10} + \gamma_{11} \cdot S_j + \upsilon_{1j}$



Decreasing homogeneity pressure of religiosity

- For non-secular societies strong **homogeneity pressure** for more religious people
  - $r_1 > r_2 \Longrightarrow P(|c_1 - E(C|R = r_1 \pm \gamma)| - |c_2 - E(C|R = r_2 \pm \gamma)|) > 0$
  - V(C|R) **monotonic decreasing**
- In more secularized societies **lessened** authority **for decoupled universal norms**
  - V(C|R) approaching **constant** for max(S)
- Taken together implies **decreasing negative influence of R on Var(C) for higher S**:
  - V(C|R,S) = v($l_0 + l_1 R + l_2\, SR$) with $\boldsymbol{l_1|S = min < 0}$, $\boldsymbol{l_2 > 0}$ & $l_2\,(\max(S) - \min(S)) < -l_1$
  - V(C|R) varies by country!

MELS for secularization, religiosity and variability of moral attitudes

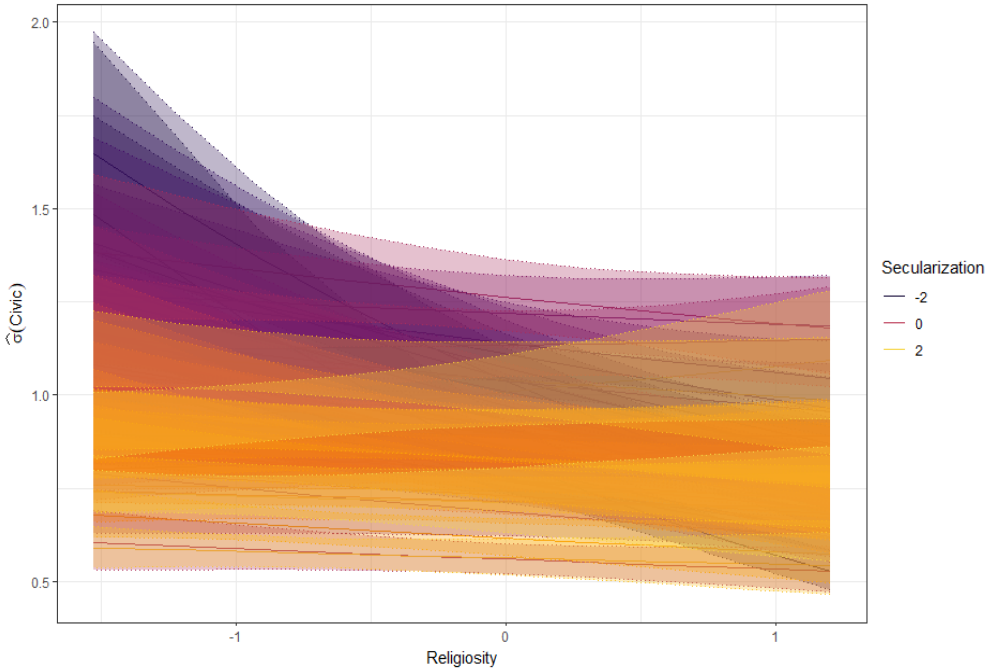
- Conditional residual variance in a a bayesian MELS with brms/Stan (Lester et al. 2019) & weakly informative standard priors
- Individual Level
  - $log(\sigma_{ij}) = l_{0j} + l_{1j} \cdot \mathrm{R}_{ij} + \mathrm{Controls}$
- Country Level
  - $l_{0j} = \mathrm{l}_{00} + \mathrm{l}_{01} \cdot \mathrm{S}_j + \nu_{0j}$
  - $l_{1j} = \mathrm{l}_{10} + \mathrm{l}_{11} \cdot \mathrm{S}_j + \nu_{1j}$
- Controls for gender, age and edu

	Estimate	95%-CI
Effects on mean		
Religiosity	-0.07 (0.04)	[-0.16; 0.02]
Secularization	0.15 (0.01)	[0.12; 0.17]
Religiosity*Secularization	-0.07 (0.01)	[-0.09; -0.04]
Effects on log. standard deviation		
Religiosity	**-0.07 (0.03)**	[-0.10; -0.05]
Secularization	**-0.07 (0.03)**	[-0.13; -0.02]
Religiosity*Secularization	**0.05 (0.01)**	[0.02; 0.08]

(Estimations based on EVS (2011) with 48891 persons from 42 country-regions)

