

Empirical prediction intervals applied to short-term mortality forecasts and excess deaths

Jonas Schöley
*joint work with
Ricarda Duerst*

 @jschoeley

 0000-0002-3340-8518

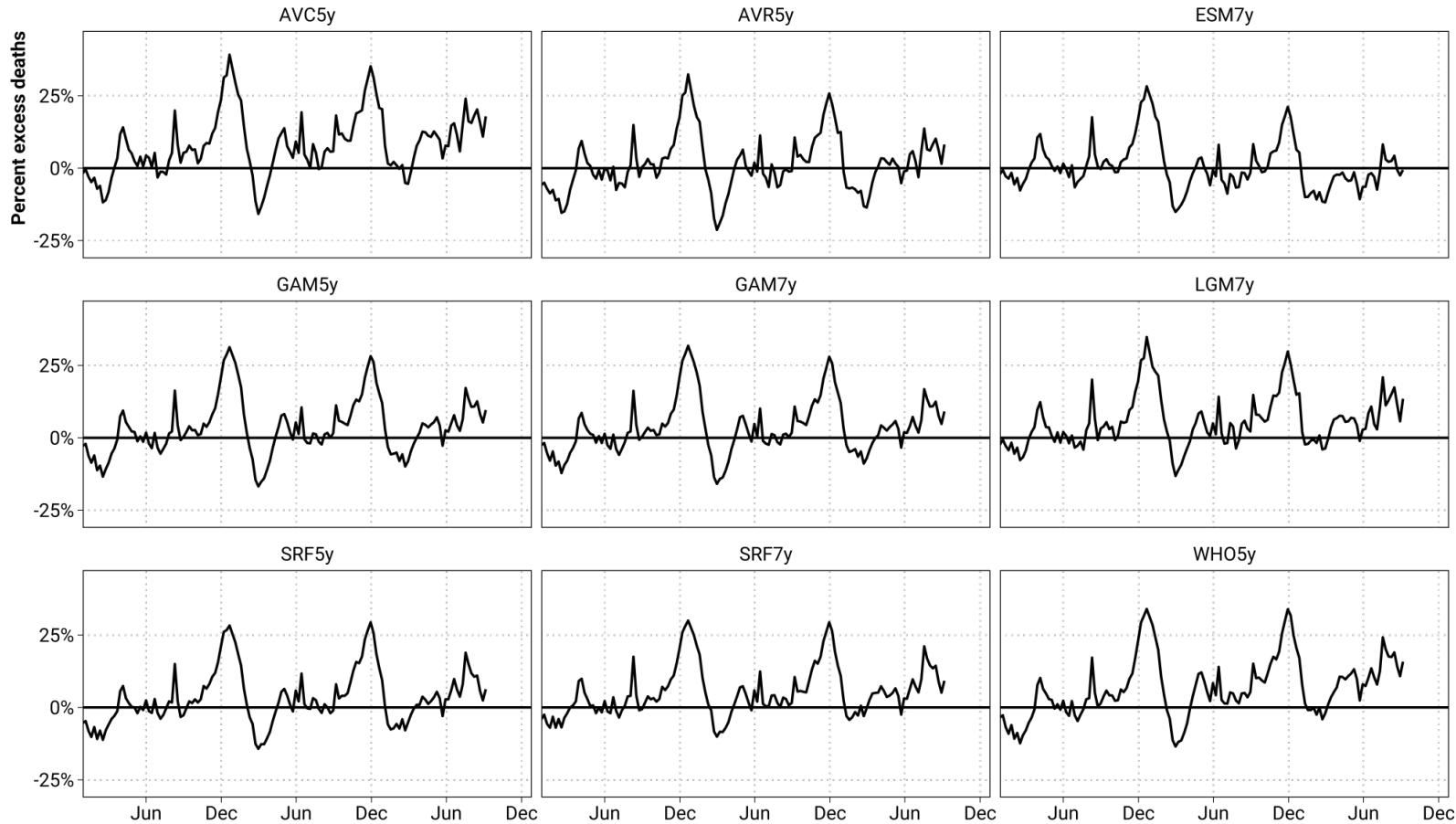
 schoeley@demogr.mpg.de



MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH

Excess deaths as a forecasting challenge

Percent excess deaths Germany 2020w1 through 2022 under different baseline models



Excess deaths as a forecasting challenge

How many deaths in a week had C19 not occurred?

Excess deaths as a forecasting challenge

How many deaths in a week had C19 not occurred?

$$\text{Excess} = \text{Observed} - \text{Expected}$$

Excess deaths as a forecasting challenge

How many deaths in a week had C19 not occurred?

$$\text{Excess} = \text{Observed} - \text{Expected}$$

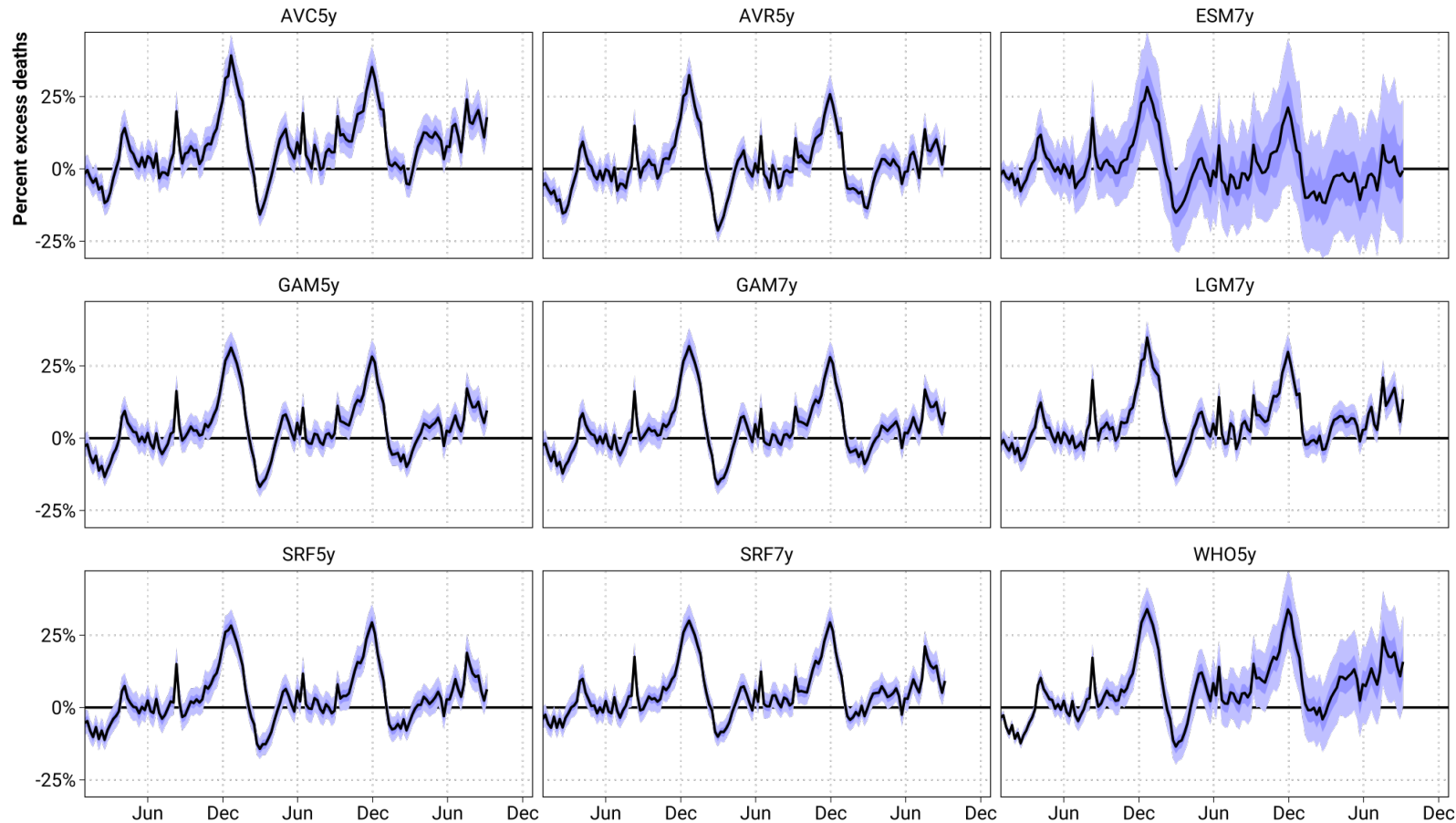
Excess deaths as a forecasting challenge

How many deaths in a week had C19 not occurred?

