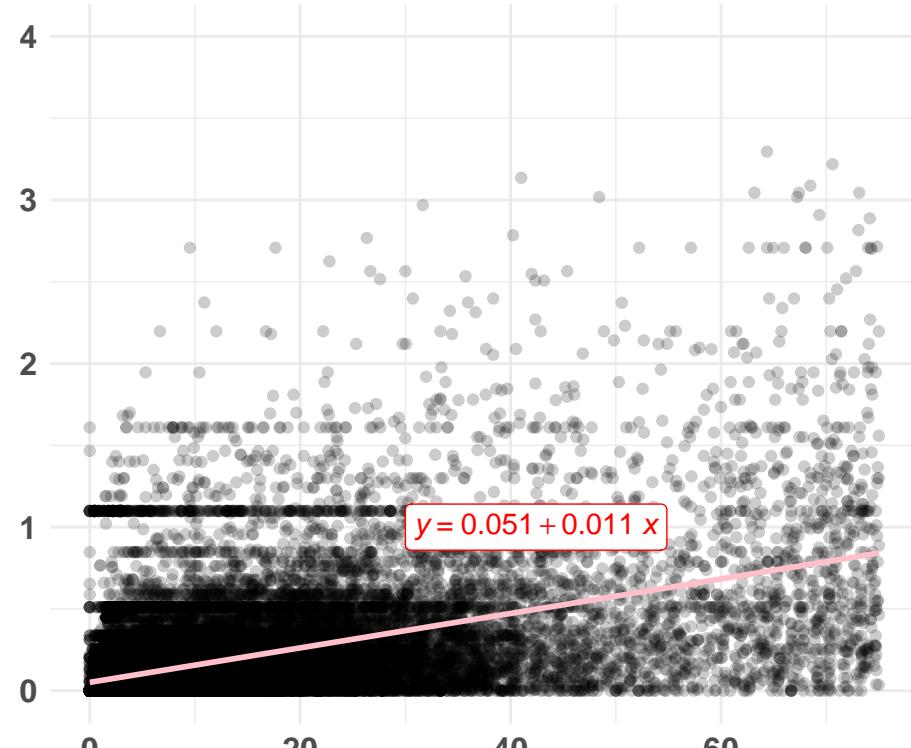
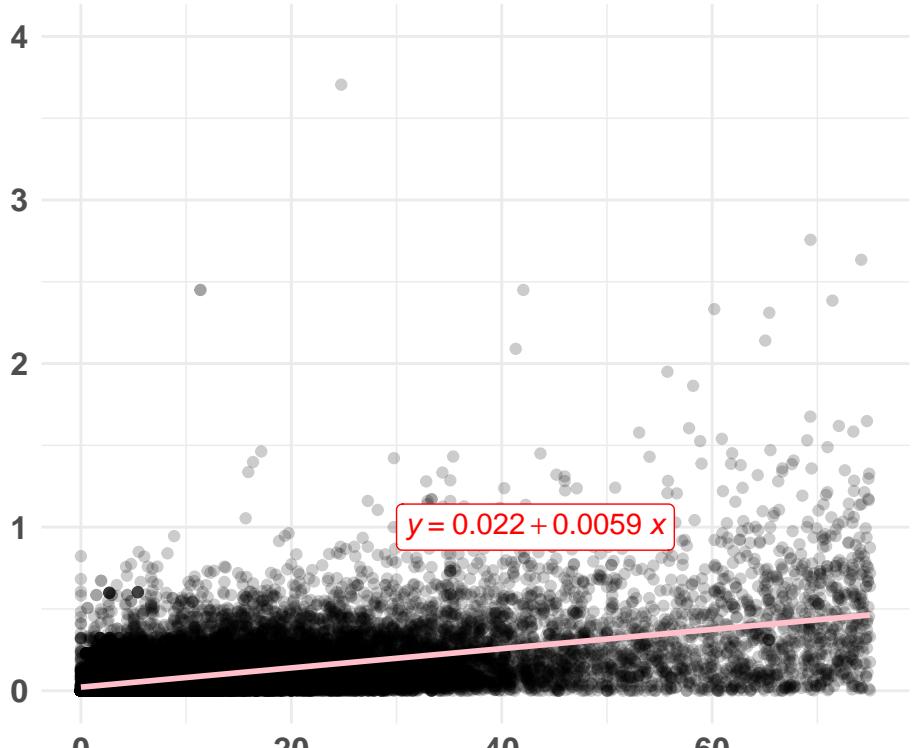


Differences in document-level (n=42504) sentiment polarity and valence between OCR processed versions of 'Old Books' and ground truth. Scores calculated with Quanteda's 'LSD 2015' and 'ANEW' dictionaries. Y axis is absolute difference in polarity/valence points from ground truth score. X axis cropped.

