# APPENDIX

### *Out-of-Sample Forecast Performance for Monthly Futures-Based Price Forecasts*

As presented in section 5, we estimate Equations 1, 2, 4, and 5 and calculate percentage errors from these forecasts to actual prices using Equation 6. These percentage errors are then used to evaluate the out-of-sample predictive performance of each alternative model to that of Hoffman's benchmark model (see Tables A1).

The test of bias, conducted based on the MPE statistic (see Table A1), show that both the Hoffman benchmark model and alternative models occasionally generate biased forecasts. Specifically, Hoffman model over-predicts cotton farm prices at short forecast horizons (1–5 months ahead), while the forecasts from alternative model specifications under-predict prices at mid to longer horizons (mostly 6–12 months ahead). This pattern is also evident in Figure A1 Panel A.

The dynamics of the MAPE and RMSPE statistics across forecast horizons and models (see Panel A of Figures A2 and A3, and Table A1) show that the errors from Model 1, moving average with basis deviation term, escalate very quickly starting from a one-month ahead horizon, and exceed Hoffman's benchmark forecast errors significantly. Model 2 forecast errors tend to be smaller than the Hoffman model errors for all forecast horizons when MAPE is considered a forecast performance metric. Similarly, when using the RMSPE statistic, Model 2 produces lower prediction errors than Hoffman's benchmark model, except for six– and seven–months ahead horizons. Model 3 forecast errors tend to be perform similar to that of Model 2 both for MAPE and RMSPE statistics. Short-term forecasts have smaller errors than longer-term forecasts across all models.

In fact, moving to regression models for cotton spot price projection resulted in up to 29%[8] accuracy gain for one-month ahead forecasts generated using Model 2 with an average gain of 12% across all forecast horizons. Model 3 performance was also satisfactory, with a maximum gain in accuracy of about 28% for 12-month ahead forecasts and an average gain in accuracy of 12% across all forecast horizons. On the other hand, Model 1's overall performance is relatively poor, with an average loss in accuracy of 81% across all forecast horizons. These findings indicate that regression-based specifications (i.e., Models 2 and 3) that relax the fixed parameter assumptions implied in the Hoffman model lead to significant improvements in cotton price forecast accuracy. Such improvement in regression-based approaches can partially be attributed to the model's capability to adjust its coefficients according to specific market conditions.

### *Modified Diebold-Mariano (MDM) and Encompassing Test*

The results discussed above showed that there are differences in accuracy between the proposed and the benchmark forecasts. The MDM test results presented in Table A1 indicate whether these differences are statistically significant. These results show that Model 2 leads to a significant increase in forecast accuracy in 11 out of 12 forecast horizons, whereas Model 3 produces significantly more accurate projections at shorter (1 – 2 months) forecast horizons. On the other hand, Model

[8] This value is calculated as $(1-\text{RMSPE-ratio}) * 100\%$.

1 results in a statistically significant deterioration in accuracy in almost all forecast horizons (10 out of 12).

In addition to accuracy, we test whether forecasts generated using the benchmark model contain all information provided in the alternative specifications. The results of the encompassing test for futures-based monthly price forecasts, shown in Table A1, reject this hypothesis for all three alternative models at all forecast horizons. This finding suggests that the forecasts from our proposed models include additional information missing from the Hoffman model and, therefore, offer an improvement over it.

***Sensitivity Analysis for Futures-Based Price Estimates***

As discussed in sections 1 and 2, cotton futures prices spiked and reached an all-time high in 2011 due to perceived supply shortages in the global market. This resulted in a significant deviation of basis from its historical average, creating an outlier in our dataset. To evaluate the sensitivity of our results to such unanticipated market conditions, we re-estimated the models after removing futures prices from the outlier period of August 2010 to August 2011. By the construction of our dataset, the 2011 futures prices are used to forecast farm prices both in 2011 and 2012. For instance, the 12-month ahead forecast for February 2012 will use the March 2012 futures contract prices observed in February 2011. Therefore, before conducting sensitivity analysis, we removed data from September 2010 until August 2012 (24 observations). Overall, the results of the analysis presented in Tables A2 as well as in Figures A1 – A3 Panel B are consistent with that of the full sample. As before, regression-based Models 2 and 3 produced more accurate and stable forecasts, as shown by lower RMSPE and positive and significant MDM statistics.