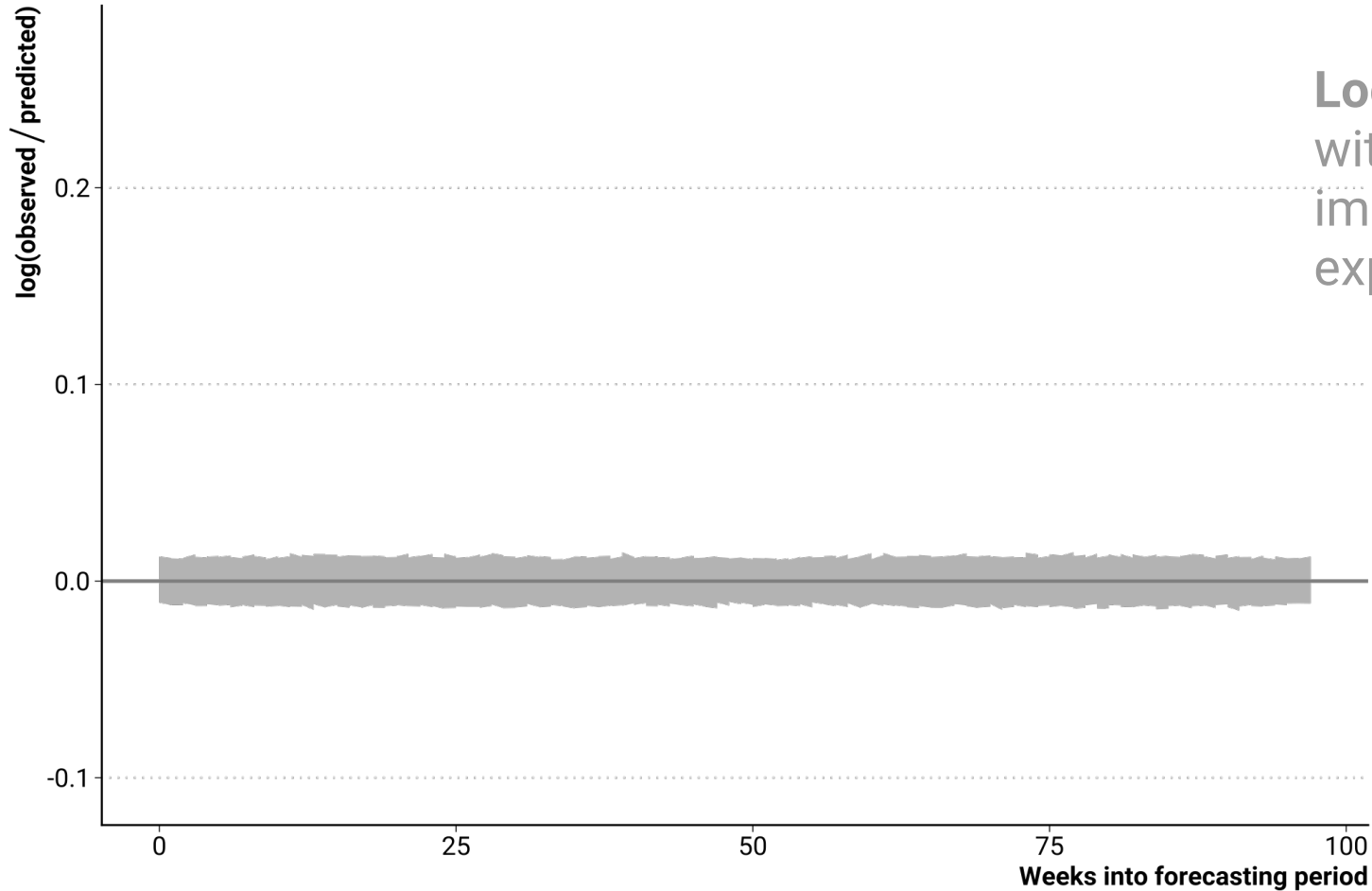
$$\text{Excess} = \text{Observed} - \text{Expected}$$

Excess deaths as a forecasting challenge

Percent excess deaths Germany 2020w1 through 2022 under different baseline models

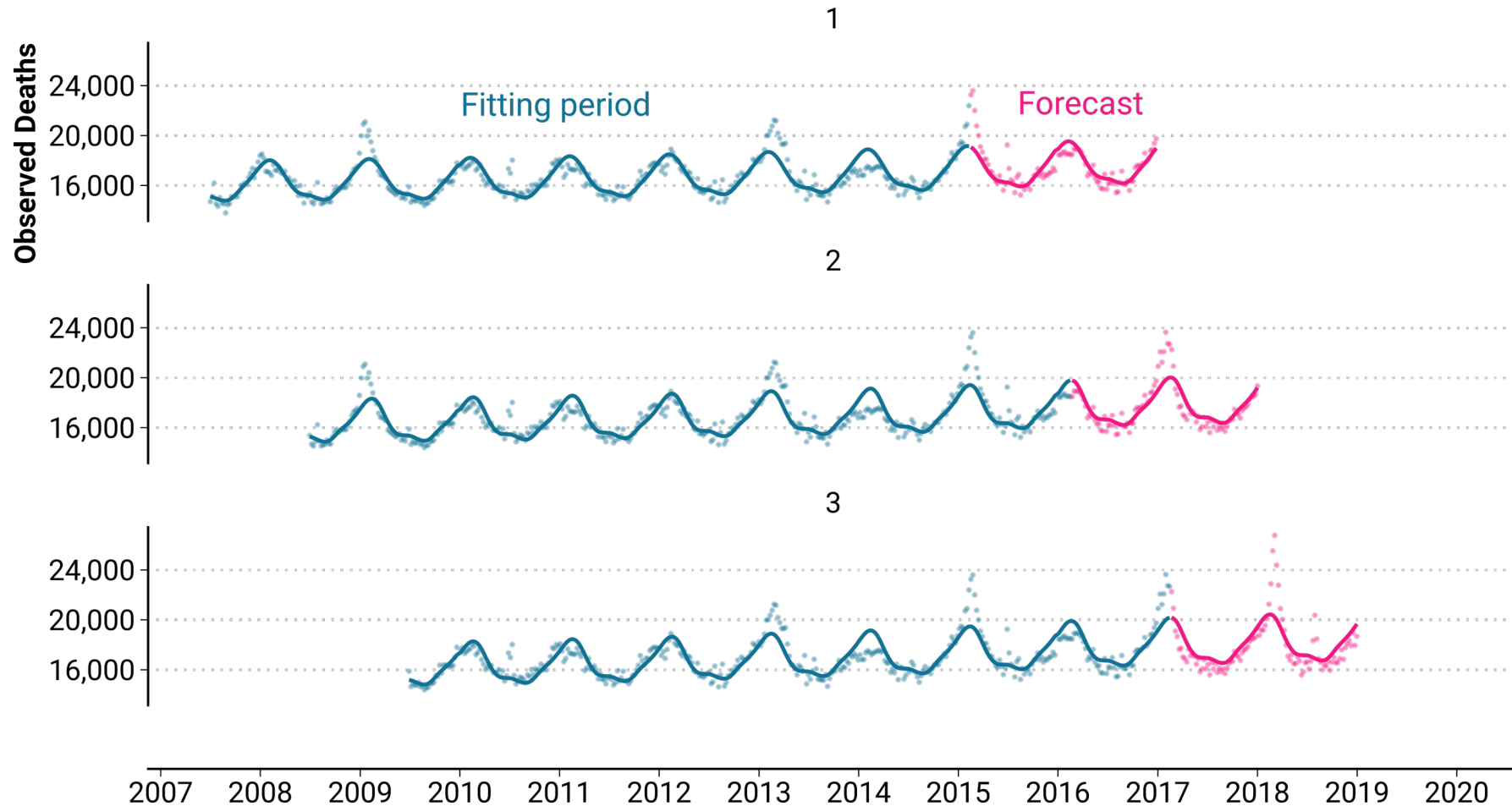


Excess deaths as a forecasting challenge

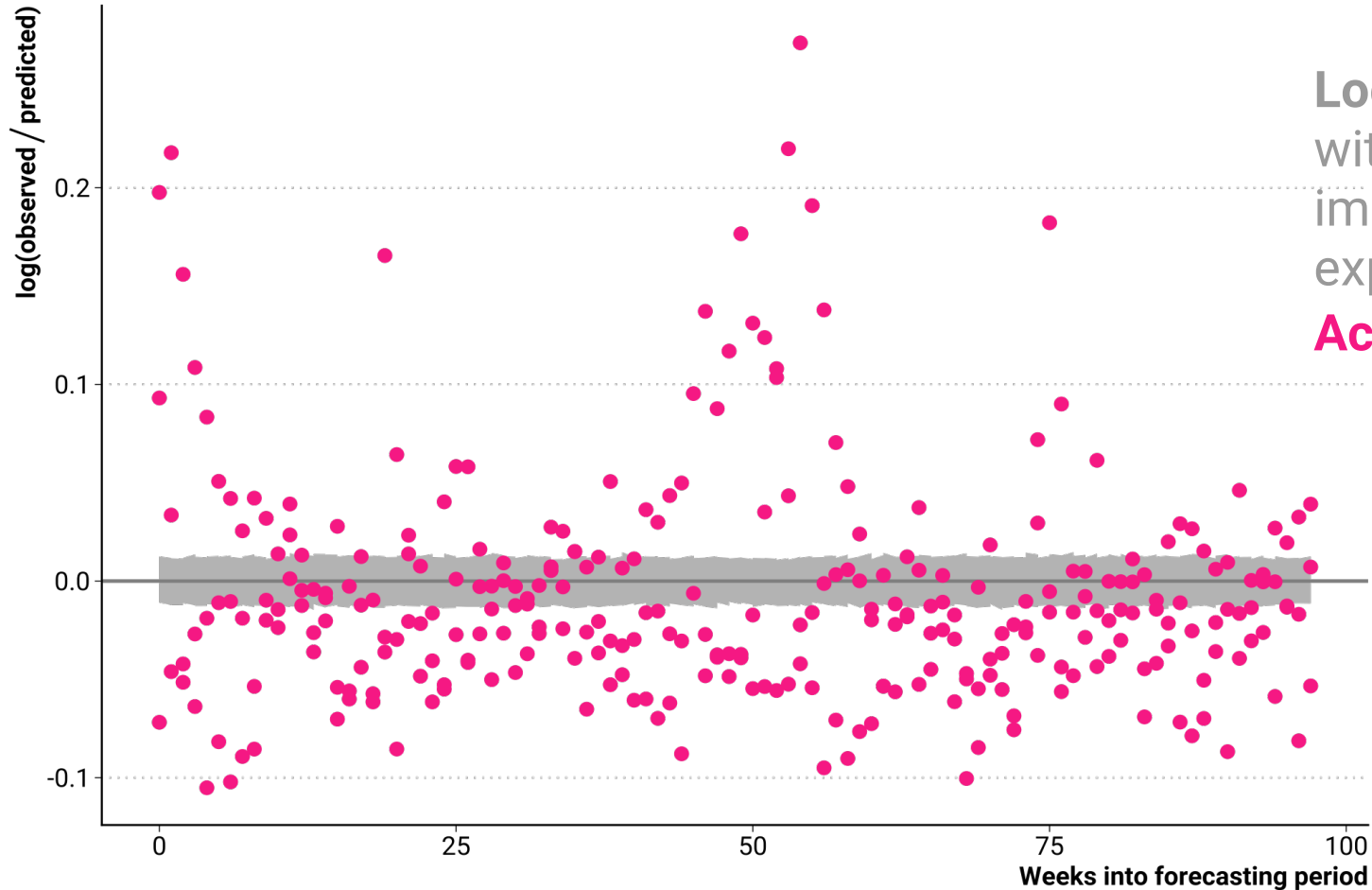


Log prediction error region with **90% nominal coverage** implied by Poisson-GAM expected deaths model

