

# Mincer-Zarnovitz Quantile and Expectile Regressions for Forecast Evaluations under Asymmetric Loss Functions

Working Paper Series 14-01 | January 2014

Guler, Kemal      Pin T. Ng      Zhijie Xiao

January 30, 2014

## Abstract

Forecast is pervasive in all areas of applications in business and daily life and, hence, evaluating the accuracy of a forecast is important for both the generators and consumers of forecasts. There are two aspects in forecast evaluation: (1) measuring the accuracy of past forecasts using some summary statistics and (2) testing the optimality properties of the forecasts through some diagnostic tests. On measuring the accuracy of a past forecast, we illustrate that the summary statistics used should match the loss function that was used to generate the forecasts. If there is strong evidence that an asymmetric loss function has been used in the generation of a forecast, then a summary statistic that corresponds to that asymmetric loss function should be used in assessing the accuracy of the forecast instead of the popular *RMSE* or *MAE*. On testing the optimality of the forecasts, we demonstrate how the quantile regressions and expectile regressions set in the prediction-realization framework of Mincer and Zarnowitz (1969) can be used to recover the unknown parameter that controls the potentially asymmetric loss function used in generating the past forecasts. Finally, we apply the prediction-realization framework to the Federal Reserve's economic growth forecast and forecast sharing in a PC manufacturing supply chain. We find that the Federal Reserve's values over prediction are approximately 1.5 times more costly than under prediction. We also find that the PC manufacturer weighs positive forecast errors (under forecasts) about four times as costly as negative forecast errors (over forecasts).

## 1 Introduction

Forecast is pervasive in all areas of business and daily life. Weather forecasts are important for planning our day-to-day activities. Farmers rely on them for the planting and harvesting of crops while airlines need them to make decisions that maintain safety in the skies. The insurance industry relies on them for informed pricing and capital decisions. Corporations use forecasting to predict their future financial needs, production planning, human resource planning, etc. Forecasting is used by investors to value companies and their securities. The startup of a new business requires forecasting of the demand for the product, expected shares in the market, the capacity of competitor, the amount and sources for funds, etc. In supply chain management, businesses have to synchronize the ordering of supplies to meet the forecasted demand of its customers. In government policy decisions, economic forecasts are very important for determining the appropriate monetary policy/fiscal policy. In the healthcare industry, forecasting can be used to target disease management or device personalized health care based on predicted risk.

As a result, evaluating the accuracy of a forecast is important for both the generators and consumers of forecasts. However, there are abundant evidence that the forecasts being generated are inconsistent with the realizations of the forecasted values. Silver (2012) discusses the weather industry's bias toward forecasting more precipitation than will actually occur, what meteorologists call "wet bias". Using 121 responses to a 26-question mail questionnaire sent to the highest ranking financial officer in 500 firms on the Fortune 500 listing, Pruitt and Gitman (1987) found that capital budgeting forecasts are optimistically biased by people with work experience. Ali, et al. (1991) find that analysts set overly optimistic forecasts of the next period's annual earnings per shares. Lee et al. (1997) and Cohen et al. (2003) provide ample evidence of overoptimistic forecasts across industries ranging from electronics and semiconductors to medical equipment and commercial aircraft in the supply chains. In terms of economic variables forecasts, Capistrán (2008) provides evidence that the Federal Reserve's inflation forecasts systematically under-predicted before Paul Volcker appointment as Chairman and systematically over-predicted afterwards until the second quarter of 1998.

Do these forecast biases signify suboptimal forecast performance? In the traditional sense, the over-prediction or under-prediction bias are indications of suboptimal forecasts. However, the traditional tests for forecast optimality rely, typically, on the assumption of symmetric mean-square error loss function. Under this square error loss, over prediction and under prediction are weighted equally, and optimal forecasts imply that the observed forecast errors will have a zero bias and are uncorrelated with variables in the forecasters' information set. However, strong arguments can be provided for the rationale that forecasters may not have adopted a symmetric errors loss function. For example, in firms' forecasting of sales, over prediction will result in over inventory and increased insurance costs, and tied up capital while under prediction leads to loss of goodwill, reputation, and current and future sales. Firms may decide that the cost of loss of goodwill is much higher than increased insurance costs and, hence, weigh the under prediction error more than the over prediction error.

For money managers of banks, over predicting the value-at-risk ties up more capital than necessary while under predicting leads to regulatory penalties and the need for increased capital provisions. They may conclude that the cost of increased capital provisions is higher than the cost of tied up capital and decide to weigh the under prediction error more than the error of over prediction. It might be particularly costly for the Federal Reserves to over predict GDP growth when growth is already slow, signaling a false recovery, which could lead to an overly tight monetary policy at exactly the wrong time. The cost of over forecast is not always the same as that of under forecast. The dissatisfaction that people have when the weatherman forecasts a sunny day but it turns out to be a rainy day and, hence, ruin a picnic party, is way higher than when it is forecasted to be rainy but turns out to be sunny and, hence, people enjoy an unexpected nice day out. This can be the explanation for the wet bias and illustrate the asymmetric loss function used by the weatherman. Keane and Runkle (1990, p. 719) argue that: "If forecasters have differential costs of over- and underprediction, it could be rational for them to produce biased forecasts. If we were to find that forecasts are biased, it could still be claimed that forecasters were rational if it could be shown that they had such differential costs." Varian (1974), Waud (1976), Zellner (1986), Christoffersen and Diebold (1997), and Patton and Timmermann (2007) all have argued that the presence of forecast bias is not necessary an indication of suboptimal forecast.

Elliott and Timmermann (2008) summarize that forecast evaluation usually comprises of two separate, but related, tasks: (1) measuring the accuracy of past forecasts using some summary statistics and (2) testing the optimality properties of the forecasts through some diagnostic tests. On the first task of assessing the accuracy of a forecast, the most popular summary measures have been the sample mean-square error ( $MSE$ ), the sample root-mean-square error ( $RMSE$ ) and the sample mean-absolute error ( $MAE$ ). However, the  $MSE$  and  $RMSE$  metrics are appropriate only if the loss function used in the forecaster's decision making is the symmetric mean-square error while the  $MAE$  is appropriate if the loss function is the symmetric mean-absolute error as will be elaborated in details in Section 2. Numerous articles have argued for the likelihood of and addressed the issues related to an asymmetric loss functions being used by forecasters (see e.g., Granger, 1969 & 1999; Varian, 1974; Granger and Newbold, 1986; Zellner, 1986; Ito, 1990; West, Edison and Cho, 1993; Weiss, 1996; Christoffersen and Diebold, 1997; Batchelor and Peel, 1998; Granger and Pesaran, 2000; Artis and Marcellino, 2001; Pesaran and Skouras, 2002; Carpistran, 2006; Patton and Timmermann, 2007; and Elliott and Timmermann, 2008). We demonstrate in this article that if there is strong evidence that an asymmetric loss function has been used in the generation of a forecast, then a summary statistic that corresponds to that asymmetric loss function, in particular the sample root-mean-weighted-square error ( $RMWSE$ ) or the sample mean-weighted-absolute error ( $MWAE$ ), both of which will be defined in Section 2, should be used in assessing the accuracy of the forecast instead of the popular  $RMSE$  or  $MAE$ .

The second task of testing the optimality properties of a forecast relies on the

knowledge of the loss function being used and the underlying data generating process (DGP) that generates the future values of the predicted variable, both of which are, unfortunately, unknown in typical situations. There are a few families of popular loss function specifications used in the literature. Elliott, Komunjer and Timmermann (2005) propose a flexible family of loss functions which subsumes the asymmetric lin-lin piecewise linear loss function, in which the *MAE* loss is a special case, and the asymmetric quad-quad loss function, which nests the *MSE* loss. It can be readily shown that the optimal forecasts for the lin-lin loss function are the conditional quantiles while those for the quad-quad loss functions are the conditional expectiles. Varian (1974), Zellner (1986), and Christoffersen and Diebold (1997) use the linex loss while Christoffersen and Diebold (2006) adopt the sign loss function. An important problem in testing the optimality properties of a forecast is the issue of determining the unknown loss function based on a sequence of observed forecasts. Elliott and Timmermann (2008) suggest that one can estimate the unknown parameter that controls the degree of asymmetry of the lin-lin and quad-quad loss functions through the first order condition of the risk adopted by the forecasters. We demonstrate how this can be accomplished using the quantile regressions and expectile regressions set naturally in the prediction-realization framework of Mincer and Zarnowitz (1969). With the estimated asymmetric parameter that controls the asymmetry of the loss functions, one can then choose between the *RMSE* and *RMWSE*, or the *MAE* and *MWAE* as the appropriate metric in evaluating forecast accuracy.

The rest of the paper is organized as follows. In Section 2, we provide arguments and evidence that the metric used in measuring the accuracy of a forecast should be chosen according to the loss function being used when the forecast was performed. We illustrate how the unknown parameter that controls the degree of asymmetry in a loss function used by a forecaster can be recovered using the Mincer-Zarnowitz prediction-realization framework in Section 3. Section 4 illustrates applications of the Mincer-Zarnowitz prediction-realization approach to assess the Federal Reserve’s Greenbook forecast and sales forecasts of an electronic component manufacturer to try to recover the parameter of the possibly asymmetric loss function used in their forecast decisions.

## 2 Matching the Accuracy Measurements for Forecasts with the Loss Functions

We first introduce some basic notations before illustrating why the summary measures used for forecast accuracy should be determined by the loss functions used by the forecasters. Let  $Y$  be the random variable to be forecasted,  $Z = \{Z_t\}_{t=1}^T = \{Y_t, X_t\}_{t=1}^T$  be the vector of relevant variables (data) used in the  $h$ -period ahead forecast of  $\{Y_t\}_{t=T+1}^{T+h}$ ,  $\hat{Y} = f(Z, \theta)$  be the point forecast, in which  $\theta$  is the unknown vector of parameters in the underlying forecast model, and  $\mathcal{L}(f, Y, Z)$  be the loss function that maps the forecast, outcome and data

into the real line. The forecast decision is to choose  $f(Z, \theta)$  that minimizes the risk

$$\begin{aligned} R(\theta, f) &= E_{Y,Z} [\mathcal{L}(f(Z, \theta), Y, Z)] \\ &= \int_z \int_y \mathcal{L}(f(z, \theta), y, z) p_Y(y|z, \theta) p_Z(z|\theta) dy dz \\ &= \int_z E_Y [\mathcal{L}(f(Z, \theta), Y, Z) | Z, \theta] dz \end{aligned}$$

where  $p_Y(y|z, \theta)$  and  $p_Z(z|\theta)$  are the conditional probability density functions. The classical forecast minimizes

$$E_Y [\mathcal{L}(f(Z, \theta), Y, Z) | Z, \theta] = \int_y \mathcal{L}(f(z, \theta), y, z) p_Y(y|z, \theta) dy \quad (1)$$

given  $Z$  and  $\theta$ . We define the forecast error as  $e = (Y - \hat{Y}) = [Y - f(z, \theta)]$ .

A few of the popular loss functions are

1. Square loss:  $\mathcal{L}(e) = e^2$
2. Absolute loss:  $\mathcal{L}(e) = |e|$
3. Lin-lin loss:  $\mathcal{L}(e) = 2[\tau + (1 - 2\tau)I(e < 0)]|e|$  for  $0 < \tau < 1$  of which the absolute loss is a special case with  $\tau = 0.5$
4. Quad-quad loss:  $\mathcal{L}(e) = 2[\omega + (1 - 2\omega)I(e < 0)](e^2)$  for  $0 < \omega < 1$  of which the squared loss is a special case with  $\omega = 0.5$

Elliott and Timmermann (2008) demonstrate that a forecast that is considered as good using one measure may not be good according to a different measure. The most popular forecasts accuracy measures are the sample  $MSE = \frac{\sum_{i=1}^K e_i^2}{K}$  and  $MAE = \frac{\sum_{i=1}^K |e_i|}{K}$ , which are the sample counterparts of the expected square loss and the expected absolute loss, respectively. Hence, it is natural to use the sample  $MSE$ , or root-mean-square error,  $RMSE = \sqrt{\frac{\sum_{i=1}^K e_i^2}{K}}$ , as a measure of forecast accuracy when the forecasts are generated using the symmetric square loss function and use the sample mean-absolute error,  $MAE$ , for the symmetric absolute loss function. However, one should use the sample mean-weighted-absolute error,

$$MWAE = \frac{\sum_{i=1}^K 2[(w) + (1 - 2w)I(e_i < 0)]|e_i|}{K}$$

as the metric for forecast accuracy for asymmetric lin-lin loss function, and use the sample mean-weighted-square error,

$$MWSE = \frac{\sum_{i=1}^K 2[(w) + (1 - 2w)I(e_i < 0)](e_i^2)}{K}$$

or root mean-weighted-square error,

$$RMWSE = \sqrt{\frac{\sum_{i=1}^K 2[(w) + (1 - 2w) I(e_i < 0)] (e_i^2)}{K}}$$

as the metric for the quad-quad loss function since these metrics are the corresponding sample counterparts of the respective expected loss functions.

