

**Appendices for: What to do about prediction errors in automated content  
analysis**

Nathan TeBlunthuis<sup>1</sup>, Valerie Hase<sup>2</sup>, and Chung-hong Chan<sup>3</sup>

<sup>1</sup>Northwestern University

<sup>2</sup>LMU Munich

<sup>3</sup>GESIS - Leibniz-Institut für Sozialwissenschaften

**Abstract**

*Keywords:* Content Analysis; Machine Learning; Classification Error; Attenuation Bias; Simulation; Computational Methods; Big Data; AI;

Appendices for: What to do about prediction errors in automated content analysis

Additional plots from Simulations 1 and 2

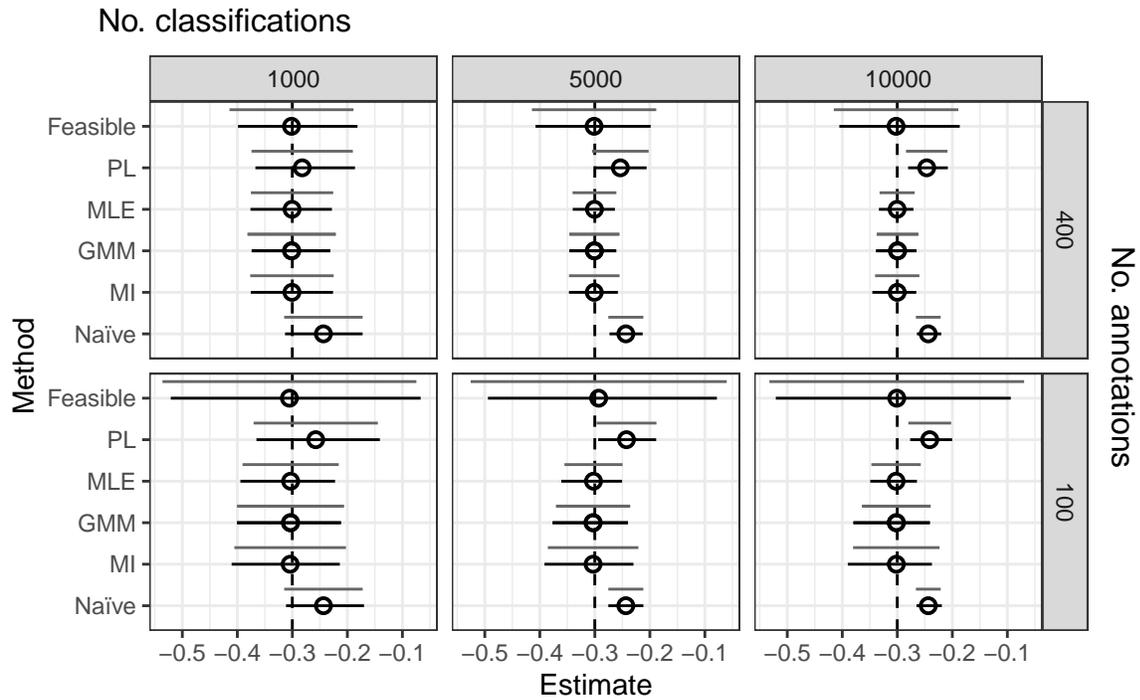


Figure 1

*Estimates of  $B_z$  in multivariate regression with  $X$  measured using machine learning and model accuracy independent of  $X$ ,  $Y$ , and  $Z$ . All methods obtain precise and accurate estimates given sufficient validation data.*