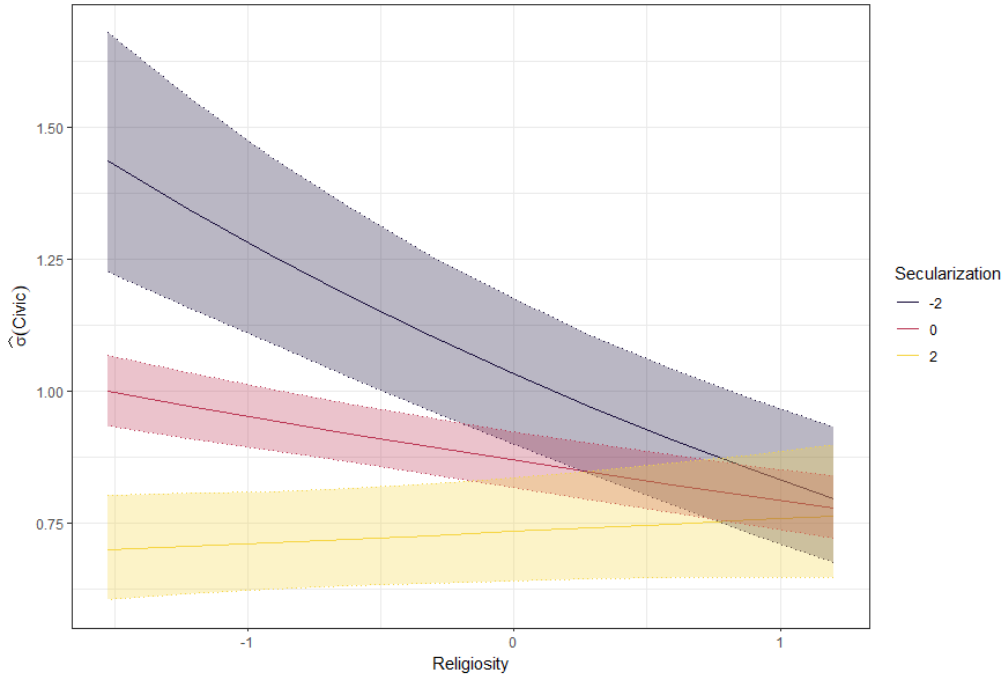
Variability decreases with religiosity only in non-secular context

- Model converges with standardized IVs
  - Rhat <= 1.01
  - Effective sample size >= 4045
- Individual level
  - $log(\sigma_{ij}) = l_{0j} + l_{1j} \cdot \mathrm{R}_{ij} + \mathrm{Controls}$
- Country level
  - $l_{0j} = -0.05 - 0.07 \cdot \mathrm{S}_{ij} + \nu_{0j}$
  - $l_{1j} = -0.07 + 0.06 \cdot \mathrm{S}_{ij} + \nu_{1j}$