Table A1: Out-of-Sample Forecast Accuracy Evaluation for a Full Sample Analysis (January 2005 - December 2023)

| Evaluation Criteria | Forecast Horizon 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Hoffman Model** | | | | | | | | | | | |
| MPE | -2.497$^{**}$ | -2.441$^{**}$ | -2.110$^{**}$ | -1.860$^{**}$ | -1.563$^{*}$ | -1.278 | -1.193 | -1.131 | -1.049 | -1.075 | -0.936 | -0.870 |
| MAPE | 9.877 | 9.634 | 9.434 | 9.901 | 10.429 | 11.172 | 11.759 | 12.489 | 13.097 | 13.788 | 14.362 | 14.996 |
| RMSPE | 15.482 | 14.846 | 13.931 | 13.697 | 14.039 | 14.681 | 15.314 | 16.099 | 16.936 | 17.854 | 18.557 | 19.298 |
| | **Model 1: BDEV Moving Average Model** | | | | | | | | | | | |
| MPE | 0.865 | 1.025 | 1.609 | 2.151 | 2.872 | 3.426$^{*}$ | 4.168$^{*}$ | 4.578$^{**}$ | 5.557$^{**}$ | 4.134$^{*}$ | 3.334$^{*}$ | 6.112$^{**}$ |
| MAPE | 10.131 | 10.792 | 12.182 | 12.925 | 14.456 | 15.812 | 17.081 | 18.458 | 19.756 | 18.984 | 18.651 | 21.724 |
| RMSPE | 18.329 | 19.941 | 21.350 | 22.934 | 26.967 | 28.021 | 32.240 | 33.812 | 39.682 | 31.983 | 27.307 | 46.714 |
| MDM | -0.230 | -0.835 | -2.178$^{**}$ | -2.193$^{**}$ | -2.449$^{**}$ | -2.867$^{***}$ | -2.788$^{***}$ | -2.830$^{***}$ | -2.675$^{***}$ | -2.838$^{***}$ | -3.230$^{***}$ | -2.219$^{**}$ |
| Encompassing | 0.691$^{***}$ | 0.637$^{***}$ | 0.521$^{***}$ | 0.447$^{***}$ | 0.387$^{***}$ | 0.326$^{***}$ | 0.280$^{***}$ | 0.231$^{***}$ | 0.206$^{***}$ | 0.205$^{***}$ | 0.191$^{***}$ | 0.196$^{***}$ |
| | **Model 2: Regression Model with Nearby Contract** | | | | | | | | | | | |
| MPE | 0.570 | 0.649 | 0.691 | 0.929 | 1.231 | 2.906$^{**}$ | 3.112$^{***}$ | 3.339$^{***}$ | 2.821$^{***}$ | 3.506$^{***}$ | 4.198$^{***}$ | 4.797$^{***}$ |
| MAPE | **8.343** | **8.193** | **8.404** | **8.808** | **9.345** | **10.571** | **10.966** | **11.369** | **11.784** | **12.230** | **12.668** | **12.934** |
| RMSPE | **11.010** | **10.680** | **10.671** | **10.857** | **11.407** | 17.344 | 16.396 | **15.835** | **14.851** | **15.540** | **16.136** | **16.649** |
| MDM | 2.026$^{**}$ | 2.021$^{**}$ | 1.650 | 1.810$^{*}$ | 1.837$^{*}$ | 1.956$^{*}$ | 1.931$^{*}$ | 2.074$^{**}$ | 1.713$^{*}$ | 1.813$^{*}$ | 1.891$^{*}$ | 2.189$^{**}$ |
| Encompassing | 0.942$^{***}$ | 0.952$^{***}$ | 0.933$^{***}$ | 0.904$^{***}$ | 0.888$^{***}$ | 0.856$^{***}$ | 0.834$^{***}$ | 0.797$^{***}$ | 0.722$^{***}$ | 0.683$^{***}$ | 0.644$^{***}$ | 0.637$^{***}$ |
| | **Model 3: Regression Model with BDEV term** | | | | | | | | | | | |
| MPE | 0.655 | 1.003 | 1.032 | 1.306$^{*}$ | 1.708$^{**}$ | 2.924$^{***}$ | 3.058$^{***}$ | 3.301$^{***}$ | 2.983$^{***}$ | 3.484$^{***}$ | 3.822$^{***}$ | 4.310$^{***}$ |
| MAPE | **8.406** | **8.236** | **8.718** | **9.100** | **9.779** | **10.824** | **11.471** | **11.895** | **12.114** | **12.651** | **13.238** | **13.685** |
| RMSPE | **11.204** | **11.029** | **11.196** | **11.278** | **11.935** | 15.121 | 15.485 | **15.557** | **15.152** | **15.905** | **16.622** | **17.233** |
| MDM | 2.070$^{**}$ | 2.118$^{**}$ | 1.267 | 1.402 | 1.138 | 1.258 | 1.051 | 1.292 | 1.252 | 1.237 | 1.132 | 1.256 |
| Encompassing | 0.900$^{***}$ | 0.896$^{***}$ | 0.851$^{***}$ | 0.819$^{***}$ | 0.816$^{***}$ | 0.850$^{***}$ | 0.832$^{***}$ | 0.777$^{***}$ | 0.675$^{***}$ | 0.626$^{***}$ | 0.567$^{***}$ | 0.552$^{***}$ |

