Right on target! Or is it?
The role of distributional shape in variance targeting

Stanislav Anatolyev* Stanislav Khrapov†
New Economic School New Economic School

February 14, 2015

Abstract

Estimation of GARCH models can be simplified by augmenting QML estimation by variance targeting which reduces the degree of parameterization and facilitates estimation. We compare the two approaches and investigate, via simulations, how non-normality features of the return distribution affect the quality of estimation of the volatility equation and corresponding volatility and value-at-risk predictions. We find that most GARCH coefficients and associated predictions are more precisely estimated when no variance targeting is employed. Bias properties are exacerbated for heavier-tailed distribution of standardized returns while the distributional asymmetry has little or moderate impact. The sample size has also a more favorable effect on estimation precision when no variance targeting is used.

Keywords: GARCH; variance targeting; non-normality; heavy tails; skewness; QML estimation.

JEL codes: C22; C53; C58

*Corresponding author. Address: Stanislav Anatolyev, New Economic School, 100A Novaya Street, Skolkovo, Moscow, 143026 Russia. E-mail: sanatoly@nes.ru.

†Address: New Economic School, 100A Novaya Street, Skolkovo, Moscow, 143026 Russia. E-mail: khrapovs@gmail.com

1 Introduction

The technique of variance targeting for GARCH models has been proposed in Engle & Mezrich (1996) in order to reduce the degree of parameterization, especially in multivariate contexts (see Pedersen & Rahbek, 2014 and Francq et al., 2014), and facilitate computation of quasi-maximum likelihood (QML) estimates. Variance targeting becomes popular in empirical work, and gets incorporated into econometric software (e.g., the `rugarch` package in R).[1]

Beside serving the main purpose, variance targeting has other merits, such as superiority of long-term volatility and value-at-risk predictions when the variance model is misspecified (Francq et al., 2011). However, variance targeting does not come at no cost. One of its disadvantages over 'regular' QML estimation is the two-step nature which results in efficiency losses (Francq et al., 2011) and a need for standard error corrections (Kristensen & Linton, 2004). Francq et al. (2011) quantify efficiency losses in a stylized ARCH model and find out that they are negligible when conditional heteroskedasticity is weak, but for some parameters they may be dramatic when it is large enough to make the unconditional kurtosis huge.

More generally, heaviness of tails plays important role in variance targeting. While Francq et al. (2011) prove asymptotic normality of estimates under variance targeting when the unconditional kurtosis is finite, Vaynman & Beare (2014) develop asymptotic theory for the case of infinite unconditional kurtosis when the rate of convergence and asymptotic distribution cease to be standard. A quick look at estimated parameters when the distribution is heavy tailed reveals drastic differences in values they can take depending on whether variance targeting is applied or not. For example, five arbitrary runs of our simulation procedure described below yields for one of parameters whose true value is unity the following estimates: {0.50, 0.54, 0.34, 0.32, 0.43} with variance targeting and {4.74, 0.67, 1.70, 2.02, 0.63} without variance targeting.[2] One can see that while the estimates are severely biased downward though not highly dispersed in the first case, they tend to take very different values, sometimes having little to do with the true value, in the second case. These differences, of course, pass over to other output from the GARCH model analysis.

In this paper, we study the differences in properties of parameter estimates, as well as some practical measures pertaining to GARCH modeling, that result in either using or abstaining from the variance targeting technique. We primarily focus on the influence of the shape of the conditional distribution of returns on discrepancies between the outcomes of the 'regular' QML procedure and the one augmented by variance targeting. Particularly, we simulate data from a GARCH model

[1] Alexios Ghalanos (2014) Introduction to the rugarch package. Available at www.maths.bris.ac.uk/R/web/packages/rugarch/vignettes/Introduction_to_the_rugarch_package.pdf

[2] Using the notation that follows, this experiment corresponds to the case $\alpha + \beta = 0.99$, $\beta = 0.8$, $\nu = 5$, $\lambda = 0$, $T = 2000$. The parameter of interest is σ^2 with the true value of 1.

with standardized returns following the Skewed Student distribution allowing a variety of indicators of thick-tailedness and asymmetry, compute biases for various statistics of parameter estimates, and track their dependence on these indicators.