The logic of out-of-sample validation

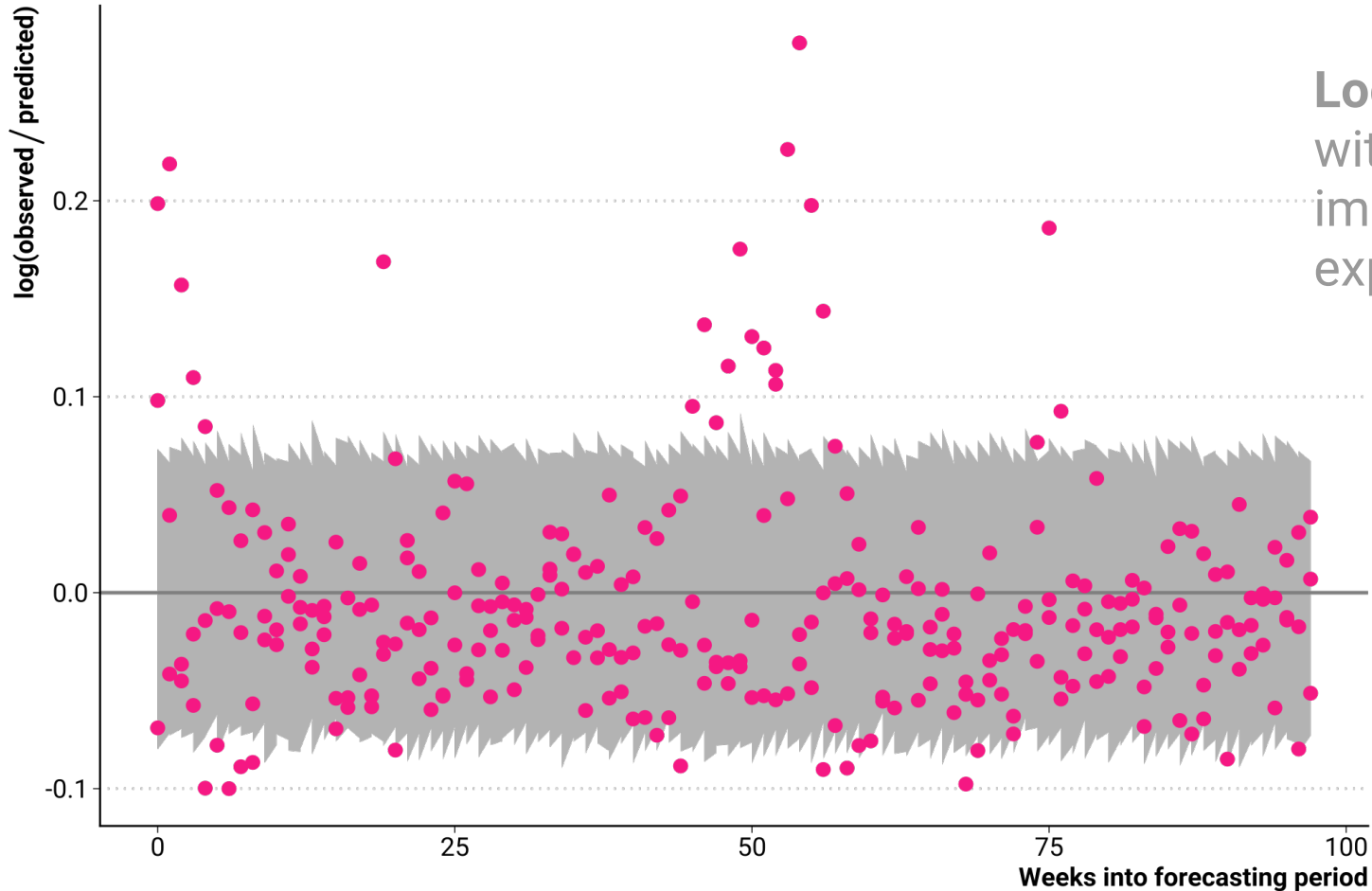


Excess deaths as a forecasting challenge



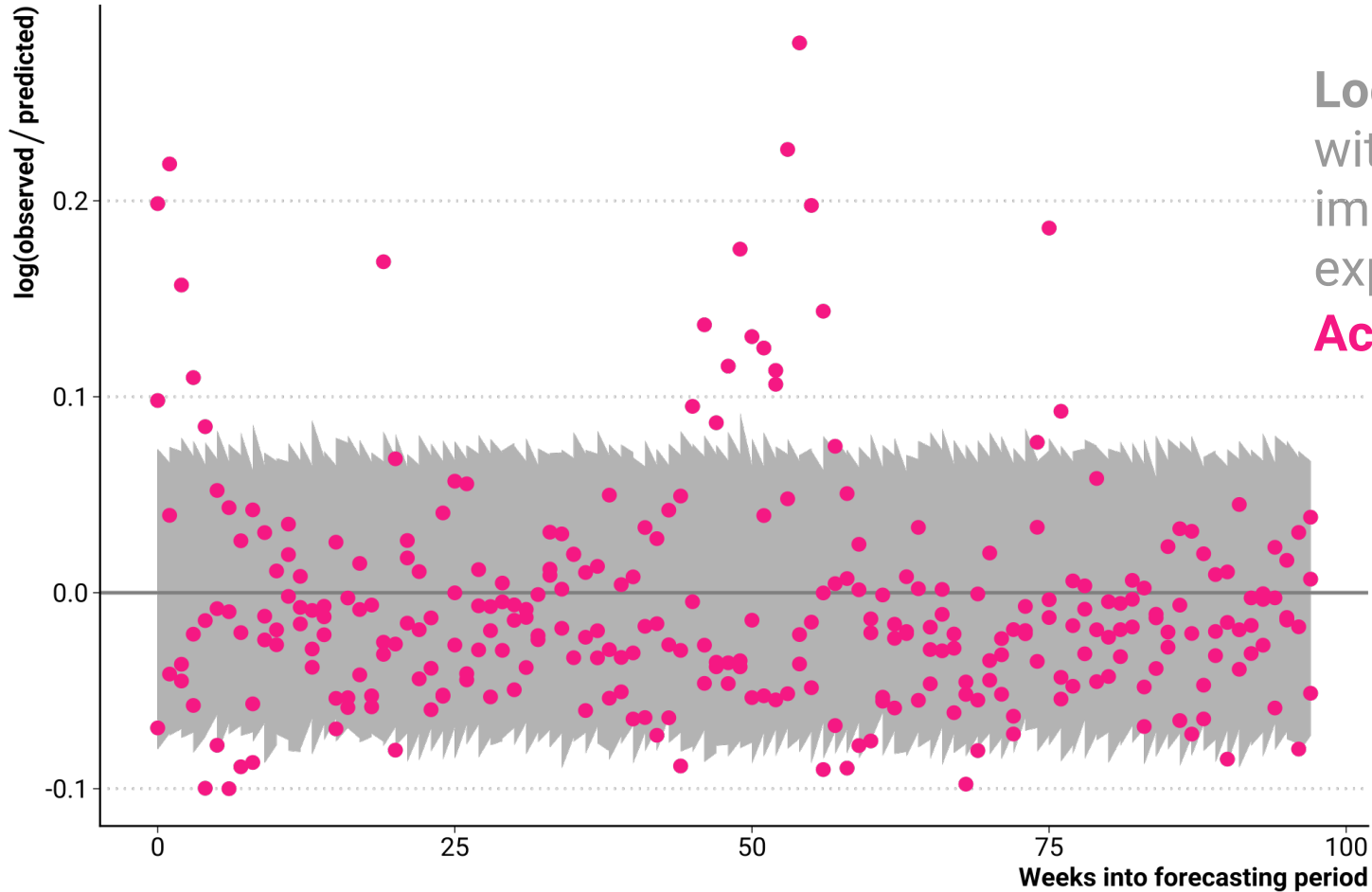
Log prediction error region
with 90% nominal coverage
implied by Poisson-GAM
expected deaths model
Actual coverage ~22%

Excess deaths as a forecasting challenge



Log prediction error region with 90% nominal coverage implied by NegBin-GAM expected deaths model

Excess deaths as a forecasting challenge



Log prediction error region
with 90% nominal coverage
implied by NegBin-GAM
expected deaths model
Actual coverage ~87%

Cross-validation

The **prediction error** of your model is estimated from out of sample validation on known data.

The logic of out-of-sample validation

Cross-validation

The **prediction error** of your model is estimated from out of sample validation on known data.

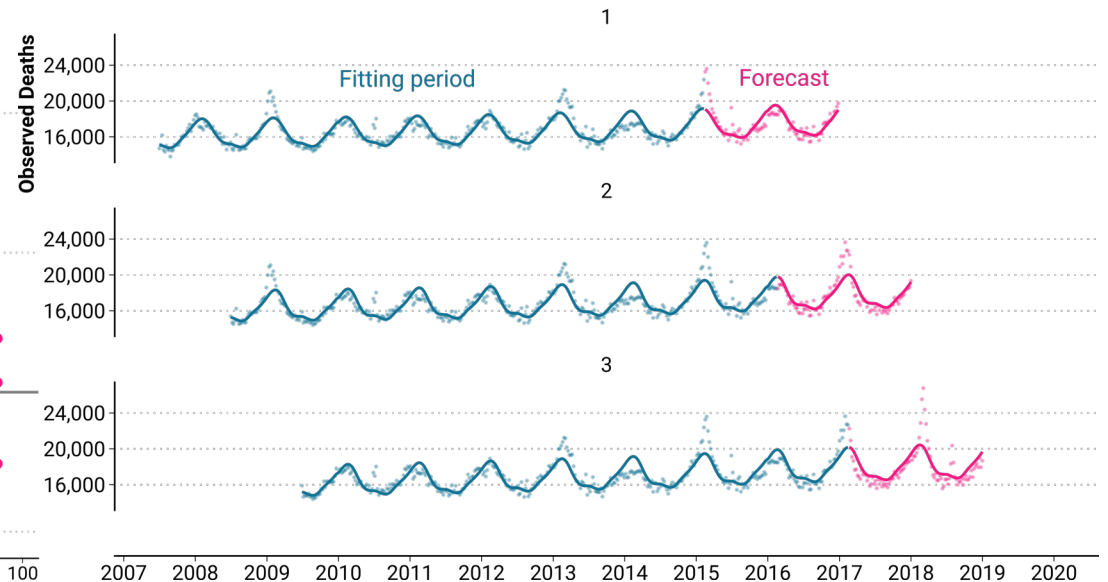
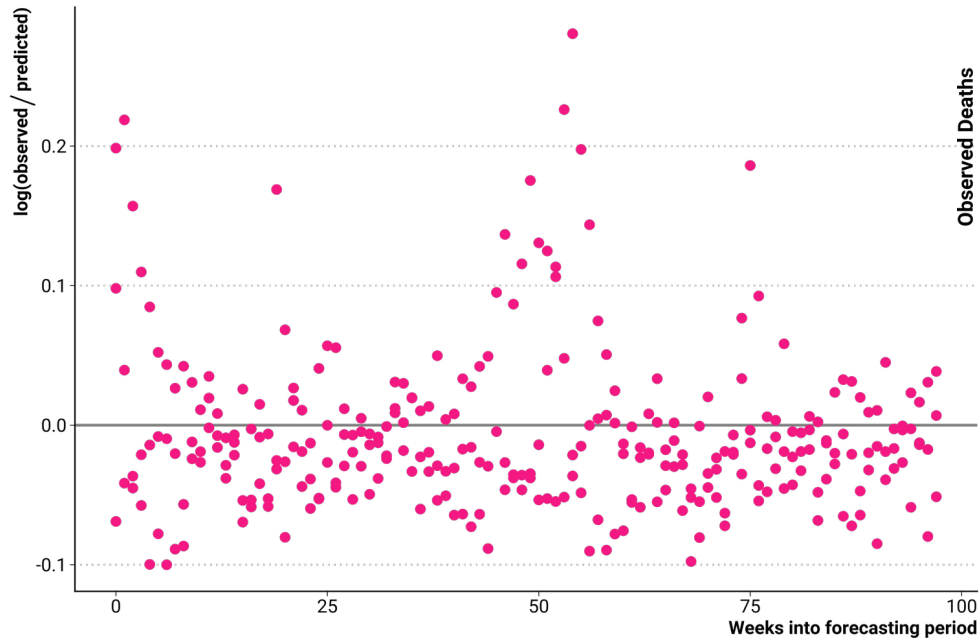
Empirical prediction intervals