## 2.1 Simulated Examples

We perform a few simulations to illustrate that the summary measures used for forecasts accuracy should match the loss functions used in generating the forecasts. Realizations of the forecasted variable  $Y$  are generated by the following data generating processes. (See Koenker and Xiao, 2006.):

**DGP1** Stationary AR(1)

$$\begin{aligned} Q_{y_t}(\tau|y_{t-1}) &= \alpha_0(\tau) + \alpha_1 y_{t-1} \\ &= (2 + 3\Phi^{-1}(\tau)) + 0.6y_{t-1} \end{aligned}$$

where  $Q_{y_t}(\tau|y_{t-1})$  is the conditional quantile function of  $Y$  given  $y_{t-1}$ . The realizations are generated using

$$y_t = (2 + 3\Phi^{-1}(u_t)) + 0.6y_{t-1} \text{ where } u_t \sim \text{i.i.d. } U[0, 1]$$

with  $u_t$  being generated from an independently and identically distributed uniform distribution,  $U[0, 1]$ .

**DGP2** Stationary AR(1)

$$\begin{aligned} Q_{y_t}(\tau|y_{t-1}) &= \alpha_0(\tau) + \alpha_1 y_{t-1} \\ &= (2 + 3\Phi^{-1}(\tau)) + 0.95y_{t-1} \end{aligned}$$

$$y_t = (2 + 3\Phi^{-1}(u_t)) + 0.95y_{t-1} \text{ where } u_t \sim \text{iid } U[0, 1]$$

**DGP3** Near Unit Root

$$\begin{aligned} Q_{y_t}(\tau|y_{t-1}) &= \alpha_0(\tau) + \alpha_1(\tau) y_{t-1} \\ &= (2 + 3\Phi^{-1}(\tau)) + \min\left\{\frac{3}{4} + \tau, 1\right\} y_{t-1} \end{aligned}$$

$$y_t = (2 + 3\Phi^{-1}(u_t)) + \min\left\{\frac{3}{4} + u_t, 1\right\} y_{t-1}$$

where  $u_t \sim \text{iid } U[0, 1]$

**DGP4** Near Unit Root

$$\begin{aligned}
Q_{y_t}(\tau|y_{t-1}) &= \alpha_0(\tau) + \alpha_1(\tau)y_{t-1} \\
&= (2 + 3\Phi^{-1}(\tau)) + \\
&\quad [I(2 + 3\Phi^{-1}(\tau) \geq 0) + 0.8I(2 + 3\Phi^{-1}(\tau) < 0)]y_{t-1}
\end{aligned}$$

$$\begin{aligned}
y_t &= (2 + 3\Phi^{-1}(u_t)) + \\
&\quad [I(2 + 3\Phi^{-1}(u_t) \geq 0) + 0.8I(2 + 3\Phi^{-1}(u_t) < 0)]y_{t-1} \\
&\text{where } u_t \sim \text{iid } U[0, 1]
\end{aligned}$$

**DGP5** Near Unit Root

$$\begin{aligned}
Q_{y_t}(\tau|y_{t-1}) &= \alpha_0(\tau) + \alpha_1(\tau)y_{t-1} \\
&= (2 + 3\Phi^{-1}(\tau)) + [0.95I(2 + 3\Phi^{-1}(\tau) \geq 0) \\
&\quad + 0.8I(2 + 3\Phi^{-1}(\tau) < 0)]y_{t-1}
\end{aligned}$$

$$\begin{aligned}
y_t &= (2 + 3\Phi^{-1}(u_t)) + [0.95I(2 + 3\Phi^{-1}(u_t) \geq 0) \\
&\quad + 0.8I(2 + 3\Phi^{-1}(u_t) < 0)]y_{t-1} \\
&\text{where } u_t \sim \text{iid } U[0, 1]
\end{aligned}$$

The simulations are performed with an in-sample size of  $n = 100$  with the number of 1-period ahead rolling forecasts used in the hold-out sample set at  $K = 400$ . A random sample of  $N = n + K$  realizations from one of the above DGPs is generated. Starting from  $T = n, \dots, N - 1$ , 1-period ahead forecast on  $Y_{T+1}$  is performed using rolling quantile regression of Koenker and Bassett (1978) with  $\tau \in (0.1, 0.2, \dots, 0.8, 0.9)$  or expectile regression in Newey and Powell (1987) with  $\omega \in (0.1, 0.2, \dots, 0.8, 0.9)$  as well as arima  $(p, d, q)$  based on the data  $Z = \{Z_t\}_{t=T-n+1}^T = \{Y_t\}_{t=T-n+1}^T$ . The quantile regression minimizes the sample *MWAE* of the residuals as its loss function while the expectile regression minimizes the sample *RMWSE*. Hence, the *MWAE* should be the appropriate metric to use in measuring the accuracy of forecasts performed using the quantile regressions while *RMWSE* should be used as the metric when forecasts are performed using expectile regressions.

Table (1) presents the sample *RMSE* and *MWAE* for the various DGPs when the quantile regressions and the arima are used to generate the forecast with the asymmetric lin-lin loss function. We can see from Table (1) that the smallest *MWAE* (highlighted as the boxed numbers) of the quantile regression

forecasts occurs at the weight  $w$  that matches the corresponding  $\tau$  of the quantile regressions used in generating the forecasts in general with the exception of the near unit root DGP3, DGP4 and DGP5 where the smallest  $MWAE$  occurs at  $w = 0.3$  for  $\tau = 0.4$ . Hence, if  $MAE$  ( $MWAE$  with  $w = 0.5$ ) is used in assessing the accuracy of the forecasts, only the forecasts performed using the quantile regression with  $\tau = 0.5$  which corresponds to the symmetric lin-lin loss with  $\tau = 0.5$  will be deemed as being optimal. When the forecasters use any of the asymmetric loss function with  $\tau \neq 0.5$ , their forecast performance will be deemed as suboptimal using the  $MAE$  metric. However, if the  $MWAE$  metric with the weight  $w$  that corresponds to the asymmetric parameter  $\tau$  of the lin-lin loss function is used, instead, all the quantile regression forecasts performed using the different degrees of asymmetry determined by  $\tau$  will be deemed as optimal. The smallest  $RMSE$  occurs at  $\tau = 0.5$  for the two stationary DGPs and at  $\tau = 0.4$  for the three near unit root DGPs. Similarly, only the quantile regression forecasts with  $\tau$  close to 0.5 are considered as optimal using the  $RMSE$  metric. The forecasts performed using the arima model is comparable to those performed using the quantile regression with  $\tau$  close to 0.5 when  $RMSE$  is used.

Table (2) presents the sample  $RMSE$  and  $MWSE$  when the expectile regressions are used to generate the forecasts with the asymmetric quad-quad loss function. Again, the smallest  $MWSE$  of the expectile regression forecasts occurs at the weight  $w$  that matches the corresponding  $\omega$  used in generating the forecasts. The results support our recommendation that the summary measures used to measure forecasts accuracy should match the loss functions used in generating the forecasts.

To use the correct weights  $w$  in the  $MWAE$  or  $MWSE$  metrics, one will need to know the weights of the asymmetric loss function used by a forecaster, which are typically unknown. Fortunately, this asymmetric parameter of the loss function can be recovered through the MZ quantile/regression approach that we will introduce next.

### 3 Recovering the Loss Functions via the Mincer-Zarnowitz Prediction-Realization Analysis

In this section, we illustrate how we can recover (back-out) the loss functions used by a forecaster using the quantile regression or expectile regression set in the framework introduced in Mincer and Zarnowitz (1969). If the risk expressed in Equation (1) is differentiable, the first order condition for the minimization of the risk becomes

$$E_Y \left[ \mathcal{L}'(f(Z, \theta), Y, Z) | Z, \theta \right] = \int_y \frac{d}{df} \mathcal{L}(f(z, \theta), y, z) p_Y(y|z, \theta) dy = 0 \quad (2)$$

As mentioned in Elliott and Timmermann (2008), the generalized forecast errors,  $\mathcal{L}'(f(Z, \theta), Y, Z)$ , should be unpredictable and follow a martingale difference sequence given all the information utilized to generate the forecasts.



Table 1: Sample *MWAE* and *RMSE* for the quantile regression and arima fits under the various DGPs. The boxed numbers correspond to the  $\tau$  that yields the smallest *MWAE*.

DGPs	$\tau$	MWAE									RMSE
		w									
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
DGP1	0.1	<b>1.16</b>	1.92	2.68	3.45	4.21	4.97	5.74	6.50	7.26	5.00
	0.2	1.27	<b>1.80</b>	2.34	2.87	3.40	3.94	4.47	5.00	5.54	4.17
	0.3	1.58	1.91	<b>2.25</b>	2.58	2.91	3.25	3.58	3.92	4.25	3.62
	0.4	2.10	2.24	2.37	<b>2.51</b>	2.65	2.78	2.92	3.06	3.19	3.28
	0.5	2.67	2.65	2.63	2.60	<b>2.58</b>	2.56	2.53	2.51	2.49	3.22
	0.6	3.42	3.23	3.04	2.86	2.67	<b>2.48</b>	2.29	2.11	1.92	3.34
	0.7	4.30	3.95	3.61	3.26	2.92	2.57	<b>2.23</b>	1.89	1.54	3.66
	0.8	5.65	5.10	4.55	4.00	3.45	2.90	2.35	<b>1.80</b>	1.25	4.24
	0.9	7.58	6.77	5.97	5.16	4.36	3.55	2.75	1.95	<b>1.14</b>	5.18
		arima	2.61	2.59	2.58	2.56	2.55	2.53	2.51	2.50	2.48
DGP2	0.1	<b>1.00</b>	1.74	2.48	3.22	3.97	4.71	5.45	6.19	6.93	4.81
	0.2	1.11	<b>1.61</b>	2.10	2.60	3.09	3.59	4.08	4.57	5.07	3.87
	0.3	1.36	1.69	<b>2.02</b>	2.35	2.68	3.01	3.34	3.67	4.00	3.39
	0.4	1.72	1.92	2.11	<b>2.31</b>	2.51	2.70	2.90	3.09	3.29	3.14
	0.5	2.30	2.33	2.36	2.39	<b>2.43</b>	2.46	2.49	2.52	2.56	3.02
	0.6	2.90	2.79	2.68	2.57	2.46	<b>2.35</b>	2.24	2.13	2.02	3.04
	0.7	3.83	3.55	3.27	2.99	2.71	2.43	<b>2.16</b>	1.88	1.60	3.32
	0.8	5.44	4.91	4.38	3.85	3.32	2.79	2.26	<b>1.73</b>	1.20	4.00
	0.9	7.16	6.40	5.64	4.88	4.12	3.36	2.60	1.84	<b>1.08</b>	4.82
		arima	2.38	2.39	2.39	2.40	2.41	2.42	2.43	2.44	2.45
DGP3	0.1	<b>3.71</b>	5.77	7.83	9.89	11.95	14.01	16.07	18.13	20.19	13.62
	0.2	4.45	<b>5.29</b>	6.13	6.96	7.80	8.63	9.47	10.30	11.14	9.11
	0.3	5.52	5.40	5.27	5.14	5.01	4.89	4.76	4.63	4.50	7.39
	0.4	5.81	5.52	<b>5.22</b>	<b>4.93</b>	4.64	4.35	4.05	3.76	3.47	7.43
	0.5	6.23	5.82	5.40	4.99	<b>4.58</b>	4.16	3.75	3.34	2.93	7.57
	0.6	6.92	6.35	5.79	5.22	4.65	<b>4.08</b>	3.51	2.95	2.38	7.81
	0.7	7.81	7.07	6.34	5.60	4.87	4.13	<b>3.39</b>	2.66	1.92	8.15
	0.8	9.45	8.47	7.49	6.51	5.53	4.55	3.57	<b>2.59</b>	1.62	8.80
	0.9	11.67	10.40	9.13	7.86	6.59	5.31	4.04	2.77	<b>1.50</b>	9.68
		arima	5.43	5.39	5.35	5.31	5.26	5.22	5.18	5.14	5.10
DGP4	0.1	<b>1.97</b>	3.66	5.34	7.03	8.71	10.40	12.08	13.77	15.45	10.22
	0.2	2.21	<b>3.54</b>	4.86	6.19	7.52	8.84	10.17	11.50	12.82	8.92
	0.3	3.77	4.25	4.73	5.22	5.70	6.18	6.67	7.15	7.63	6.79
	0.4	4.89	4.77	<b>4.66</b>	<b>4.54</b>	4.42	4.30	4.19	4.07	3.95	5.81
	0.5	5.45	5.15	4.85	4.55	<b>4.25</b>	3.95	3.64	3.34	3.04	5.97
	0.6	6.26	5.78	5.30	4.81	4.33	<b>3.85</b>	3.37	2.89	2.41	6.25
	0.7	7.36	6.68	5.99	5.31	4.62	3.94	<b>3.25</b>	2.57	1.88	6.70
	0.8	8.60	7.73	6.85	5.98	5.10	4.23	3.35	<b>2.48</b>	1.60	7.25
	0.9	10.86	9.68	8.51	7.33	6.16	4.98	3.81	2.63	<b>1.46</b>	8.24
		arima	4.67	4.66	4.66	4.65	4.65	4.64	4.64	4.63	4.63
DGP5	0.1	<b>1.59</b>	2.91	4.22	5.54	6.85	8.17	9.48	10.80	12.11	8.00
	0.2	1.74	<b>2.81</b>	3.88	4.94	6.01	7.07	8.14	9.20	10.27	7.08
	0.3	3.13	3.39	3.65	3.91	4.17	4.43	4.69	4.95	5.21	4.91
	0.4	3.66	3.65	<b>3.64</b>	<b>3.63</b>	3.63	3.62	3.61	3.60	3.59	4.55
	0.5	4.19	4.03	3.86	3.69	<b>3.52</b>	3.35	3.18	3.01	2.85	4.62
	0.6	5.00	4.65	4.30	3.95	3.59	<b>3.24</b>	2.89	2.54	2.19	4.86
	0.7	6.14	5.59	5.03	4.48	3.92	3.36	<b>2.81</b>	2.25	1.70	5.31
	0.8	7.29	6.55	5.82	5.09	4.36	3.62	2.89	<b>2.16</b>	1.43	5.80
	0.9	9.10	8.13	7.15	6.18	5.21	4.23	3.26	2.29	<b>1.31</b>	6.66
		arima	3.75	3.73	3.70	<b>3.68</b>	3.65	3.63	3.60	3.58	3.56

Table 2: Sample *MWSE* and *RMSE* for the expectile regression and arima fits under the various DGPs. The boxed numbers correspond to the  $\omega$  that yields the smallest *MWSE*.