It turns out that for most parameters and associate predictions – with a notable exception of unconditional variance – the bias under variance targeting is larger both in median and interpercentile terms, sometimes by fewfold. This tendency is typically exacerbated for heavier-tailed distribution of standardized returns while the distributional asymmetry has little or moderate impact. A sample size also has more favorable effect on estimation precision when no variance targeting is used. For the unconditional variance mentioned above as an exception, the median bias is larger under variance targeting as well, but for this parameter the estimator dispersion may be very high because of a long right tail when no variance targeting is employed. Thus, we conclude that, at least as long as no estimates of unconditional variance are involved in the statistic of interest, one should avoid variance targeting provided that its computational benefits are not overwhelming.

The paper is organized as follows. In Section 2, we describe the simulation design including the model, data generating process, estimation methods and statistics of interest. In Section 3, we show and analyze the results of simulation experiments. Section 4 concludes.

2 Simulation design

2.1 Model and estimates

The conditional variance model starts from the decomposition of returns

$$r_t = \sqrt{h_t}\varepsilon_t,$$

where the conditional mean is suppressed, h_t is conditional variance, and ε_t is IID standardized return having distribution $\mathcal{D}$. To reach interesting conclusions it is sufficient to consider the simple GARCH(1,1) model Bollerslev (1986)

$$h_t = \omega + \alpha r_{t-1}^2 + \beta h_{t-1}. \tag{1}$$

The GARCH equation is assumed to be stable in the sense that $\alpha + \beta < 1$. The unconditional variance equals

$$\sigma^2 = \frac{\omega}{1-\alpha-\beta}. \tag{2}$$

Estimation of parameters of the GARCH model in the formulation (1) is done via normal QML, with the corresponding likelihood for one observation equalling

$$\ell_t^{NT}(\omega,\alpha,\beta) = -\frac{1}{2}\left(\ln\left(\omega + \alpha r_{t-1}^2 + \beta h_{t-1}\right) + \frac{r_t^2}{\omega + \alpha r_{t-1}^2 + \beta h_{t-1}}\right).$$

Because of (2), given the estimates the estimate $\hat{\omega}^{NT}$, $\hat{\alpha}^{NT}$, $\hat{\beta}^{NT}$ of ω, α, β, the estimate of σ^2 is formed as

$$\left(\hat{\sigma}^2\right)^{NT} = \frac{\hat{\omega}^{NT}}{1 - \hat{\alpha}^{NT} - \hat{\beta}^{NT}}.$$

The machinery of variance targeting works as follows (Engle & Mezrich, 1996). An alternative reparameterization of (1) is, using (2),

$$h_t = (1 - \alpha - \beta)\,\sigma^2 + \alpha r_{t-1}^2 + \beta h_{t-1}.$$

The parameter σ^2 can be pre-estimated directly from the sample by the method of moments:

$$\left(\hat{\sigma}^2\right)^{T} = \frac{1}{T}\sum_{t=1}^{T} r_t^2.$$

The parameters α and β are estimated given this estimate, with the corresponding likelihood for one observation equalling

$$\ell_t^{T}(\alpha,\beta) = -\frac{1}{2}\left(\ln\left((1-\alpha-\beta)\left(\hat{\sigma}^2\right)^{T} + \alpha r_{t-1}^2 + \beta h_{t-1}\right) + \frac{r_t^2}{(1-\alpha-\beta)\left(\hat{\sigma}^2\right)^{T} + \alpha r_{t-1}^2 + \beta h_{t-1}}\right)$$

whose maximization yield estimates $\hat{\alpha}^T$, $\hat{\beta}^T$ of α, β. Given these, the estimate of ω is formed as

$$\hat{\omega}^T = \left(1 - \hat{\alpha}^T - \hat{\beta}^T\right)\left(\hat{\sigma}^2\right)^T.$$

2.2 Data generation

As a matter of normalization, in all DGPs we set $\sigma^2 = 1$; then $\omega = 1 - \alpha - \beta$. We consider two values for the feedback parameter β: 0.8 and 0.9, and two values for the persistence parameter $\alpha + \beta$: 0.95 and 0.99. These values approximately match GARCH parameters of daily stock and exchange rate returns (e.g., Bollerslev, 1987).