The **distribution of prediction errors** of your model is estimated from out of sample validation on known data.

The logic of out-of-sample validation

Empirical prediction intervals

The **distribution of prediction errors** of your model is estimated from out of sample validation on known data.



The logic of out-of-sample validation

Cross-validation

The **prediction error** of your model is estimated from out of sample validation on known data.

Empirical prediction intervals

The **distribution of prediction errors** of your model is estimated from out of sample validation on known data.

You're model is only as precise as has been demonstrated on a comparable forecasting challenge.

Some history

1971 Williams & Goodman

A Simple Method for the Construction of Empirical Confidence Limits for Economic Forecasts

Journal of the American Statistical Association

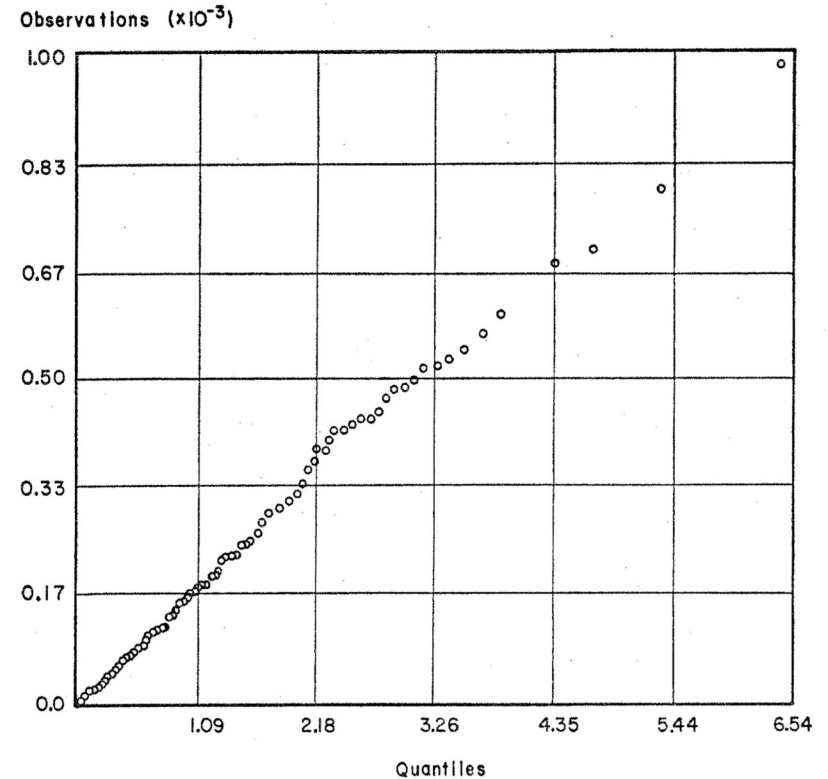
doi.org/10.2307/2284223

$$y_t = \gamma_1 t + \gamma_2 t^2 + \sum_{i=1}^{12} \beta_i m_{it} + e_t$$

A simple method for the construction of empirical confidence intervals for time series forecasts is described. The procedure is to go through the series making a forecast from each point in time. **The comparison of these forecasts with the known actual observations will yield an empirical distribution of forecasting errors. This distribution can then be used to set confidence intervals for subsequent forecasts.**

The technique appears to be particularly useful when the mechanism generating the series cannot be fully identified from the available data or when limits based on more standard considerations are difficult to obtain.

GAMMA PLOT OF ACTUAL FORECAST ERRORS, N=127
GRAND RAPIDS BUSINESS TELEPHONES



Some history

1983 Stoto

The Accuracy of Population Projections

Journal of the American Statistical Association

doi.org/10.1080/01621459.1983.10477916

$$P_T = P_0 \exp \int_0^T r(t) dt .$$

The average growth rate over the projection period is

$$\bar{r} = \frac{1}{T} \int_0^T r(t) dt ,$$

Population projections are key elements of many planning and policy studies but are inherently inaccurate. **This study of past population projection errors provides a means for constructing confidence intervals for future projections.** We first define a statistic to measure pro-

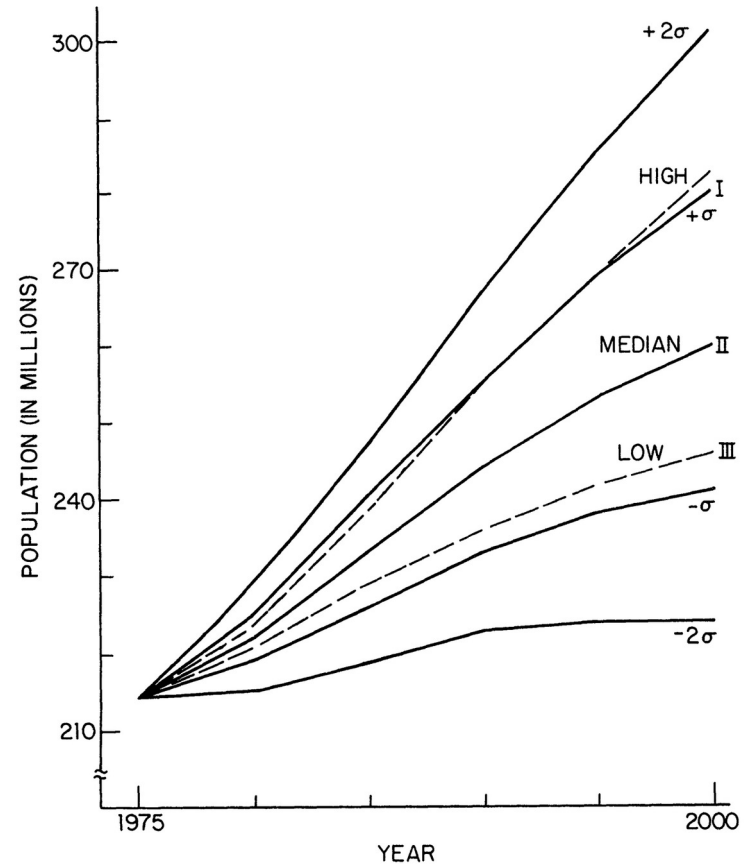


Figure 4. U.S. CENSUS PROJECTIONS (HIGH, MEDIAN, LOW) AND $\sigma = .3$ CONFIDENCE INTERVALS

Some history

1986 Cohen

*Population forecasts and confidence intervals for Sweden:
a comparison of model-based and empirical approaches*

Demography

doi.org/10.2307/2061412

1987 Smith

*Tests of Forecast Accuracy and Bias for County
Population Projections*

Journal of the American Statistical Association

doi.org/10.1080/01621459.1987.10478528

1987 Mandrakis & Hibon

*Confidence Intervals. An Empirical Investigation of the
Series in the M-Competition*

International Journal of Forecasting

[doi.org/10.1016/0169-2070\(87\)90045-8](https://doi.org/10.1016/0169-2070(87)90045-8)