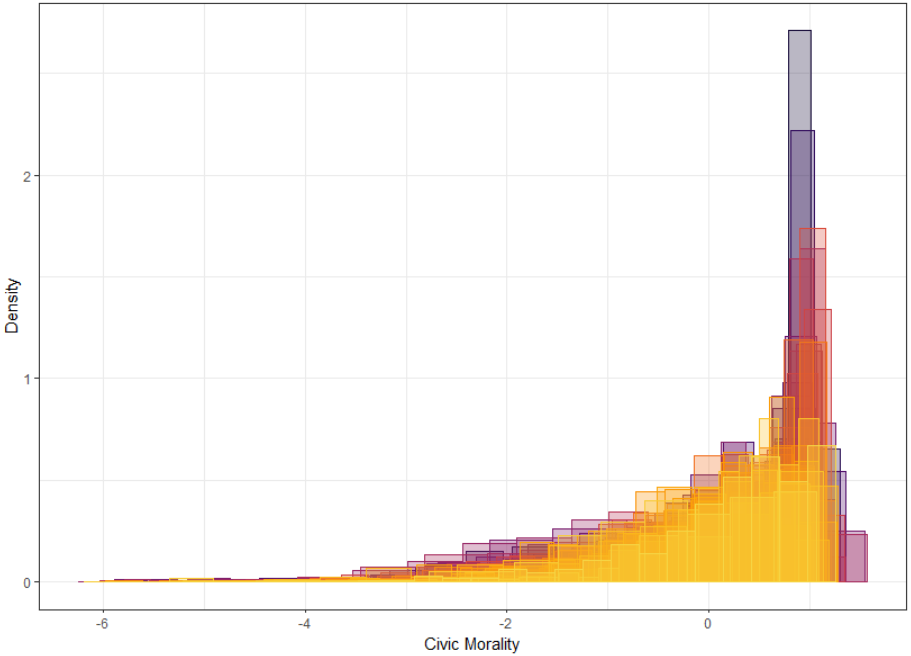


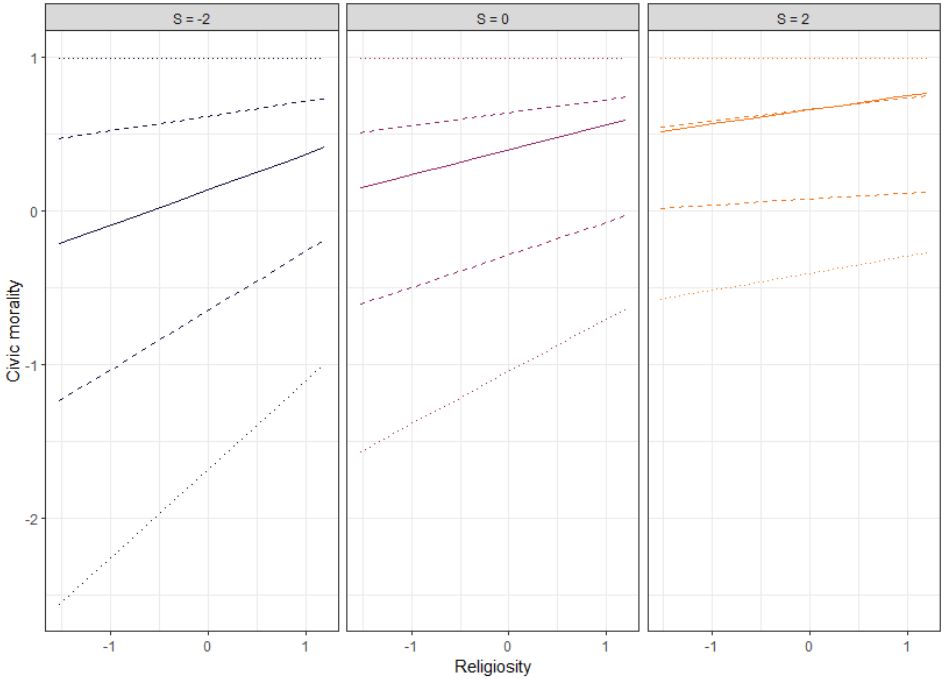
Variability decreases with religiosity only in non-secular context

- Model converges with standardized IVs
  - Rhat <= 1.01
  - Effective sample size >= 4045
- Individual level
  - $log(\sigma_{ij}) = l_{0j} + l_{1j} \cdot \mathrm{R}_{ij} + \mathrm{Controls}$
- Country level
  - $l_{0j} = -0.05 - 0.07 \cdot \mathrm{S}_{ij} + \nu_{0j}$
  - $l_{1j} = -0.07 + 0.06 \cdot \mathrm{S}_{ij} + \nu_{1j}$



Multilevel quantile model shows asymmetric CL-effect on lower quartile





Secularization changes homogenisation patterns

- Theoretical expectation only **partially supported**
  - Religiosity is **homogenizing factor** for civic morality **in less secular societies**
  - But in **more secular societies**, **non-religious people also** become **more similar**

➢ Results show **fruitfulness of theoretical deduction** and **empirical test** of variability assumption

➢ Unexpected result still informative for **partial weakness of theory**

➢ Non-symmetric relationship should inform **additional theoretization**

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Recommendations and open directions

Open directions

- Further development of **dreaded decision tree** of handling **variability**
- Integrate **multilevel dimension** directly into **typology of mechanisms**
- **Multilevel MC simulation** for
  - misspecifications
  - outliers
  - level 1 to level 2 relations
- Going beyond dispersion ...

Think beyond the mean

- Consider varying dispersion not as a statistical problem
  - It is an additional **source of differences** you might want to explain
  - If you assume **constant** dispersion, be explicit and **justify** it theoretically!
  - Use QR for asymmetries and outliers, REML-VFR with moderately small sample
  - Insights in **conditional dispersion** can **inform theoretical development**
  - **Larger range of predictions** from theory provide **more severe empirical tests**
    - **allows to distinguish better between different theoretical explanations**

Take home message

- Remember to **STAVE:**
  - **S**pecify your **T**heory
  - **A**ssess **V**ariability assumptions
  - **E**xtrapolate predictions

© Christoph Naefgen

The collaborators

daniel.gotthardt@uni-hamburg.de

christoph.naefgen@fernuni-hagen.de

anne.k.reinarz@durham.ac.uk

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Literature

Literatur

Bassett Jr, Gilbert, und Roger Koenker. 1978. „Asymptotic theory of least absolute error regression“. *Journal of the American Statistical Association* 73(363):618–22. doi: 10.2307/2286611.