As the distribution $\mathcal{D}$, we use the Skewed Student distribution of Hansen (1994) which serves as a workhorse for estimating skewed conditional distributions in empirical finance; see, for example, Jondeau & Rockinger (2003). The Skewed Student distribution with zero mean and unit variance has the following probability density function:

$$f(x|\eta,\lambda) = \begin{cases} bc\left(1+\frac{1}{\eta-2}\left(\frac{a+bx}{1-\lambda}\right)^2\right)^{-(\eta+1)/2}, & x < -a/b, \\ bc\left(1+\frac{1}{\eta-2}\left(\frac{a+bx}{1+\lambda}\right)^2\right)^{-(\eta+1)/2}, & x \geq -a/b, \end{cases}$$

where $2 < \eta < \infty$, and $-1 < \lambda < 1$. The constants a, b, and c are given by

$$a = 4\lambda c\frac{\eta-2}{\eta-1}, \quad b^2 = 1+3\lambda^2-a^2, \quad c = \frac{\Gamma\left(\frac{\eta+1}{2}\right)}{\sqrt{\pi(\eta-2)}\Gamma\left(\frac{\eta}{2}\right)}.$$

The shape parameters are η (degrees of freedom) and λ (degree of asymmetry). We vary[3] the value of the kurtosis parameter η within the set $\{30, 10, 5, 4, 3\}$ implying a wide range of tail thickness. The value of 30 makes $\mathcal{D}$ indistinguishable (when $\lambda = 0$) from the normal distribution, the value of 3 implies very heavy tails when third and fourth moments fail to exist,[4] while the intermediate values match degrees of freedom of daily stock and exchange rate returns (e.g., Bollerslev, 1987). While we do most of experiments for zero skewness to focus on the effect of heavy tails, we also vary the asymmetry parameter λ within the set $\{-0.1, -0.3, -0.5, -0.8\}$, where -0.1 approximately matches the asymmetry coefficient for the S&P500 index returns, while the other values imply severalfold exaggerated skewness. Figure 1 presents plots of the Skewed Student density corresponding to (some of) combinations of shape parameters that we employ. The density corresponding to the combination $\eta = 30$, $\lambda = 0$ (the darkest curves) is indistinguishable from the standard normal density.

We report most results for the sample length $T = 2000$ which is typical in GARCH estimation on daily data. We also analyze an effect of a sample size setting also T to 500, 1,000, 5,000 and 10,000.

The number of simulations is 1,000 (the next version of the paper will be based on 10,000 simulations). The code is written in Python and based on Andrew Patton's MatLab version[5] available here: http://public.econ.duke.edu/∼ap172/code.html.

2.3 Criteria

For the baseline and variance targeting estimation methods, we first of all track the bias in $\hat{\omega}$, estimates of the intercept parameter ω, the bias in $\hat{\beta}$, estimates of the feedback parameter β, the

[3]See Figure 1 on page 10 for examples of Skewed Student density function for varying parameters η and λ.

[4]For data with infinite higher-order moments, Hill & Renault (2012) propose robust variance targeting based in tail-trimming techniques.

[5]Python implementation is available here: https://github.com/khrapovs/skewstudent.

bias in $\hat{\alpha}$, estimates of the news impact parameter α, the bias in $\hat{\alpha}+\hat{\beta}$, estimates of the persistence parameter $\alpha+\beta$, and the bias in $\hat{\sigma}^2$, estimates of the unconditional variance σ^2:

$$\begin{aligned} B_\omega &= \hat{\omega}-\omega, \\ B_\beta &= \hat{\beta}-\beta, \\ B_\alpha &= \hat{\alpha}-\alpha, \\ B_{\alpha+\beta} &= \hat{\alpha}+\hat{\beta}-(\alpha+\beta), \\ B_{\sigma^2} &= \hat{\sigma}^2-\sigma^2. \end{aligned}$$

Second, we look at two one-step predictors. One is the one-step-ahead (logarithm of) volatility prediction given $r_t^2 = h_t = 1$ which is equal to $\hat{\omega}+\hat{\alpha}+\hat{\beta}$ and whose true value is 1 (and 0 for the logarithm):

$$B_{+1} = \ln \hat{h}_{t+1}\Big|_{r_t^2=h_t=1} - \ln h_{t+1}|_{r_t^2=h_t=1} = \ln\left(\hat{\omega}+\hat{\alpha}+\hat{\beta}\right).$$

The other is the one-period in advance 'naive' 5%-level value-at-risk (VaR) prediction given $r_t^2 = h_t = 1$ which is equal to $\sqrt{\hat{\omega}+\hat{\alpha}+\hat{\beta}}Z_{5\%}$, where $Z_{5\%} \approx -1.645$ is a 5%-quantile of the standard normal distribution:

$$B_{5\%} = \left(\sqrt{\hat{h}_{t+1}\Big|_{r_t^2=h_t=1}} - \sqrt{h_{t+1}|_{r_t^2=h_t=1}}\right) Z_{5\%} = \left(\sqrt{\hat{\omega}+\hat{\alpha}+\hat{\beta}}-1\right) Z_{5\%}.$$

The forecast is 'naive' because one uses a quantile of the normal distribution instead of that of the unknown true one. However, the forecast is proportional to that corresponding to the true distribution, and so is its bias.