1988 Smith

*Stability over time in the distribution of
population forecast errors*

Demography

doi.org/10.2307/2061544

Some history

2008 Shafer & Vovk

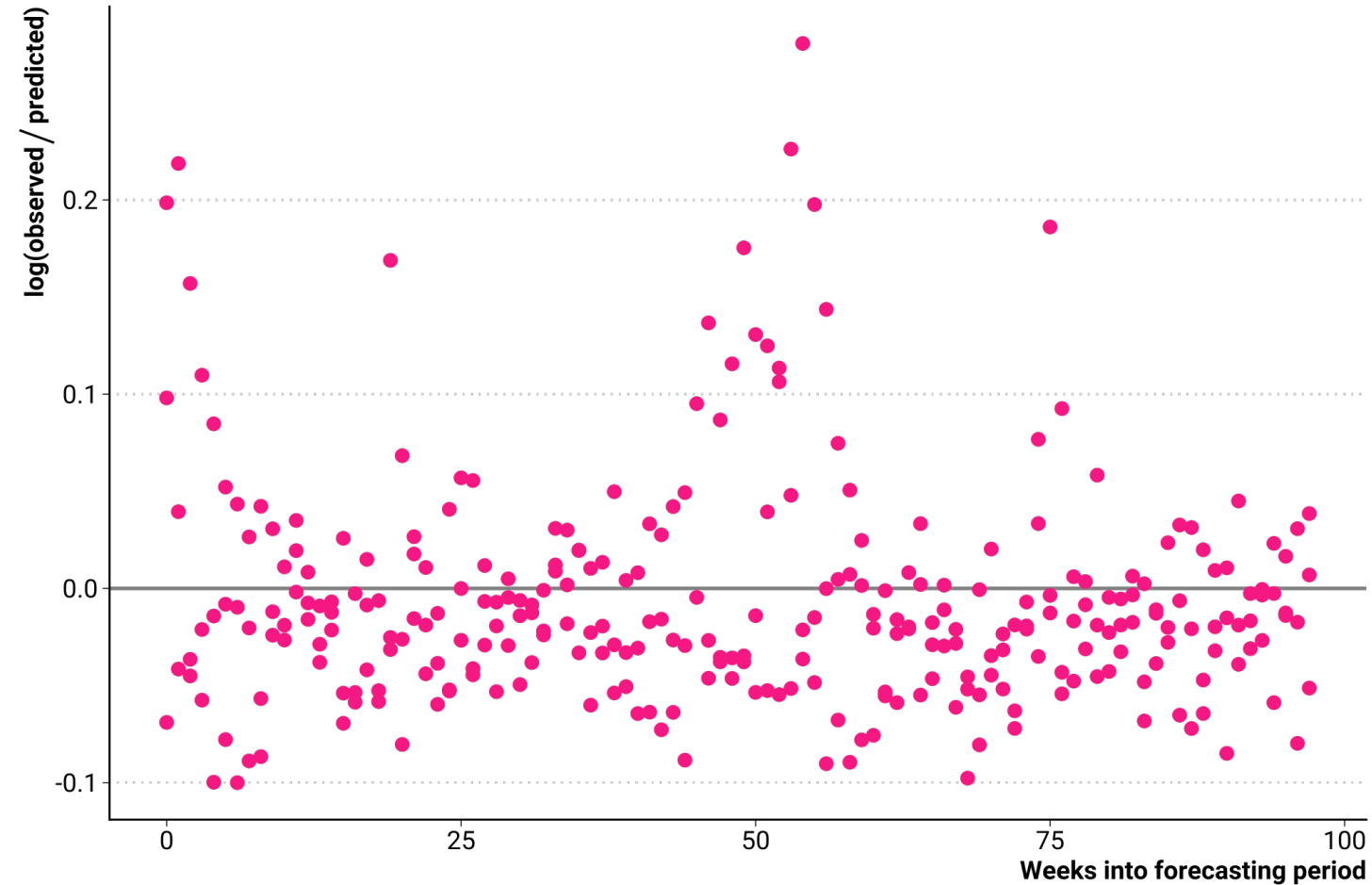
A Tutorial on Conformal Prediction

Journal of Machine Learning Research

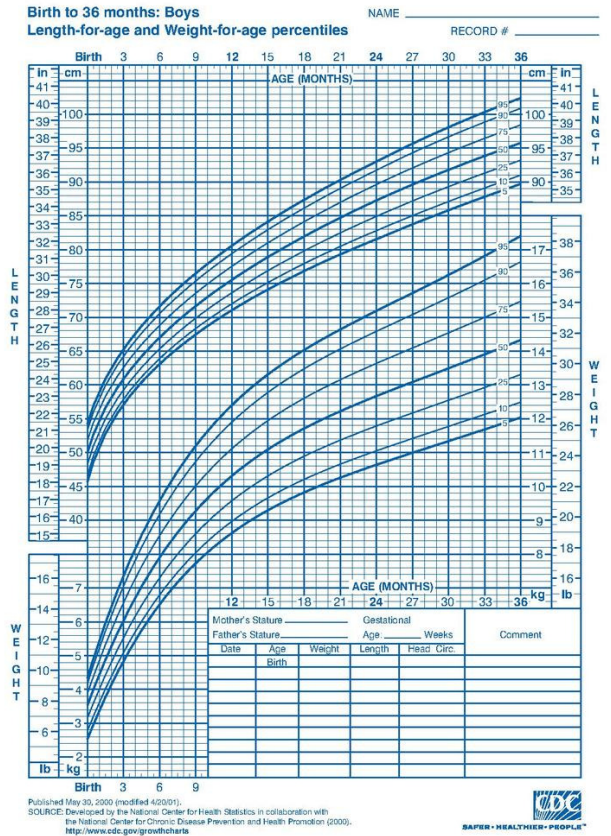
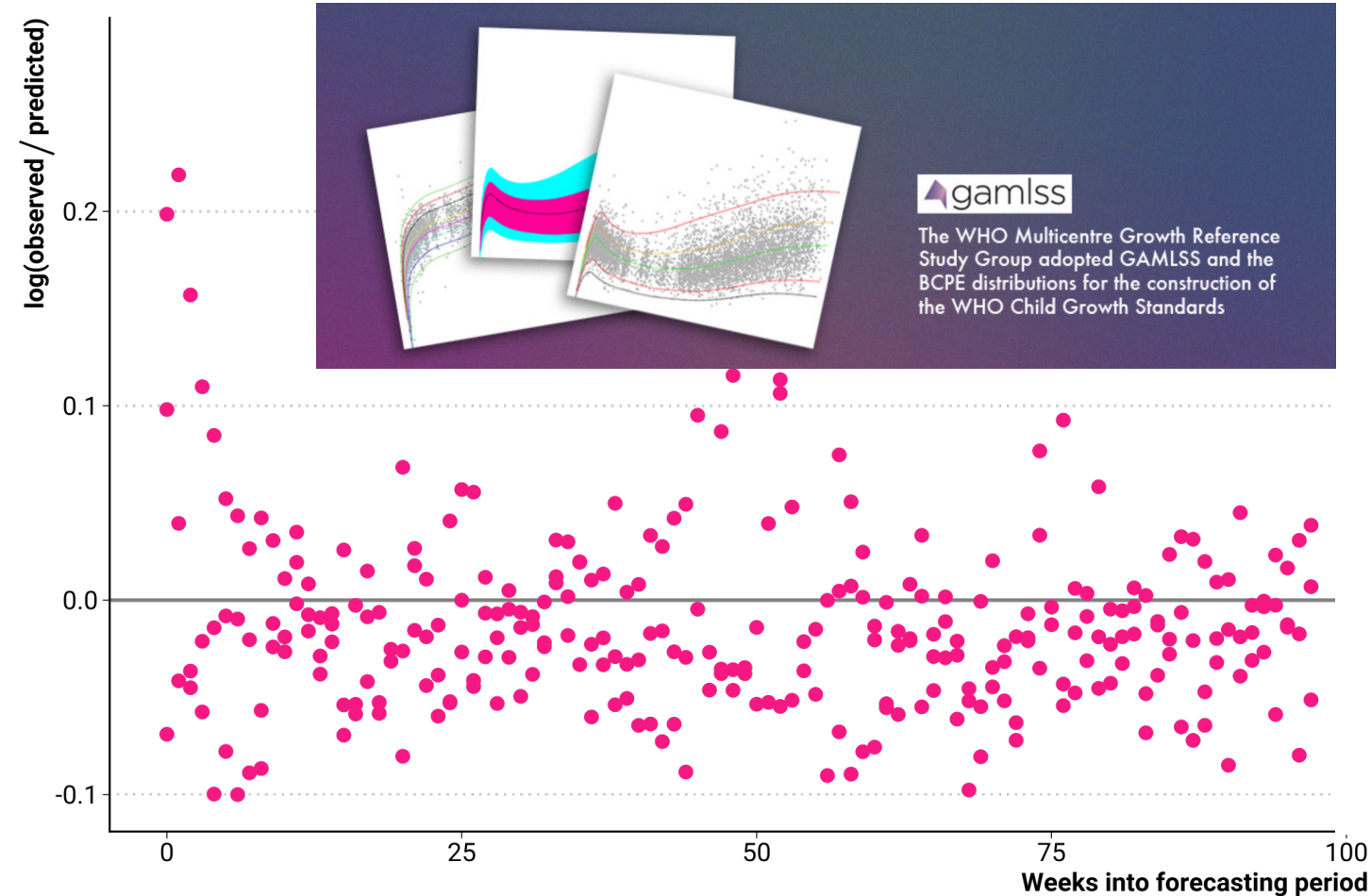
doi.org/10.5555/1390681.1390693

“Conformal prediction uses past experience to determine precise levels of confidence in new predictions.”

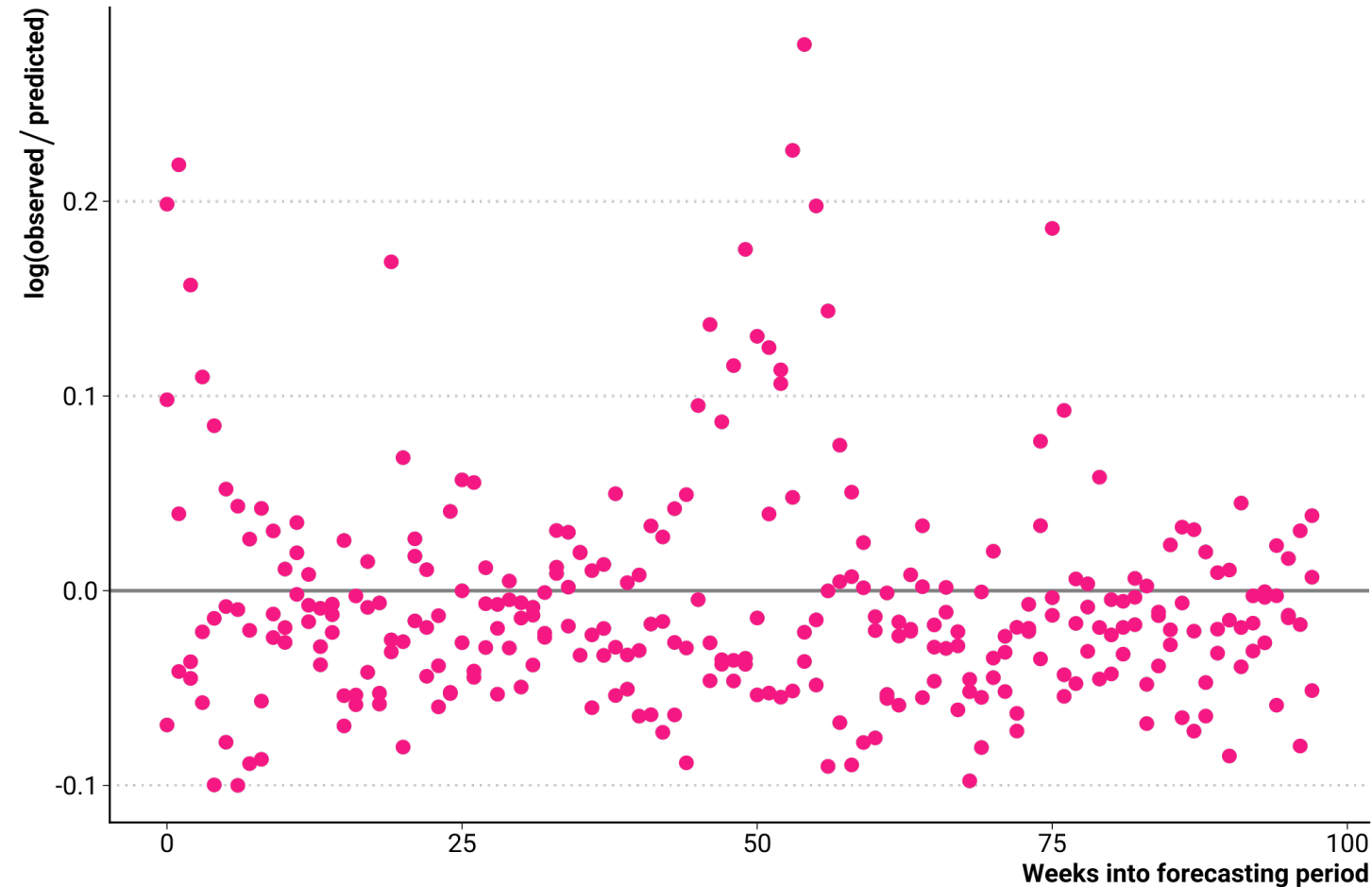
Modeling the distribution of empirical errors