Note 1: The number of observations at each horizon is 228.
Note 2: BDEV stands for basis deviation term. MPE represents mean percentage error. MAPE stands for mean absolute percentage error, RMSPE represents the root mean square percentage error, and MDM stands for the Modified Diebold-Mariano test.
Note 3: The bolded values in MAPE and RMSPE indicators show that compared to Hoffman model a specific alternative model produces lower forecast errors.
Note 4: The null hypothesis for the test of bias states that the MPEs are not statistically different from 0.
Note 5: We conduct MDM test against Hoffman benchmark. The null hypothesis for the 2-tail test states that the two forecasts have similar forecast accuracy.
Note 6: The null hypothesis for the encompassing test states that the benchmark model forecast contains all the information provided by the alternative model forecasts.
Note 7: Statistical significance of a two-tailed test is indicated with the asterisk: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

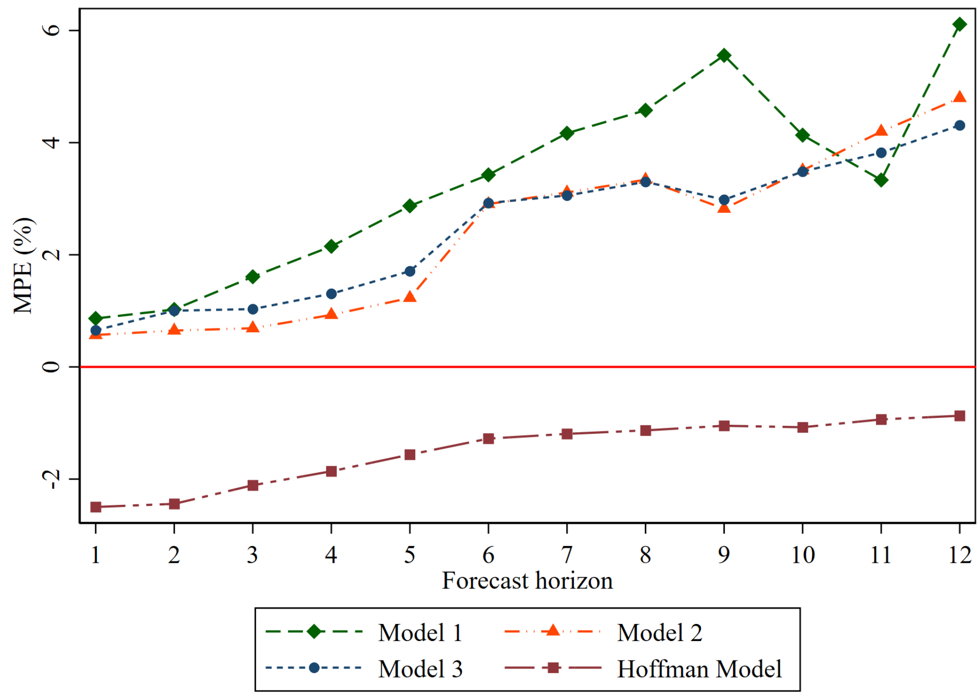
Table A2: Out-of-Sample Forecast Accuracy Evaluation for Sensitivity Analysis (without August 2010 - August 2011 futures prices)

| | Forecast Horizon | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Evaluation Criteria | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| | Hoffman Model | | | | | | | | | | | |
| MPE | -0.350 | -0.382 | -0.333 | -0.271 | -0.242 | -0.234 | -0.504 | -0.755 | -0.929 | -1.230 | -1.401 | -1.631 |
| MAPE | 7.818 | 7.865 | 8.156 | 8.815 | 9.539 | 10.575 | 11.330 | 12.226 | 13.017 | 13.758 | 14.314 | 14.853 |
| RMSPE | 10.513 | 10.382 | 10.457 | 11.187 | 12.430 | 13.758 | 14.852 | 16.033 | 17.078 | 18.104 | 18.826 | 19.476 |
| | Model 1: BDEV Moving Average Model | | | | | | | | | | | |
| MPE | -0.324 | -0.625 | -0.739 | -0.712 | -0.761 | -0.749 | -0.937 | -1.106 | -1.081 | -1.257 | -1.255 | -0.736 |
| MAPE | 8.150 | 8.615 | 9.435 | 10.201 | 11.434 | 12.476 | 12.990 | 14.055 | 14.573 | 15.086 | 15.714 | 16.636 |