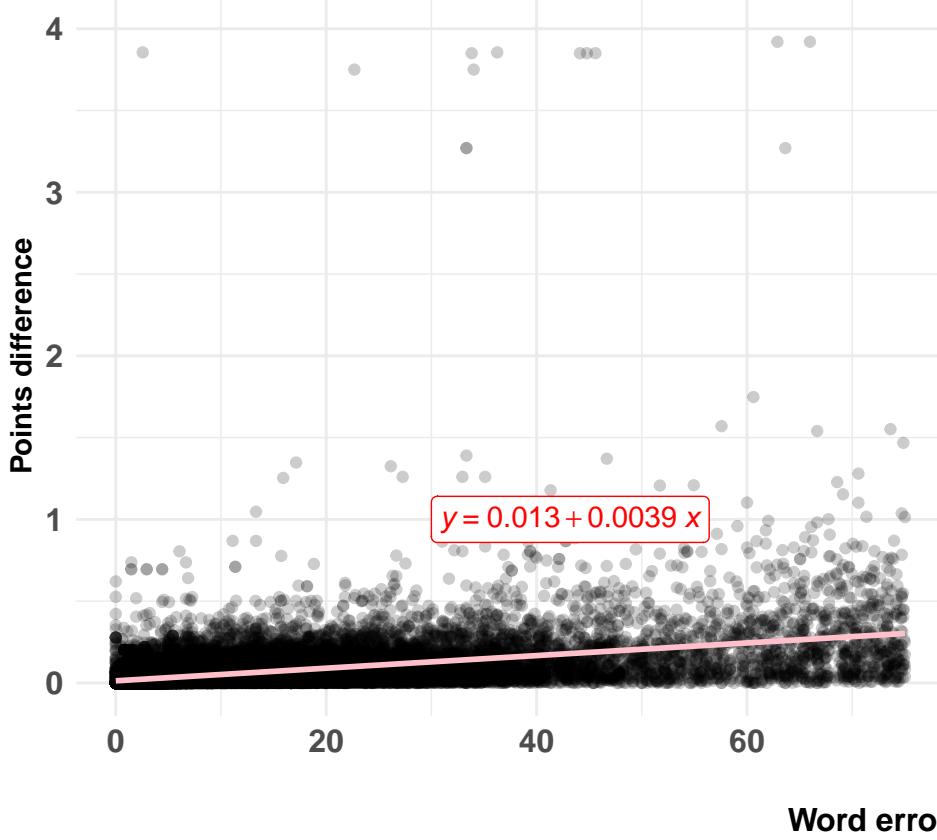
a) Polarity



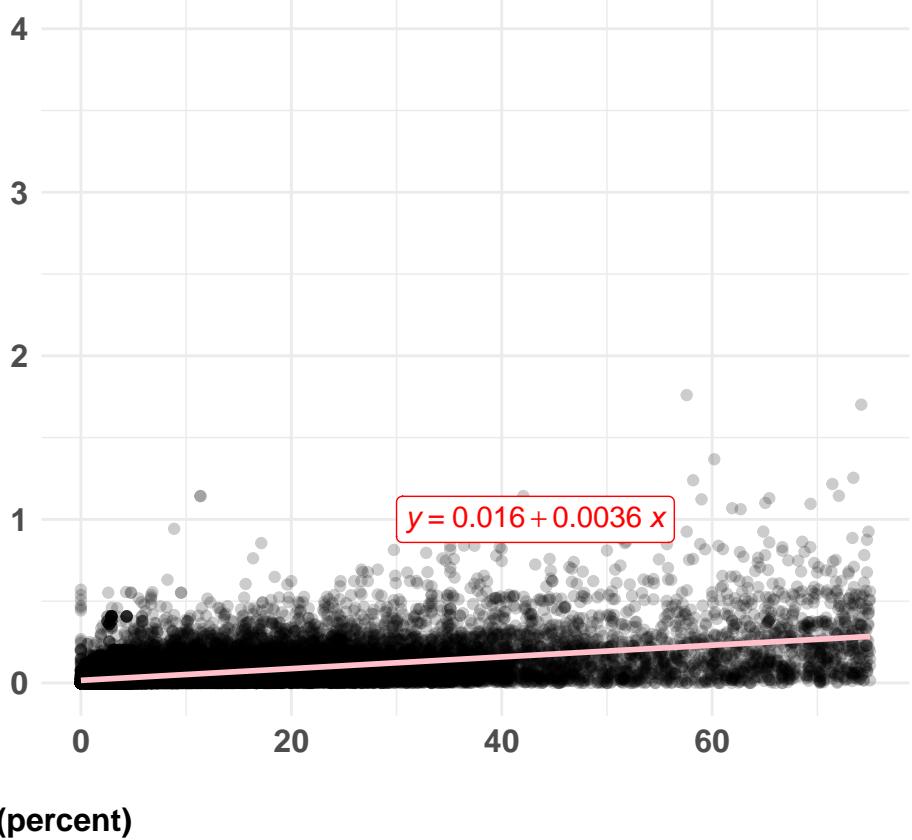
b) Valence: pleasure



c) Valence: arousal



d) Valence: dominance



Word error (percent)