

Some implications of a quartic loss function

Kevin Aretz
Lancaster University

David Peel
Lancaster University

Abstract

Motivated by a central banker with a symmetric but non-quadratic loss function, we show in this note that the approximations of two plausible loss functions of this type will include a quartic term. For skewed distributions, we establish that such a loss function implies a systematic inflation bias even when the bank targets the natural rate. Moreover, we show that the weights in an optimal combination of forecasts will differ from that under quadratic loss. We illustrate these differences using simulated data and data from the Livingston Surveys of Professional Forecasters.

Citation: Aretz, Kevin and David Peel, (2007) "Some implications of a quartic loss function." *Economics Bulletin*, Vol. 7, No. 13 pp. 1-7

Submitted: June 18, 2007. **Accepted:** August 20, 2007.

URL: <http://economicsbulletin.vanderbilt.edu/2007/volume7/EB-07G10011A.pdf>

1 Introduction

We consider some of the implications of the assumption that the loss function of a policy maker is symmetric, but non-quadratic. Two loss functions which exhibit this property have been proposed in the literature. The first is a target zone loss function where deviations outside a band from target are penalized at an increasing rate. This objective function might represent that of agents who are particularly concerned to avoid large forecast errors. This might apply to some central bankers, who have a symmetric, but non-quadratic, inflation zone target. For instance, in case of the U.K. the Bank of England Governor is obliged to write an open letter to the Chancellor, if inflation deviates more than 1.0% on either side of the 2.0% target. Boinet and Martin (2006) show that Taylor rules based on such a formulation can explain UK monetary experience. Alternatively, Bray and Goodhart (2002) suggest that the policy makers' loss function is symmetric, but may be bounded from above, as "the worst penalty which can be applied to these agents is to sack them, if they are perceived to have failed. To be publicly sacked as a failure is painful, often severely so, but the pain is finite".

We show below that a quartic term appears in the approximation of both these loss functions. As a consequence, when the distribution of the target variable is non-normal and skewed, an optimal inflation bias can occur. Consistent with the rest of the literature, we define bias as the difference between the average inflation rate and the target. Elliott and Timmermann (2004) offer a general proof of the proposition that, if forecast errors are normally distributed, the weights on the forecasts in an optimal combination will be the same under a variety of symmetric and asymmetric loss functions. We offer a more direct proof of this for the quartic loss function. Moreover, by means of a simulation exercise and data on inflation forecasts we show how the optimal bias and the weights in the optimal combination of forecasts under quartic loss differ from that under quadratic loss.

2 Analysis

Consider the target zone loss function L , as suggested by Boinet and Martin (2006):

$$L = \frac{e^{\alpha c^2} - \alpha c^2 - 1}{2\alpha^2} = \frac{1}{4}c^4 + O(c^6), \quad (1)$$

where c equals the outcome of $\tilde{c} = \tilde{\pi} - \pi^*$, with $\tilde{\pi}$ being a random variable, e.g., inflation, and π^* the forecast or target. As seen in Figure 1, this loss function describes a target zone model, and from formula (1) it can be approximated by a quartic function to order five.

Consider the Bell shaped loss function of Bray and Goodhart (2002):

$$L = \frac{1 - e^{-kc}}{k} = c^2 - 0.5kc^4 + O(c^6), \quad (2)$$

We draw this loss function in Figure 2. As can be seen from formula (2), this loss function can also be approximated to order five. For simplicity, we concentrate from now on the target loss function. Note, however, that our conclusions on the loss function proposed by Bray and Goodhart (2002) are qualitatively similar, with the important caveat that skewness of the error

term will imply an alternative bias to that in the target zone approximation.

From formula (1), the expected loss, $E(\tilde{L})$, will be minimized at:

$$E(\tilde{c}^3) = E[(\tilde{\pi} - E(\tilde{\pi})) + (E(\tilde{\pi}) - \pi^*)]^3 \quad (3)$$

$$= \sigma_3 + 3\sigma^2[E(\tilde{\pi}) - \pi^*] + [E(\tilde{\pi}) - \pi^*]^3 = 0, \quad (4)$$

where $E(\tilde{\pi})$, σ^2 , and σ_3 are the expected value, the variance, and the skewness of $\tilde{\pi}$, respectively.