DGPs	$\omega$	MWSE									RMSE
		w									
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
DGP1	0.1	<b>4.39</b>	7.03	9.68	12.33	14.98	17.62	20.27	22.92	25.56	3.87
	0.2	4.75	<b>6.40</b>	8.06	9.71	11.37	13.02	14.68	16.33	17.98	3.37
	0.3	5.57	6.61	<b>7.65</b>	8.70	9.74	10.78	11.82	12.86	13.90	3.12
	0.4	6.70	7.25	7.81	<b>8.36</b>	8.92	9.47	10.03	10.59	11.14	2.99
	0.5	8.15	8.27	8.38	8.50	<b>8.61</b>	8.73	8.84	8.96	9.07	2.93
	0.6	10.08	9.75	9.42	9.09	8.76	<b>8.42</b>	8.09	7.76	7.43	2.96
	0.7	12.80	11.96	11.12	10.28	9.44	8.60	<b>7.76</b>	6.92	6.09	3.07
	0.8	17.04	15.54	14.04	12.53	11.03	9.53	8.03	<b>6.53</b>	5.03	3.32
	0.9	25.04	22.46	19.89	17.32	14.74	12.17	9.59	7.02	<b>4.44</b>	3.84
	arima	8.66	8.70	8.74	8.79	8.83	8.87	8.92	8.96	9.00	2.97
DGP2	0.1	<b>4.25</b>	6.93	9.61	12.29	14.96	17.64	20.32	23.00	25.67	3.87
	0.2	4.65	<b>6.28</b>	7.90	9.53	11.15	12.77	14.40	16.02	17.65	3.34
	0.3	5.47	6.49	<b>7.50</b>	8.51	9.52	10.53	11.55	12.56	13.57	3.09
	0.4	6.56	7.11	7.66	<b>8.21</b>	8.75	9.30	9.85	10.40	10.95	2.96
	0.5	7.94	8.08	8.21	8.35	<b>8.49</b>	8.62	8.76	8.89	9.03	2.91
	0.6	9.73	9.46	9.18	8.91	8.63	<b>8.36</b>	8.08	7.80	7.53	2.94
	0.7	12.21	11.47	10.73	9.99	9.26	8.52	<b>7.78</b>	7.05	6.31	3.04
	0.8	16.00	14.67	13.34	12.00	10.67	9.33	8.00	<b>6.67</b>	5.33	3.27
	0.9	23.17	20.87	18.56	16.26	13.96	11.66	9.35	7.05	<b>4.75</b>	3.74
	arima	7.89	8.06	8.22	8.39	8.55	8.72	8.88	9.04	9.21	2.92
DGP3	0.1	<b>28.85</b>	41.21	53.56	65.92	78.27	90.63	102.98	115.34	127.69	8.85
	0.2	31.40	<b>36.70</b>	42.00	47.30	52.60	57.90	63.21	68.51	73.81	7.25
	0.3	36.40	37.94	<b>39.47</b>	41.01	42.54	44.07	45.61	47.14	48.67	6.52
	0.4	42.21	41.21	40.21	<b>39.20</b>	38.20	37.20	36.20	35.19	34.19	6.18
	0.5	48.43	45.54	42.64	39.74	<b>36.85</b>	33.95	31.05	28.16	25.26	6.07
	0.6	55.07	50.63	46.19	41.76	37.32	<b>32.88</b>	28.45	24.01	19.57	6.11
	0.7	62.49	56.67	50.85	45.03	39.21	33.39	<b>27.57</b>	21.75	15.93	6.26
	0.8	72.09	64.79	57.48	50.17	42.86	35.55	28.24	<b>20.93</b>	13.62	6.55
	0.9	88.22	78.77	69.31	59.86	50.40	40.95	31.50	22.04	<b>12.59</b>	7.10
	arima	47.51	44.77	42.03	39.29	36.55	33.81	31.07	28.33	25.59	6.05
DGP4	0.1	<b>38.14</b>	54.45	70.76	87.08	103.39	119.70	136.02	152.33	168.64	10.17
	0.2	46.04	<b>52.48</b>	58.93	65.37	71.82	78.26	84.71	91.15	97.60	8.47
	0.3	56.32	56.96	<b>57.60</b>	58.24	58.87	59.51	60.15	60.79	61.43	7.67
	0.4	66.46	63.37	60.27	<b>57.17</b>	54.07	50.97	47.87	44.78	41.68	7.35
	0.5	75.84	70.17	64.50	58.83	<b>53.16</b>	47.49	41.82	36.15	30.48	7.29
	0.6	84.62	77.03	69.45	61.87	54.29	<b>46.70</b>	39.12	31.54	23.95	7.37
	0.7	93.66	84.45	75.24	66.03	56.82	47.61	<b>38.40</b>	29.19	19.98	7.54
	0.8	104.58	93.70	82.82	71.94	61.06	50.18	39.30	<b>28.42</b>	17.54	7.81
	0.9	120.95	107.89	94.83	81.77	68.71	55.65	42.59	29.53	<b>16.47</b>	8.29
	arima	74.52	69.15	63.79	58.42	53.05	47.68	42.32	36.95	31.58	7.28
DGP5	0.1	<b>9.73</b>	15.40	21.06	26.72	32.38	38.04	43.70	49.36	55.02	5.69
	0.2	11.46	<b>14.49</b>	17.53	20.56	23.59	26.62	29.65	32.69	35.72	4.86
	0.3	14.33	15.63	<b>16.93</b>	18.23	19.53	20.83	22.13	23.43	24.73	4.42
	0.4	17.57	17.64	17.70	<b>17.77</b>	17.83	17.90	17.97	18.03	18.10	4.22
	0.5	20.99	20.09	19.20	18.31	<b>17.42</b>	16.52	15.63	14.74	13.84	4.17
	0.6	24.70	22.98	21.26	19.55	17.83	<b>16.11</b>	14.40	12.68	10.96	4.22
	0.7	29.11	26.59	24.06	21.53	19.01	16.48	<b>13.96</b>	11.43	8.91	4.36
	0.8	34.95	31.52	28.09	24.65	21.22	17.79	14.35	<b>10.92</b>	7.49	4.61
	0.9	44.60	39.88	35.15	30.43	25.70	20.98	16.25	11.53	<b>6.80</b>	5.07
	arima	22.52	21.40	<b>20.27</b>	19.15	18.03	16.91	15.78	14.66	13.54	4.25

This leads to the following sample analog of the orthogonality condition for a one-period ahead forecast:

$$\frac{1}{K} \sum_{t=n}^{n+K-1} \mathcal{L}'(f(z_t, \theta), y_{t+1}, z_t) = 0 \quad (3)$$

In addition,  $\mathcal{L}'(f(z_t, \theta), y_{t+1}, z_t)$  should be uncorrelated with any information used in the forecast at time  $t$ . Hence, another common orthogonality condition used is

$$\frac{1}{K} \sum_{t=n}^{n+K-1} \mathcal{L}'(f(z_t, \theta), y_{t+1}, z_t) v_t = 0 \quad (4)$$

where  $v_t$  is any function of the data,  $\{z_s\}_{s=1}^t$ , available at time  $t$ .

It is well known from the orthogonality condition in Equation (2) that the optimal forecast when minimizing the risk in Equation (1) for the symmetric square loss function is the conditional mean function,  $f(z, \theta) = E(Y|Z = z, \theta)$ , while the conditional median is the optimal forecast for the symmetric absolute loss function. Likewise, the conditional quantile function,  $f(z, \theta, \tau) = F_Y^{-1}(\tau|Z = z, \theta) = Q_\tau(Y|Z = z, \theta)$ , is the optimal forecast for the asymmetric lin-lin loss (see, e.g., Koenker, 2005, pp. 5-6) and the conditional expectile function,  $f(z, \theta) = \mu_\omega(Y|Z = z, \theta)$ , is the optimal forecast for the quad-quad loss function; see, e.g. Newey and Powell (1987, p. 823). Hence, the sample orthogonality condition in Equation (3) yields the sample least-squares regressions for the symmetric square loss function, the sample least-absolute-deviation regressions for the symmetric absolute loss function, the quantile regressions for the lin-lin loss and the expectile regressions for the quad-quad loss function.

Mincer and Zarnowitz (1969) use the forecast,  $\hat{y} = f(z, \theta)$ , as  $v_t$  in the orthogonality condition in Equation (4) for the square-error loss and this boils down to performing an ordinary least squared regression on the following linear model:

$$y_{t+1} = \alpha + \beta \hat{y}_{t+1} + \varepsilon_{t+1} \quad (5)$$

where  $\varepsilon_{t+1}$  is an error term satisfying  $E(\varepsilon_{t+1}|z_t) = 0$ . The unbiasedness and efficiency of the forecast can be evaluated by testing the intercept and slope through the joint hypothesis,

$$H_0 : \alpha = 0 \cap \beta = 1 \quad (6)$$

Optimal forecast is characterized by the upholding of  $H_0$ . Since the conditional quantile is the optimal forecast for the lin-lin loss function, we can perform the  $\tau$  quantile regression for the model in Equation (5) with  $0 < \tau < 1$ , and  $\varepsilon_{t+1}$  such that  $F_{\varepsilon_{t+1}}^{-1}(\tau) = 0$ ; the optimality of the forecast can then be evaluated using the Wald-type test on the joint hypothesis:

$$H_0 : \alpha_\tau = 0 \cap \beta_\tau = 1 \quad (7)$$

where  $\alpha_\tau$  and  $\beta_\tau$  are regression quantile estimates of the intercept and slope coefficients.

Similarly, for the quad-quad loss function, the optimality of the forecast can be tested through the similar joint hypothesis

$$H_0 : \alpha_\omega = 0 \cap \beta_\omega = 1 \quad (8)$$

using the  $\omega$ -expectile regression with  $0 < \omega < 1$ . Assuming the loss function used by a forecaster belongs to one of the flexible lin-lin or quad-quad family, we can recover (back-out) the parameter of asymmetry by selecting the  $\tau$  or  $\omega$  which fails to reject the joint hypothesis in Equation (7) or Equation (8). We coin this as the "MZ quantile/expectile regression approach".

To illustrate the MZ quantile/expectile regression approach, we first generate  $n + K = 300$  observations using DGP1 of Section 2.1. The first  $n = 100$  of these observations are used as the in-sample data which serves as the basis for future forecasts using quantile regression estimates to simulate the lin-lin loss function adopted by a forecaster. The remaining  $K = 200$  observations are used as the 200 realized observations  $\{y_{t+1}\}_{t=1}^K$  in the hold-out sample for which the forecasts are targeting. Next, nine sets of  $K = 200$  forecast observations  $\{\hat{y}_{t+1}\}_{t=1}^K$  are constructed using nine rolling quantile regression fits to the in-sample data, one for each of  $\tau_f \in \{0.1, 0.2, \dots, 0.8, 0.9\}$ , to simulate nine different lin-lin loss functions used by the forecaster, each parameterized by the asymmetry parameter  $\tau_f$ . Hence, we have 200 pairs of predictions and realizations  $\{y_{t+1}, \hat{y}_{t+1}\}_{t=1}^K$  for each of the nine  $\tau_f \in \{0.1, 0.2, \dots, 0.8, 0.9\}$ . To try to recover the actual  $\tau_f$  used in one of the nine asymmetric lin-lin loss functions that generated the forecast,  $\{\hat{y}_{t+1}\}_{t=1}^K$ , we perform nine quantile regressions with  $\tau_{MZ} \in \{0.1, 0.2, \dots, 0.8, 0.9\}$  for Equation (5) using each of the nine sets of predictions and realizations pairs  $\{y_{t+1}, \hat{y}_{t+1}\}_{t=1}^K$ . We expect the quantile regression for Equation (5) with  $\tau_{MZ}$  equals to the corresponding  $\tau_f$  that generated the forecasts to be closest to the 45° line and not reject the null hypothesis in Equation (7).