| RMSPE | 11.777 | 11.947 | 12.488 | 13.797 | 15.262 | 16.557 | 17.487 | 19.000 | 20.109 | 20.537 | 21.599 | 23.785 |
| MDM | -0.570 | -1.347 | -2.405** | -2.534** | -3.444*** | -3.342*** | -2.858*** | -2.999*** | -2.431** | -2.023** | -1.877* | -1.940* |
| Encompassing | 0.527*** | 0.499*** | 0.430*** | 0.370*** | 0.340*** | 0.313*** | 0.323*** | 0.271*** | 0.262*** | 0.303*** | 0.276*** | 0.264*** |
| | Model 2: Regression Model with Nearby Contract | | | | | | | | | | | |
| MPE | 1.597*** | 1.732*** | 1.861*** | 2.113*** | 2.439*** | 4.250*** | 4.427*** | 4.600*** | 3.906*** | 4.545*** | 5.186*** | 5.759*** |
| MAPE | **5.848** | **5.883** | **6.310** | **7.325** | **8.352** | **10.148** | **10.821** | **11.433** | **11.715** | **12.016** | **12.207** | **12.156** |
| RMSPE | **7.797** | **7.667** | **8.053** | **9.185** | **10.584** | 17.524 | 16.494 | **15.838** | **14.548** | **14.998** | **15.244** | **15.374** |
| MDM | 4.682*** | 4.672*** | 4.350*** | 3.479*** | 2.435** | 2.254** | 1.927* | 1.966* | 2.143** | 2.571** | 3.007*** | 3.655*** |
| Encompassing | 0.932*** | 0.953*** | 0.956*** | 0.930*** | 0.886*** | 0.838*** | 0.815*** | 0.784*** | 0.727*** | 0.708*** | 0.694*** | 0.692*** |
| | Model 3: Regression Model with BDEV term | | | | | | | | | | | |
| MPE | 1.564*** | 2.026*** | 1.916*** | 2.288*** | 2.946*** | 4.505*** | 4.666*** | 4.678*** | 3.960*** | 4.416*** | 4.787*** | 5.926*** |
| MAPE | **6.066** | **5.888** | **6.353** | **7.401** | **8.723** | **10.338** | **11.010** | **11.492** | **11.395** | **11.776** | **11.913** | **12.343** |
| RMSPE | **8.187** | **7.838** | **8.182** | **9.362** | **10.945** | 15.031 | 15.339 | **15.159** | **14.107** | **14.543** | **14.841** | **16.018** |
| MDM | 3.705*** | 4.362*** | 4.111*** | 3.145*** | 1.287 | 1.178 | 1.156 | 1.586 | 2.497** | 2.872*** | 3.376*** | 3.394*** |
| Encompassing | 0.881*** | 0.916*** | 0.921*** | 0.881*** | 0.808*** | 0.797*** | 0.782*** | 0.781*** | 0.756*** | 0.736*** | 0.721*** | 0.666*** |