The volatility and VaR predictions are functions of the same set of parameter estimates (of ω and $\alpha+\beta$), but they are meant to give an idea how biases of parameter estimates translate into those of practically useful measures. When r_t^2 and/or h_t are different from unity, some parameters carry more weight relative to others. For example, when current volatility is higher than the unconditional variance and the current return is smaller than its standard deviation, the loading on β will be higher than that on α; when current volatility is higher than the unconditional variance and the current return is larger than its standard deviation, the loading on the suitable linear combination of α and β will overweight the loading on ω, the unity.

3 Results

3.1 Parameter biases

Figures 2–6 show distributions of biases of parameter estimates in the following order: B_ω, B_β, B_α, $B_{\alpha+\beta}$, and B_{σ^2}. The two columns correspond to two values of the feedback parameter β, and the five rows correspond to five values of the degrees of freedom parameter η. The blue lines correspond to variance targeting, the green lines – to 'regular' QML estimation. The vertical lines of related color are placed at medians of distributions; in addition, grey vertical lines are placed at zero, the ideal value for biases. Correspondingly, Tables 1–5 contain 5%, 25%, 50%, 75% and 95% percentiles, expressed in percentages to the true parameter value, of bias distributions. We show percentiles instead of means, variances and/or mean squared errors because of possible non-existence of the second (and even first) moments for these distributions. As the patterns when $\alpha + \beta$ equals 0.95 and 0.99 are qualitatively similar, we show only the results for $\alpha + \beta = 0.99$, which implies the true value for the intercept $\omega = 0.01$. Recall also that $\sigma^2 = 1$ throughout.

From Figure 2 on page 11 and Table 1 on page 20 one learns that estimates of the intercept ω are not severely median biased: the maximal reported median bias corresponds to an about 20% deviation from the true value. The median bias is positive, smaller for the smaller value of β, and always bigger when variance targeting is used, although not by much. The shape of the bias distribution is similar in cases of variance targeting and of no variance targeting, and so is its dispersion. While the median bias increases though quite slowly with heavy-tailedness of the return distribution, the dispersion goes up significantly with it: the $[5\%, 95\%]$ interpercentile range approximately doubles when the return distribution turns from mesokurtic to strongly leptokurtic. The maximal reported 95% percentile corresponds to an estimate which is nearly triple of the true value.

Figure 3 on page 12 and Table 2 on page 20 show that the median bias of estimates of the feedback parameter β is practically non-existent, whether the return distribution is thin or thick tailed, and whether variance targeting is used or not. However, the tails of the distributions and their dispersion steadily rise with the degree of return leptokurtosity. The left tail is about twice as long as the right tail.

The behavior of the bias for the news impact parameter α illustrated in Figure 4 on page 13 and quantified in Table 3 on page 21 is in a stark contrast to those for ω and β. While with variance targeting the median bias is quite small, from 0% to 2%, irrespective of the value of β and, importantly, of the degree of return leptokurtocity, the (negative) median bias under variance targeting is larger and increases (in absolute value) with the degree of return leptokurtocity pretty fast. However, even for highly tailed return distributions the median bias is relatively moderate (maximum 15%) if contrasted to the bias dispersion which is big in both cases: the 95% percentile can exceed 80% and is bigger when no variance targeting is used.

Evidence presented in Figure 5 on page 14 and Table 4 on page 21 shows that the relative bias of estimates of the persistence parameter $\alpha + \beta$ is of a much smaller magnitude than that for ω, the median bias being of order of few percent, despite the fact that estimation of the first ingredient of the sum, α, may be subject to substantial biases. This happens because in absolute terms α is (as typically) much smaller than β whose estimation is subject to a very small median bias. The median bias is always negative, larger (in absolute value) for the smaller value of β, and increasing in the degree of return leptokurtosity. It is noticeably larger when variance targeting is used than when it is not, the discrepancy quickly increasing with heavy-tailedness of the return distribution. As far as the dispersion is concerned, it is inherited from those in estimation of both β and α in such a way that most probability weight is put on the left tail which is quite long and heavy, though the maximal reported downward bias is about just 10%. The right tail of bias distributions are bounded because of an implicitly embedded condition on $\alpha + \beta$ not to exceed (or to exceed but slightly) unity.