In the context of modelling central bank outcomes, we observe from formula (4) that, if the inflation distribution exhibits positive (negative) skewness, then a central bank will exhibit a systematic negative (positive) inflation bias even when it targets the natural rate. In this context, it is interesting that Ruge-Murcia (2000) has documented a deflationary bias of central banks who have an inflation zone target. Further prominent central bankers have suggested that their own banks do not target output above the natural rate, as assumed in the standard model (see, e.g., Svensson, 1997). For instance, see Vickers (1998, p.369) or Blinder (1998, p.48) for such sentiments. However, if the inflation distribution exhibits skewness, an inflationary bias can result under quartic loss. In this context, Goodhart's (2001) comments on the Bank of England Monetary Committee are interesting. He writes: "But in either case the existence of a skew would affect our decision on the appropriate interest rate. Unlike uncertainty and variance, skew and risk mapped directly into the interest rate decision."

Moreover, if the errors are non-normal, the optimal combination of forecasts will exhibit weights which differ from those obtained under quadratic loss. To illustrate this, assume we have two forecasts of a variable $\tilde{\pi}$, f_1 and f_2 . These forecasts have the properties that $\tilde{\pi} = f_1 + \tilde{e}$ and $\tilde{\pi} = f_2 + \tilde{v}$, where \tilde{e} and \tilde{v} are two mean zero, serially uncorrelated error terms. We denote the variances of \tilde{e} and \tilde{v} by $\sigma_{\tilde{e}}^2$ and $\sigma_{\tilde{v}}^2$, respectively, and the covariance by $\sigma_{\tilde{e}\tilde{v}}$. The optimal combination weights of these two forecasts under quartic loss minimize the following loss expectation:

$$E(\tilde{L}) = E[(\tilde{\pi} - \Pi^*)^4] = E[(\tilde{\pi} - \delta - \lambda_1 f_1 - \lambda_2 f_2)^4] = E[(\tilde{\pi}(1 - \lambda_1 - \lambda_2) - \delta + \lambda_1 \tilde{e} + \lambda_2 \tilde{v})^4] \quad (5)$$

where Π^* is the combination forecast, λ_1 and λ_2 are the combination weights and δ is the degree of bias. By inspection, $\lambda_1 + \lambda_2 = 1$.

Consequently, formula (5) simplifies to:

$$E[(-\delta + \lambda_1 \tilde{e} + (1 - \lambda_1) \tilde{v})^4]. \quad (6)$$

For simplicity but without loss of generality, we assume that the errors are non-skewed, so that $\delta = 0$.¹ As a result, we require to minimize formula (6) with respect to λ_1 . Clearly, the

¹Whenever this assumption does not hold, we also need to set the derivative w.r.t. δ equal to zero:

$$\frac{\partial E(\tilde{L})}{\partial \delta} \propto -E[(-\delta + \lambda_1 \tilde{e} + (1 - \lambda_1) \tilde{v})^3] \quad (7)$$

$$= \delta^3 - 3\delta^2 E[(\lambda_1 \tilde{e} + (1 - \lambda_1) \tilde{v})] + 3\delta E[(\lambda_1 \tilde{e} + (1 - \lambda_1) \tilde{v})^2] - E[(\lambda_1 \tilde{e} + (1 - \lambda_1) \tilde{v})^3] \quad (8)$$

$$= \delta^3 + 3\delta E[(\lambda_1 \tilde{e} + (1 - \lambda_1) \tilde{v})^2] - E[(\lambda_1 \tilde{e} + (1 - \lambda_1) \tilde{v})^3] = 0. \quad (9)$$

Clearly, this equation holds for $\delta = 0$, if \tilde{e} and \tilde{v} exhibit no skewness. Moreover, in case of normality the solution for λ_1 is identical to the one shown later in the text, as the inclusion of δ in the other first-order condition

optimal weight will depend on the relative kurtosis of the respective errors as well as other cross products. Intuitively, if one of the errors has relatively high kurtosis, one will, *ceteris paribus*, wish to give it lower weight in the combination.

