

Counterfactual time-series prediction with encoder-decoder networks (Online Appendix)

March 26, 2018

Contents

1	RNNs architecture and training	1
2	Estimates on simulated data	4
3	Estimates on Basque Country data	8
4	Estimates on California data	13
5	Estimates on West Germany data	18
6	Estimates on mayoral elections data	23

1 RNNs architecture and training

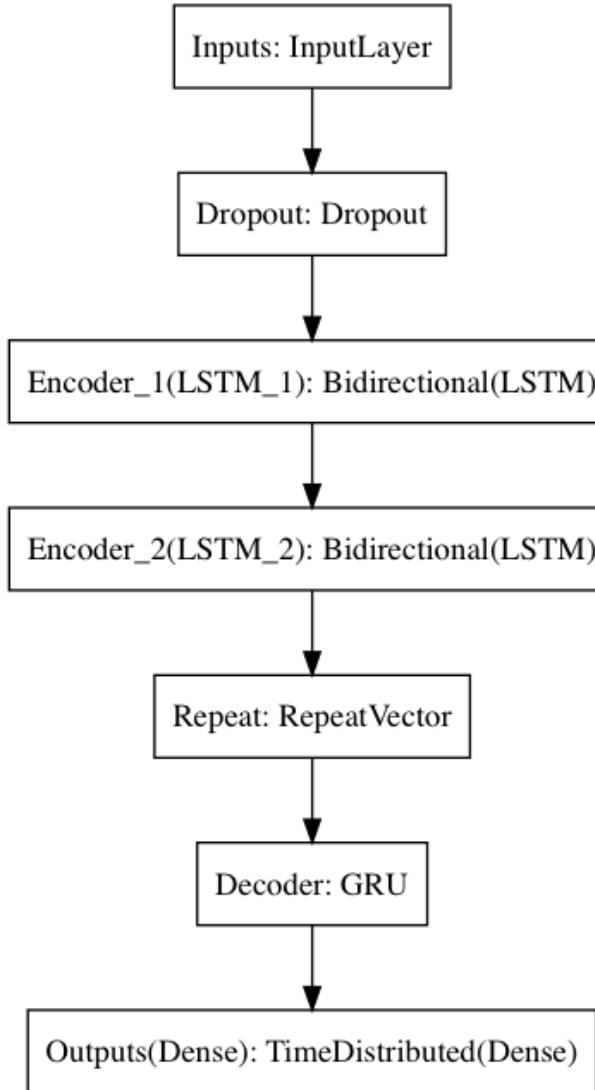


Figure 1: Encoder-decoder networks architecture. Dropout is applied to the visible input sequences, which are then fed to a two-layer bidirectional LSTM encoder. The encoder encodes the input sequences into a single vector that contains information about the entire sequence. The output of the encoder is repeated t times and fed to the single-layer GRU decoder, which translates the encoded sequence into the predicted sequence. Finally, a dense layer is applied to the decoder output to generate predictions.

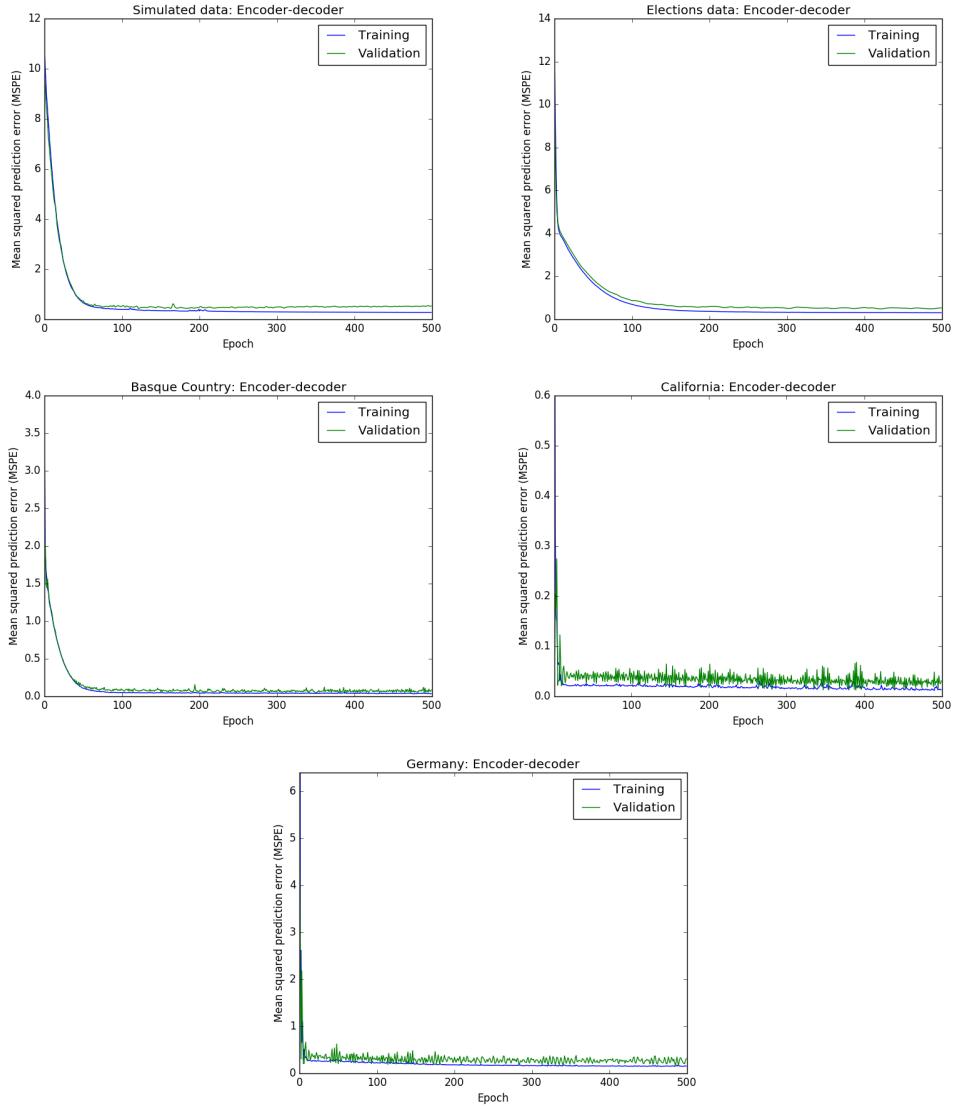


Figure 2: Evolution of encoder-decoder networks training and validation loss in terms of MSPE.

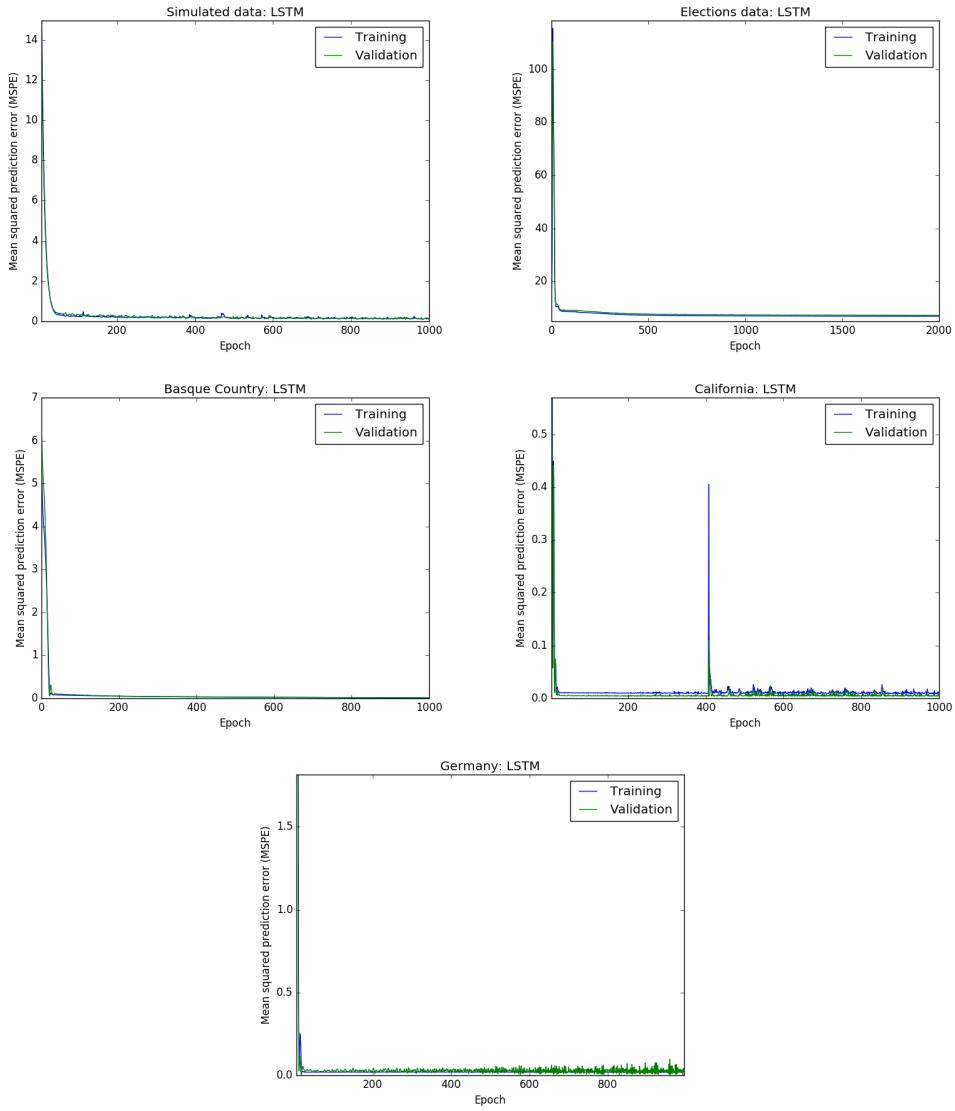


Figure 3: Evolution of LSTM training and validation loss in terms of MSPE.

2 Estimates on simulated data

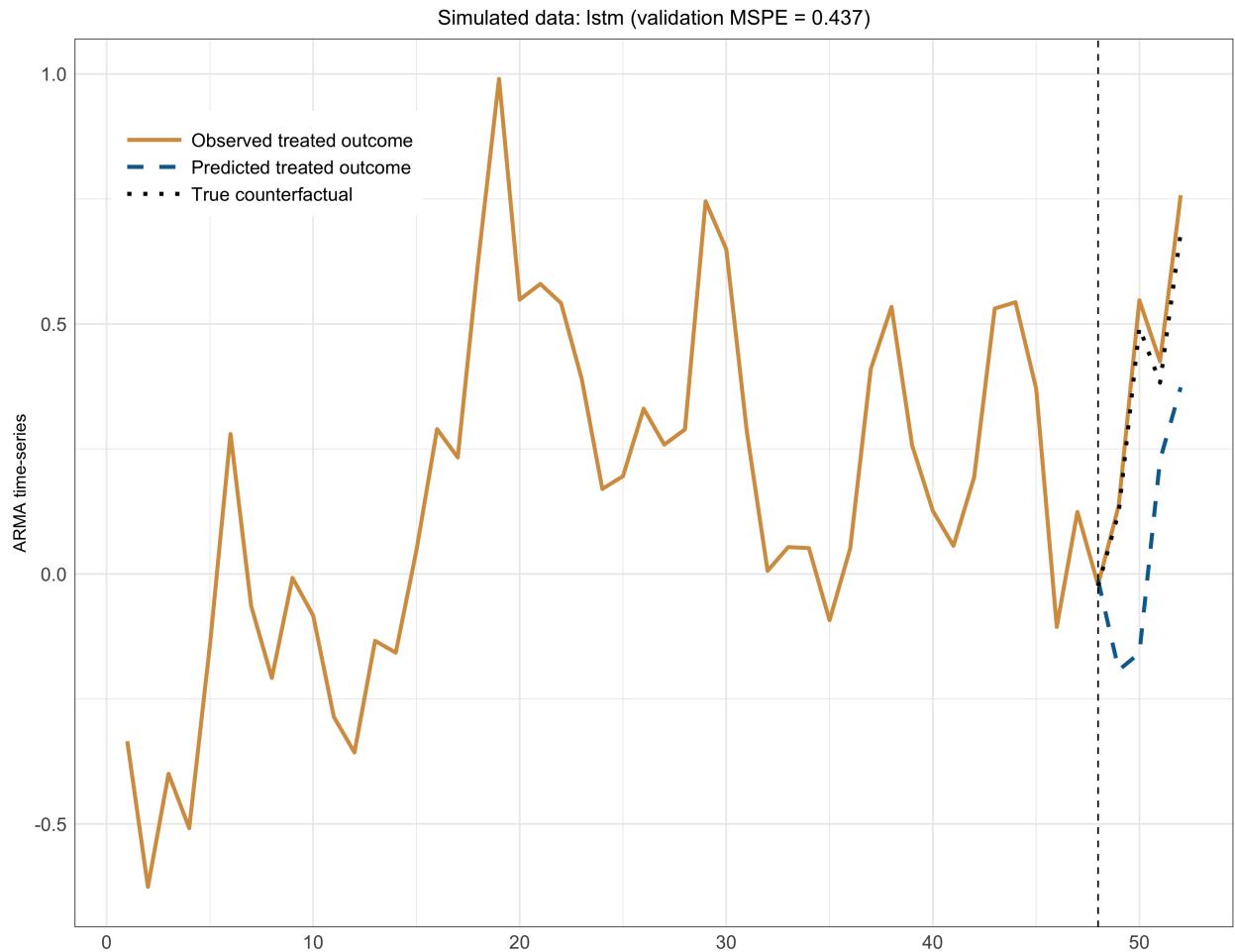


Figure 4: Observed ARMA simulated time-series and LSTM predictions. See notes to Fig. OA-1.