Last but not least, from Figure 6 on page 15 and Table 5 on page 22 one can see that estimates of the unconditional variance σ^2 exhibit severe biases.[6] Even the median bias reaches and exceeds (in absolute value) 50% in some cases. It steadily goes up with the degree of return leptokurtosity, and is noticeably larger for the smaller value of β. The bias distribution quickly shifts leftward as the degree of return leptokurtosity increases, and in the extreme case the $[5\%, 95\%]$ interpercentile range lies entirely in the negative line. Although much tighter, the bias distribution when variance targeting is used has a much larger median bias, up to twice as much for more leptokurtic return distributions. This may sound surprising, as the estimate in this case is just that of the method-of-moments (which in an IID environment would be an unbiased estimate) and does not use the GARCH model at all. Exploitation of the model imposing the connection between the model parameters in the likelihood function significantly reduces median bias, but makes the bias distribution much more dispersed, primarily in the right tail, with 95% percentile reaching multiples of the true value.

To conclude, parameter estimate biases are strongly and positively related to heaviness of tails of the return distribution. Among the model parameters, it is the unconditional variance[7] that experiences severity of estimation biasedness most, with the intercept coming next. However, that severity is qualitatively different when variance targeting is used or not.

[6] The example in the Introduction corresponds to one of these cases.

[7] As a practical matter, recall that in the GARCH model the unconditional variance is a long-run volatility forecast.

3.2 Prediction biases

Now we analyze the bias of one-period ahead (logarithm of) volatility forecast when the present volatility and present return are fixed at unity. This bias is inherited from the biases in the intercept ω and the persistence parameter $\alpha + \beta$. From the previous analysis we know that the relative median bias in the former by far exceeds that in the latter. However, the value of the latter is about 100 times larger than the value of the former, hence in absolute terms the bias of the volatility forecast is dominated by that in the persistence parameter. Figure 7 on page 16 and Table 6 on page 22 present the distributions and figures. The bias is negative and much larger (in absolute value) for the smaller value of β. The median bias rises with return leprokurtocity, moderately when no variance targeting is used, and more sharply when variance targeting is used. The discrepancy is exacerbated for the smaller value of feedback. The same is happening with the whole left tail of the bias distribution.

The bias of one-period in advance value-at-risk (VaR) is also a concave function of $\hat{\omega} + \hat{\alpha} + \hat{\beta}$ like that for the logarithmic volatility forecast, hence it exhibits similar tendencies (with a flip of the sign). In particular, it is analogously driven mostly by a bias in the persistence parameter $\alpha + \beta$. Figure 8 on page 17 and Table 7 on page 23 show that its median is always positive, larger for the smaller value of β, and increasing with the heavy-tailedness of the return distribution. The bulk of the probability mass of the bias under variance targeting is much farther from zero than under no variance targeting, and its dispersion is generally larger when variance targeting is used.

Next we analyze, using the VaR criterion as an example, the impact of skewness of the return distribution on the bias. In Figure 9 on page 18 and Table 8 on page 23 the five rows now correspond to five values of the degrees of asymmetry λ. One concludes that as the degree of distributional asymmetry varies from zero up to extreme values, the distribution of the bias change little – from a very slight rightward shift under mesokurtocity to a more pronounced though still moderate rightward shift under strong leptokurtocity. The bias distribution under variance targeting is farther from zero for all considered combinations.

Finally, we analyze, again using the VaR criterion as an example, the impact of the sample size on the bias. Figure 10 on page 19 and Table 9 on page 24 where the five rows now correspond to five values of T, reveal that when no variance targeting is used, the median bias quickly diminishes as more sample observations are available. Irrespective of the degree of heavy-tailedness of the return distribution, it goes down by a factor of about five as the sample size gets multiplied by 20. The right 95% percentile is also shrinking fast. This happens in a slower way when variance targeting is used: the factor referenced above is no larger than 3.

4 Conclusion

Variance targeting is a computationally attractive detour in estimation of GARCH models. However, our simulations indicate that in practice variance targeting may lead to bigger biases in parameter estimates and associated prediction measures than when no targeting is used in the QML procedure. Under variance targeting the bias properties deteriorate with return heavy-tailedness faster, and the bias is a bit more sensitive to return skewness and sample sizes than in the 'regular' QML. Thus, variance targeting should probably be avoided in those cases when computational burden is not prohibitive, which certainly includes the one-dimensional case.

References

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.

Bollerslev, T. (1987). A Conditionally Hetroskedastic Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics*, 69(3), 542–547.

Engle, R. F. & Mezrich, J. (1996). GARCH for Groups. *Risk*, 9, 36–40.