Modeling the distribution of empirical errors



Modeling the distribution of empirical errors



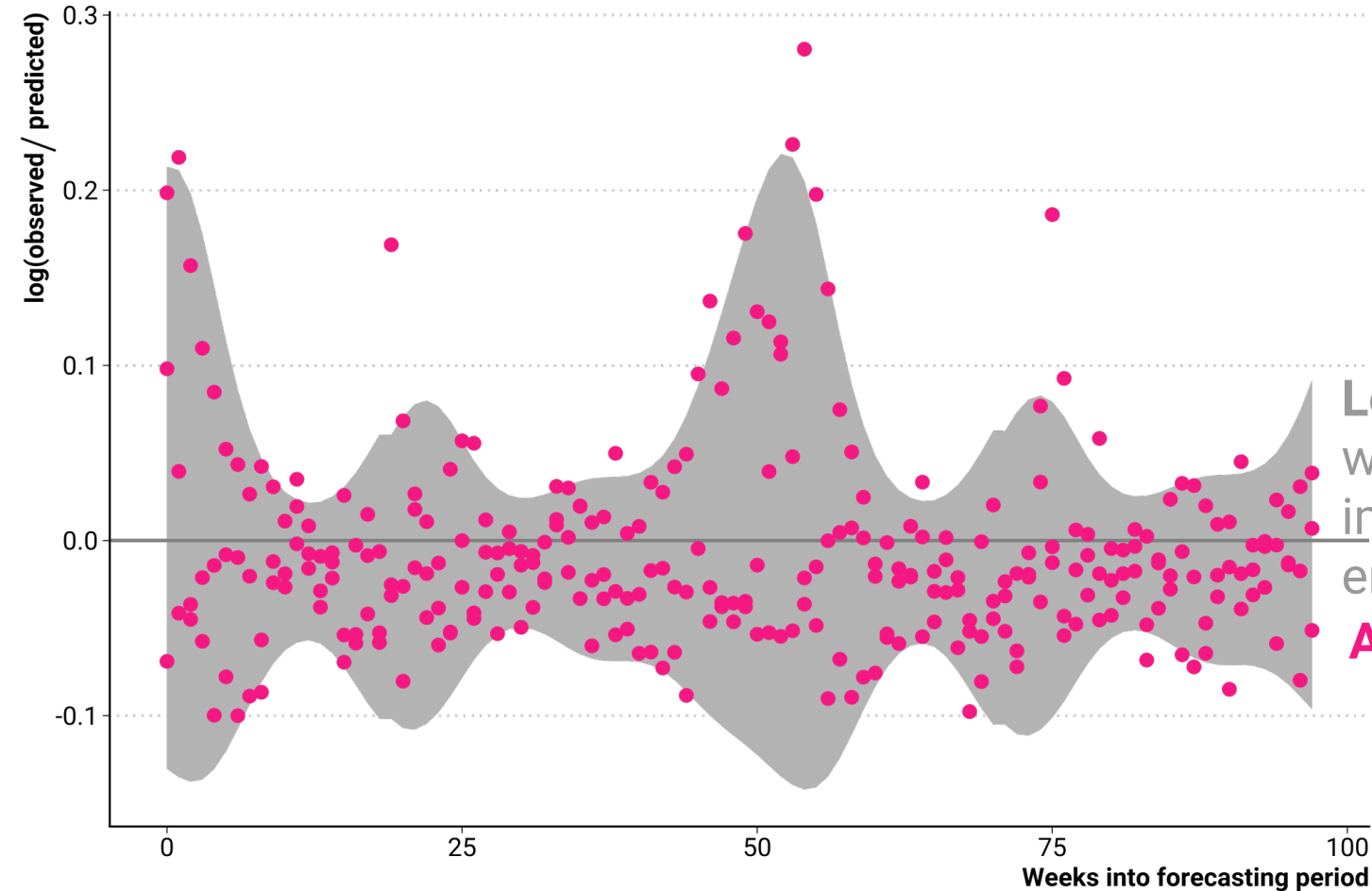
$\log(\text{observed}/\text{predicted}) \sim$
Skewed-Normal(μ_t, σ_t, ν_t)

$$\mu_t = b_{\mu 0}$$

$$\sigma_t = \exp(b_{\sigma 0} + b_{\sigma t} t + s_{\sigma}(t))$$

$$\nu_t = b_{\nu 0} + s_{\nu}(t)$$

Modeling the distribution of empirical errors



$\log(\text{observed}/\text{predicted}) \sim$
Skewed-Normal(μ_t, σ_t, ν_t)

$$\mu_t = b_{\mu 0}$$

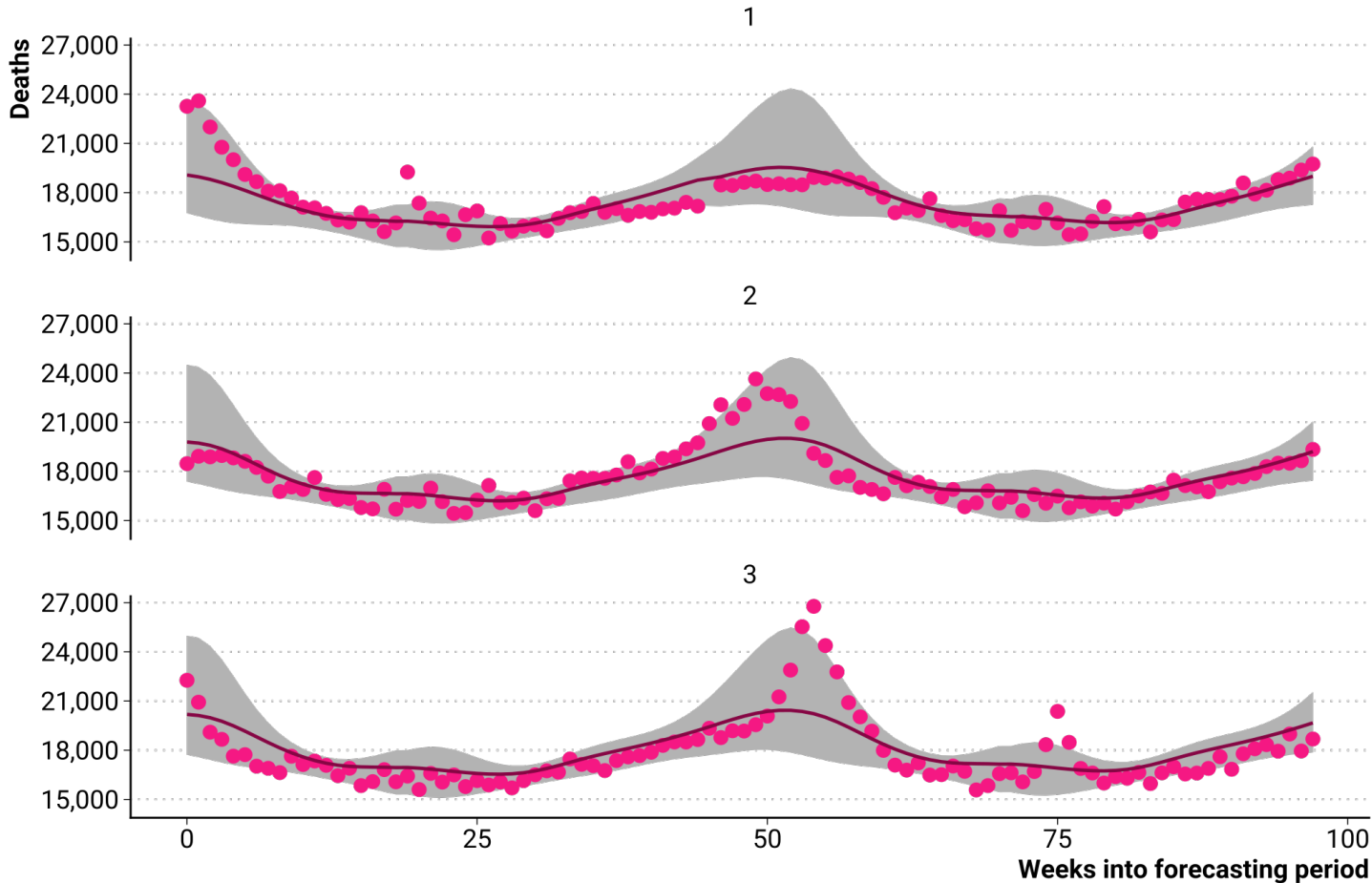
$$\sigma_t = \exp(b_{\sigma 0} + b_{\sigma t} t + s_{\sigma}(t))$$

$$\nu_t = b_{\nu 0} + s_{\nu}(t)$$

Log prediction error region
with 90% nominal coverage
implied by Skewed-Normal
empirical prediction intervals