Baumgaertner, Bert O., Rebecca C. Tyson, und Stephen M. Krone. 2016. „Opinion strength influences the spatial dynamics of opinion formation“. *The Journal of Mathematical Sociology* 40(4):207–18. doi: 10.1080/0022250X.2016.1205049.

Braumoeller, Bear F. 2006. „Explaining Variance; Or, Stuck in a Moment We Can't Get Out Of“. *Political Analysis* 14(3):268–90. doi: 10.1093/pan/mpj009.

Culpepper, Steven Andrew. 2010. „Studying Individual Differences in Predictability With Gamma Regression and Nonlinear Multilevel Models“. *Multivariate Behavioral Research* 45(1):153–85. doi: 10.1080/00273170903504885.

Dobbelaere, Karel. 2002. *Secularization: An Analysis at Three Levels*. Brüssel: Presses interuniversitaires européennes.

Downs, George W., und David M. Rocke. 1979. „Interpreting heteroscedasticity“. *American Journal of Political Science* 23(4):816–28. doi: 10.2307/2110809.

Dzhafarov, Ehtibar N., und Janne V. Kujala. 2012. „Selectivity in probabilistic causality: Where psychology runs into quantum physics“. *Journal of Mathematical Psychology* 56(1):54–63. doi: 10.1016/j.jmp.2011.12.003.

EVS. 2011. *European Values Study 2008: Integrated Dataset (EVS 2008)*. Köln: GESIS Data Archive.

Geraci, Marco, und Matteo Bottai. 2014. „Linear Quantile Mixed Models“. *Statistics and Computing* 24(3):461–79. doi: 10.1007/s11222-013-9381-9.

Ghiselli, Edwin E. 1960. „The Prediction of Predictability“. *Educational and Psychological Measurement* 20(1):3–8. doi: 10.1177/001316446002000101.

Haaf, Julia M., und Jeffrey N. Rouder. 2019. „Some Do and Some Don't? Accounting for Variability of Individual Difference Structures“. *Psychonomic Bulletin & Review* 26(3):772–89. doi: 10.3758/s13423-018-1522-x.

Hardt, Nino. 2011. „Heterogenität im Antwortverhalten bei Kundenzufriedenheitsbefragungen - ein bayesianischer Mixture-Modell-Ansatz“. Dissertation, Katholische Universität Eichstätt-Ingolstadt, Ingolstadt.

Hedeker, Donald, Robin J. Mermelstein, und Hakan Demirtas. 2008. „An Application of a Mixed-Effects Location Scale Model for Analysis of Ecological Momentary Assessment (EMA) Data“. *Biometrics* 64(2):627–34. doi: 10.1111/j.1541-0420.2007.00924.x.

Holm, Anders, Mette Ejrnæs, und Kristian Karlson. 2015. „Comparing Linear Probability Model Coefficients across Groups“. *Quality & Quantity* 49(5):1823–34. doi: 10.1007/s11135-014-0057-0.

King, Gary, und Margaret E. Roberts. 2015. „How robust standard errors expose methodological problems they do not fix, and what to do about it“. *Political Analysis* 23(2):159–79. doi: 10.1093/pan/mpu015.

Lang, Jonas W. B., Paul D. Bliese, und Alex de Voogt. 2018. „Modeling consensus emergence in groups using longitudinal multilevel methods“. *Personnel Psychology* 71(2):255–81. doi: 10.1111/peps.12260.

Leamer, Edward E. 2010. „Tantalus on the Road to Asymptopia“. *Journal of Economic Perspectives* 24(2):31–46. doi: 10.1257/jep.24.2.31.

Literatur

Lester, Houston F., Kristin L. Cullen-Lester, und Ryan W. Walters. 2019. „From Nuisance to Novel Research Questions: Using Multilevel Models to Predict Heterogeneous Variances“. *Organizational Research Methods* 24(2):342–88. doi: 10.1177/1094428119887434.