Simulated data: BSTS (training MSPE = 0.07)

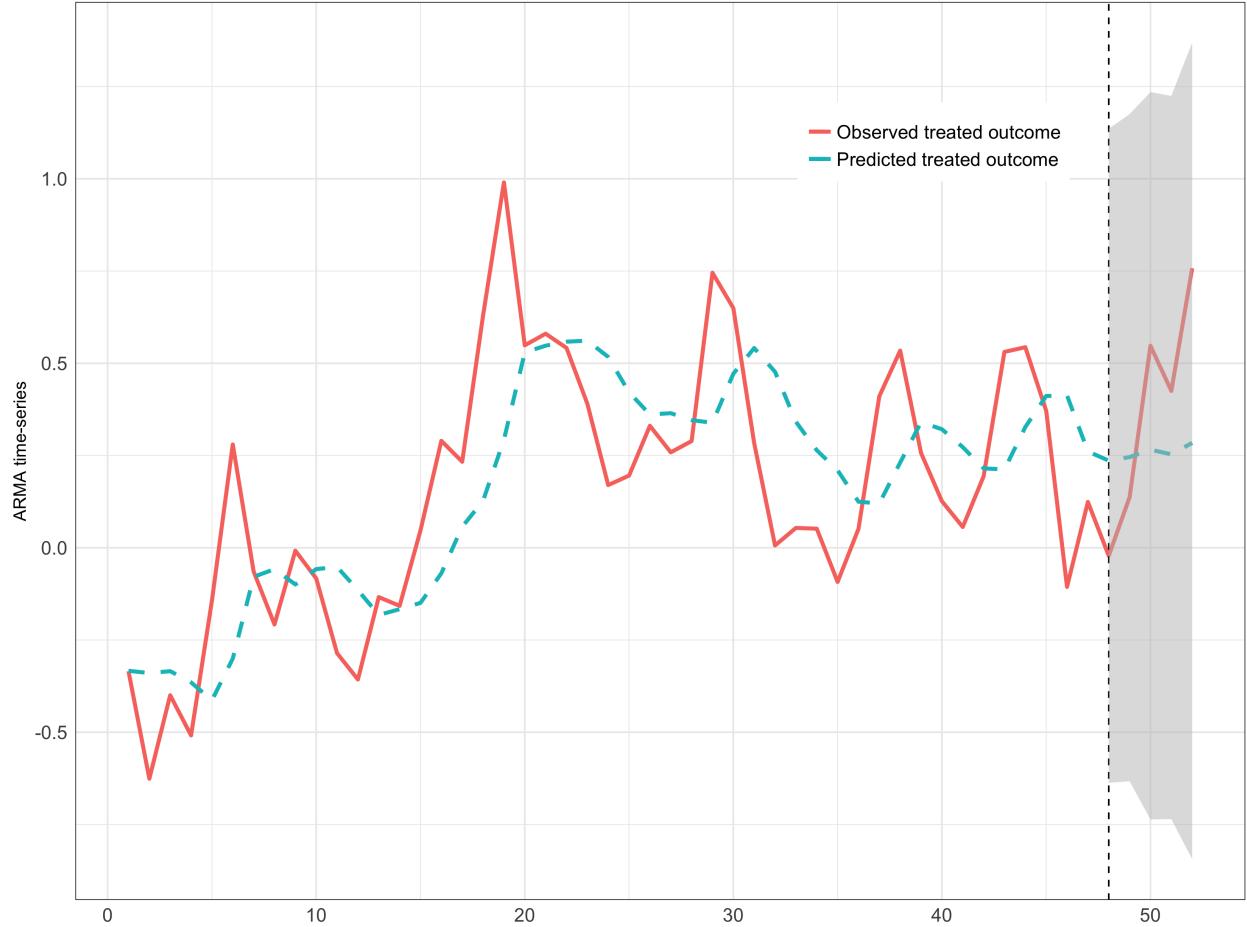


Figure 5: Observed ARMA simulated time-series and BSTS predictions. BSTS model with semilocal linear trend and seasonal components and spike-and-slab priors is trained with 10,000 MCMC samples with the first 1,000 samples discarded as burn-in. Predictions are obtained by averaging across MCMC draws and 95% credible intervals (shaded region) are obtained from the distribution of MCMC draws. The dashed vertical line indicates the start of the intervention.

Simulated data: Synthetic control (training MSPE = 0.05)

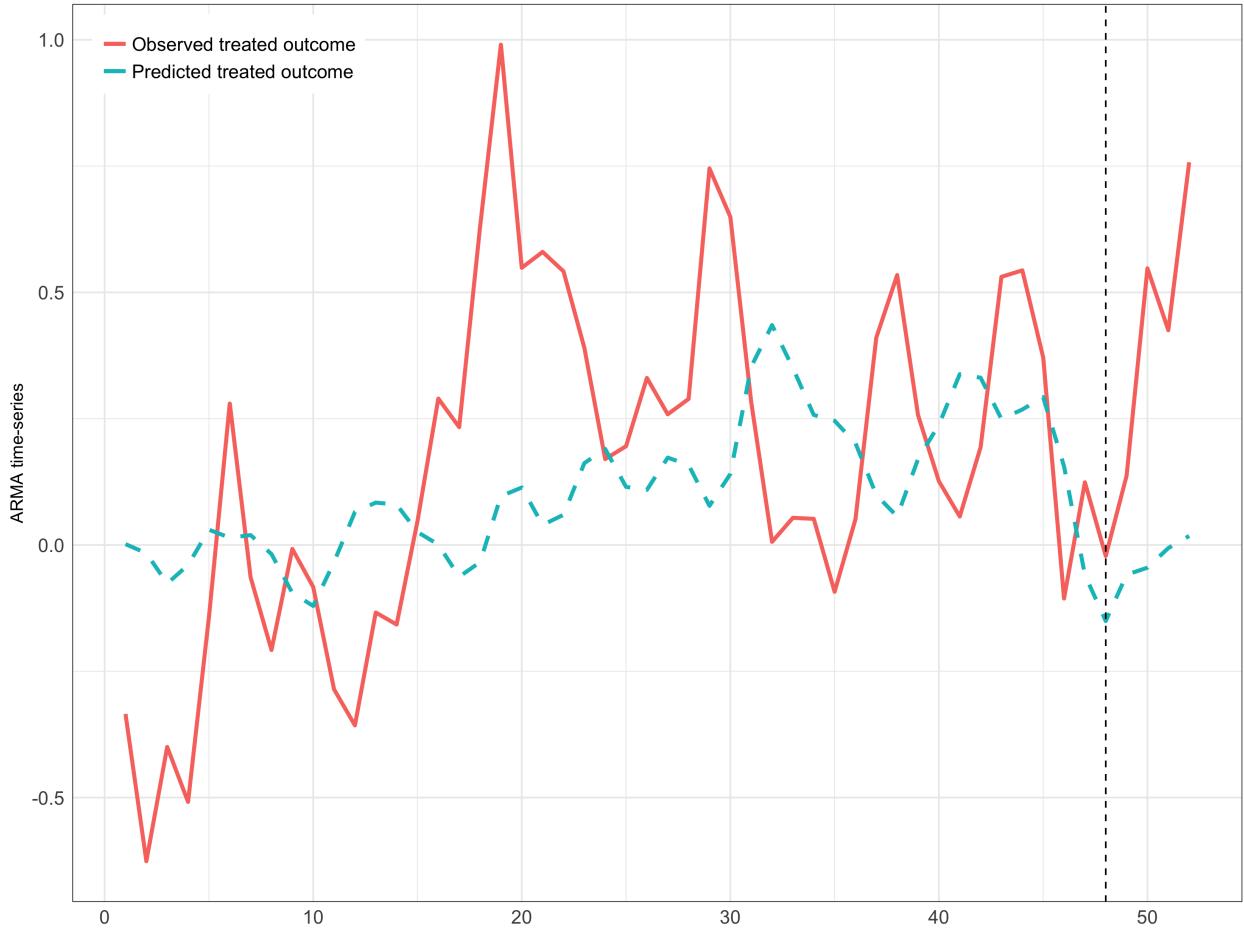


Figure 6: Observed ARMA simulated time-series and synthetic control predictions. Synthetic control is constructed using input means over various periods during the pre-period as covariates. The reported training loss is the mean squared prediction error (MSPE) over time-steps 43 to 47. The dashed vertical line indicates the start of the intervention.

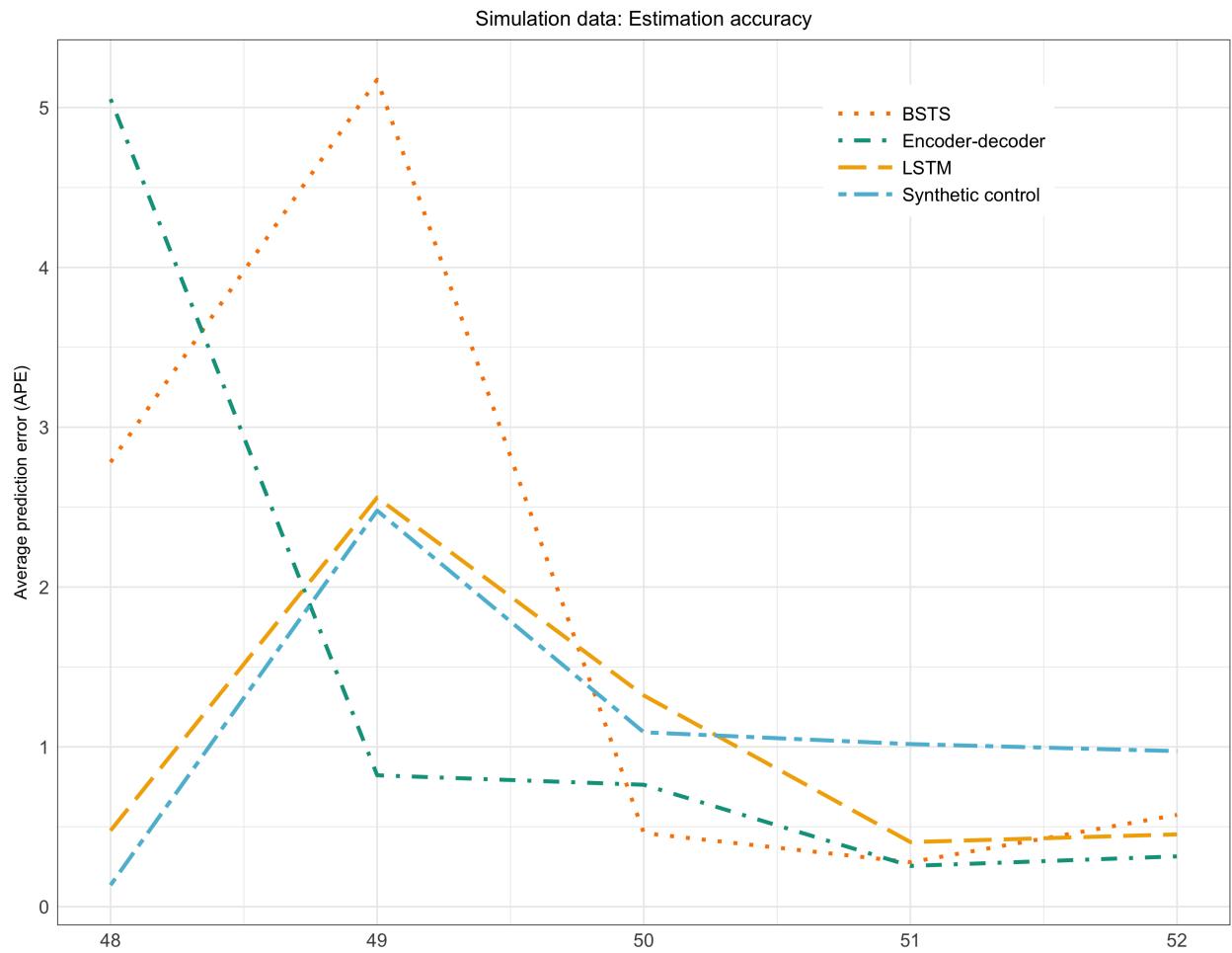


Figure 7: Estimation accuracy in ARMA simulated data in terms of absolute percentage error (APE).

3 Estimates on Basque Country data

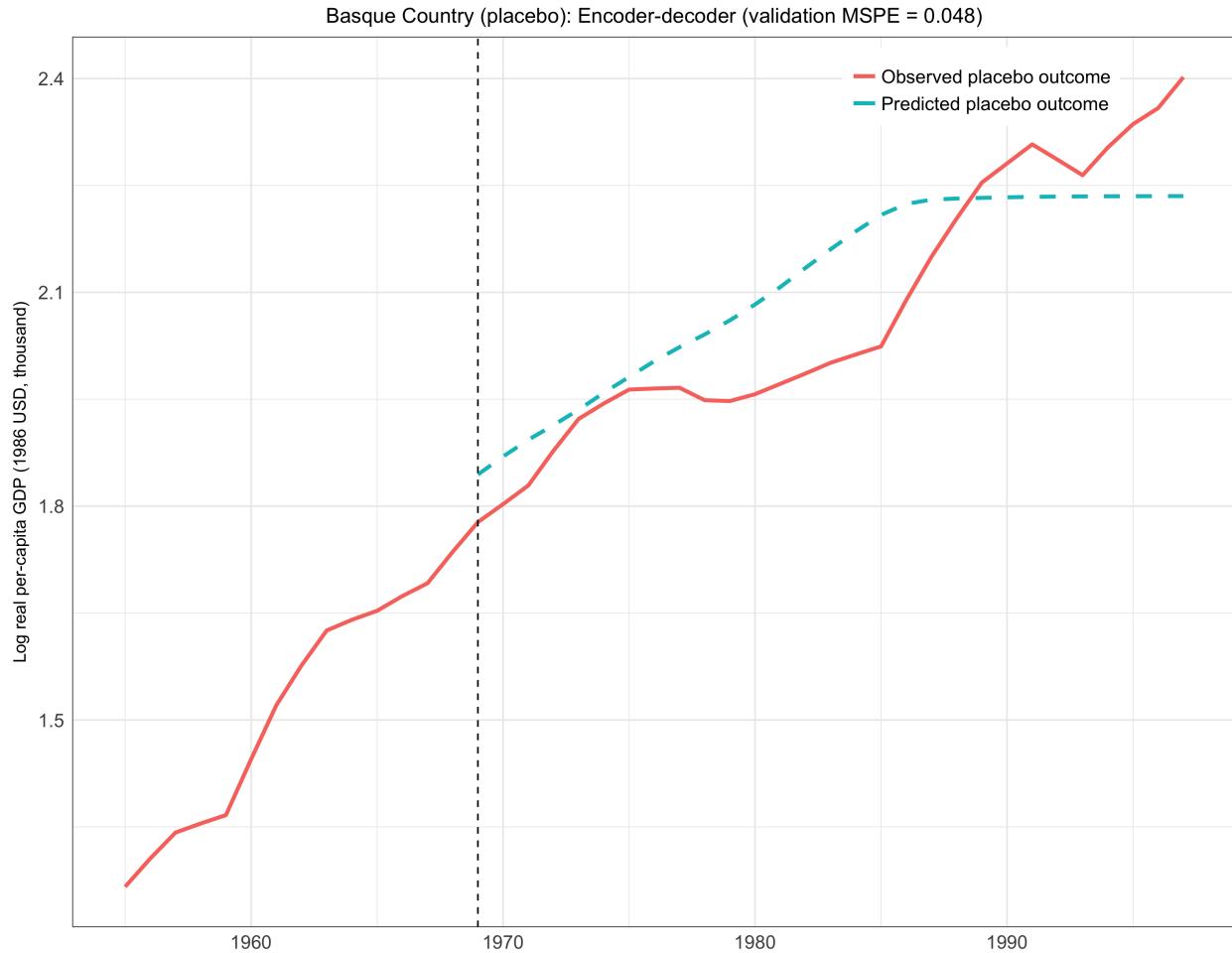


Figure 8: Observed placebo and encoder-decoder predicted outcomes in Basque Country data. Validation MSPE is calculated on a held-out sample. The dashed vertical line indicates the start of the placebo intervention.

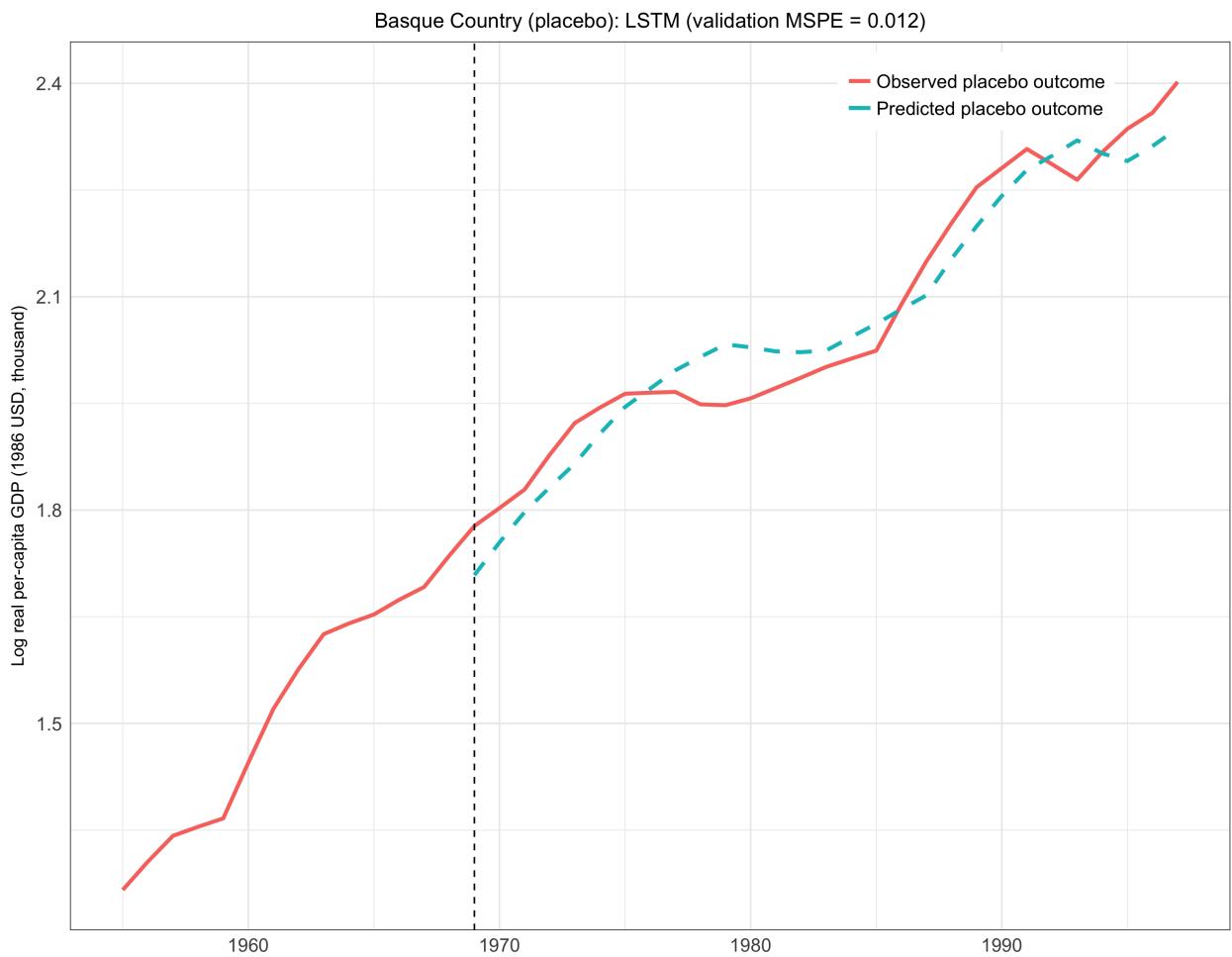


Figure 9: Observed placebo and LSTM predicted outcomes in Basque Country data. See notes to Fig. OA-8.

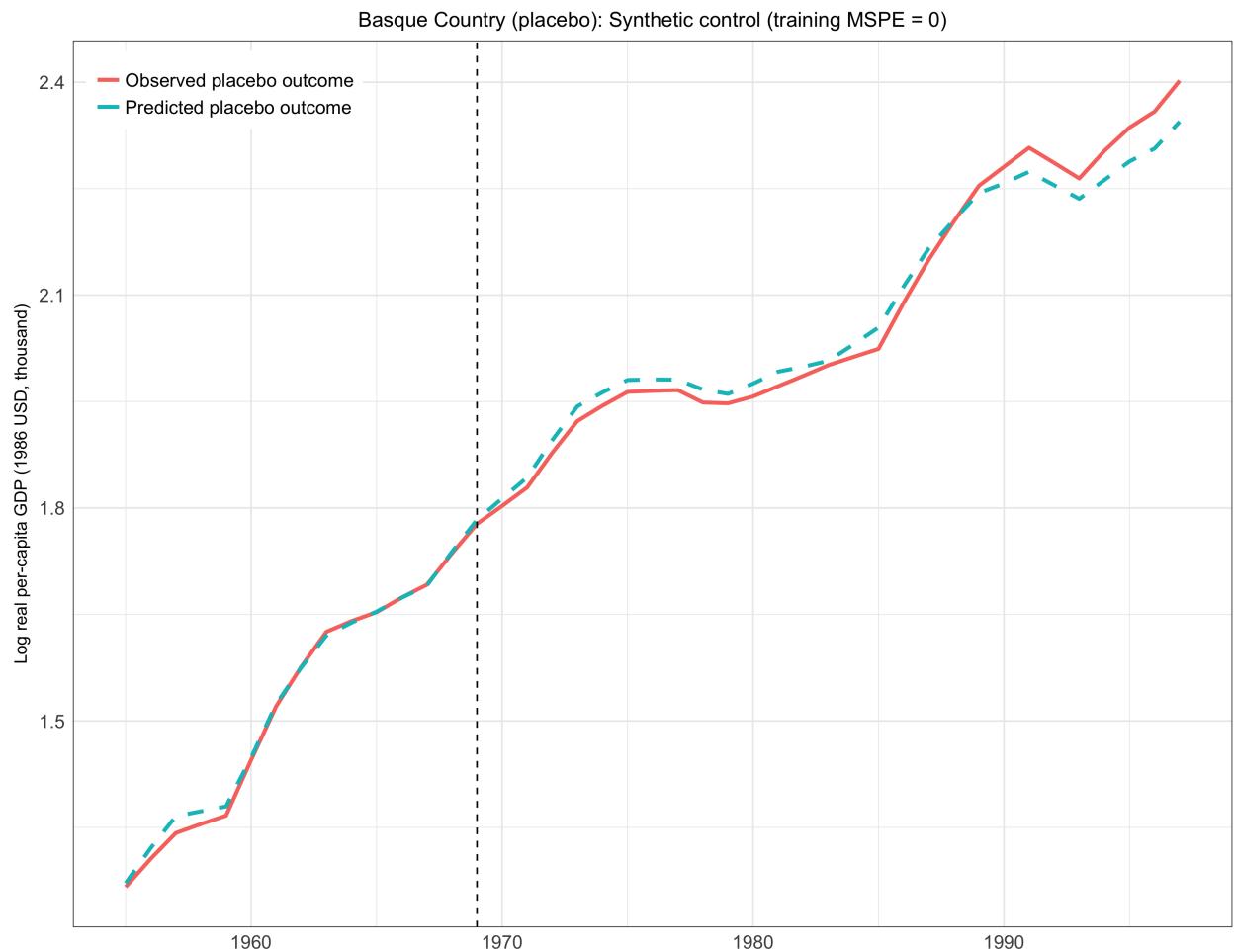


Figure 10: Observed placebo and synthetic control predicted outcomes in Basque Country data. The reported training loss is the mean squared prediction error (MSPE) over the period of 1960 to 1968. The dashed vertical line indicates the start of the placebo intervention.

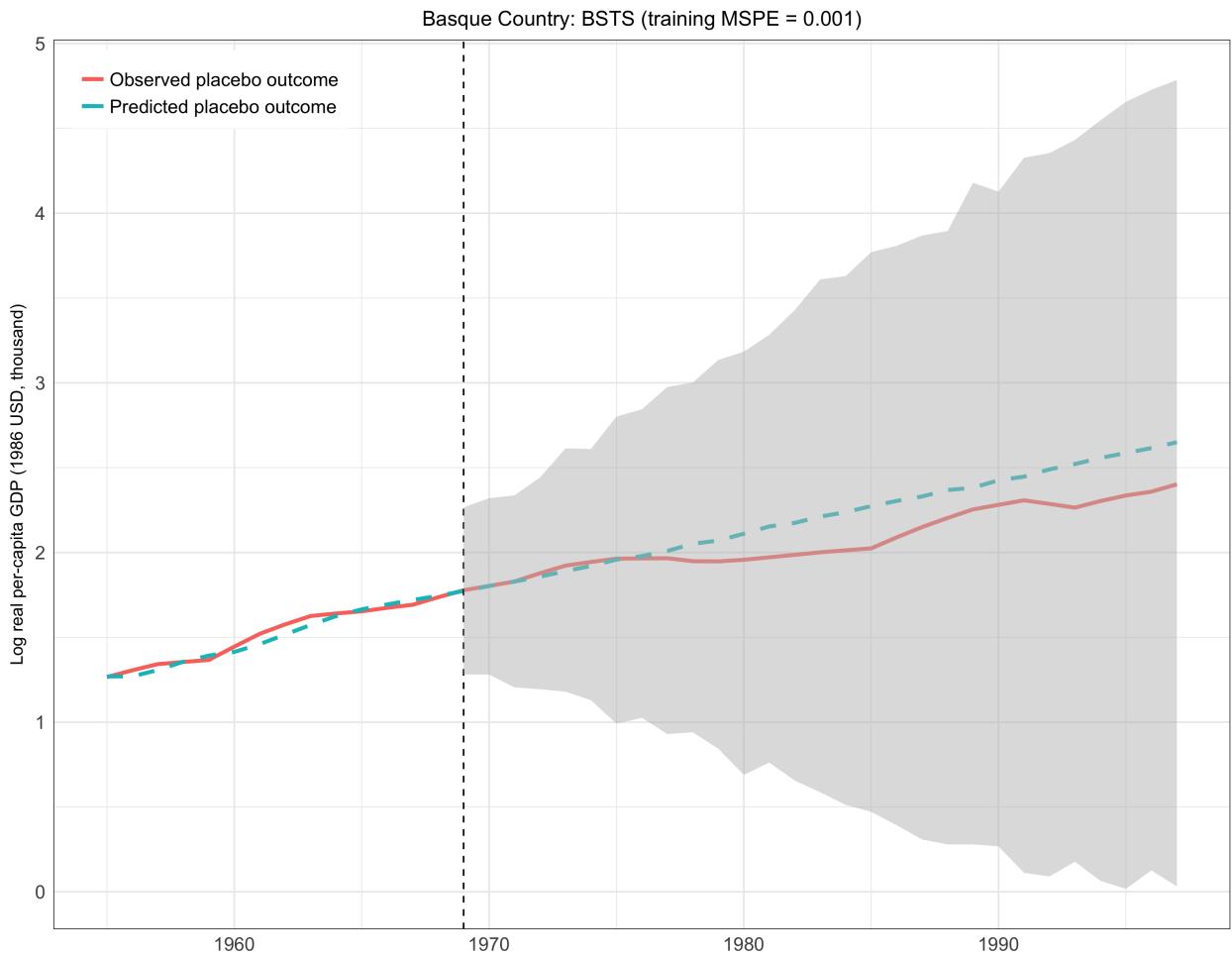


Figure 11: Observed placebo and BSTS predicted outcomes in Basque Country data. See notes to Fig. OA-5.

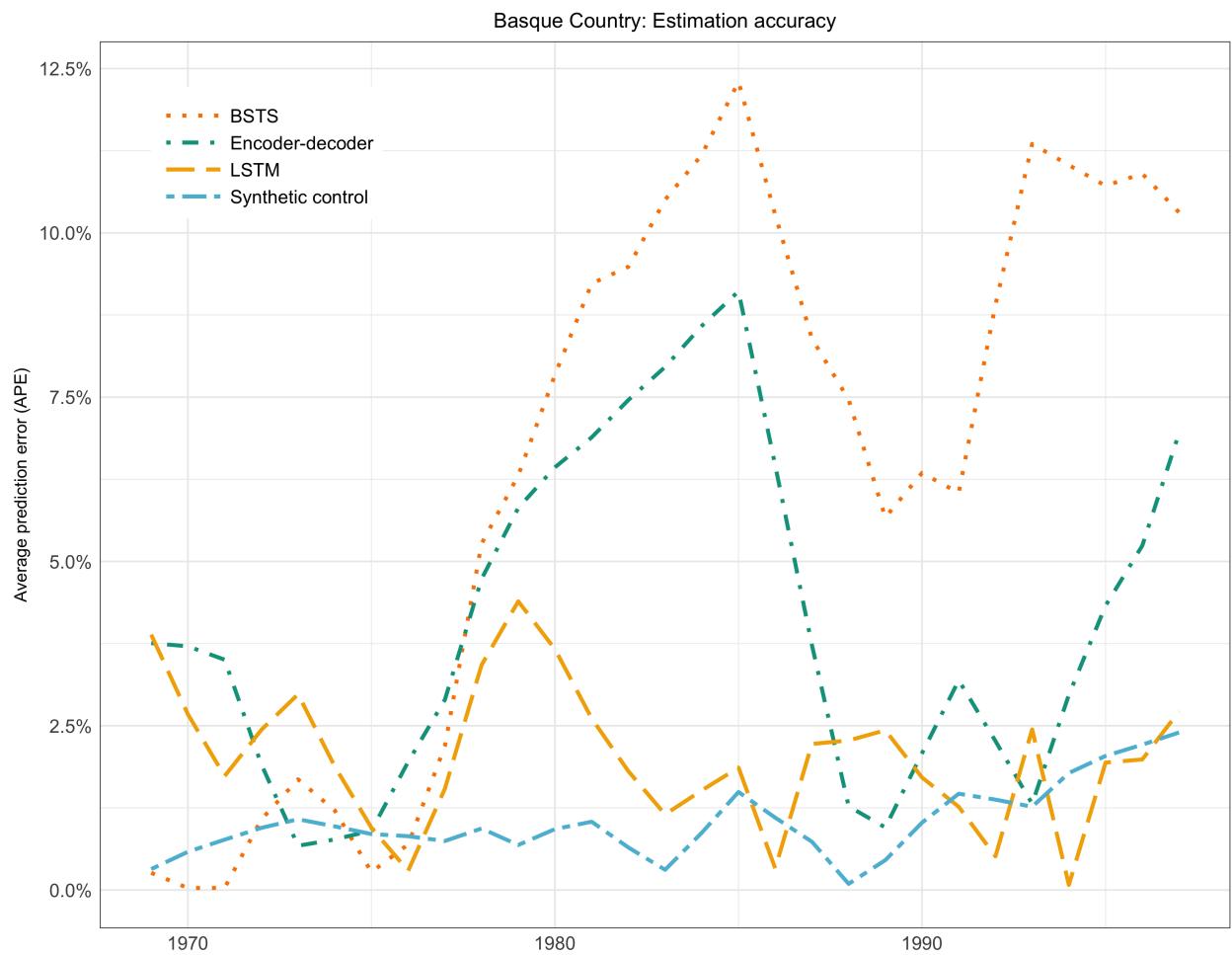


Figure 12: Estimation accuracy in terms of absolute percentage error (APE) for each post-period time-step in Basque County data. Predictions are plotted in Figs. OA-8, OA-9, OA-10, and OA-11.

4 Estimates on California data



Figure 13: Observed placebo and synthetic control predicted outcomes in California data. The reported training loss is the mean squared prediction error (MSPE) over the period of 1975 to 1988. The dashed vertical line indicates the start of the placebo intervention.

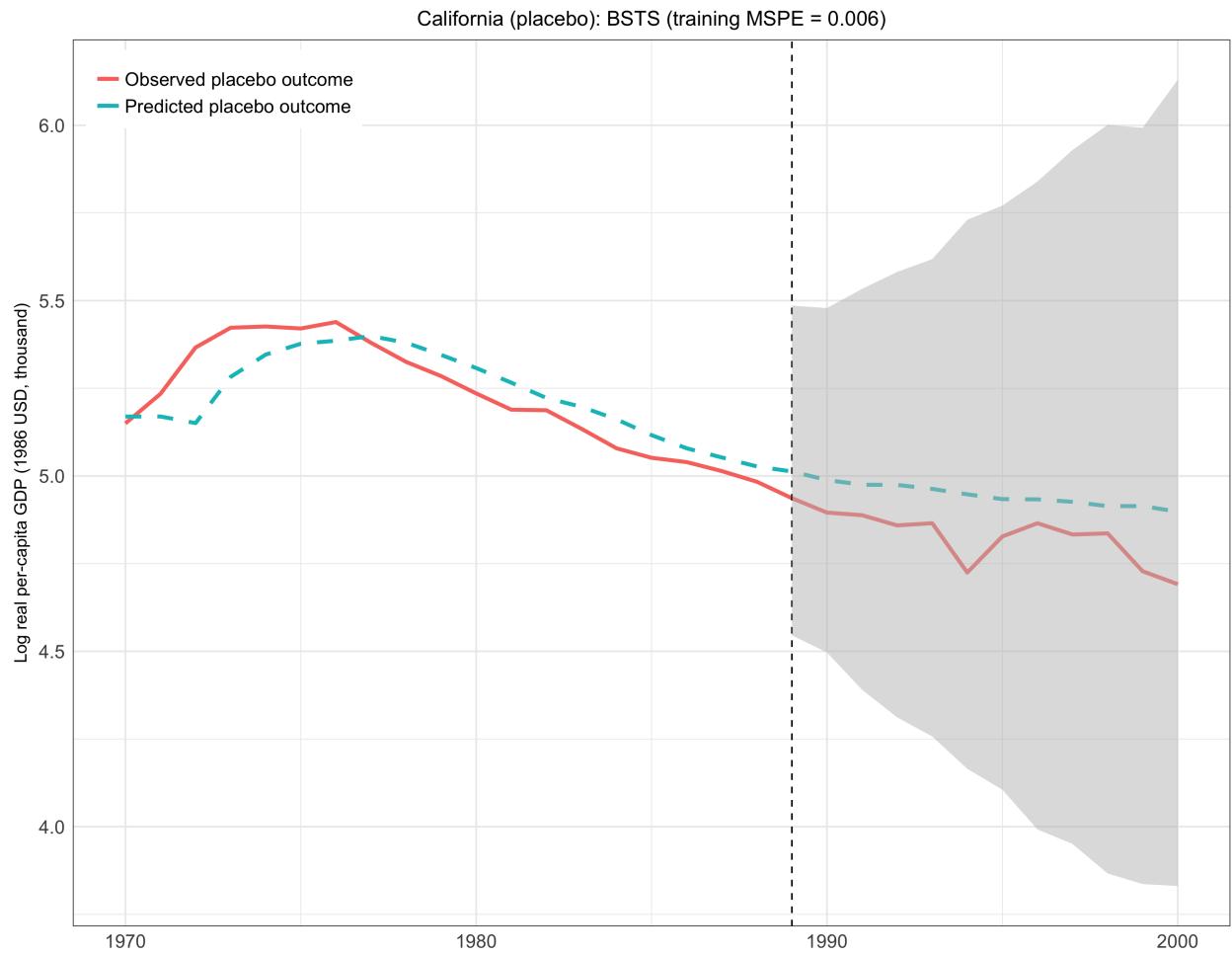


Figure 14: Observed placebo and BSTS predicted outcomes in California data. See notes to Fig. OA-5.

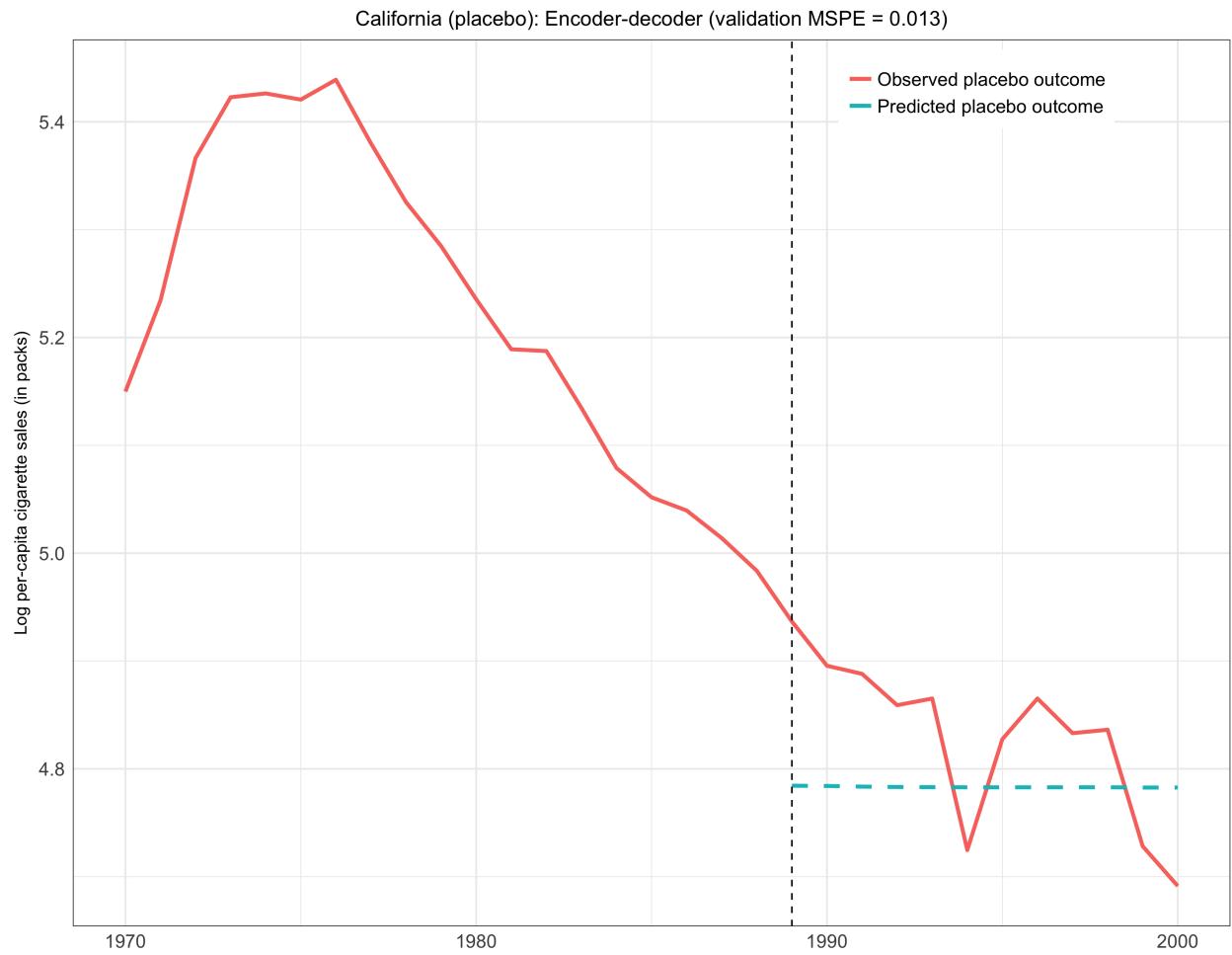


Figure 15: Observed placebo and encoder-decoder predicted outcomes in California data. See notes to Fig. OA-8.

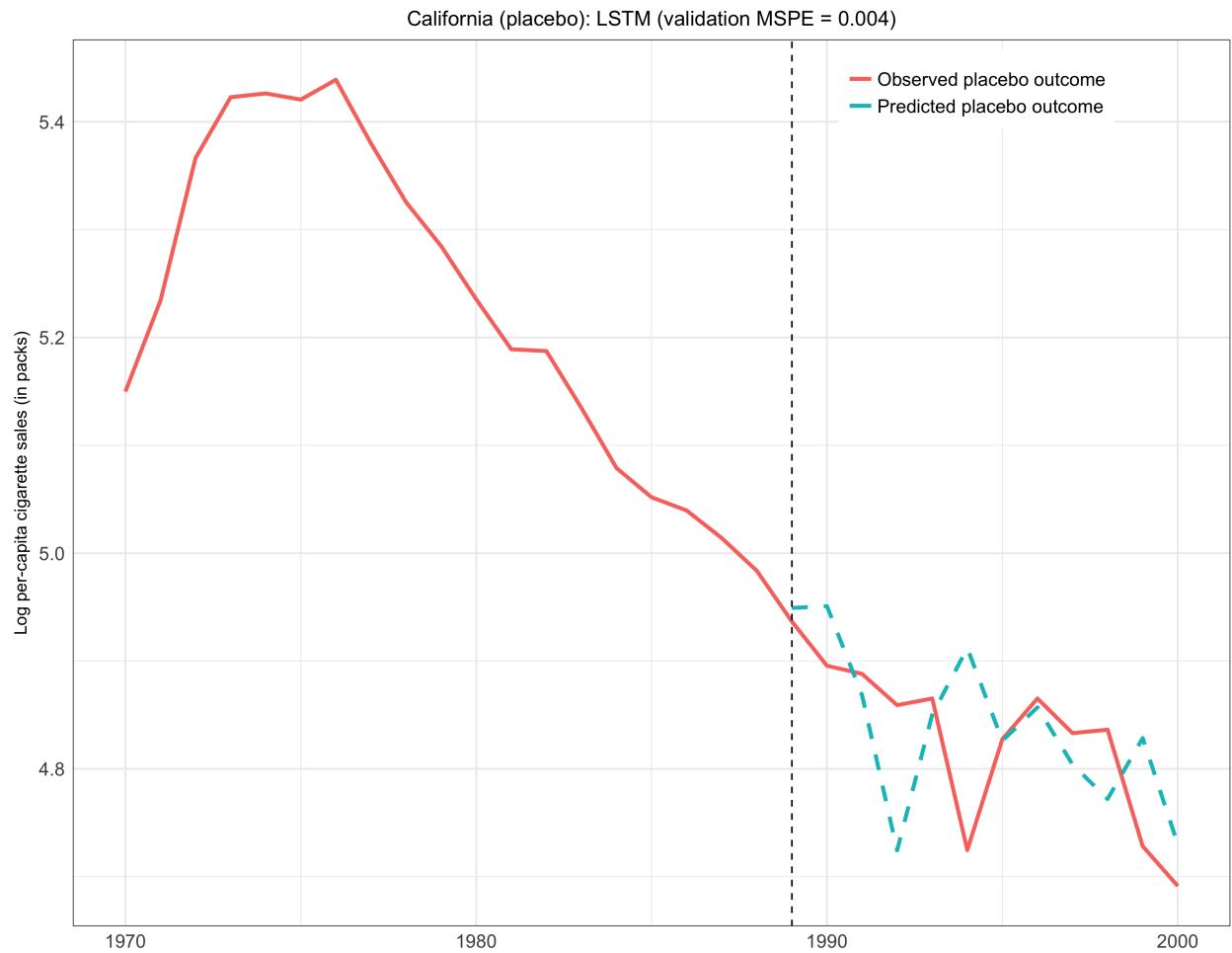


Figure 16: Observed placebo and LSTM predicted outcomes in California data. See notes to Fig. OA-8.

California: Estimation accuracy

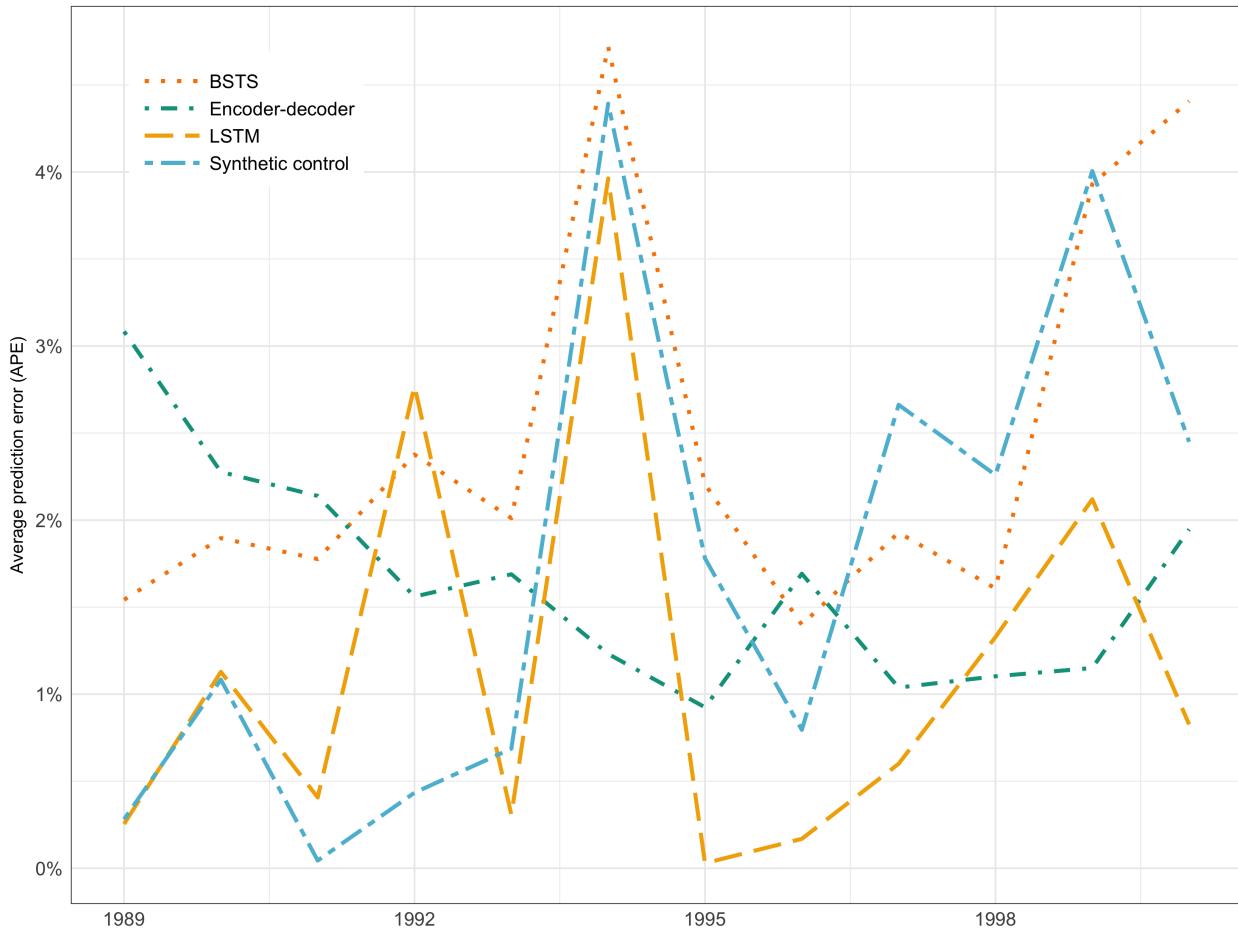


Figure 17: Estimation accuracy in terms of absolute percentage error (APE) for each post-period time-step in California data. Predictions are plotted in Figs. OA-14, OA-13, OA-15, and OA-16.

5 Estimates on West Germany data

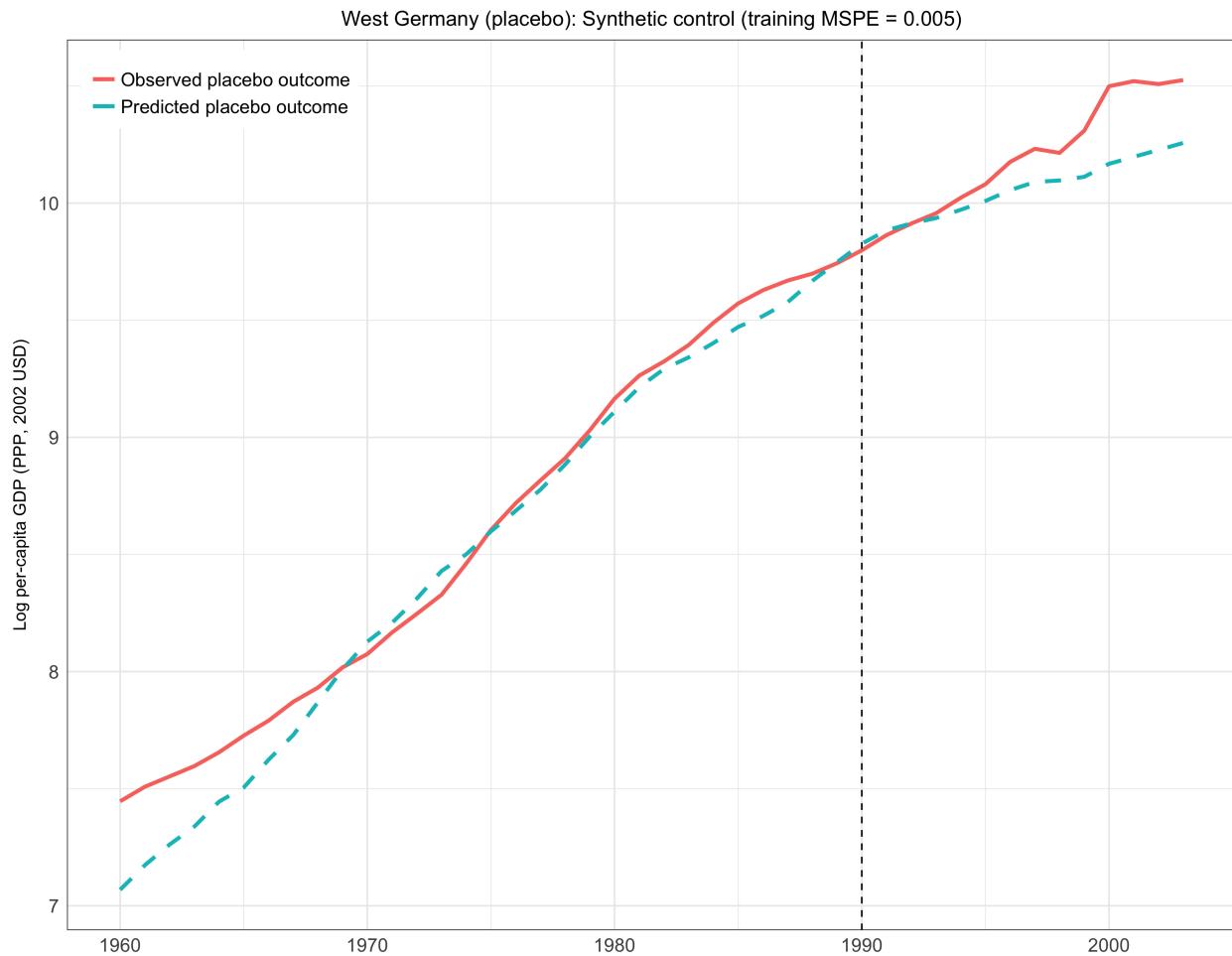


Figure 18: Observed placebo and synthetic control predicted outcomes in West Germany data. The reported training loss is the mean squared prediction error (MSPE) over the period of 1981 to 1989. The dashed vertical line indicates the start of the placebo intervention.

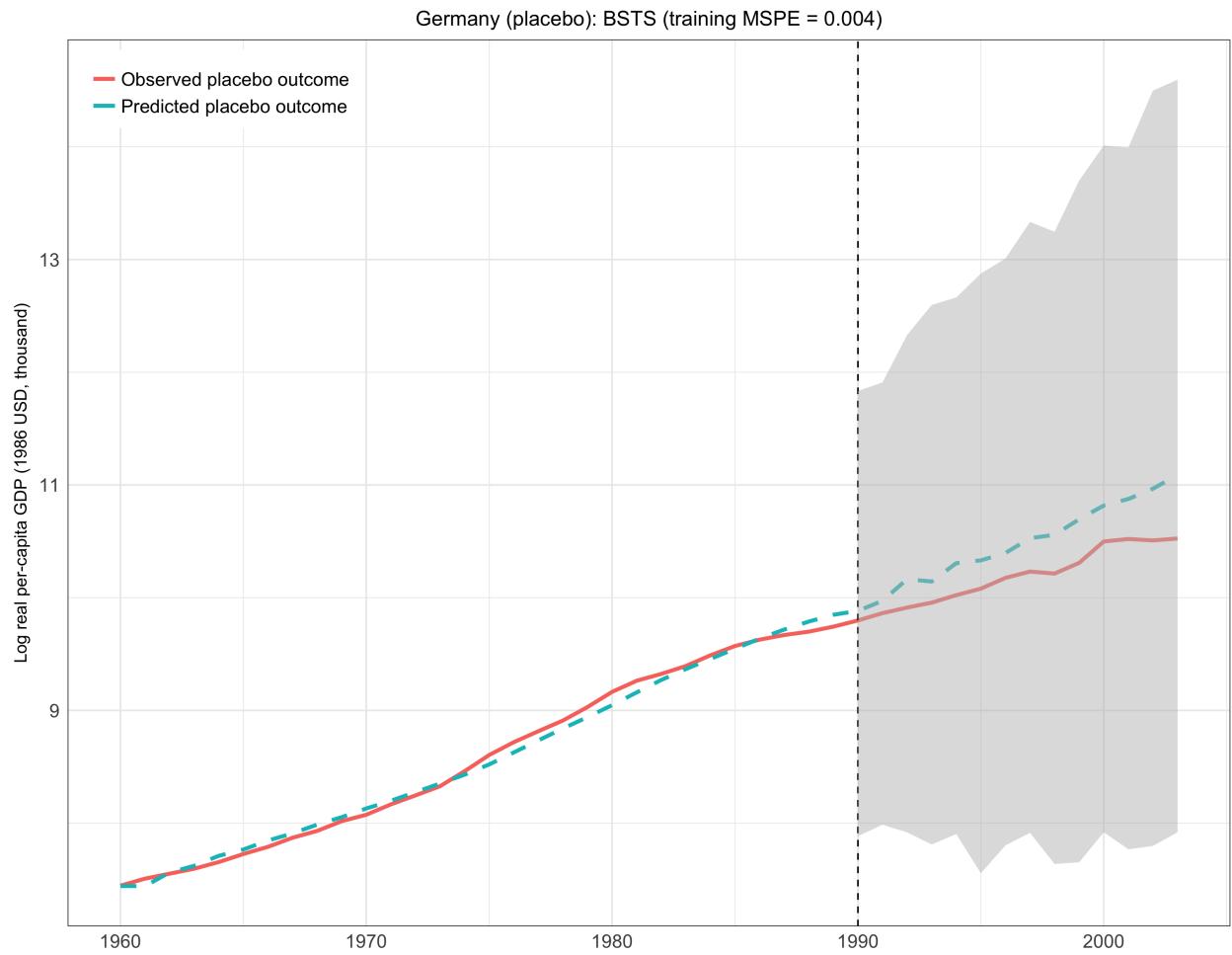


Figure 19: Observed placebo and BSTS predicted outcomes in West Germany data. See notes to Fig. OA-5.

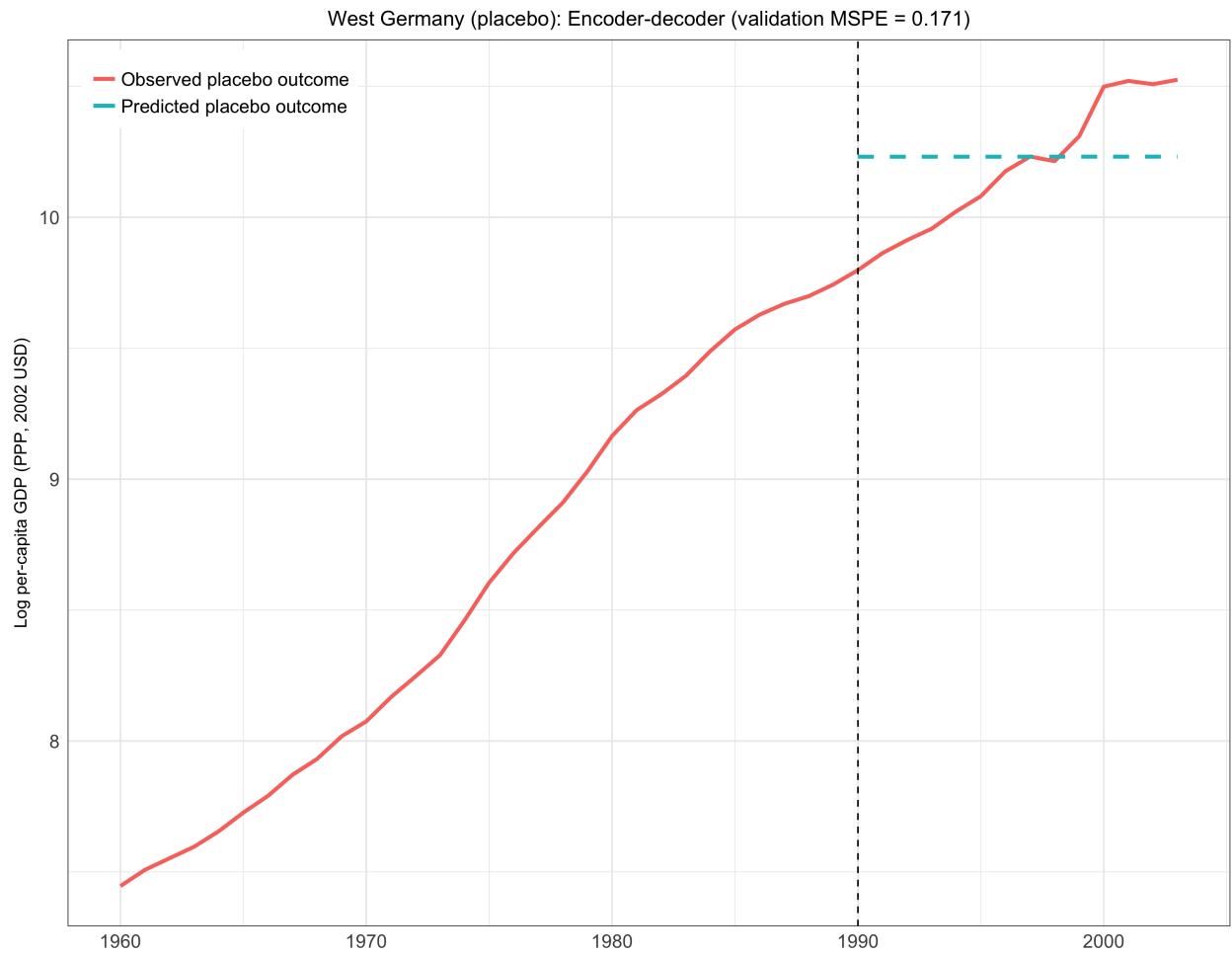


Figure 20: Observed placebo and encoder-decoder predicted outcomes in West Germany data. See notes to Fig. OA-8.

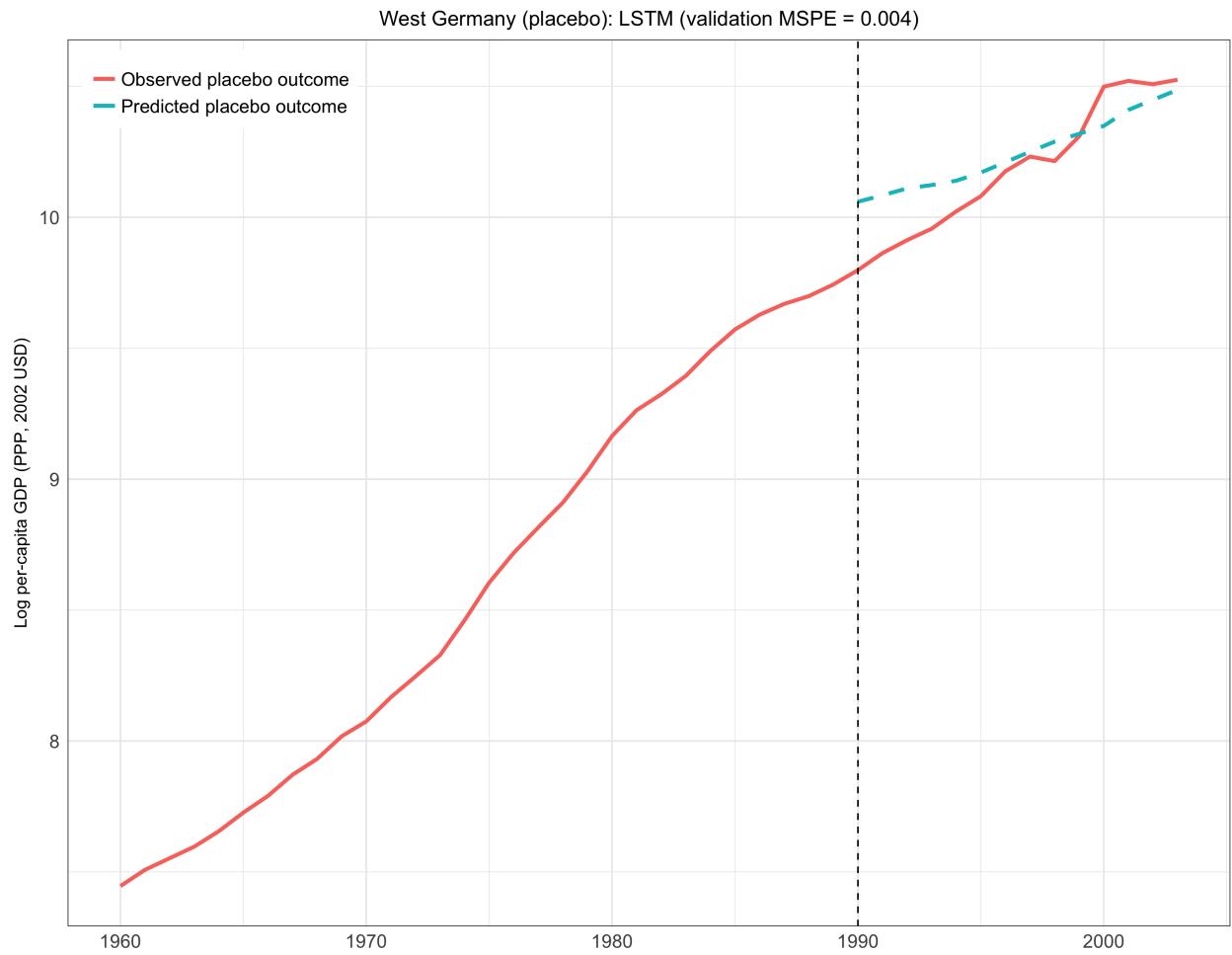


Figure 21: Observed placebo and LSTM predicted outcomes in West Germany data. See notes to Fig. OA-8.

West Germany: Estimation accuracy

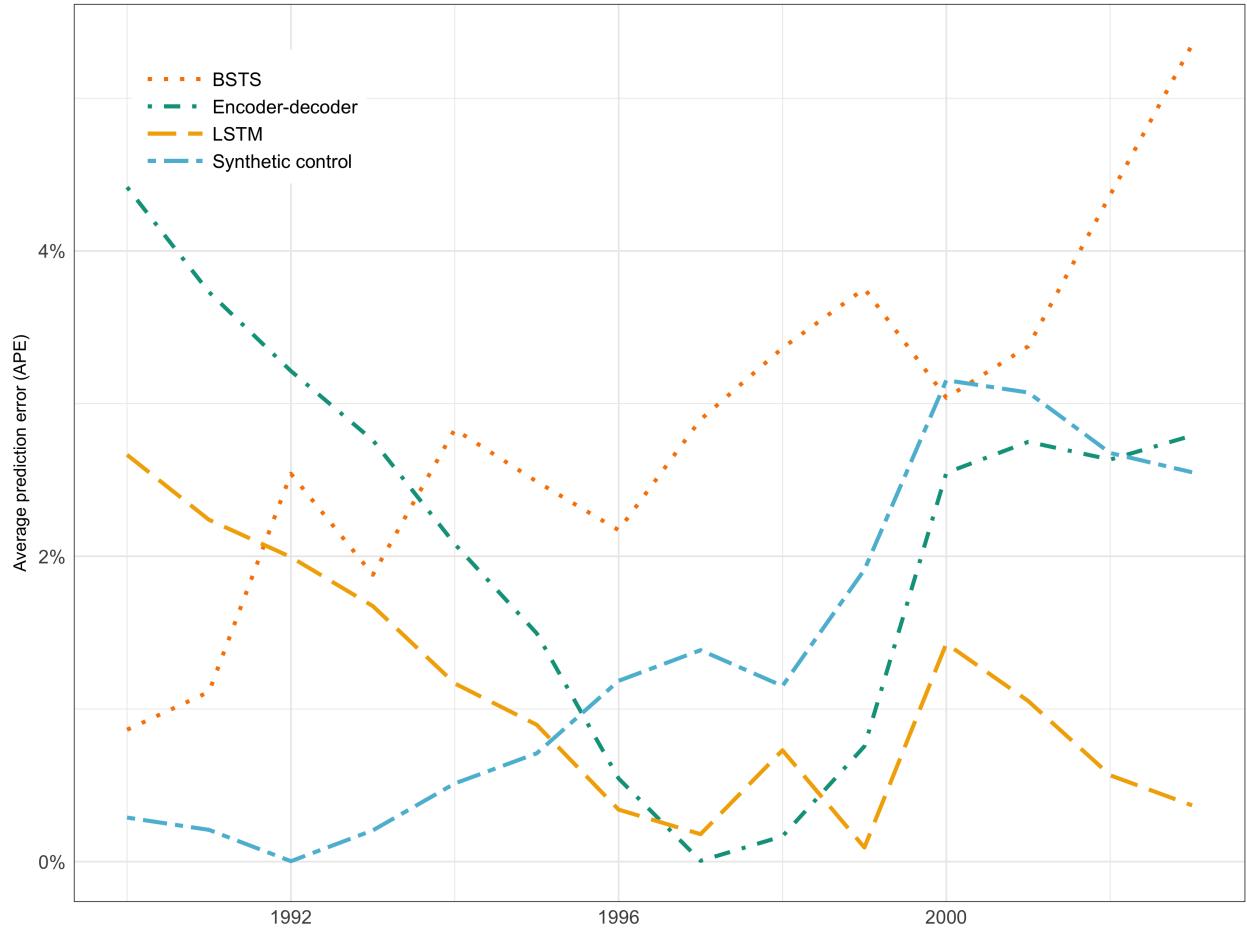


Figure 22: Estimation accuracy in terms of absolute percentage error (APE) for each post-period time-step in West Germany data. Predictions are plotted in Figs. OA-19, OA-18, OA-20, and OA-21.

6 Estimates on mayoral elections data

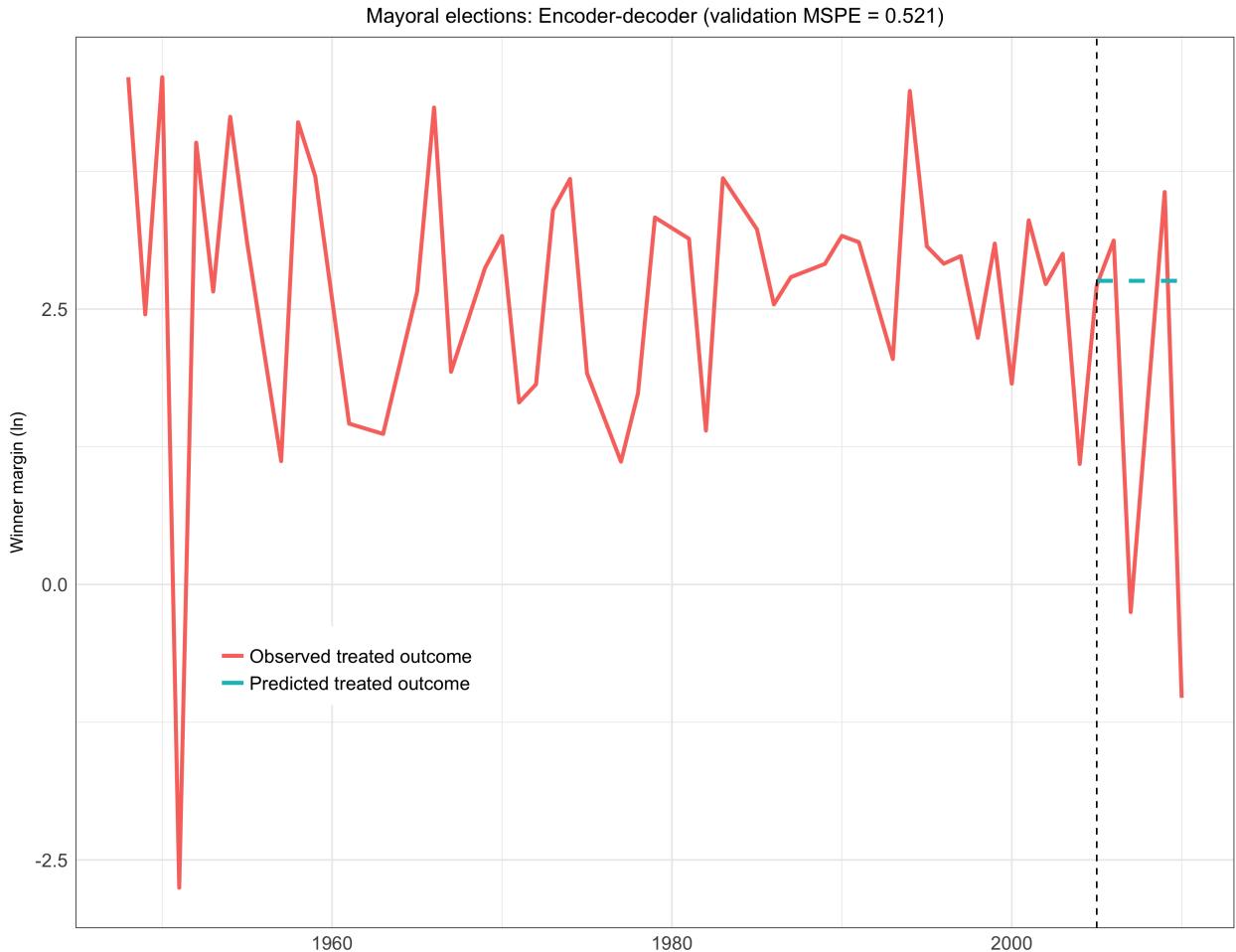


Figure 23: Observed and encoder-decoder predicted winner margins in Panagopoulos and Green treated cities. Validation MSPE is calculated on a held-out sample. The dashed vertical line is the start of the test set (2005) and the dotted vertical line represents the 2006 mayoral election.

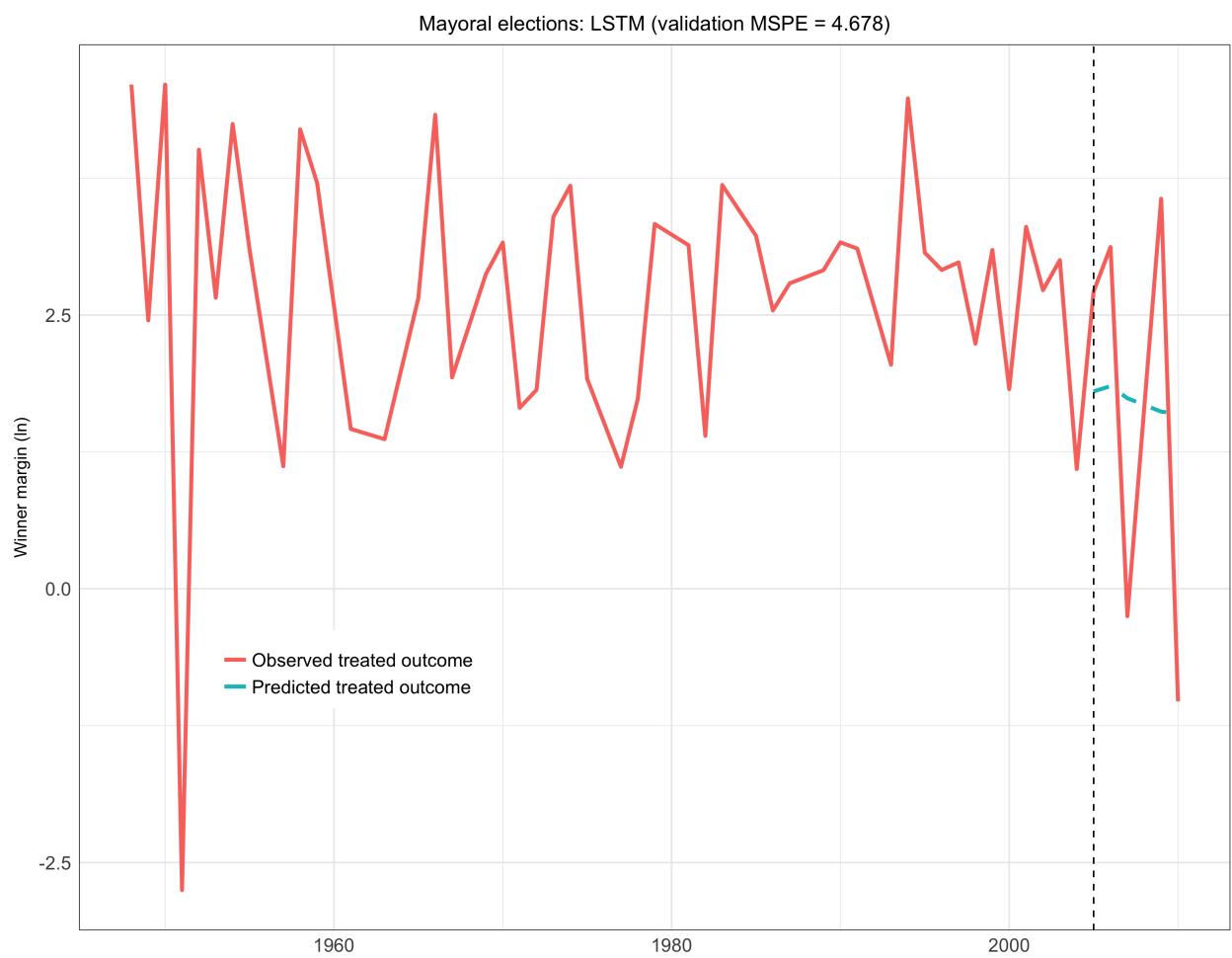


Figure 24: Observed and LSTM predicted winner margins in Panagopoulos and Green treated cities. See notes to Fig. OA-23.

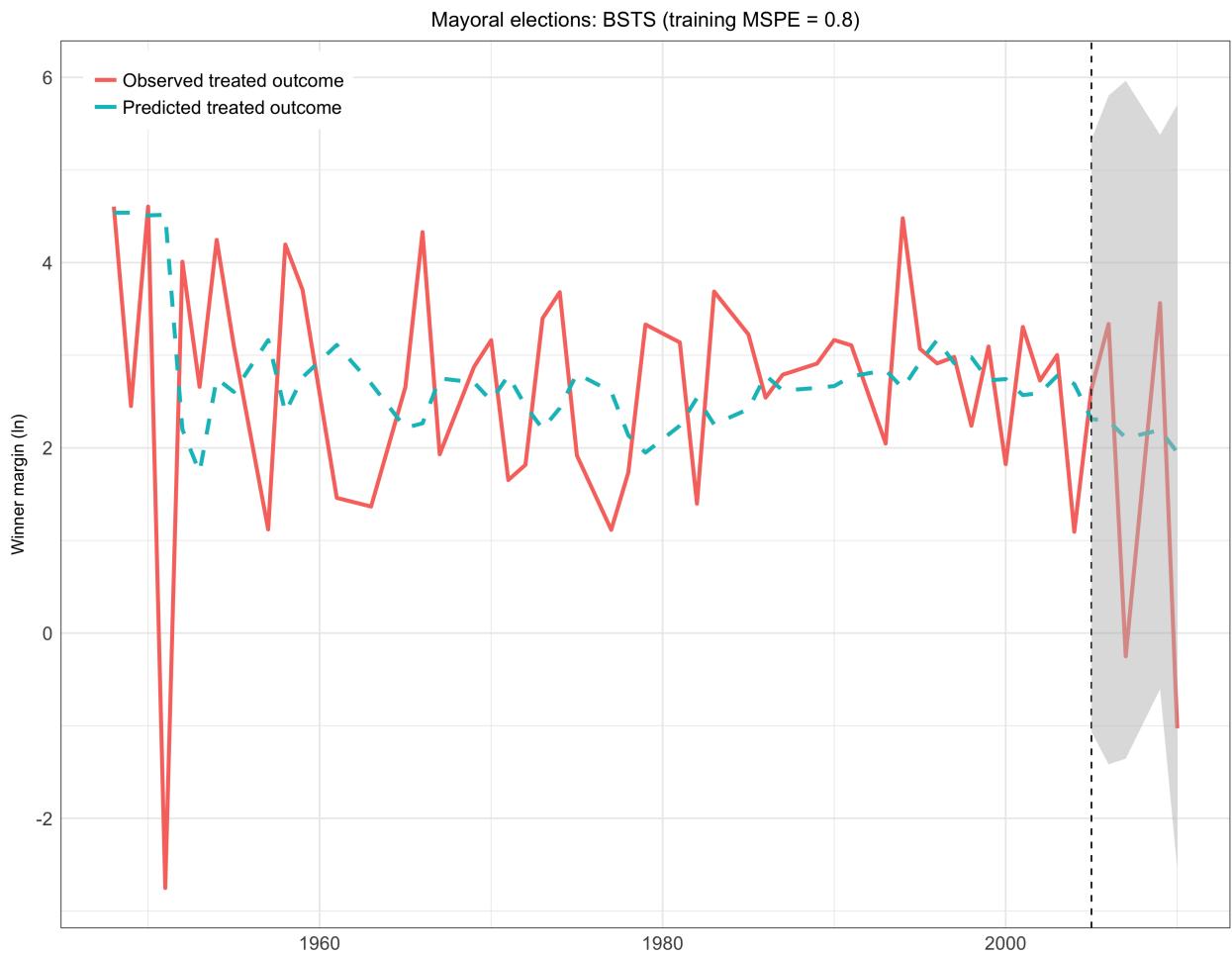


Figure 25: Observed and BSTS predicted winner margins in Panagopoulos and Green treated cities. The model predicts the time-series of the mean winner margin in Panagopoulos and Green treated cities using only winner margins in non-treated cities as predictors. See notes to Fig. OA-5.

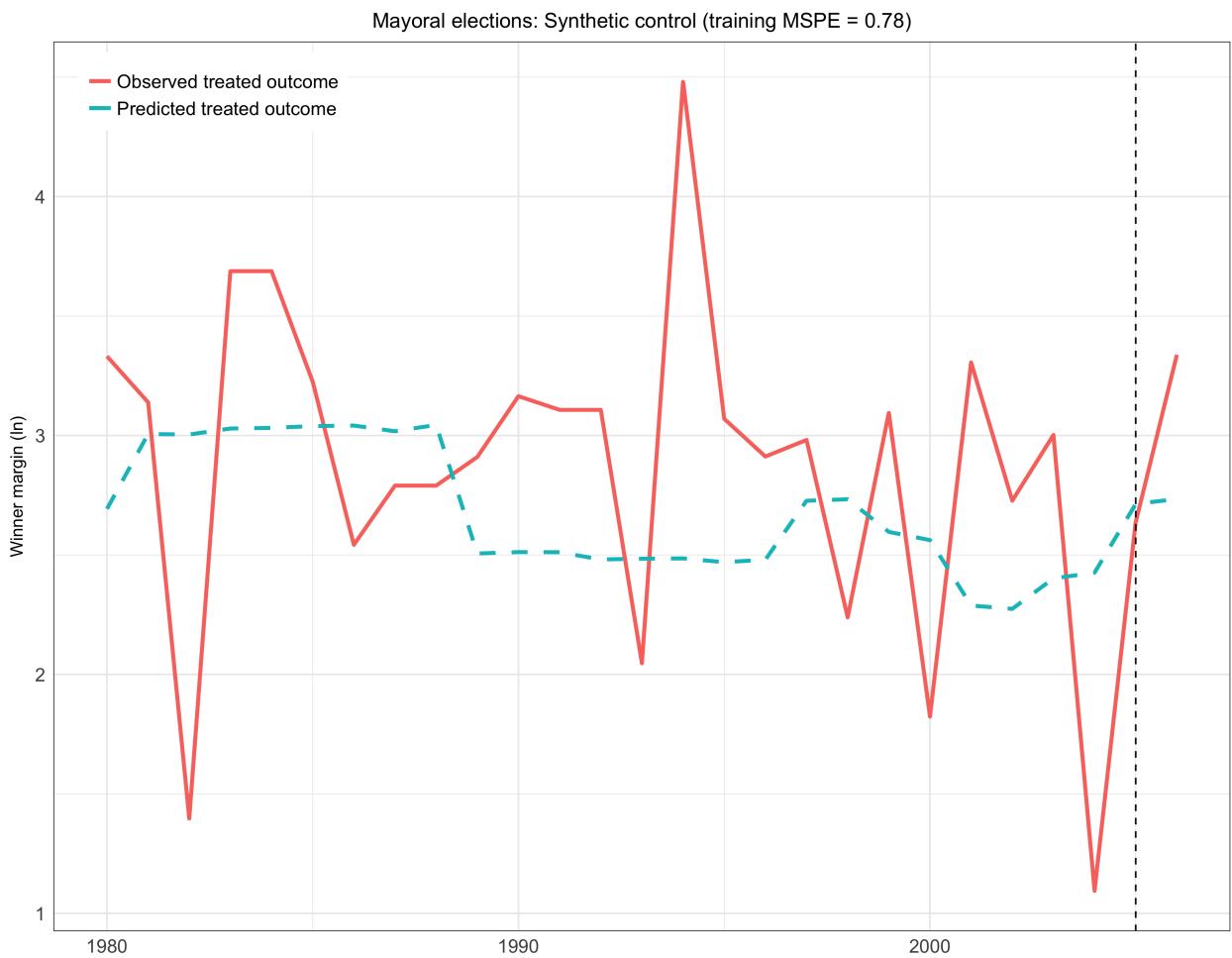


Figure 26: Observed and synthetic control predicted winner margins in Panagopoulos and Green treated cities. Synthetic control is constructed using pretreatment covariates including log vote total and winner margin means over the pre-period. The reported training loss is the mean squared prediction error (MSPE) over the period of 2000 to 2004.