Francq, C., Horvath, L., & Zakoian, J.-M. (2011). Merits and Drawbacks of Variance Targeting in GARCH Models. *Journal of Financial Econometrics*, 9(4), 619–656.

Francq, C., Horvath, L., & Zakoian, J.-M. (2014). Variance Targeting Estimation of Multivariate GARCH Models. MPRA Paper No. 57794.

Hansen, B. E. (1994). Autoregressive Conditional Density Estimation. *International Economic Review*, 35(3), 705–730.

Hill, J. B. & Renault, E. (2012). Variance Targeting for Heavy Tailed Time Series. University of North Carolina at Chapel Hill, working paper.

Jondeau, E. & Rockinger, M. (2003). Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements. *Journal of Economic Dynamics and Control*, 27(10), 1699–1737.

Kristensen, D. & Linton, O. (2004). Consistent Standard Errors for Target Variance Approach to GARCH Estimation – Solution. *Econometric Theory*, 20(5), 990–993.

Pedersen, R. S. & Rahbek, A. (2014). Multivariate Variance Targeting in the BEKK-GARCH Model. *Econometrics Journal*, 17, 24–55.

Vaynman, I. & Beare, B. K. (2014). Stable Limit Theory for the Variance Targeting Estimator. In Y. Chang, T. B. Fomby, & J. Y. Park (Eds.), *Essays in Honor of Peter C. B. Phillips (Advances in Econometrics, Volume 33)* (pp. 639–672). Emerald Group Publishing Limited.

Appendix

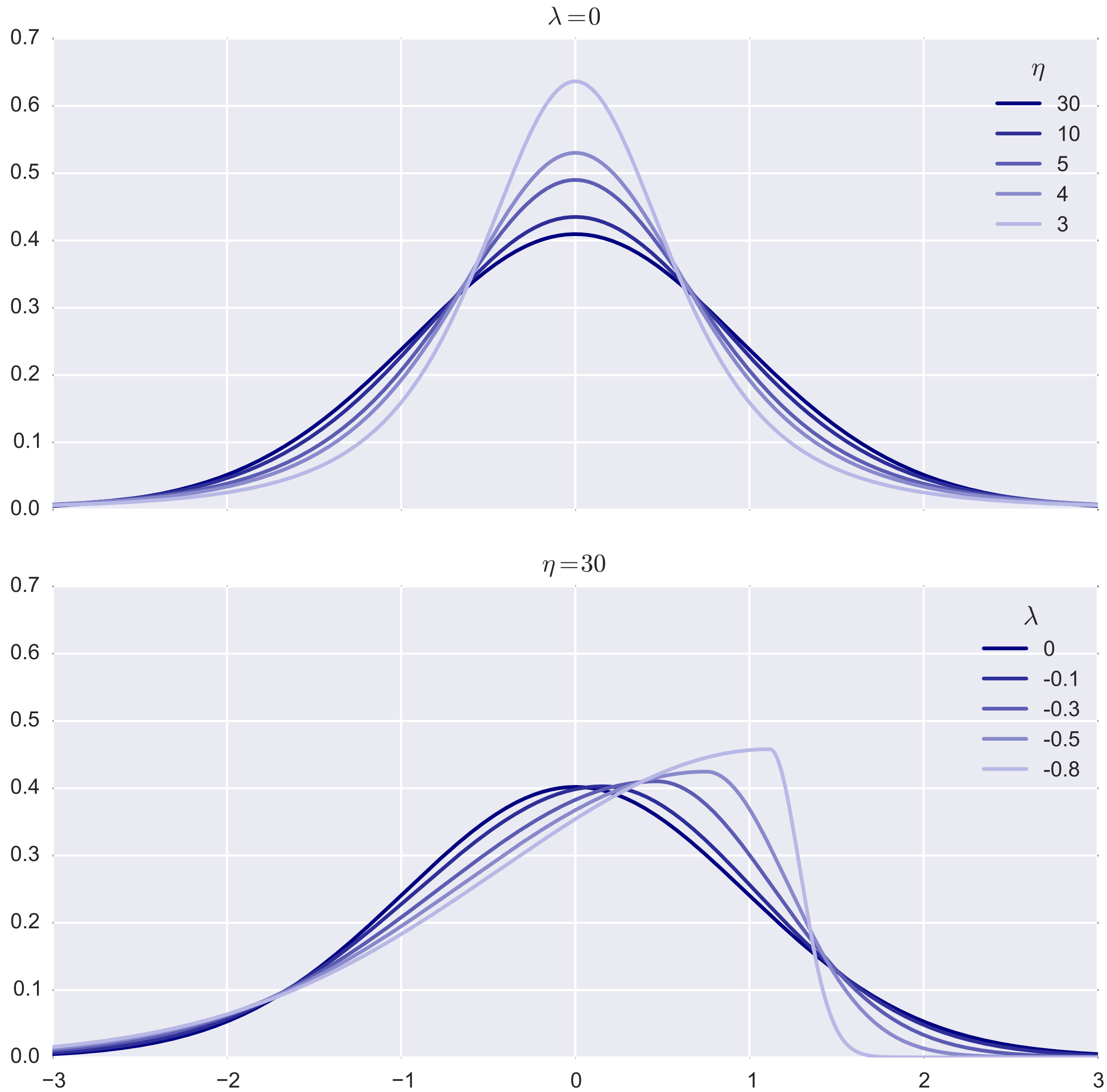

Figure 1: Skewed Student distribution of Hansen (1994)

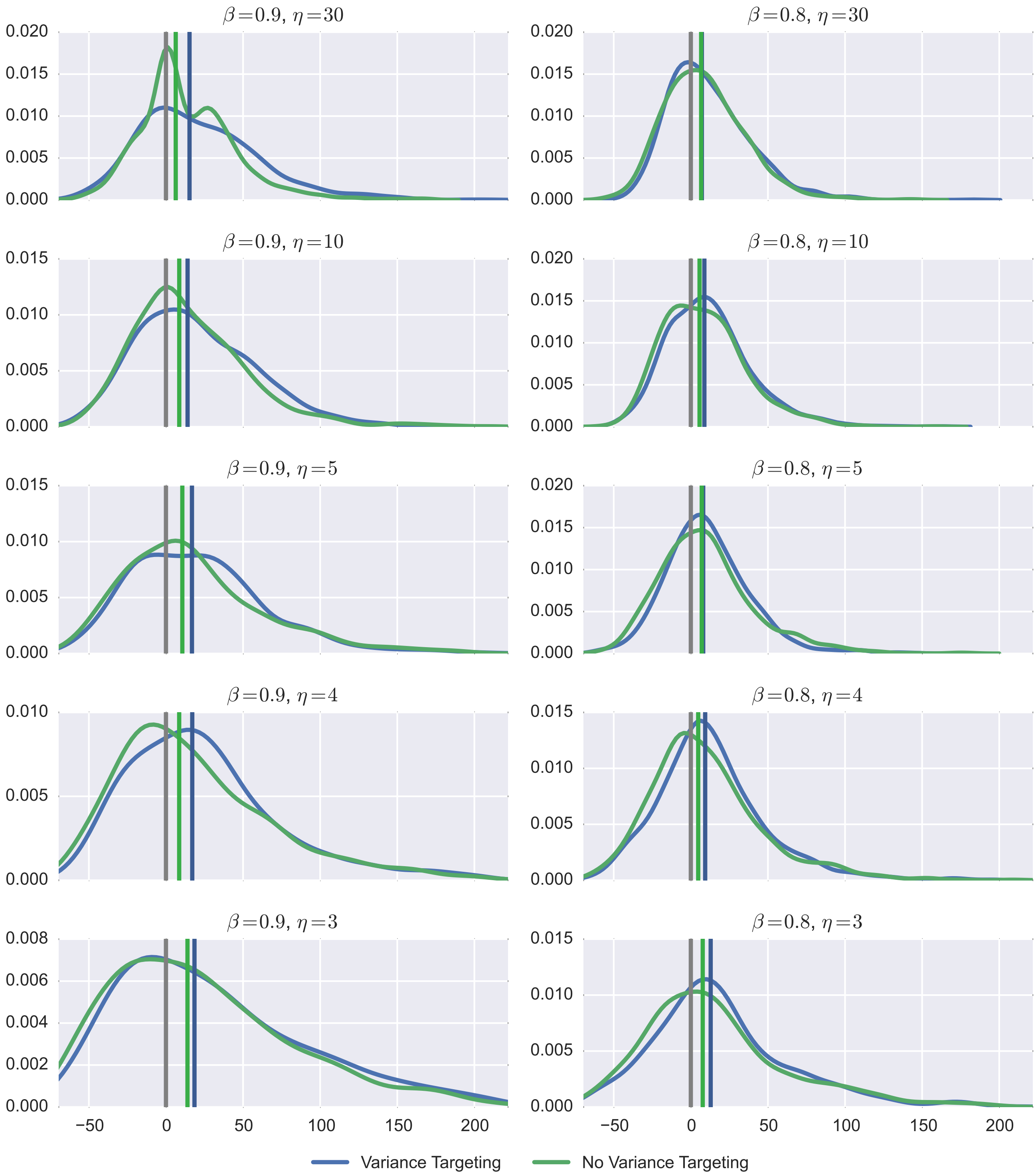

Figure 2: Intercept bias, B_ω, for varying β and η.
Horizontal axis is in percentage of the true parameter value. Vertical solid lines are medians and zero (gray). Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