Lindley, Dennis V., und Adrian FM Smith. 1972. „Bayes estimates for the linear model“. *Journal of the Royal Statistical Society: Series B (Methodological)* 34(1):1–18. doi: 10.1111/j.2517-6161.1972.tb00885.x.

Martin, Stephen R., Donald R. Williams, und Philippe Rast. 2019. „Measurement invariance assessment with Bayesian hierarchical inclusion modeling“. doi: 10.31234/osf.io/qbdjt.

Mattes, André, Felice Tavera, Anja Ophey, Mandy Roheger, Robert Gaschler, und Hilde Haider. 2021. „Parallel and serial task processing in the PRP paradigm: A drift–diffusion model approach“. *Psychological Research* 85(4):1529–52. doi: 10.1007/s00426-020-01337-w.

Nelder, J. A., und R. W. M. Wedderburn. 1972. „Generalized Linear Models“. *Journal of the Royal Statistical Society. Series A (General)* 135(3):370–84. doi: 10.2307/2344614.

Nunner-Winkler, Gertrud. 1996. „Ein Plädoyer für einen eingeschränkten Universalismus“. S. 126–44 in *Zur Bestimung der Moral. Philosophische und sozialwissenschaftliche Beiträge zur Moralforschung*, herausgegeben von W. Edelstein und G. Nunner-Winkler. Frankfurt a.M.: Suhrkamp.

Ostroff, Cheri, und C. Ashley Fulmer. 2014. „Variance as a construct: Understanding variability beyond the mean“. S. 185–210 in *The nature of work: Advances in psychological theory, methods, and practice*, herausgegeben von K. Ford, J. R. Hollenbeck, und A. M. Ryan. Washington, DC, US: American Psychological Association.

Rosenberg, Andrew S., Austin J. Knuppe, und Bear F. Braumoeller. 2017. „Unifying the Study of Asymmetric Hypotheses“. *Political Analysis* 25(3):381–401. doi: 10.1017/pan.2017.16.

Smyth, Gordon K. 1989. „Generalized Linear Models with Varying Dispersion“. *Journal of the Royal Statistical Society: Series B (Methodological)* 51(1):47–60. doi: 10.1111/j.2517-6161.1989.tb01747.x.

Smyth, Gordon K. 2002. „An efficient algorithm for REML in heteroscedastic regression“. *Journal of Computational and Graphical Statistics* 11(4):836–47. doi: 10.1198/106186002871.

Snijders, Tom A. B., und Roel J. Bosker. 2012. *Multilevel analysis an introduction to basic and advanced multilevel modeling*. 2nd edition. Los Angeles, London, New Delhi, Singapore, Washington D.C.: Sage.

Viechtbauer, Wolfgang, und José Antonio López-López. 2022. „Location-Scale Models for Meta-Analysis“. *Research Synthesis Methods* 13(6):697–715. doi: 10.1002/jrsm.1562.

Wenz, Sebastian E. 2019. „What Quantile Regression Does and Doesn’t Do: A Commentary on Petscher and Logan (2014)“. *Child Development* 90(4):1442–52. doi: 10.1111/cdev.13141.

Western, Bruce, und Deirdre Bloome. 2009. „Variance Function Regressions for Studying Inequality“. *Sociological Methodology* 39(1):263–326. doi: 10.1111/j.1467-9531.2009.01222.x.

White, Halbert. 1980. „A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity“. *Econometrica* 48(4):817. doi: 10.2307/1912934.

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Backup Slides

Outlier Scenarios

- We specificy 10% outliers via a mixture of normal distributions with constant overall standard distribution for *S4*, while we force influential outliers in *S5* with a multivariate normal distribution for 10% of the cases:

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} \sim MVN\left(\begin{matrix} 8 \\ 4 \end{matrix}, \sigma_{xy}\right) \quad \sigma^2_{xy} = \begin{pmatrix} 0{,}5^2 & 0 \\ 0 & 0{,}5^2 \end{pmatrix}$$

- Even vertical outliers in *S4* **highly disturb the tails of the estimates** of the **variance** based **models** while QR is only affected by many influential outliers (*S5*)

Current Limitations of Simulation Study

- **Simulation Uncertainty** (Monte-Carlo error) higher than expected
  - More complex scenarios will need even more simulation runs
- **Differentiation** between **heterogeneous effects** (interactions) and **heteregenoues variability** not fully explored
- **Comparability** especially of **parameters** between **QR** and **VFR** limited
- **Extension** to **multilevel models** and **implications** for **variance misspecification** planned