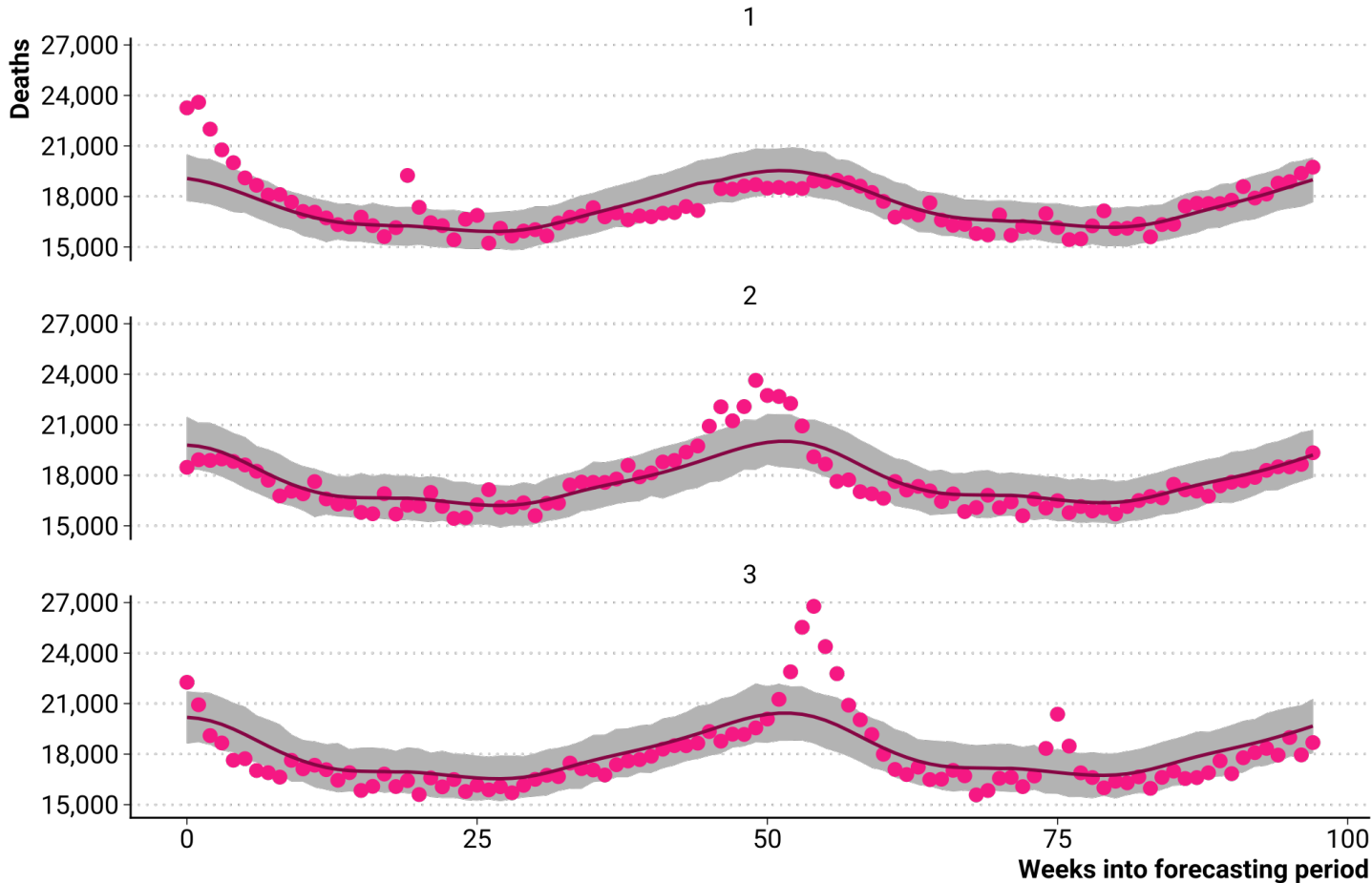
Actual coverage ~92%

Modeling the distribution of empirical errors



Prediction interval with
90% nominal coverage
via Skewed-Normal
empirical error model
Actual coverage ~92%

Modeling the distribution of empirical errors



Prediction interval with
90% nominal coverage
via Negative-Bin.
GAM model

Actual coverage ~87%

Evaluating the prediction interval

6.2 Interval Score

Interval forecasts form a crucial special case of quantile prediction. We consider the classical case of the central $(1 - \alpha) \times 100\%$ prediction interval, with lower and upper endpoints that are the predictive quantiles at level $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$. We denote a scoring rule for the associated interval forecast by $S_\alpha(l, u; x)$, where l and u represent for the quoted $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles. Thus, if the forecaster quotes the $(1 - \alpha) \times 100\%$ central prediction interval $[l, u]$ and x materializes, then his or her score will be $S_\alpha(l, u; x)$. Putting $\alpha_1 = \frac{\alpha}{2}$, $\alpha_2 = 1 - \frac{\alpha}{2}$, $s_1(x) = s_2(x) = 2\frac{x}{\alpha}$, and $h(x) = -2\frac{x}{\alpha}$ in (42), and reversing the sign of the scoring rule, yields the negatively oriented *interval score*,

$$\begin{aligned} S_\alpha^{\text{int}}(l, u; x) \\ = (u - l) + \frac{2}{\alpha}(l - x)\mathbb{1}\{x < l\} + \frac{2}{\alpha}(x - u)\mathbb{1}\{x > u\}. \quad (43) \end{aligned}$$

Gneiting & Raftery (2007). Strictly Proper Scoring Rules. [10.1198/016214506000001437](https://doi.org/10.1198/016214506000001437)

Evaluating the prediction interval

6.2 Interval Score

Interval forecasts form a crucial special case of quantile prediction. We consider the classical case of the central $(1 - \alpha) \times 100\%$ prediction interval, with lower and upper endpoints that are the predictive quantiles at level $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$. We denote a scoring rule for the associated interval forecast by $S_\alpha(l, u; x)$, where l and u represent for the quoted $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles. Thus, if the forecaster quotes the $(1 - \alpha) \times 100\%$ central prediction interval $[l, u]$ and x materializes, then his or her score will be $S_\alpha(l, u; x)$. Putting $\alpha_1 = \frac{\alpha}{2}$, $\alpha_2 = 1 - \frac{\alpha}{2}$, $s_1(x) = s_2(x) = 2\frac{x}{\alpha}$, and $h(x) = -2\frac{x}{\alpha}$ in (42), and reversing the sign of the scoring rule, yields the negatively oriented *interval score*,

$$\begin{aligned} S_\alpha^{\text{int}}(l, u; x) \\ = (u - l) + \frac{2}{\alpha}(l - x)\mathbb{1}\{x < l\} + \frac{2}{\alpha}(x - u)\mathbb{1}\{x > u\}. \quad (43) \end{aligned}$$

Prediction interval with
90% nominal coverage

Coverage NB ~87%
Interval Score NB ~5979

Gneiting & Raftery (2007). Strictly Proper Scoring Rules. [10.1198/016214506000001437](https://doi.org/10.1198/016214506000001437)

Evaluating the prediction interval

6.2 Interval Score

Interval forecasts form a crucial special case of quantile prediction. We consider the classical case of the central $(1 - \alpha) \times 100\%$ prediction interval, with lower and upper endpoints that are the predictive quantiles at level $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$. We denote a scoring rule for the associated interval forecast by $S_\alpha(l, u; x)$, where l and u represent for the quoted $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles. Thus, if the forecaster quotes the $(1 - \alpha) \times 100\%$ central prediction interval $[l, u]$ and x materializes, then his or her score will be $S_\alpha(l, u; x)$. Putting $\alpha_1 = \frac{\alpha}{2}$, $\alpha_2 = 1 - \frac{\alpha}{2}$, $s_1(x) = s_2(x) = 2\frac{x}{\alpha}$, and $h(x) = -2\frac{x}{\alpha}$ in (42), and reversing the sign of the scoring rule, yields the negatively oriented *interval score*,

$$S_\alpha^{\text{int}}(l, u; x) \\ = (u - l) + \frac{2}{\alpha}(l - x)\mathbb{1}\{x < l\} + \frac{2}{\alpha}(x - u)\mathbb{1}\{x > u\}. \quad (43)$$