Figure 1 shows the results of the nine  $\tau_{MZ} \in \{0.1, 0.2, \dots, 0.8, 0.9\}$  quantile regression fits for Equation (5) for four selected  $\tau_f \in \{0.2, 0.4, 0.6, 0.8\}$  quantile regressions used in generating the forecasts. Each plot presents the nine  $\tau_{MZ}$  quantile regression fits for Equation (5) for each of the four  $\tau_f \in \{0.2, 0.4, 0.6, 0.8\}$  quantile regression predictions. The red line represents the 45° line, the green line is the ordinary least-squares regression line, the blue line is the  $\tau_{MZ} = \tau_f$  quantile regression line when the recovered  $\tau_{MZ}$  matches exactly the  $\tau_f$  that is used in generating the forecast while the grey lines are the  $\tau_{MZ}$  quantile regression fits for  $\tau_{MZ} \neq \tau_f$ . Figure 2 shows similar results of expectile regressions for Equation (5) for  $\omega_f \in \{0.2, 0.4, 0.6, 0.8\}$  to illustrate the case when the quad-quad loss function instead of the lin-lin loss function is used in generating the forecasts. Again the red line represents the 45° line, the green line is the ordinary least-squares regression line, the blue line is the  $\omega_{MZ} = \omega_f$  expectile regression line while the grey lines are the  $\omega_{MZ}$  expectile regression fits for  $\omega_{MZ} \neq \omega_f$ . Figure 3 and Figure 4 show the results for the quantile regressions and expectile regressions, respectively, for DGP4.

We can see that both the  $\tau_{MZ}$ -quantile regressions and  $\omega_{MZ}$ -expectile regres-

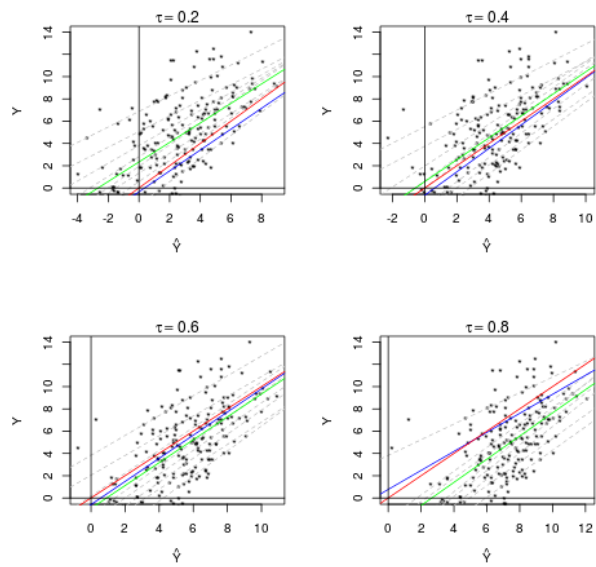


Figure 1: A single realization of the prediction-realization quantile regressions for DGP1.

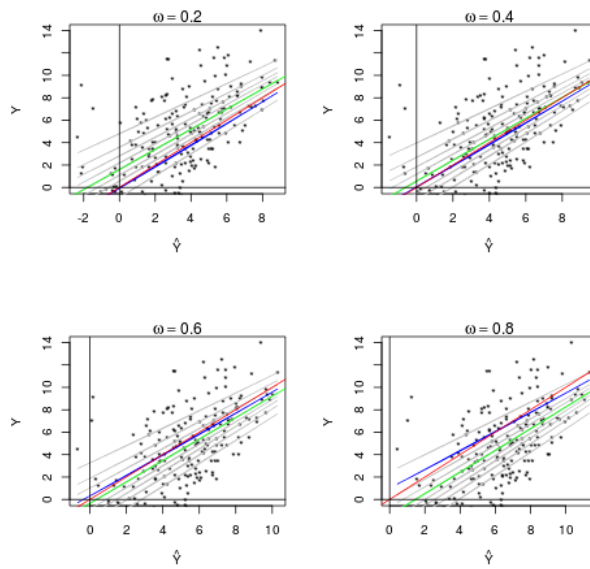


Figure 2: A single realization of the prediction-realization expectile regressions for DGP1.

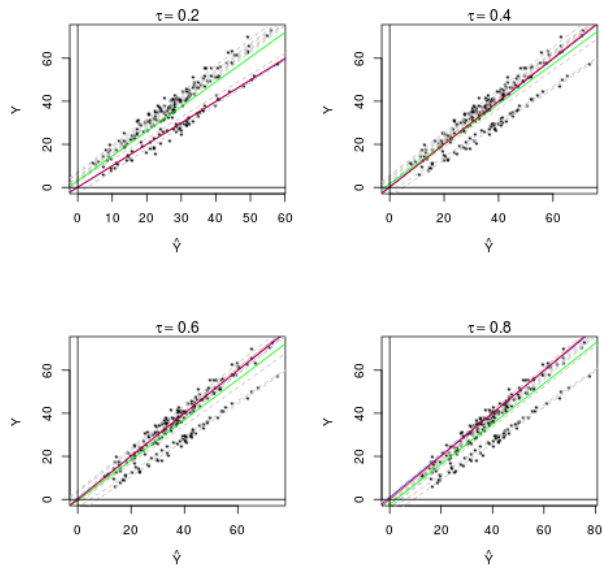


Figure 3: A single realization of the prediction-realization quantile regressions for DGP4.

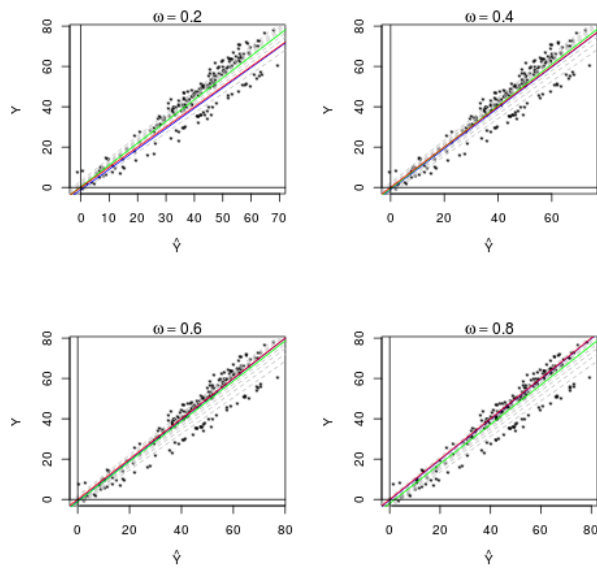


Figure 4: A single realization of the prediction-realization expectile regressions for DGP4.



sions of  $y_{t+1}$  on  $\hat{y}_{t+1}$  tract the  $45^\circ$  line closely when  $\tau_{MZ} = \tau_f$  or  $\omega_{MZ} = \omega_f$  while the ordinary least-squares regression line deviates from the  $45^\circ$  line in general except when  $\tau_f = 0.5$  or  $\omega_f = 0.5$  which are not shown in the figures.

### 3.1 Simulations of Mincer-Zarnowitz Prediction-Realization Optimality Test

To see how the MZ quantile/expectile regression approach performs when applied to the lin-lin and quad-quad loss function, we perform simulations using the model specified in Equation (5) for the various DGPs depicted in Section 2.1. The size of the in-sample data used in the first-stage estimation of the various  $\tau_f (\omega_f) \in [0.1, 0.2, \dots, 0.8, 0.9]$  quantile (expectile) regressions are  $n = [100, 200, 400]$ . These estimated  $\tau_f (\omega_f)$ -quantile (expectile) regressions are subsequently used to perform forecasts for the  $K = N - n = [100, 200, 400]$  hold-out periods.

Each of the subsequent plots in Figure 5 to Figure 9 show the percentage of the  $NMC = 100$  Monte Carlo replications in which the  $H_0$  in (7) is not rejected for the various combinations of  $n$ ,  $K$ ,  $\tau_{MZ}$ ,  $\tau_f$ , and DGPs for the MZ quantile regression approach.

In general, the percentage of non-rejection is the highest when  $\tau_{MZ} = \tau_f$  as we will expect. The sample size  $n$  used in the first-stage in-sample estimation does not have as significant an impact on the percentage of "correct" guess of the weight,  $\tau_f$ , used in the forecasts in the lin-lin loss function as the sample size  $K$  in the second hold-out stage. This is to be expected for the consistency of the test in the second-stage relies on the sample size  $K$  in the second-stage. Results are qualitatively similar for the MZ expectile regression approach. Hence, the MZ quantile/expectile approach that finds the  $\tau_{MZ} (\omega_{MZ})$  quantile (expectile) regression for which  $H_0$  in (7) ((8)) is not being rejected provides a reliable way to recover the weight,  $\tau_f (\omega_f)$ , used in the lin-lin (quad-quad) loss function of the forecaster.

### 3.2 Robustness of the Mincer-Zarnowitz Prediction-Realization Optimality Test to Model Misspecifications

To study how robust is the MZ quantile/expectile regression approach to model misspecifications, we performed two simulations based on the DGPs introduced in Section 2.1: (1) The DGPs are augmented with an additional linear term  $\gamma x_t$  where  $\gamma = 1$ ,  $x_t \sim \chi^2(10)$  is iid from a chi-square distribution with ten degrees of freedom and (2) the DGPs are augmented with an additional nonlinear term  $\gamma x_t^2$  where  $\gamma = 1$ ,  $x_t \sim \chi^2(10)$ . The percentages of non-rejection of  $H_0$  in (7) for quantile regressions without the  $x_t$  in the true models (misspecified model) and those with the augmented  $x_t$  (correctly specified model) for the augmented DGP1 are presented in Figure 10 and Figure 11, respectively while those for the augmented DGP3 are presented in Figure 12 and Figure 13, respectively. The results for the augmented DGP2, DGP4 and DGP5 are similar and, hence, are

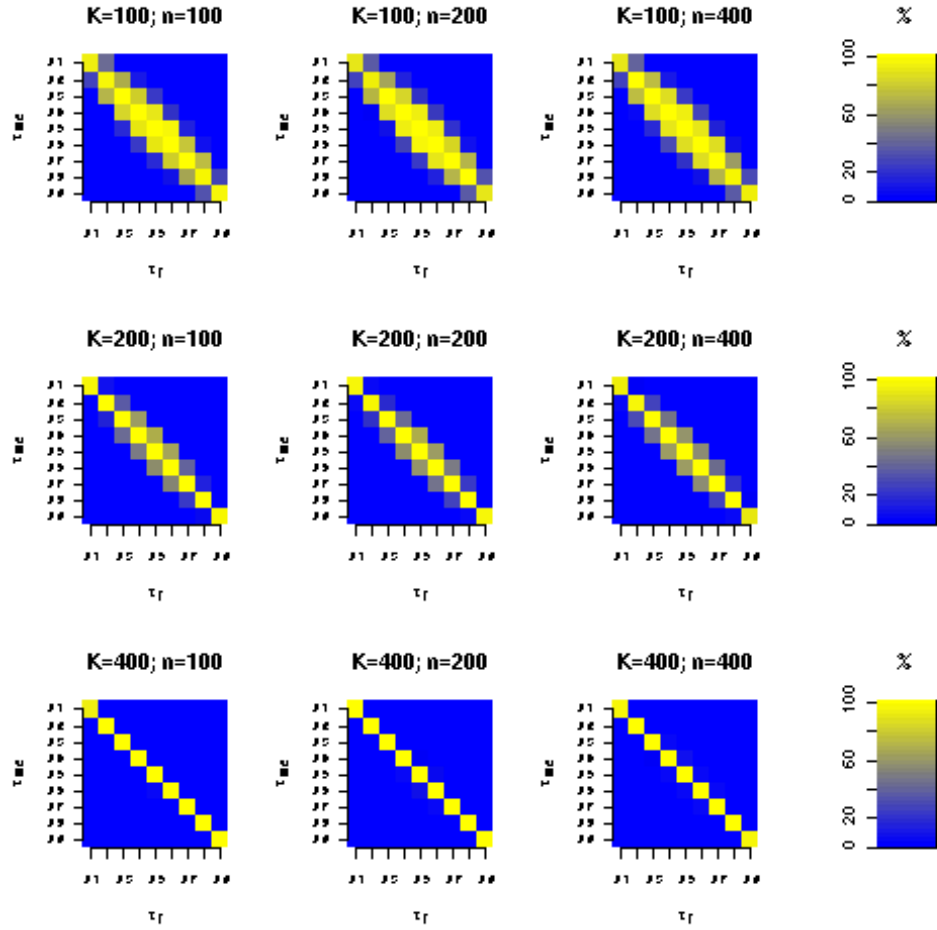


Figure 5: Percentage of the total number of replications ( $NMC$ ) that fails to reject  $H_0$  for  $DGP1$ .

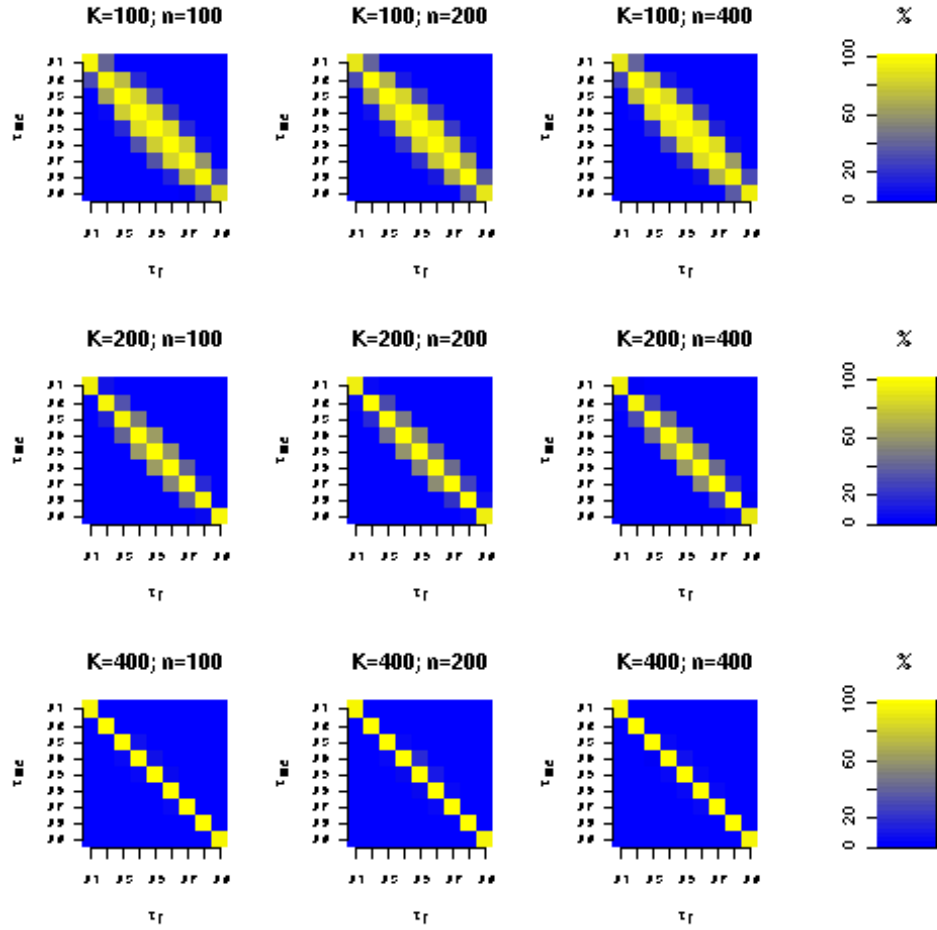


Figure 6: Percentage of the total number of replications ( $NMC$ ) that fails to reject  $H_0$  for  $DGP2$ .

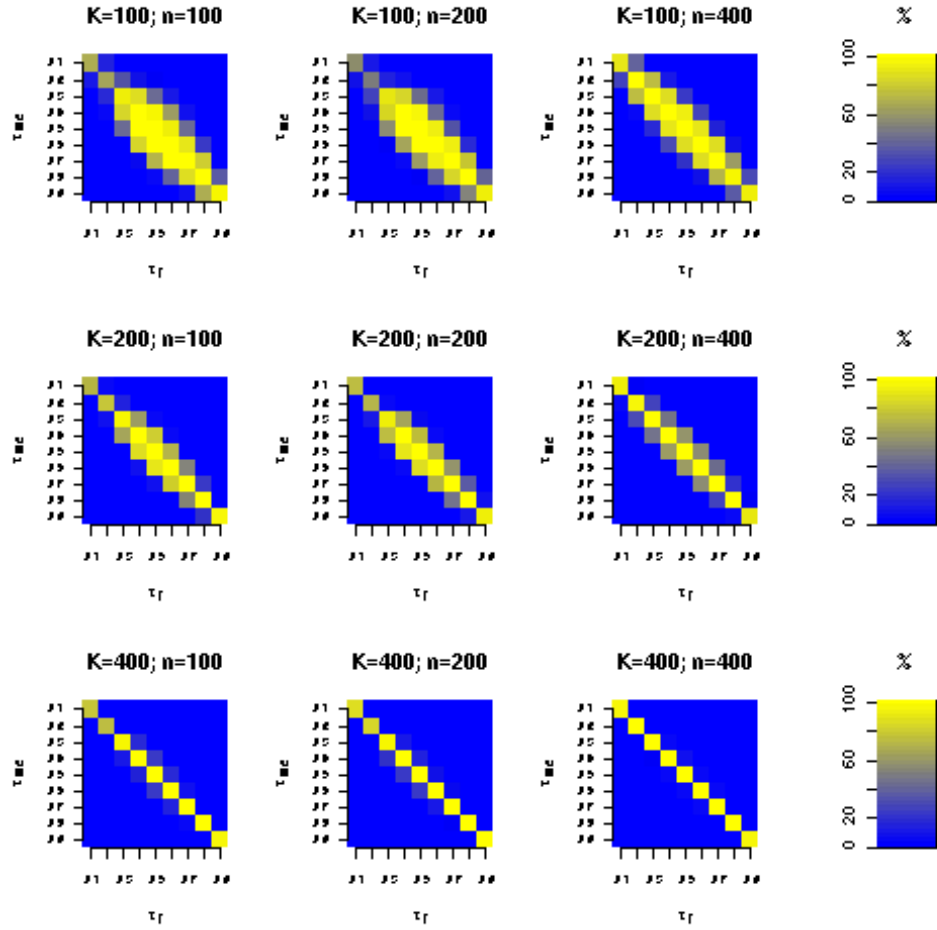


Figure 7: Percentage of the total number of replications ( $NMC$ ) that fails to reject  $H_0$  for  $DGP3$ .

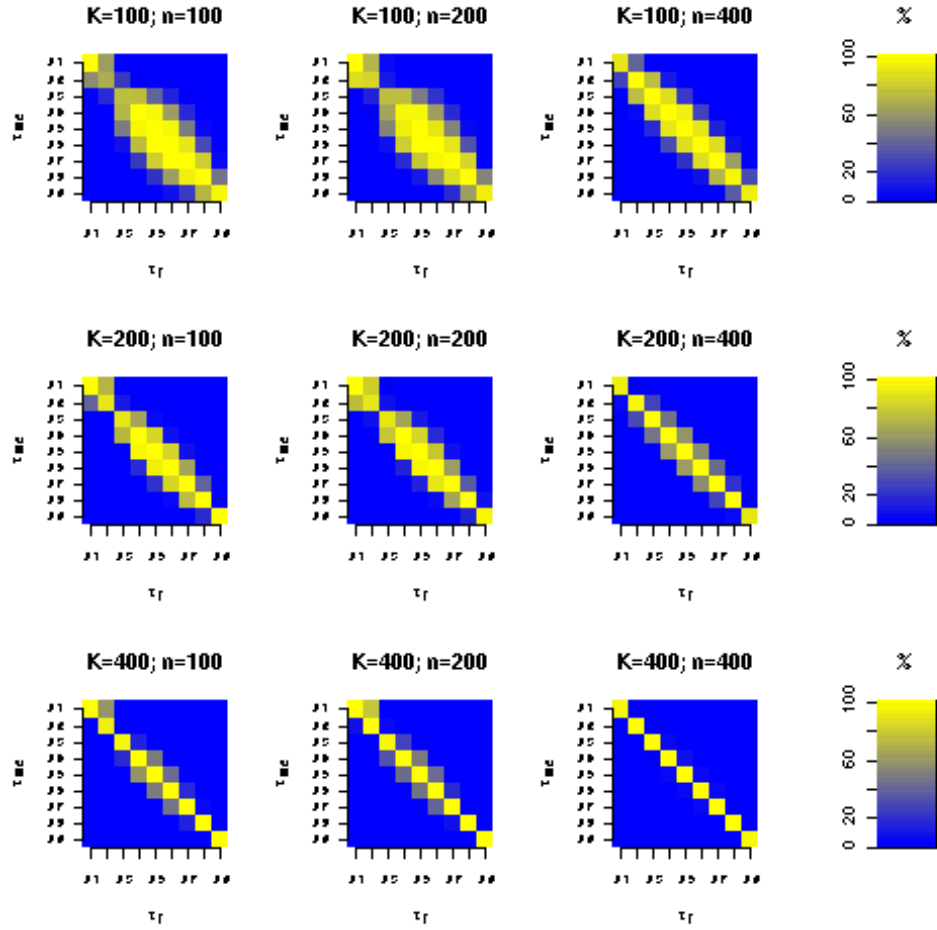


Figure 8: Percentage of the total number of replications ( $NMC$ ) that fails to reject  $H_0$  for  $DGP4$ .

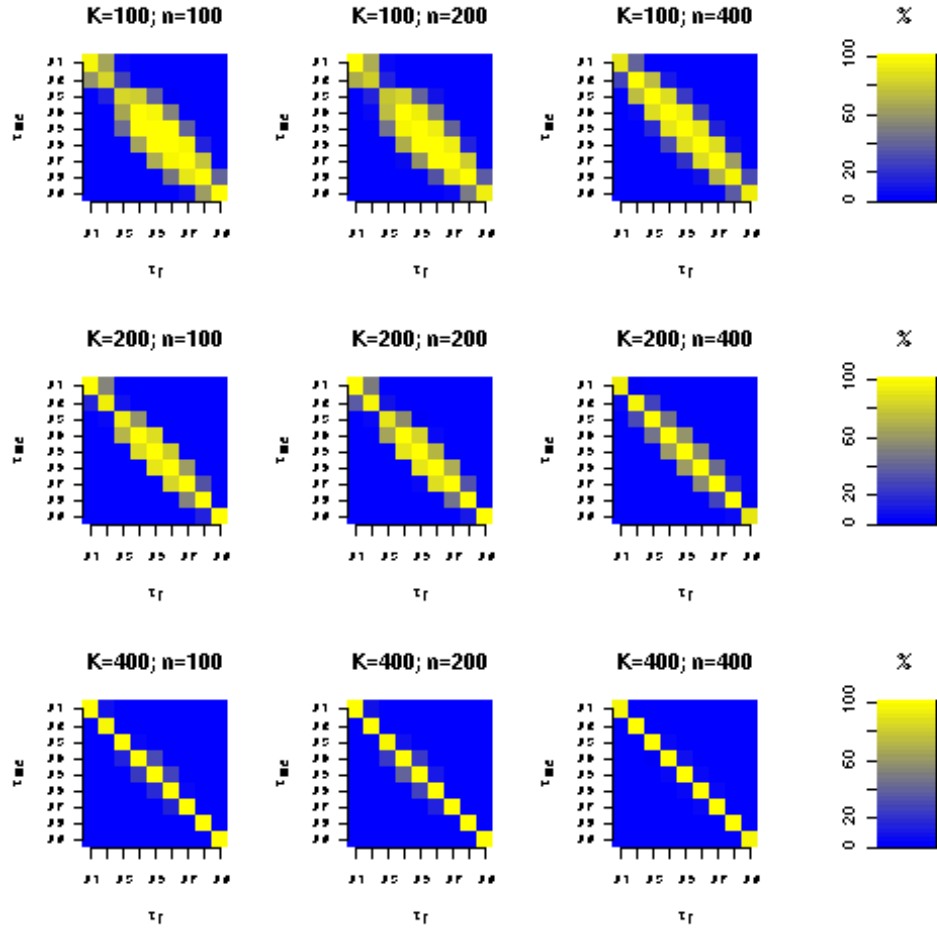


Figure 9: Percentage of the total number of replications ( $NMC$ ) that fails to reject  $H_0$  for  $DGP5$ .

not presented here. The results for the expectile regressions are qualitatively similar and are, hence, not reported here either.

For the second simulation where DGP1 is augmented with the additional non-linear term, the results are presented in Figure 14 and Figure 15, respectively, for the misspecified and correctly specified stage 1 quantile regressions while Figure 16 and Figure 17, respectively, present those for DGP3.

We can see that the MZ quantile/expectile regression approach is quite robust to misspecifications in the regression equation in terms of missing variables for the high percentages of non-rejection cluster around the diagonals when  $\tau_{MZ} = \tau_f$  and  $\omega_{MZ} = \omega_f$ .

## 4 Applications

In this section, we apply the MZ quantile/expectile regression approach to try to recover the asymmetric parameter in the potentially asymmetric loss functions used in the Federal Reserve's forecasting of economic variables and the forecasting of an electronic component manufacturer's demand in its supply chain.