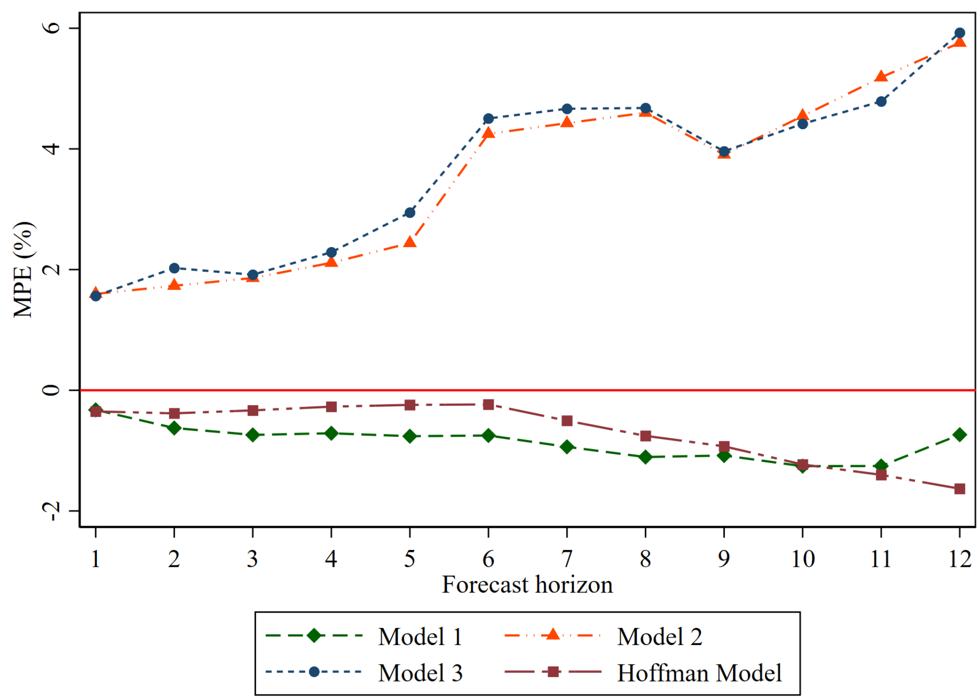
Note 1: For sensitivity analysis, we chose to drop futures prices of the outlier time frame of August 2010 - August 2011.
Note 2: The number of observations at each horizon is 204.
Note 3: BDEV stands for basis deviation term. MPE represents mean percentage error. MAPE stands for mean absolute percentage error, RMSPE represents the root mean square percentage error, and MDM stands for the Modified Diebold-Mariano test.
Note 4: The bolded values in MAPE and RMSPE indicators show that compared to Hoffman model a specific alternative model produces lower forecast errors.
Note 5: The null hypothesis for the test of bias states that the MPEs are not statistically different from 0.
Note 6: We conduct MDM test against Hoffman benchmark. The null hypothesis for the 2-tail test states that the two forecasts have similar forecast accuracy.
Note 7: The null hypothesis for the encompassing test states that the benchmark model forecast contains all the information provided by the alternative model forecasts.
Note 8: Statistical significance of a two-tailed test is indicated with the asterisk: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure A1: Mean Percentage Errors (MPE) of Farm Price Forecasts by Forecast Horizon