Prediction interval with
90% nominal coverage

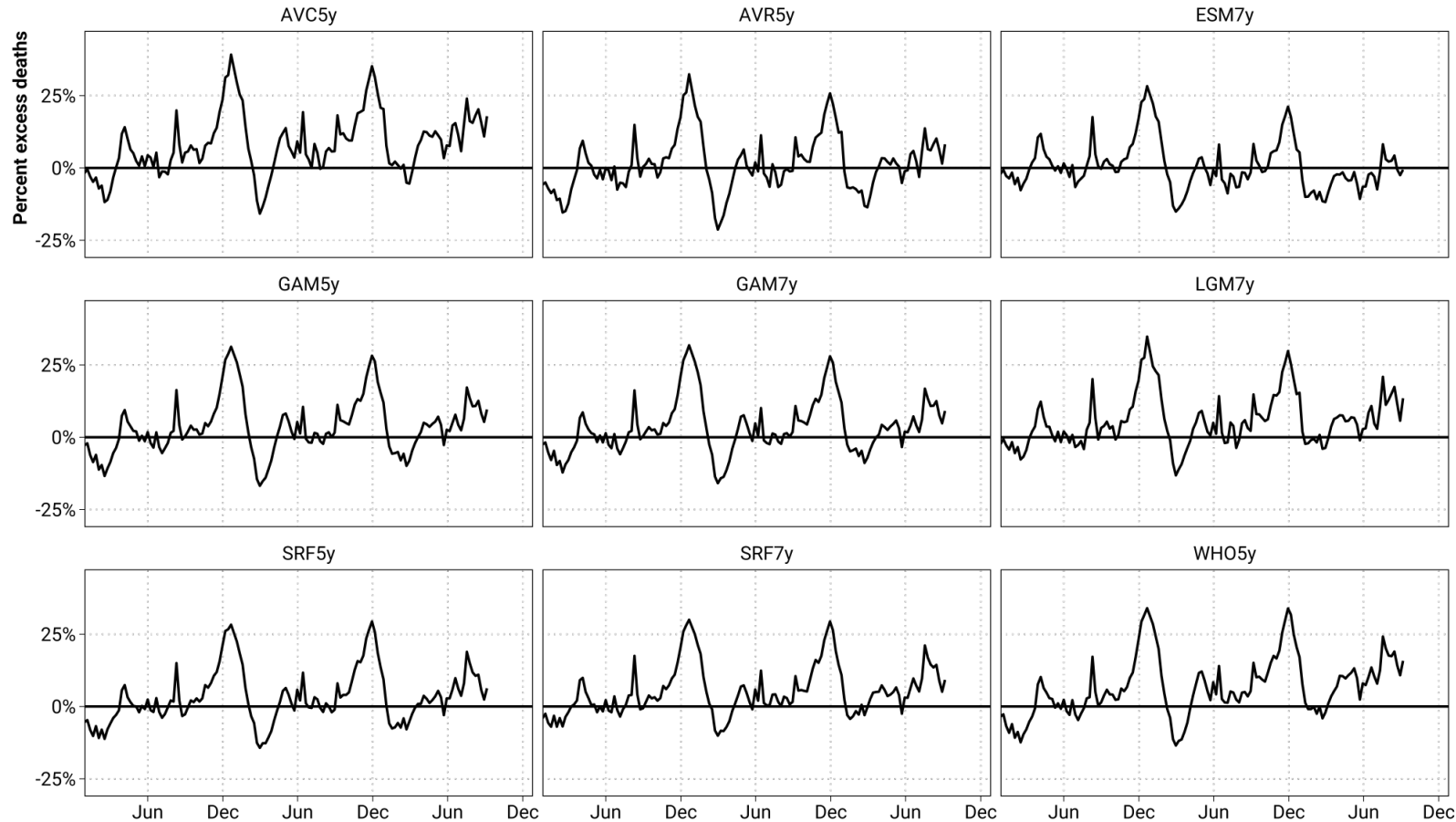
Coverage NB ~87%
Interval Score NB ~5979

Coverage SNO ~92%
Interval Score SNO ~4815

Gneiting & Raftery (2007). Strictly Proper Scoring Rules. [10.1198/016214506000001437](https://doi.org/10.1198/016214506000001437)

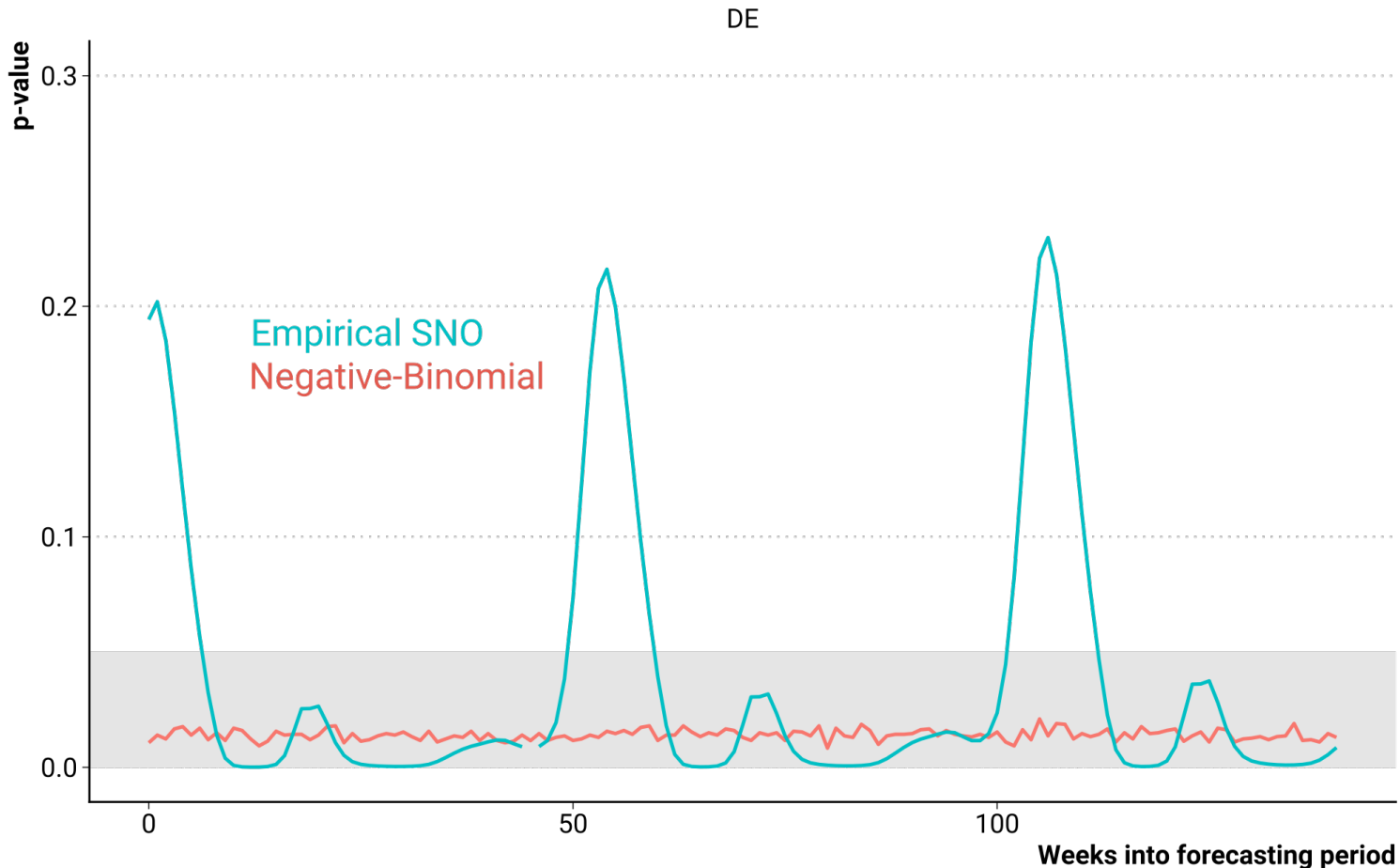
Excess deaths as a forecasting challenge

Percent excess deaths Germany 2020w1 through 2022 under different baseline models



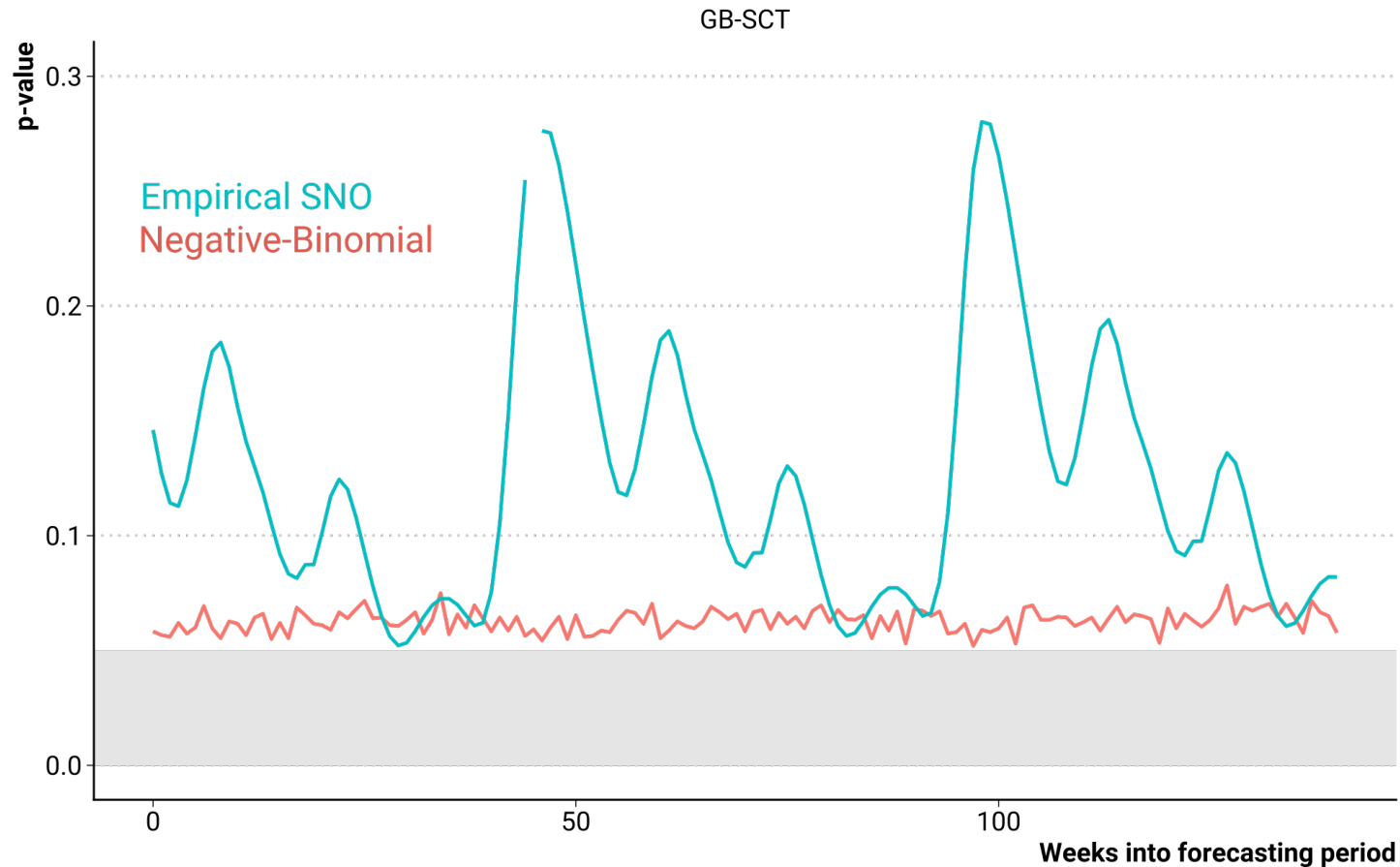
Excess deaths as a forecasting challenge

p-value of 10% excess deaths given H0: "continuation of past trends"



Excess deaths as a forecasting challenge

p-value of 10% excess deaths given H0: "continuation of past trends"



Reproducible analysis

github.com/jschoeley

Jonas Schöley

 @jschoeley

 0000-0002-3340-8518

 schoeley@demogr.mpg.de



MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH