**The impact of structural changes on forecasting performance: an example**

In this supplementary material, we how the impact of structural change problem on forecasting performance and how we can mitigate the problem using the Estimation Window Combining (EWC) method and the Intercept Correction (IC) method using an example based on simulation. First, we assume that the price variable has values of 2.99 for most of the weeks but occasionally get reduced to 2.29 or 1.99, which is usual in the retailer context. We also assume the following unobserved true product demand/sales[[1]](#footnote-1):

, , when

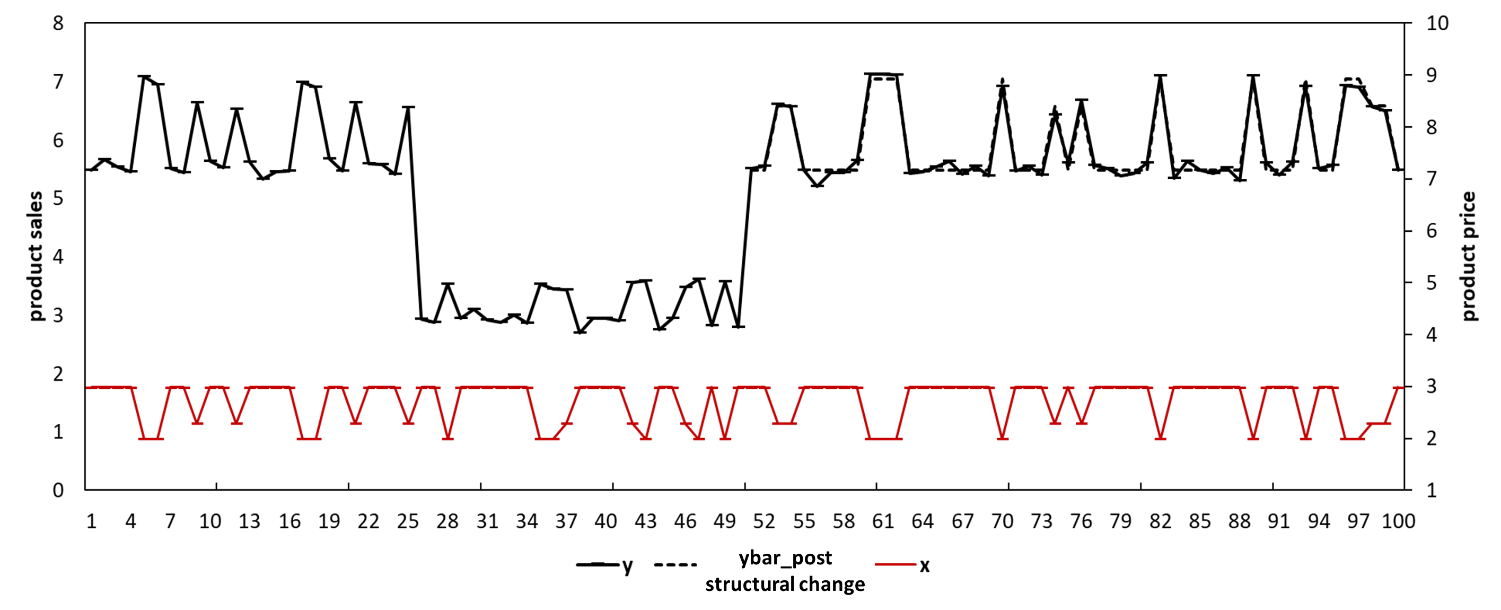
, , when

, , when

(1)

where and represent the product sales and the price at week *t*, and is the error term. This suggests two structural changes occurring at week 25 and 50 respectively. The sales and price are shown in Figure 1[[2]](#footnote-2).

Figure 1. The predictions and forecasts by [[3]](#footnote-3)



Suppose now we have the data from week 1 to week 75 and we want to forecast the product sales for the time period from week 76 to week 100. We could develop a congruent model (i.e., ) which overlook the structural change problem. If we know the date of the structural change, we may estimate this model exclusively using the post-structural change data (e.g., data from week 51 to week 75) and generate unbiased forecasts. We refer this model as Figure 1 shows its predictions/forecasts (i.e., “ybar\_post structural change”). Table 1 shows its forecasting performance (e.g., MAE= 0.09, MSE= 0.01, MAPE= 1.5%, and SMAPE= 1.5%).

However, the date of the structural change is normally unknown. Thus, the model tends to be estimated using all the available data (i.e., from week 1 to week 75). We refer this model as Figure 2 shows its predictions/forecasts (i.e., ybar\_1). Table 1 shows its forecasting performance (e.g., MAE= 0. 949, MSE= 0. 961, MAPE= 15.8%, and SMAPE= 17.2%). The model is thus subject to downwards forecast bias for the time period between week 76 to week 100, and gets outperformed by

Figure 2. The predictions and forecasts by

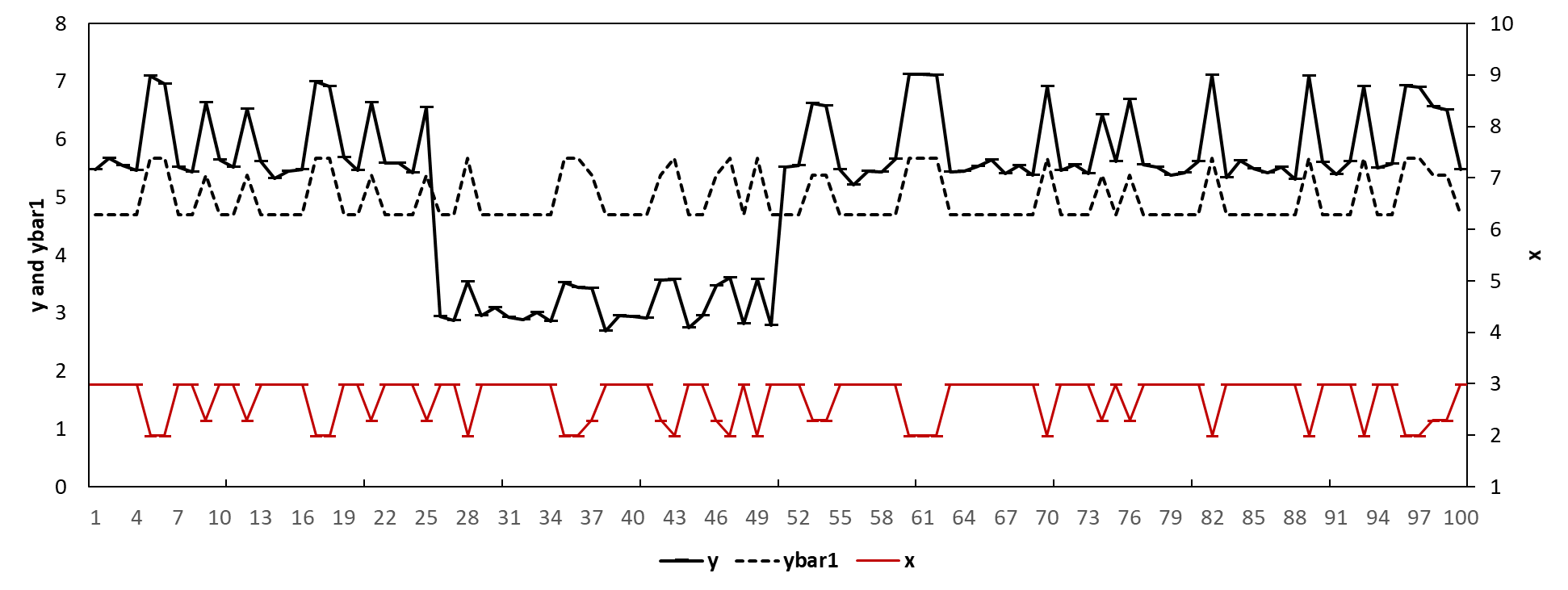
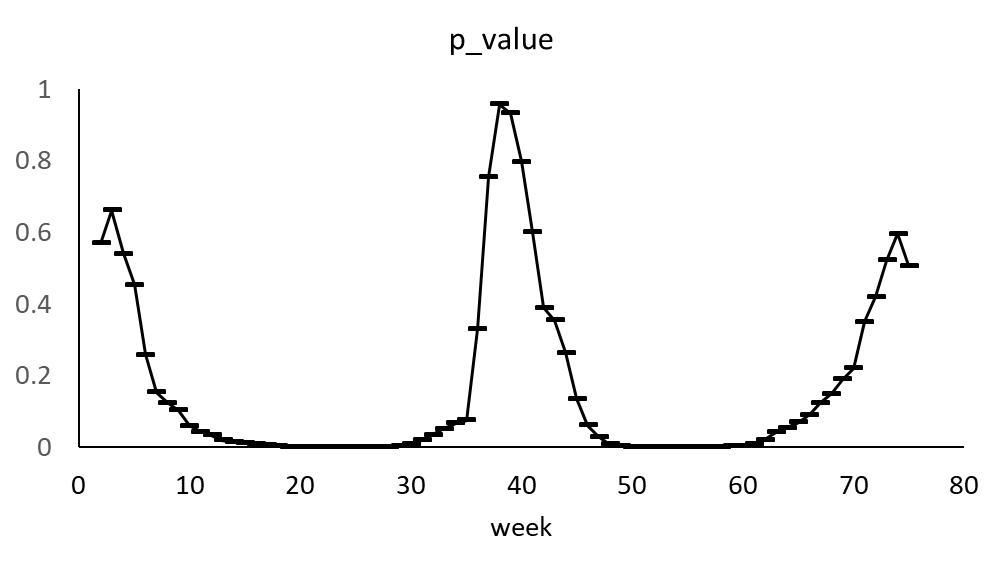


Table 1. The forecasting performance of different models in the simulation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | MAE | MSE | MAPE | SMAPE |
| The model estimated with all available data | 0.95 | 0.96 | 15.8% | 17.2% |
| The model estimated with post-structural change data | 0.09 | 0.01 | 1.5% | 1.5% |
| The model with intercept correction | 0.20 | 0.07 | 3.2% | 3.2% |
| The model with estimation window combining | 0.86 | 0.80 | 14.2% | 15.4% |

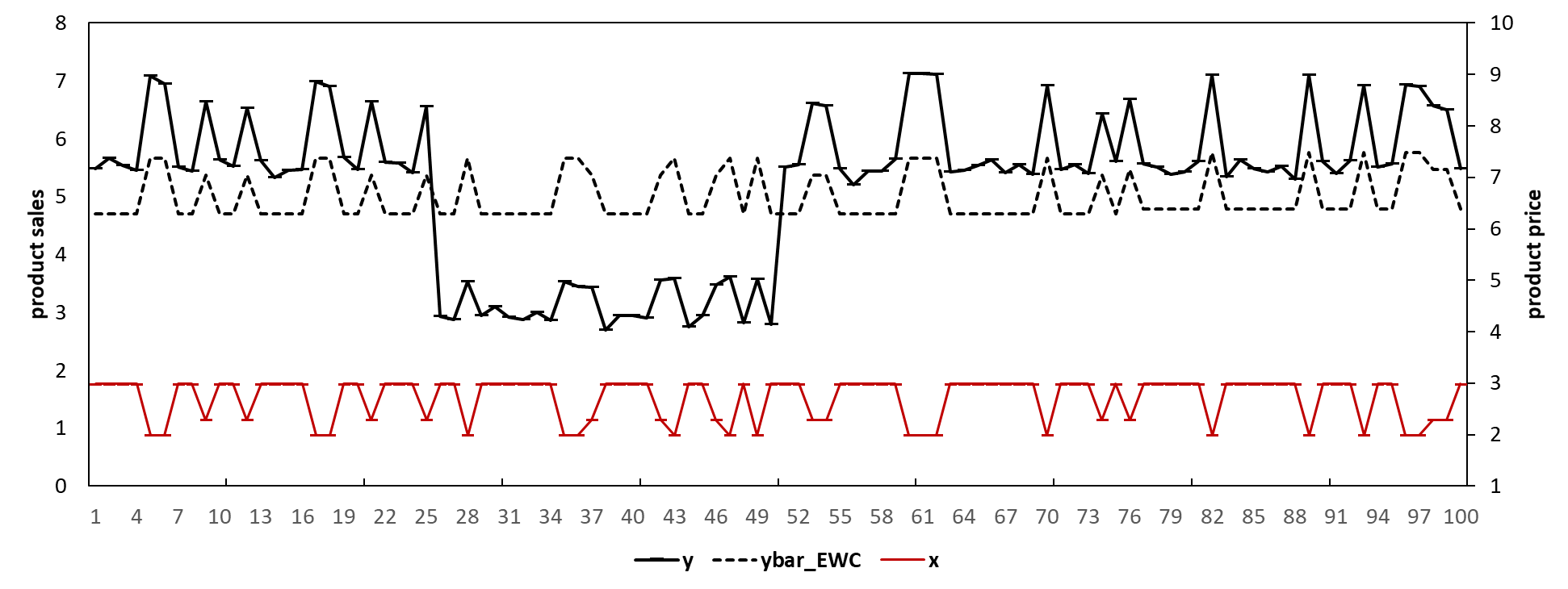
To mitigate the forecast bias and potentially improve the forecasting performance, we may implement the Estimation Window Combining (EWC) method and the Intercept Correction (IC) discriminantly based on if we detect the presence of any structural change using a sequential Chow (1960) test. For instance, we conduct the Chow (1960) tests for the model for multiple times and each time we assume the structural change occurs at one of the weeks. We reject the null hypothesis of no structural change if we detect structural change at any of the weeks. Figure 3 shows the p-values of the sequential Chow test where we conduct the Chow test each time assuming the structural change occurs at week 5, 6, … until 70. Figure 3 indicates the presence of structural change for the model (e.g., very likely occurring at week 25 and week 50 as the p-value of the Chow test for these two weeks are the lowest and close to zero). Alternative tests (e.g., those considering multiple breaks, heteroskedasticity, and unit roots etc.) are available but require additional priori knowledge and assumptions such as the number and the locations of potential structural changes (Andrews, 1993; Andrews & Ploberger, 1994; Bai & Perron, 1998, 2003). We choose the sequential Chow test because we only need to know if any structural change is present and also the test is easy to implement.

Figure 3 p-values of the sequential Chow test for each observation



We then implement the EWC method by combining the forecasts generated by the congruent model (i.e., ) but with different estimation windows. We estimate the model using the data [1, 75], and generate the first set of forecasts (referred to as ). We then estimate the model using the data [2, 75], and generate the second set of forecasts (referred to as ), and so forth. In this example, we arbitrarily choose to generate 60 sets of forecasts and we calculate the final forecasts as their average value. We refer this model as Figure 4 shows the forecasts (e.g., ybar\_EWC). Table 1 shows its forecasting performance (e.g., 0. 86 for MAE, 0. 80 for MSE, 14.2% for MAPE, and 15.4% for SMAPE). outperforms .

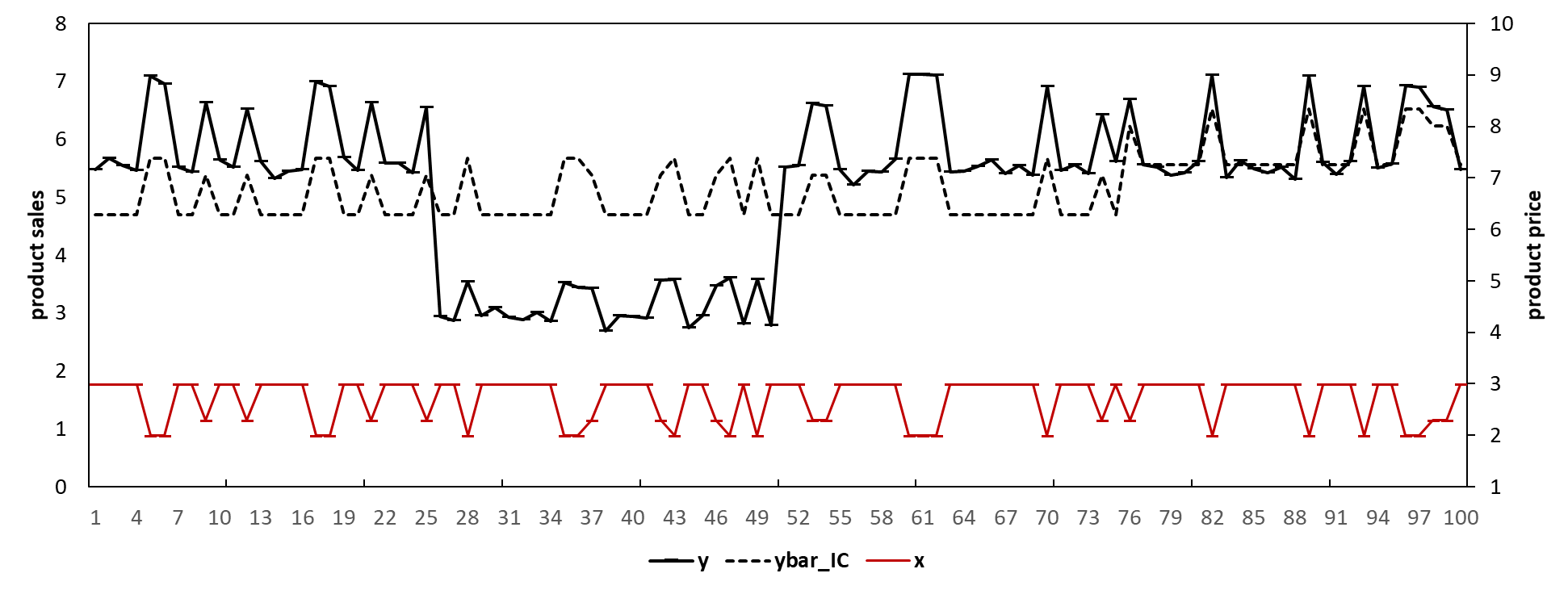
Figure 4. The predictions and forecasts by



In Figure 4, the black dashed line for the period [1, 75] represents the predictions by The black dashed line for the period [76, 100] represents the predictions by .

We can also implement the intercept correction (IC) method for the congruent model. We may estimate the forecast bias as the average value of an *ad hoc* number (e.g., we choose four in this example) of the residuals close to the forecast origin. e.g., where is the estimated forecast bias. We calculate the final forecasts by adding the estimated bias back to the forecasts generated by . We refer to this model as Figure 5 shows its predictions/forecasts (e.g., ybar\_IC). Table 1 shows its forecasting performance (e.g., MAE= 0.2, MSE= 0.07, MAPE= 3.2%, and SMAPE= 3.2%). outperforms .

Figure 5 The predictions and forecasts by



Reference:

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica, 61*, 825-851.

Andrews, D. W. K., & Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica, 62*, 1383-1414.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica, 66*, 47- 78.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural-change models. *Journal of Applied Econometrics, 18*, 1-22.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica, 28*(3).

1. We assume that there is no out-of-stock situation here. [↑](#footnote-ref-1)
2. In practice, the data for retailer product sales at the SKU level are of high varaitions and the change in the effect of marketing activities tends to be gradual rather than in an abrup way. Thus the data are unlikely to be as obvious as shown in Figure 1. However, here our purpose is only to show an example to illustrate the rationales. [↑](#footnote-ref-2)
3. In Figure 1, we highlight the period before the first structural change (e.g., week [1,25]) in blue. We highlight the period after the second structural change but before the forecast origin (e.g., week [51, 75]) in yellow. We highlight the period between the two structural changes (e.g., [26, 50]) in green, and we highlight the forecast period (e.g., week [76, 100]) in red. [↑](#footnote-ref-3)