### 4.1 The Federal Reserve's Greenbook Forecast

The report "Current Economical and Financial Conditions: Summary and Outlook" prepared by the Board of Governors of the Federal Reserve System for the Federal Open Market Committee Meeting is viewed by many as the authoritative forecast of the important economic variables such as GDP, inflation rate, unemployment rate, etc. It is also known as the "Greenbook". Many studies have been done using the Greenbook forecasts and found that the forecasts were not optimal. The question though is whether the Greenbook forecasts are really suboptimal or they are in fact optimal for an asymmetric loss function used by the Federal Reserves.

Similar to Patton and Timmerman (2007), we use the real *GDP* and the Federal Reserve's forecast from the first quarter of 1969 to the first quarter of 2000. Figure 18 presents the recovered asymmetric parameter  $\tau$  used in the Federal Reserve's Greenbook forecasts using the quantile regression approach while Figure 19 shows the asymmetric parameter  $\omega$  recovered using the expectile regression. The ordinary least squares regression reject the null hypothesis in (6) and, hence, the Greenbook's forecast is deemed not optimal if we assume the Federal Reserves actually used the square loss decision function. The MZ quantile regression approach fails to reject the null in (7) for  $\tau \in \{.30, .35, .40, .45\}$  while the expectile regression fails to reject the null in (8) for  $\omega \in \{.35, .40, .45\}$ . Using the quantile regression estimates, the ratio of optimal loss on negative errors to positive errors,  $\mathcal{L}(-e)/\mathcal{L}(e) = (1 - \tau)/\tau$  is between 1.22 and 2.33 while the ratio derived from the expectile regression estimates falls between 1.22 and 1.86. Patton and Timmerman's (2007) average estimated ratio is 1.44 with

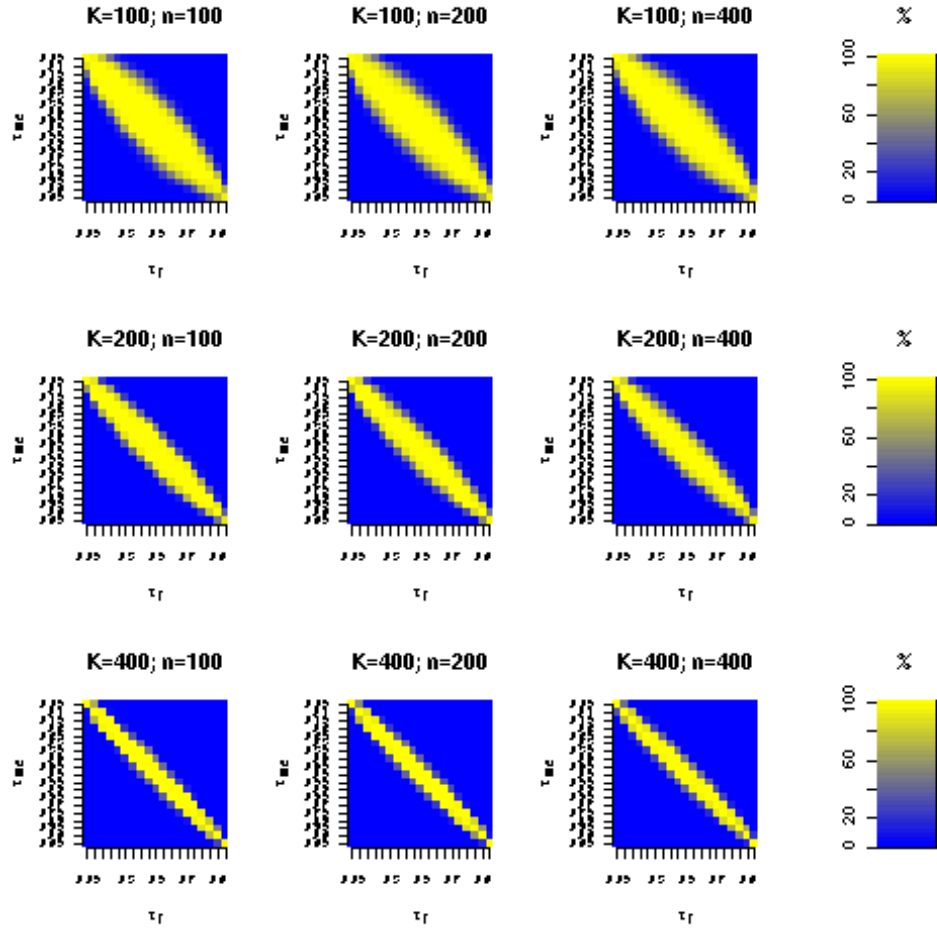


Figure 10: Percentage of non-reject of  $H_0$  for the misspecified quantile regressions for DGP1 augmented with an linear term  $\gamma x_t$ .



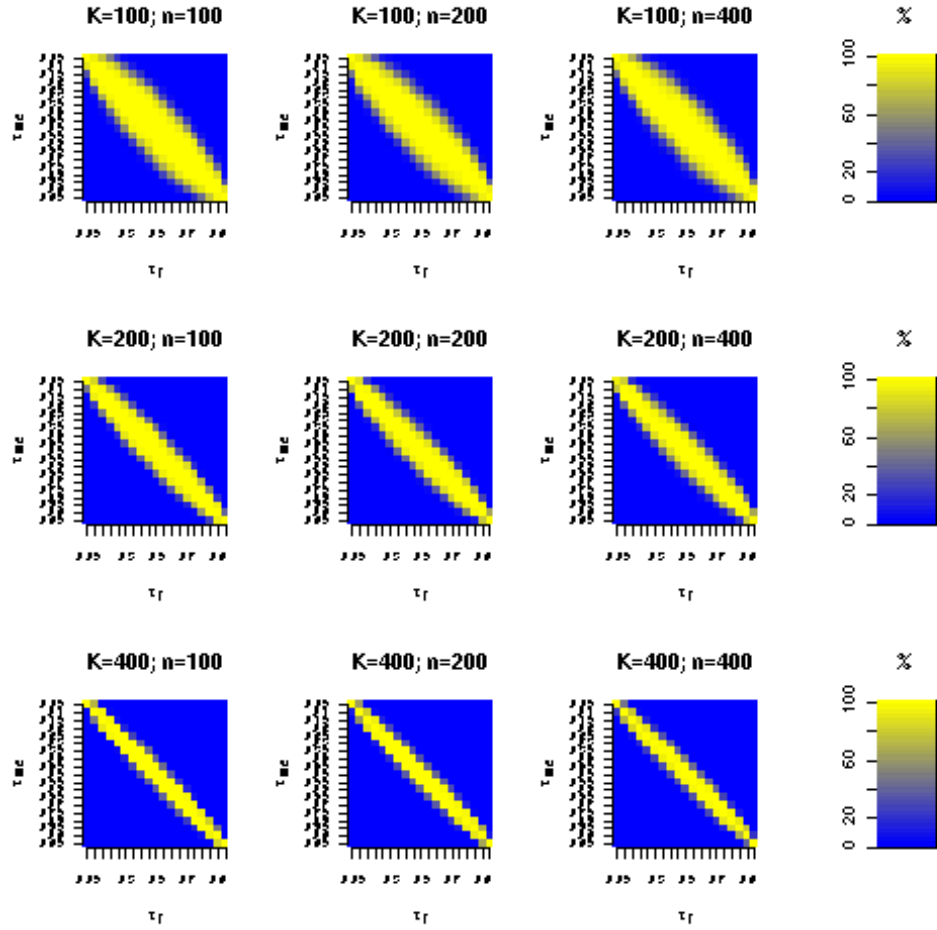


Figure 11: Percentage of non-reject of  $H_0$  for the correctly specified quantile regressions for DGP1 augmented with an linear term  $\gamma x_t$ .

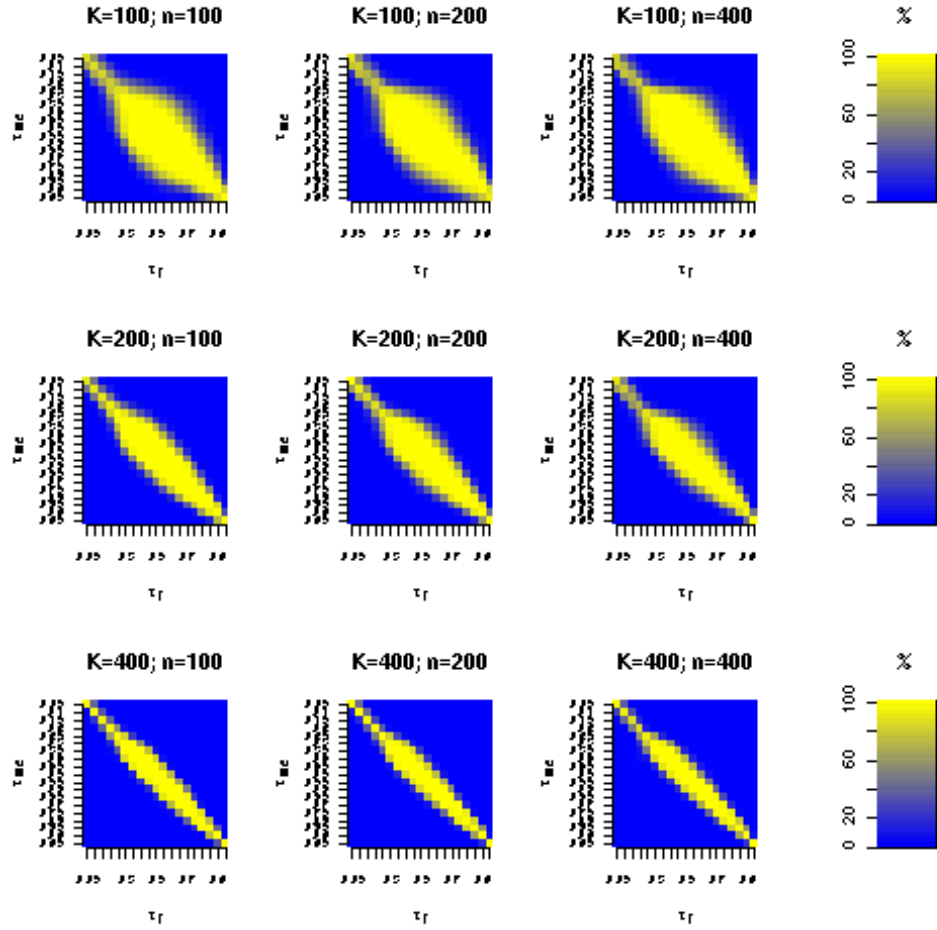


Figure 12: Percentage of non-reject of  $H_0$  for the incorrectly specified quantile regressions for DGP3 augmented with an linear term  $\gamma x_t$ .

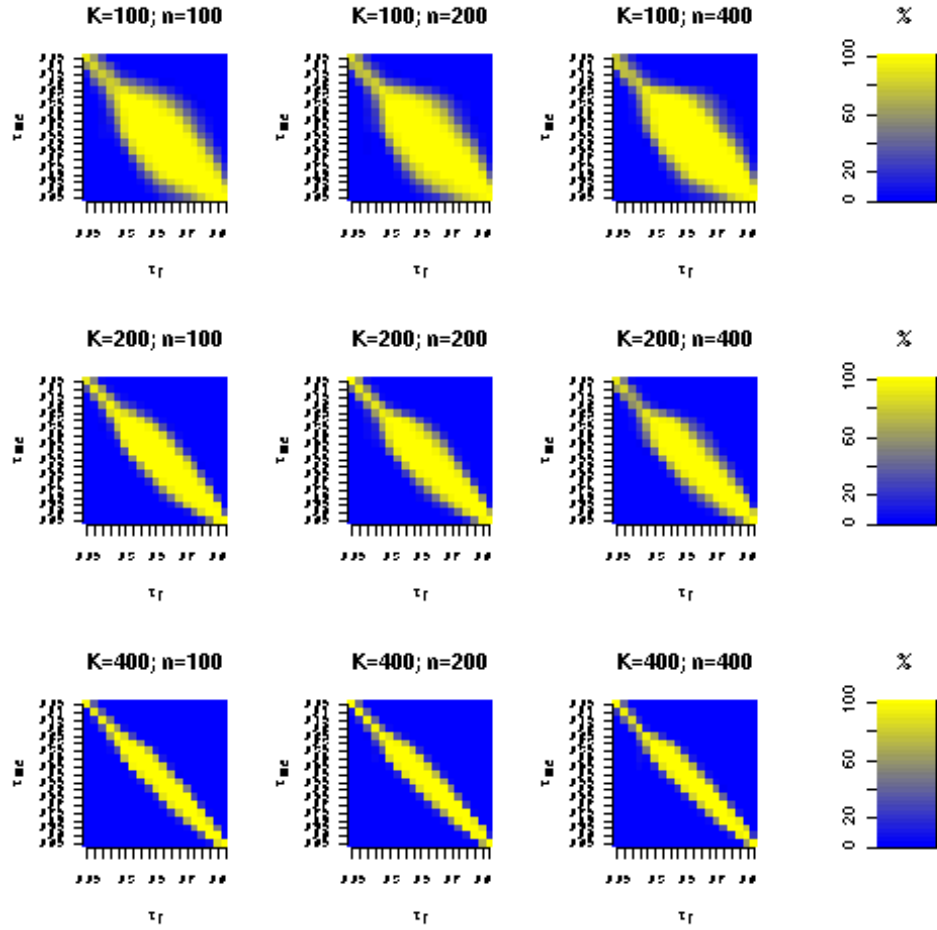


Figure 13: Percentage of non-reject of  $H_0$  for the correctly specified quantile regressions for DGP3 augmented with an linear term  $\gamma x_t$ .

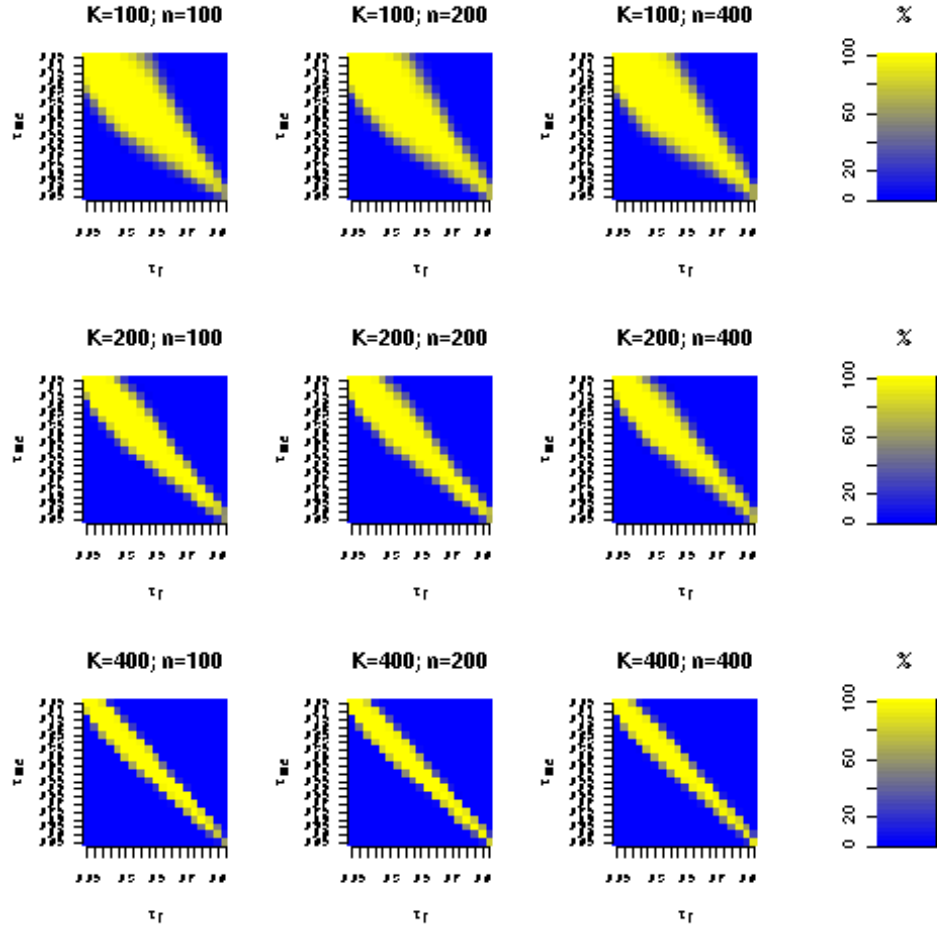


Figure 14: Percentage of non-reject of  $H_0$  for the misspecified quantile regressions for DGP1 augmented with a nonlinear term  $\gamma x_t^2$ .

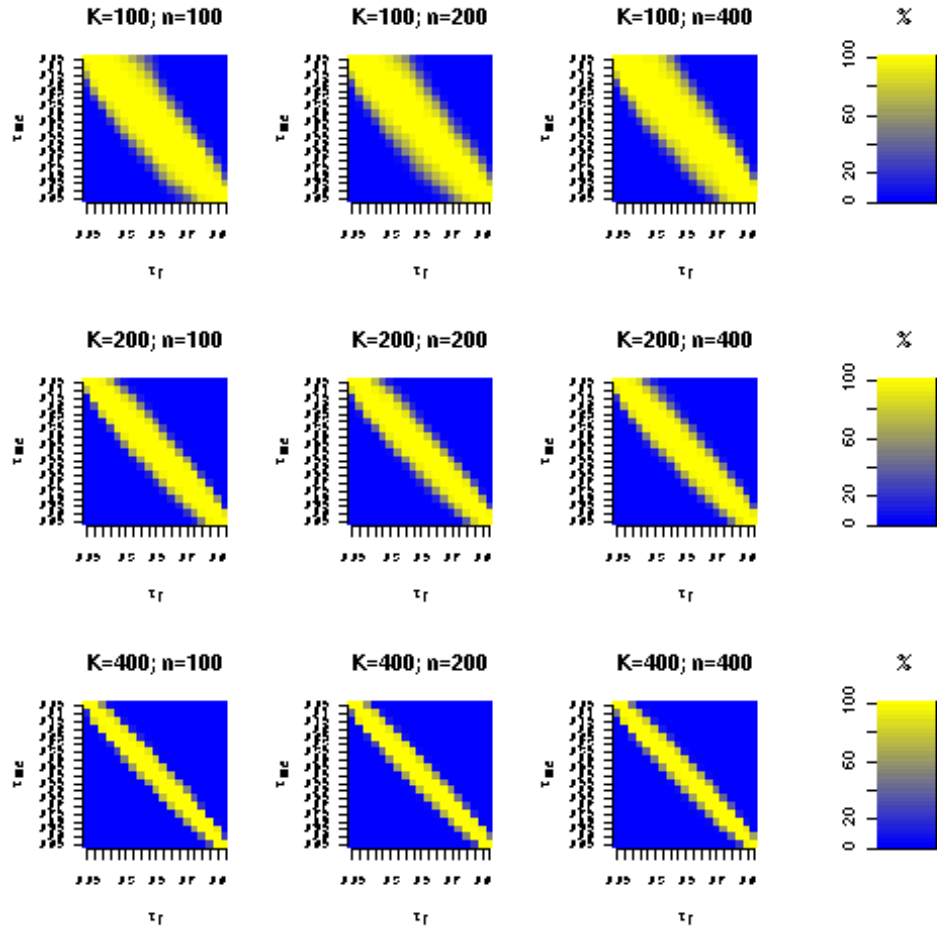


Figure 15: Percentage of non-reject of  $H_0$  for the correctly specified quantile regressions for DGP1 augmented with a nonlinear term  $\gamma x_t^2$ .

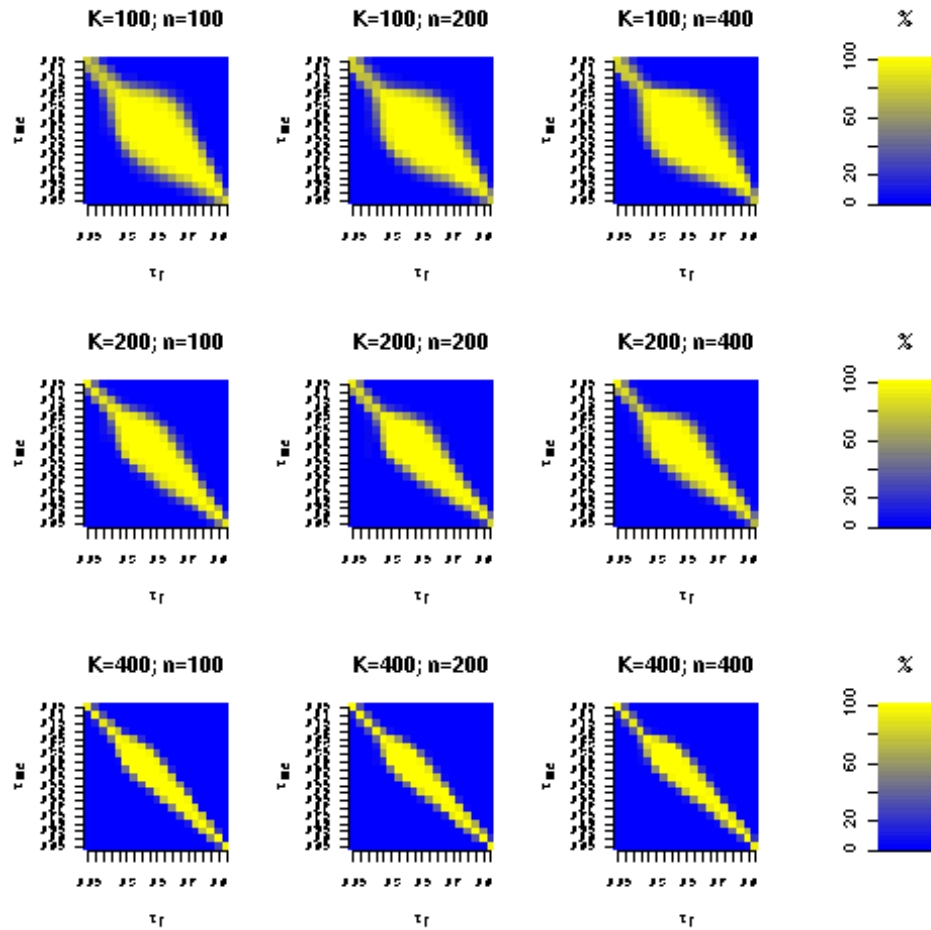


Figure 16: Percentage of non-reject of  $H_0$  for the misspecified quantile regressions for DGP3 augmented with a nonlinear term  $\gamma x_t^2$ .

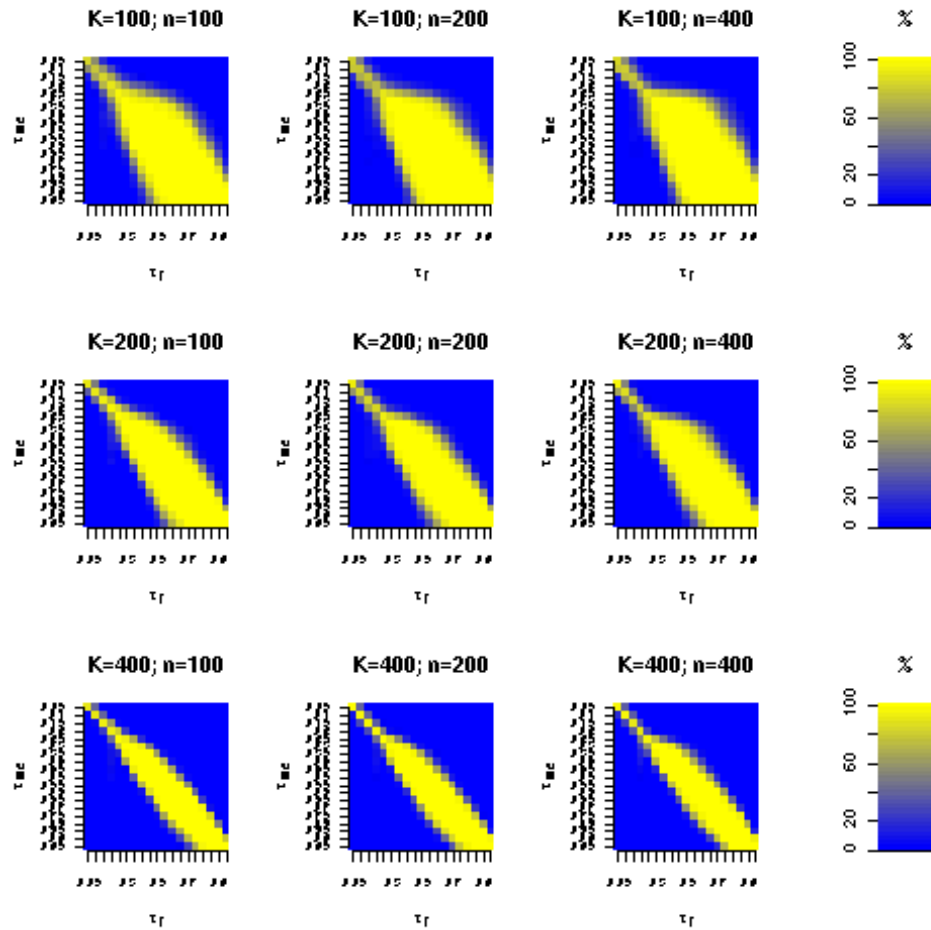


Figure 17: Percentage of non-reject of  $H_0$  for the correctly specified quantile regressions for DGP3 augmented with a nonlinear term  $\gamma x_t^2$ .