Under quadratic loss, the optimal combination, which was first obtained by Bates and Granger (1969), is given by:

$$\lambda_1 = \frac{\sigma_{\tilde{v}}^2 - \sigma_{\tilde{e}\tilde{v}}}{\sigma_{\tilde{e}}^2 + \sigma_{\tilde{v}}^2 - 2\sigma_{\tilde{e}\tilde{v}}} \quad (10)$$

Now consider the same scenario under a quartic loss function. Intuition suggests that, while in general the optimal combination weights, λ_1 and λ_2 , should be different than under quadratic loss, we should still obtain the above outcome, if the error terms are normally distributed. This arises, as in the latter case the optimal combination weights do not depend on moments above the second. Elliott and Timmermann (2004) proof this claim in a slightly different and more general setting. We offer a more direct proof of this claim in the following.

Differentiating the expected loss (6) with respect to λ_1 and setting to zero (with $\delta = 0$), we obtain:

$$\frac{dE(\tilde{L})}{d\lambda_1} \propto E[(\lambda_1\tilde{e} + (1 - \lambda_1)\tilde{v})^3(\tilde{e} - \tilde{v})] = E[(\lambda_1\tilde{e} + (1 - \lambda_1)\tilde{v})^3\tilde{e}] - E[(\lambda_1\tilde{e} + (1 - \lambda_1)\tilde{v})^3\tilde{v}]. \quad (11)$$

As \tilde{e} and \tilde{v} are mean zero and as $E(\tilde{x}\tilde{y}) = E(\tilde{x})E(\tilde{y}) + cov(\tilde{x}, \tilde{y})$, this can be written as:

$$cov[(\lambda_1\tilde{e} + (1 - \lambda_1)\tilde{v})^3, \tilde{e}] - cov[(\lambda_1\tilde{e} + (1 - \lambda_1)\tilde{v})^3, \tilde{v}] = 0. \quad (12)$$

Using a multivariate version of Stein's lemma, which states that

$$cov[\tilde{x}, h(\tilde{y}, \tilde{z})] = E[h_{\tilde{y}}(\tilde{y}, \tilde{z})]cov(\tilde{x}, \tilde{y}) + E[h_{\tilde{z}}(\tilde{y}, \tilde{z})]cov(\tilde{x}, \tilde{z}), \quad (13)$$

if \tilde{x} , \tilde{y} , and \tilde{z} are multivariate normal and the function h is differentiable in its two arguments (see, e.g., Balvers, 2001), we can rewrite equation (12) as:

$$3E[(\lambda_1\tilde{e} + (1 - \lambda_1)\tilde{v})^2][\lambda_1\sigma_{\tilde{e}}^2 + (1 - \lambda_1)\sigma_{\tilde{e}\tilde{v}} - \lambda_1\sigma_{\tilde{e}\tilde{v}} - (1 - \lambda_1)\sigma_{\tilde{v}}^2] = 0 \quad (14)$$

Solving the term in parenthesis for λ_1 yields a solution identical to that shown in equation (10). Note that our conclusion also applies to the approximation of the bounded loss function suggested by Bray and Goodhart (2002). In this case, the first-order condition can be written as:

$$2[\lambda_1\sigma_{\tilde{e}}^2 + (1 - \lambda_1)\sigma_{\tilde{e}\tilde{v}} - \lambda_1\sigma_{\tilde{e}\tilde{v}} - (1 - \lambda_1)\sigma_{\tilde{v}}^2] - 6kE[(\lambda_1\tilde{e} + (1 - \lambda_1)\tilde{v})^2][\lambda_1\sigma_{\tilde{e}}^2 + (1 - \lambda_1)\sigma_{\tilde{e}\tilde{v}} - \lambda_1\sigma_{\tilde{e}\tilde{v}} - (1 - \lambda_1)\sigma_{\tilde{v}}^2] = 0. \quad (15)$$

Obviously, the former solution also sets this equation equal to zero.

does not alter the second term in parentheses on the left-hand side of (14). If the combination forecast error is instead positively (negatively) skewed, then $\delta > (<) 0$.

3 Simulation and Real Data Examination

We analyze both simulated data with varying degrees of non-normality and real inflation forecasts to examine the differences that the assumption of quartic loss relative to quadratic loss can make. Following Elliott and Timmermann (2004), we simulate data using a mixture of normals model with two states of nature. The mean realizations of the three simulated variables, i.e., the variable to forecast, $\tilde{\pi}$, and the two forecasts, f_1 and f_2 , are zero in the first and 0.5 in the second state of nature. A property of their method is that we can easily control the degree of non-normality the forecast errors exhibit by varying the switching probability between the two states of nature. More specifically, while setting $p = 0$ generates normally-distributed forecast errors, increasing p leads to more negatively skewed forecast errors with more positive excess kurtosis.

Table 1 indicates that, since both forecast errors are mean zero by construction, the optimal bias under quadratic loss is also always zero. When $p > 0$, the optimal bias under quartic loss is negative to counterbalance the effect of extreme realizations caused by negative skewness. Similarly, the optimal combination weights are only equivalent in case of normality. With increasing p , moments above the second become progressively more important. Thus, the optimal combination weights can differ by up to 17% between quadratic and quartic loss. When $p > 0$, the variance of the absolute forecast combination error is smaller under quartic than under quadratic loss.

While Table 1 suggests that under non-normality the optimal bias and combination weights under quadratic and quartic loss can differ substantially, we now ask whether the forecast errors of *real* economic variables show sufficient non-normality to justify this distinction. We thus investigate 6-month and 12-month inflation forecasts from the Livingston Surveys of Professional Forecasters. Since there is a large number of individual forecasters, we only consider pairs of forecasts obtained from survey participants whose ID numbers follow each other, given we obtain at least 40 observations. Most forecasts are neither unbiased nor efficient. We thus considered outcomes based on ‘corrected’ forecast errors, which are the residuals from OLS regressions of $\tilde{\pi}$ on f .

Our results, available on request, show that *real* inflation forecast errors can exhibit substantial skewness and kurtosis. As a result, the optimal combination weights often differ by up to 20% to 30%. However, the difference in the optimal bias hardly ever exceeds 1%. Finally, for both forecasting horizons the range of the absolute combination forecast error is smaller under quartic than under quadratic loss.

4 Conclusions

For central bankers who have a symmetric but non-quadratic loss function, we show that skewness of inflation can create a deflationary bias. In this regard, Goodhart’s (2001) comments that skewness matters for policy decisions are interesting. Second, we also illustrate that due to the quartic term the weights in the optimal combination of forecasts will in general differ from that under quadratic loss depending on the higher moments of the distribution of forecast errors. We examine these differences employing simulated and real data from the Livingston Survey of Professional Forecasters.

References

- Balvers, R. J. (2001) *Foundations of asset pricing*, West Virginia University Teaching Script.
- Bates, J. M. and C. W. J. Granger (1969) “The combination of forecasts” *Operations Research Quarterly* **20**, 451–468.
- Blinder, A. S. (1998) *Central Banking in Theory and Practice*, The MIT Press: Cambridge, MA.
- Boinet, V. and C. Martin (2006) “Targets, zones and asymmetries: A flexible nonlinear model of recent UK monetary policy” forthcoming in *Oxford Economic Papers*.
- Bray, M. and C. A. E. Goodhart (2002) “You might as well be hung for a sheep as a lamb: The loss function of an agent” forthcoming in the *Manchester School Journal*.
- Elliott, G. and A. Timmermann (2004) “Optimal forecast combination under general loss functions and forecast error distributions” *Journal of Econometrics* **122**, 47–79.
- Goodhart, C. A. E. (2001) “Monetary transmission lags and the formulation of the policy decision on interest rates” *Federal Reserve Bank of St. Louis* **3**, 165–182.
- Ruge-Murcia, F. J. (2000) “Uncovering financial market beliefs about inflation targets” *Journal of Applied Econometrics* **15** (3), 483–512.
- Svensson, L. (1997) “Optimal inflation targets, ‘conservative’ central banks, and linear inflation contracts” *American Economic Review* **87**, 98–114.
- Vickers, J. (1998) “Inflation targeting in practice: The UK experience” *Bank of England Quarterly Bulletin* **38** (4), 368–375.

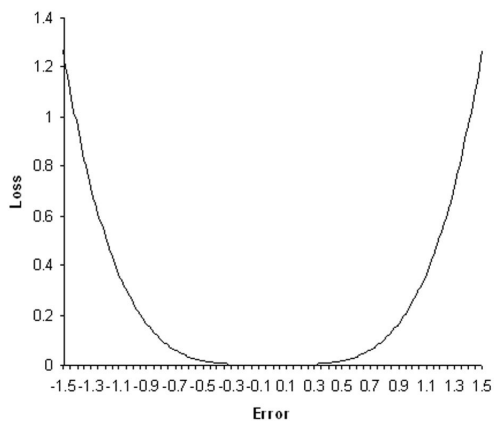


Figure 1: Approximation of the Boinet and Martin (2006) Target Zone Loss Function