Panel A: Full Sample




Panel B: Sensitivity Analysis (excludes 2011 futures price data)

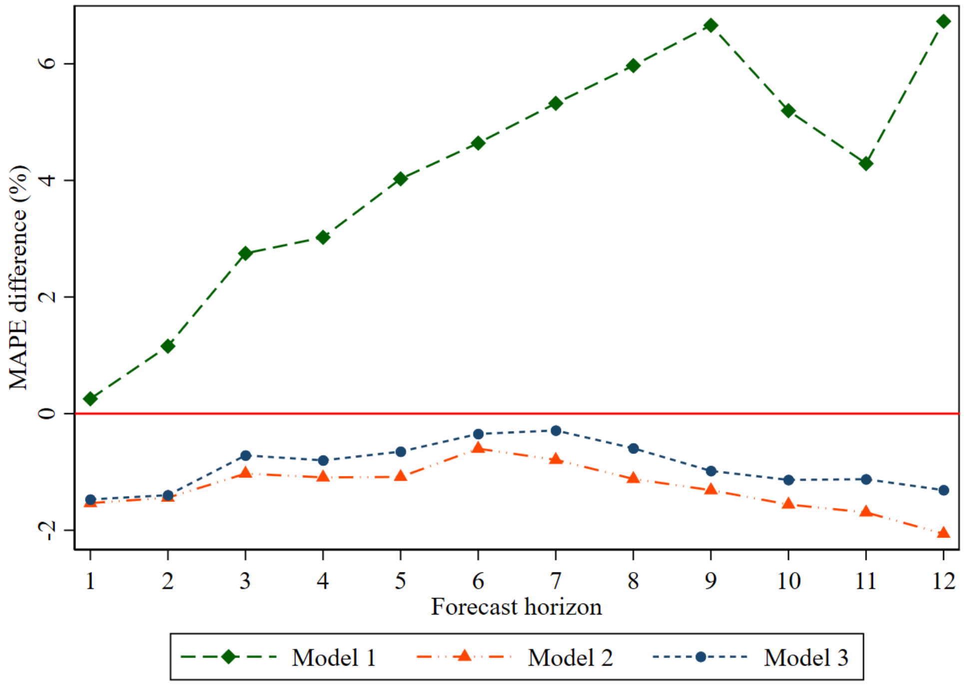



Note 1: The figure above plots the accuracy statistics of the MPE criterion (detailed in Equation 9) across models and horizons. Within each horizon, we test if the MPEs are significantly different from zero. Respective models are as detailed in Section 3.

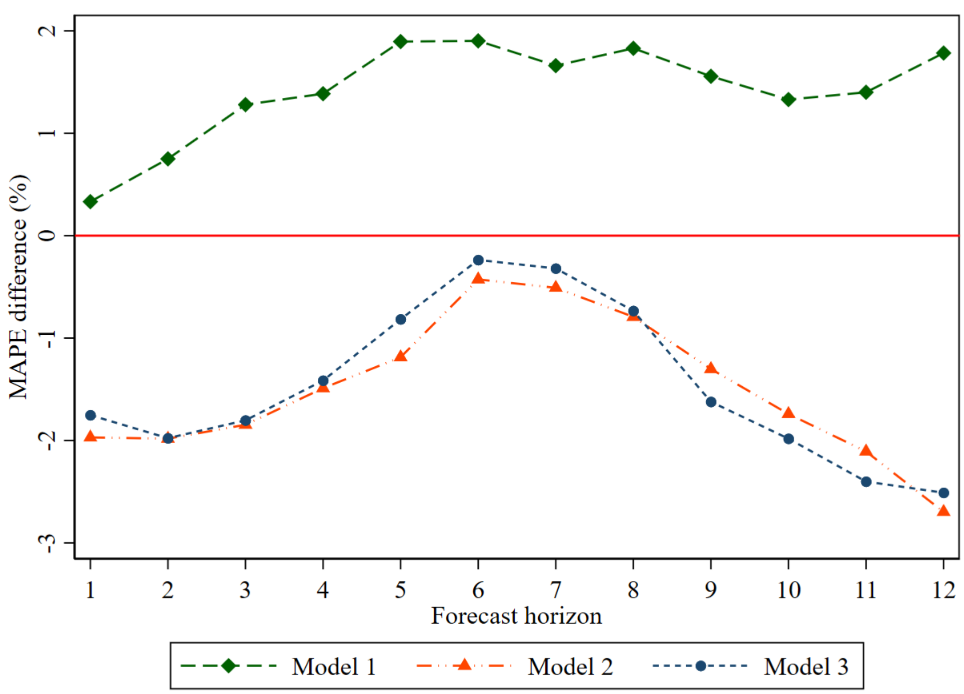
Note 2: Model 1 represents the moving average approach with basis deviation term. Model 2 represents the regression model with nearby futures prices. Model 3 represents the regression model with basis deviation term.

Figure A2: Mean Absolute Percentage Errors (MAPE) Difference of Farm Price Forecasts by Forecast Horizon

Panel A: Full Sample




Panel B: Sensitivity Analysis (excludes 2011 futures price data)