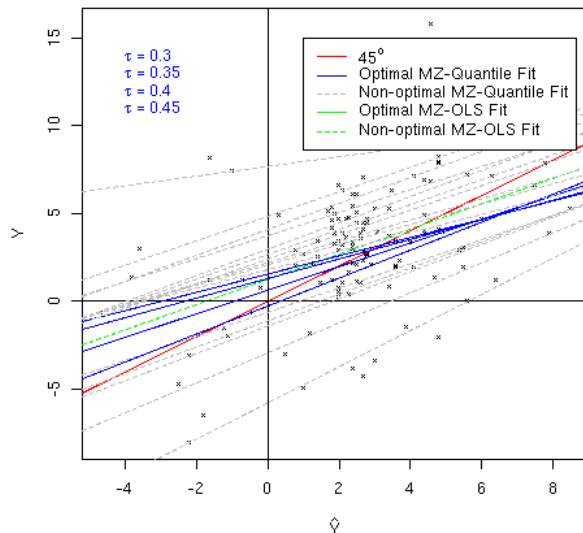


Figure 18: Mincer-Zarnowitz quantile regression estimates of the asymmetric parameter used in the Federal Reserve’s Greenbook forecast.

a minimum of 0.52 and a maximum of 2.76 when the loss function does not depend on the forecasts, and the average is 3.48, 1.97, and 1.26 for the 0.25, 0.5, and 0.75 quantiles of the real GDP growth when the loss function depends on the forecasts. So the general picture of the inference we arrive at using the MZ quantile/expectile regression approach is consistent with what Patton and Timmermman (2007) have found that the Federal Reserves appears to value over prediction roughly 1.5 times more costly than under prediction.

## 4.2 An Electronic Component Manufacturer Supply Chain

We have data on the actual shipment ( $y_{t+h}$ ) and one-month through twelve-month ahead ( $h = 1, \dots, 12$ ) forecasts ( $\hat{y}_{t+h}$ ) for 106 stock-keeping units (SKU) in 19 product families (FAM) processed at 4 global business units (GBU) and 6 distribution hubs (HUB) across 3 regions (REG) over the globe of an electronic component manufacturer in the U.S. We only analyze the combinations of SKU, FAM, GBU, HUB and REG with at least 30 usable observations. There are altogether 90 such combinations. Among these combinations, the incomplete observations with zero actual shipment are dropped from the analysis even though there are positive forecasts.



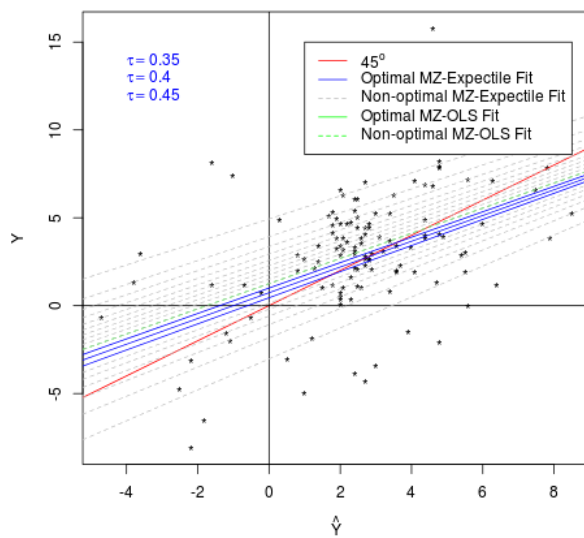


Figure 19: Mincer-Zarnowitz expectile regression estimates of the asymmetric parameter used in the Federal Reserve’s Greenbook forecast.

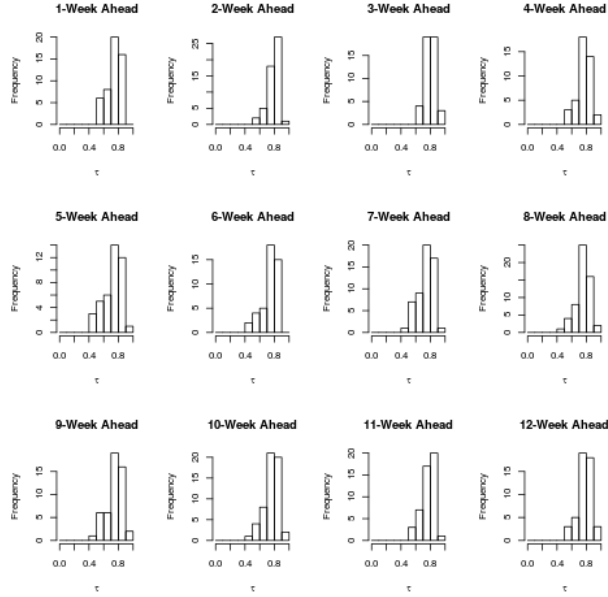


Figure 20: Histogram of the asymmetric parameter of the lin-lin loss function recovered by the MZ quantile regression approach.

The recovered asymmetric parameter of the lin-lin loss function and quad-quad loss function at the one-month through twelve-month forecast horizons using the MZ quantile/expectile approach are presented in the histograms in Figure 20 and Figure 21, respectively. We can see from both figures that the recovered asymmetric parameters for both loss functions have values scattered around  $\tau = 0.8$  and  $\omega = 0.8$  for all the different forecast horizons. The global median of the median  $\tau$  at each of the twelve forecast horizons is 0.8 and so is the global median of the median  $\omega$  across the various horizons. The global mean of the mean  $\tau$  at the various forecast horizons is 0.78 while the global mean of the mean  $\omega$  is 0.76. This suggests that the manufacturer weighs positive forecast errors (under forecasts) about four times ( $\mathcal{L}(e)/\mathcal{L}(-e) = \tau/(1-\tau) = 0.8/0.2 = 4$ ) as costly as negative forecast errors (over forecasts). Under forecasts will lead to unfulfilled order, which could lead to loss of goodwill, reputations, and loss of current and future sales while over forecasts will result in over inventory, higher insurance costs, and tied up capital. In this case, the manufacturer views under forecasts more costly than over forecasts.

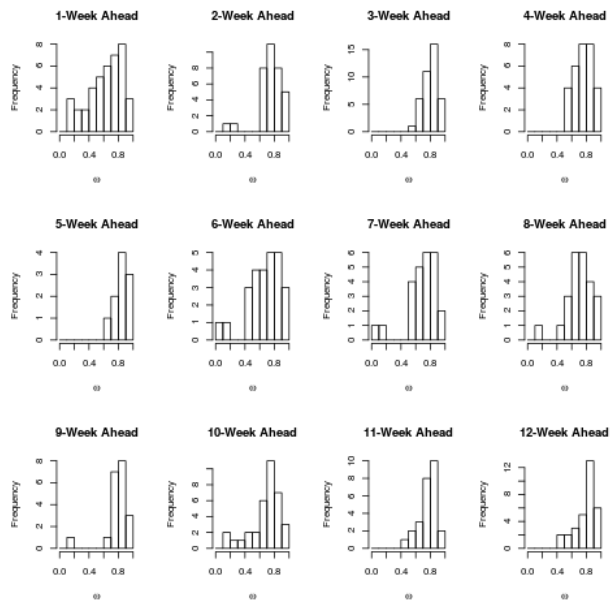


Figure 21: Histogram of recovered asymmetric parameter of the quad-quad loss function recovered by the MZ expectile regression approach.

## 5 Conclusions

Forecast is ubiquitous in all areas of daily life. Typical summary metrics used in measuring forecast performance are the sample root-mean-square error and the mean-absolute error. However, these popular summary matrices are appropriate only for the symmetric mean-square error and mean-absolute error loss functions used by the forecasters. We have argued and demonstrated that when forecasters use an asymmetric loss function such as the lin-lin or quad-quad lost functions, the appropriate summary metrics for measuring forecast performance should, instead, be the sample mean-weighted-absolute error and mean-weighted-square error, respectively, that reflect the different weights assigned to over and under prediction. However, to correctly utilize the sample mean-weighted-absolute error and mean-weighted-square error, one will need to know the weights for over and under prediction that a forecaster assigned when performing the forecasts. These weights that characterize the asymmetric loss functions can be recovered in the recommended MZ quantile/expectile regression approach.

On the evaluation of forecast optimality, we provide theoretical justification for extending Mincer and Zarnowitz (1969) prediction-realization framework that is based on the ordinary least-squares regression to one that uses the quantile regressions and expectile regressions. Simulation results demonstrating the efficacy of the MZ quantile/expectile regression approach are provided and show that this approach is robust to specific forms of model misspecification in the data generating process.

The MZ quantile/expectile regression approach is then applied to Federal Reserve's Greenbook forecast and the forecast of an electronic component manufacturer's demand in its supply chain. We have found that the Federal Reserve appeared to value over prediction roughly 1.5 times more costly than under prediction. We have also found that the electronic component manufacturer appeared to weigh under forecast about four times as costly as over forecast.

## References

- [1] Ali, Ashiq, April Klein, and James Rosenfeld (1992). "Analysts' Use of Information about Permanent and Transitory Earnings Components in Forecasting Annual EPS." *The Accounting Review*, 67(1):183-198.
- [2] Artis, Michael, and Massimiliano Marcellino. (2001). "Fiscal Forecasting: The Track Record of the IMF, OECD and EC." *Econometrics Journal*, 4(1): S20-36.
- [3] Batchelor, Roy, and David A. Peel. (1998). "Rationality Testing under Asymmetric Loss." *Economics Letters*, 61(1): 49-54 .
- [4] Capistran, Carlos Carmona. 2006. "Bias in Federal Reserve Inflation Forecasts: Is the Federal Reserve Irrational or Just Cautious?" Banco de Mexico Working Paper, no.2006-14.

- [5] Christoffersen, Peter F., and Francis X. Diebold. (1997). "Optimal Prediction under Asymmetric Loss." *Econometric Theory*, 13(6): 808–17.
- [6] Cohen, M., T. Ho, Z. Ren and C. Terwiesch. (2003). "Measuring imputed cost in the semiconductor equipment supply chain." *Management Science*, 49(12): 1653–1670.
- [7] Elliott, Graham, Ivana Komunjer, and Allan Timmermann. (2005). "Estimation and Testing of Forecast Rationality under Flexible Loss." *Review of Economic Studies*, 72(4): 1107–25.
- [8] Elliott, Graham and Allan Timmermann. (2008). "Economic Forecasting." *Journal of Economic Literature*, 46(1): 3–56.
- [9] Granger, Clive W. J. (1969). "Prediction with a Generalized Cost of Error Function." *Operations Research*, 20(2): 199–207.
- [10] Granger, Clive W. J. (1999). "Outline of Forecast Theory Using Generalized Cost Functions." *Spanish Economic Review*, 1(2): 161–73.
- [11] Granger, Clive W. J., and Paul Newbold. (1986). *Forecasting Economic Time Series*. Second edition. Orlando, Fla.; London; Sydney and Toronto: Harcourt, Brace, Jovanovich; Academic Press.
- [12] Granger, Clive W. J., and M. Hashem Pesaran. (2000). "Economic and Statistical Measures of Forecast Accuracy." *Journal of Forecasting*, 19(7): 537–60.
- [13] Ito, Takatoshi. (1990). "Foreign Exchange Rate Expectations: Micro Survey Data." *American Economic Review*, 80(3): 434–49.
- [14] Koenker, Roger and Gilbert Bassett. (1978). "Regression Quantiles." *Econometrica*, 46(1): 33–50.
- [15] Koenker, Roger and Zhijie Xiao. (2006). "Quantile Autoregression." *Journal of the American Statistical Association*, 101: 980–990.
- [16] Lee, H., V. Padmanabhan and S. Whang. (1997). "Information distortion in a supply chain: The bullwhip effect." *Management Science*, 43(4): 546–588.
- [17] Mincer, Jacob, and Victor Zarnowitz. (1969). "The Evaluation of Economic Forecasts." In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, ed. J. Mincer. New York: National Bureau of Economic Research, 14–20.
- [18] Newey, W. K. and J. L. Powell. (1987). "Asymmetric least squares estimation and testing." *Econometrica*, 55:819–847.

- [19] Patton, Andrew J., and Allan Timmermann. (2007). “Properties of Optimal Forecasts under Asymmetric Loss and Nonlinearity.” *Journal of Econometrics*, 140(2): 884–918.
- [20] Pesaran, M. Hashem, and Spyros Skouras. (2002). Decision-Based Methods for Forecast Evaluation. In *A Companion to Economic Forecasting*, ed. M. Clements and D. Hendry. Malden, Mass. and Oxford: Blackwell, 241–67.
- [21] Pruitt, Stephen W. and Lawrence J. Gitman. (1987). “Capital Budgeting Forecast Biases: Evidence from the Fortune 500.” *Financial Management*, 16(1): 46-51.
- [22] Silver, Nate. (2012). *The Signal and the Noise: Why So Many Predictions Fail - but Some Don't*. Penguin Books.
- [23] Varian, Hal R. (1974). “A Bayesian Approach to Real Estate Assessment.” In *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, ed. S. E. Fienberg and A. Zellner. Amsterdam: North-Holland, 195–208.
- [24] Weiss, Andrew A. (1996). “Estimating Time Series Models Using the Relevant Cost Function.” *Journal of Applied Econometrics*, 11(5): 53 9–60.
- [25] West, Kenneth D., Hali J. Edison, and Dongchul Cho. (1993). “A Utility-Based Comparison of Some Models of Exchange Rate Volatility.” *Journal of International Economics*, 35 (1–2):23–45 .
- [26] Zellner, Arnold. (1986). “Bayesian Estimation and Prediction Using Asymmetric Loss Functions.” *Journal of the American Statistical Association*, 81(394): 44 6–51.