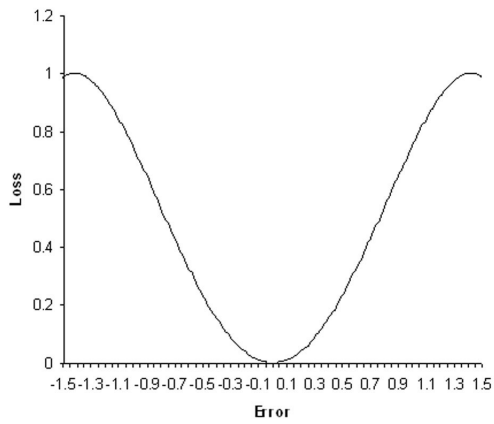


Figure 2: Approximation of the Bray and Goodhart (2002) Bounded Loss Function

Table 1: Simulation

This table shows the optimal weights, δ and λ_1 , for the simulated variables, $\tilde{\pi}$, f_1 and f_2 , under both quadratic (quad) and quartic (quar) loss. The simulated variables are generated through a mixture of normals model, with means equal to zero in the first and equal to 0.5 in the second state. The two covariance matrices are:

$$\sigma_1^2 = \begin{bmatrix} 1.000 & & \\ 0.200 & 2.500 & \\ 0.150 & 0.125 & 2.000 \end{bmatrix}, \quad \sigma_2^2 = \begin{bmatrix} 0.100 & & \\ 0.050 & 0.500 & \\ 0.300 & 0.200 & 3.200 \end{bmatrix}$$

The switching probability between the two states of nature equals p . In total, we generate 1,000,000 values for each random variable in each set. The forecast errors are the residuals from OLS regressions of $\tilde{\pi}$ onto each forecast. We also report the skewness and the kurtosis of the forecast errors (fct err). Finally, we show the variance of the absolute (Var(abs)) and the minimum (Min) and maximum (Max) of the raw forecast combination error under both loss functions.

p	Optimal weights (δ and λ_1)						Skewness		Kurtosis		Combination forecast error					
	quad		quar		quad-quar		Fct	Fct	Fct	Fct	Var(abs)		Min		Max	
	δ	λ_1	δ	λ_1	δ	λ_1	err ₁	err ₂	err ₁	err ₂	quad	quar	quad	quar	quad	quar
0.0	0.00	0.59	0.00	0.59	0.00	0.00	0.00	3.00	3.00	35.57	35.57	-4.93	-4.93	5.16	5.16	
0.1	0.00	0.53	-0.04	0.56	0.04	-0.03	-0.12	-0.12	3.16	3.17	33.85	33.61	-4.70	-4.66	4.55	4.60
0.2	0.00	0.49	-0.08	0.54	0.08	-0.05	-0.24	-0.25	3.40	3.43	32.42	31.51	-4.65	-4.57	4.92	4.99
0.3	0.00	0.46	-0.11	0.52	0.11	-0.06	-0.36	-0.38	3.69	3.74	30.98	29.24	-4.85	-4.73	4.28	4.41
0.4	0.00	0.43	-0.14	0.50	0.14	-0.07	-0.50	-0.53	4.12	4.21	29.55	26.93	-4.74	-4.60	4.40	4.52
0.5	0.00	0.41	-0.17	0.49	0.17	-0.08	-0.65	-0.69	4.68	4.83	27.56	24.34	-4.57	-4.39	4.49	4.65
0.6	0.00	0.38	-0.20	0.47	0.20	-0.09	-0.80	-0.87	5.48	5.75	24.84	21.47	-5.00	-4.79	4.40	4.63
0.7	0.00	0.35	-0.21	0.45	0.21	-0.10	-0.97	-1.08	6.60	7.08	21.16	18.29	-4.84	-4.60	4.56	4.78
0.8	0.00	0.30	-0.22	0.43	0.22	-0.13	-1.16	-1.33	8.15	9.11	16.36	14.56	-4.99	-4.80	3.83	4.04
0.9	0.00	0.22	-0.20	0.39	0.20	-0.17	-1.26	-1.55	10.17	12.40	10.19	9.93	-4.61	-4.46	3.76	3.92