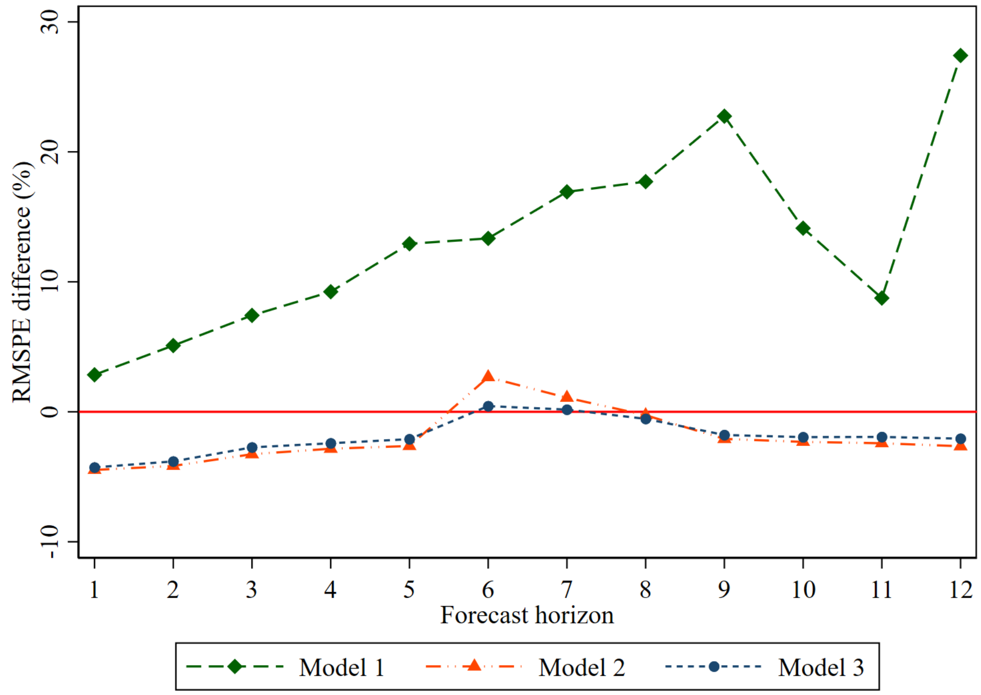



Note 1: This figure shows the differences in MAPE statistics between alternative model forecasts and Hoffman's model projections. A negative (positive) MAPE difference indicates that the alternative model produces more (less) accurate forecast than the Hoffman model.

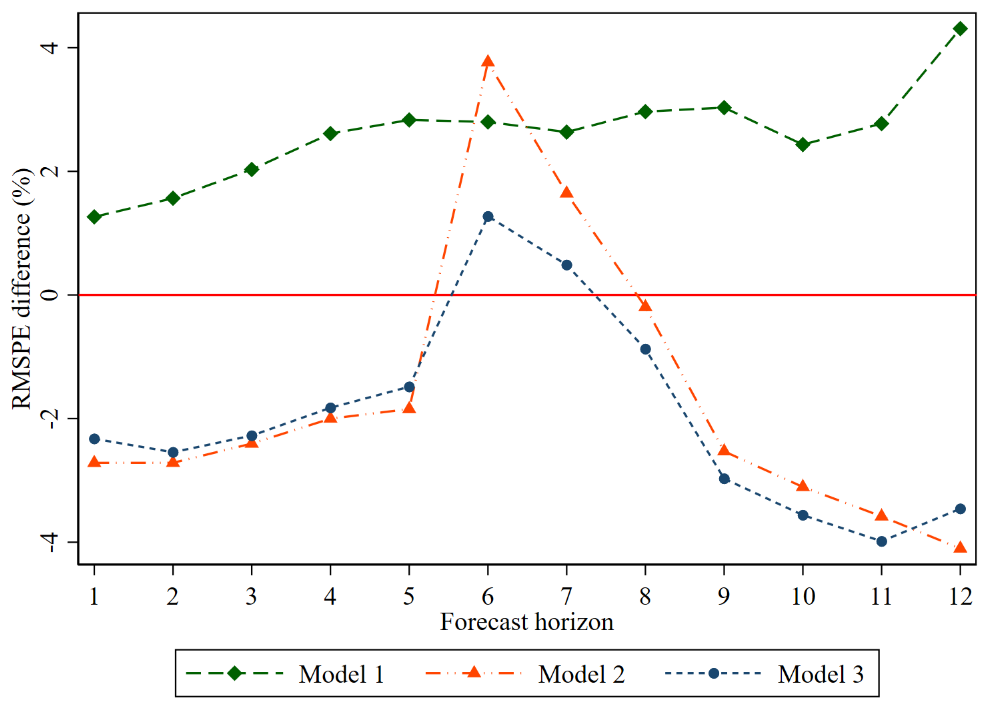
Note 2: Model 1 represents the moving average approach with basis deviation term. Model 2 represents the regression model with nearby futures prices. Model 3 represents the regression model with basis deviation term.

Figure A3: Root Mean Squared Percentage Errors (RMSPE) Difference of Farm Price Forecasts by Forecast Horizon

Panel A: Full Sample




Panel B: Sensitivity Analysis (excludes 2011 futures price data)




Note 1: This figure shows the differences in RMSPE statistics between alternative model forecasts and Hoffman's model projections. A negative (positive) RMSPE difference indicates that the alternative model produces more (less) accurate forecast than the Hoffman model.

Note 2: Model 1 represents the moving average approach with basis deviation term. Model 2 represents the regression model with nearby futures prices. Model 3 represents the regression model with basis deviation term.