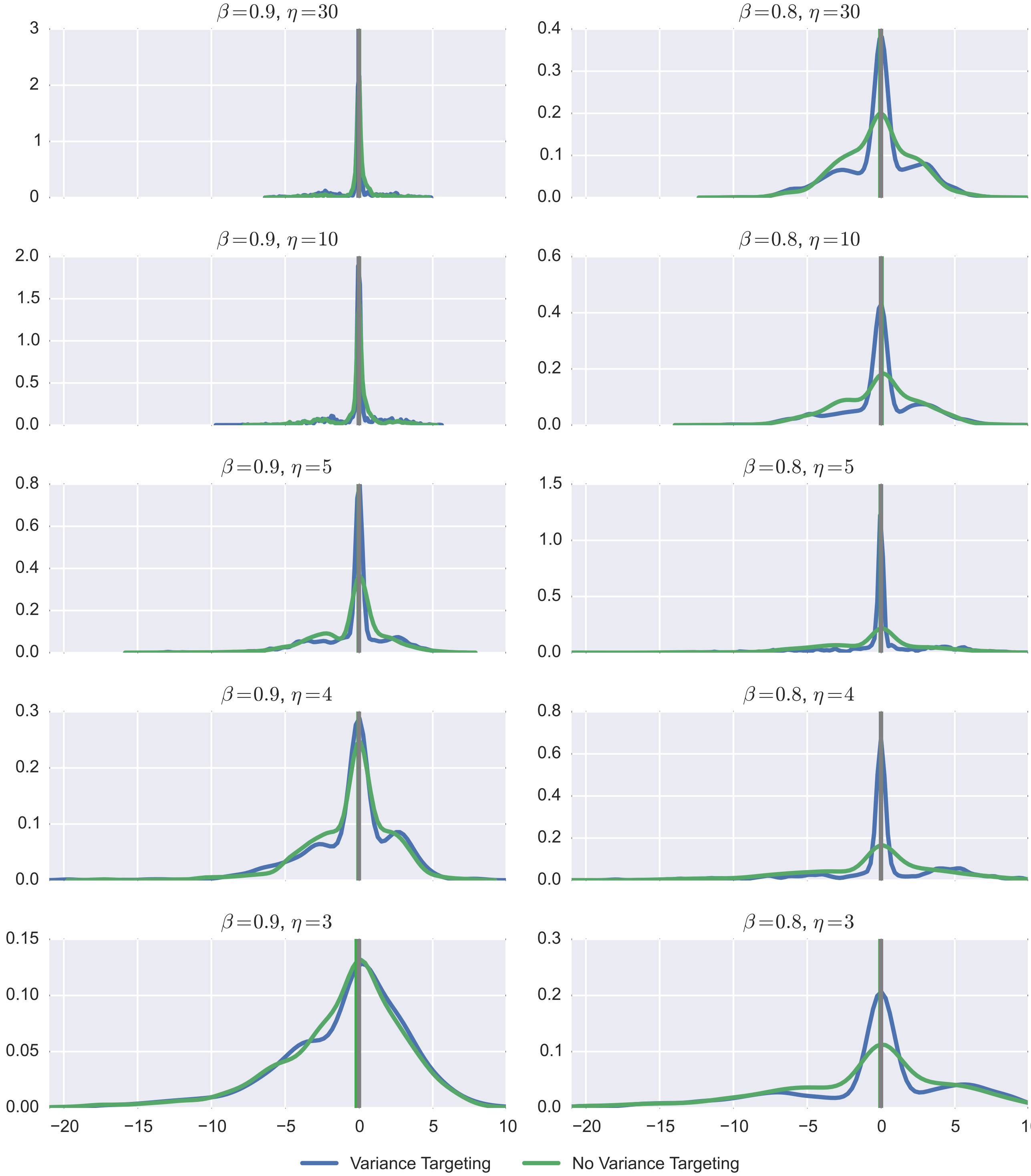

Figure 3: Feedback bias, B_β, for varying β and η.
Horizontal axis is in percentage of the true parameter value. Vertical solid lines are medians and zero (gray). Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

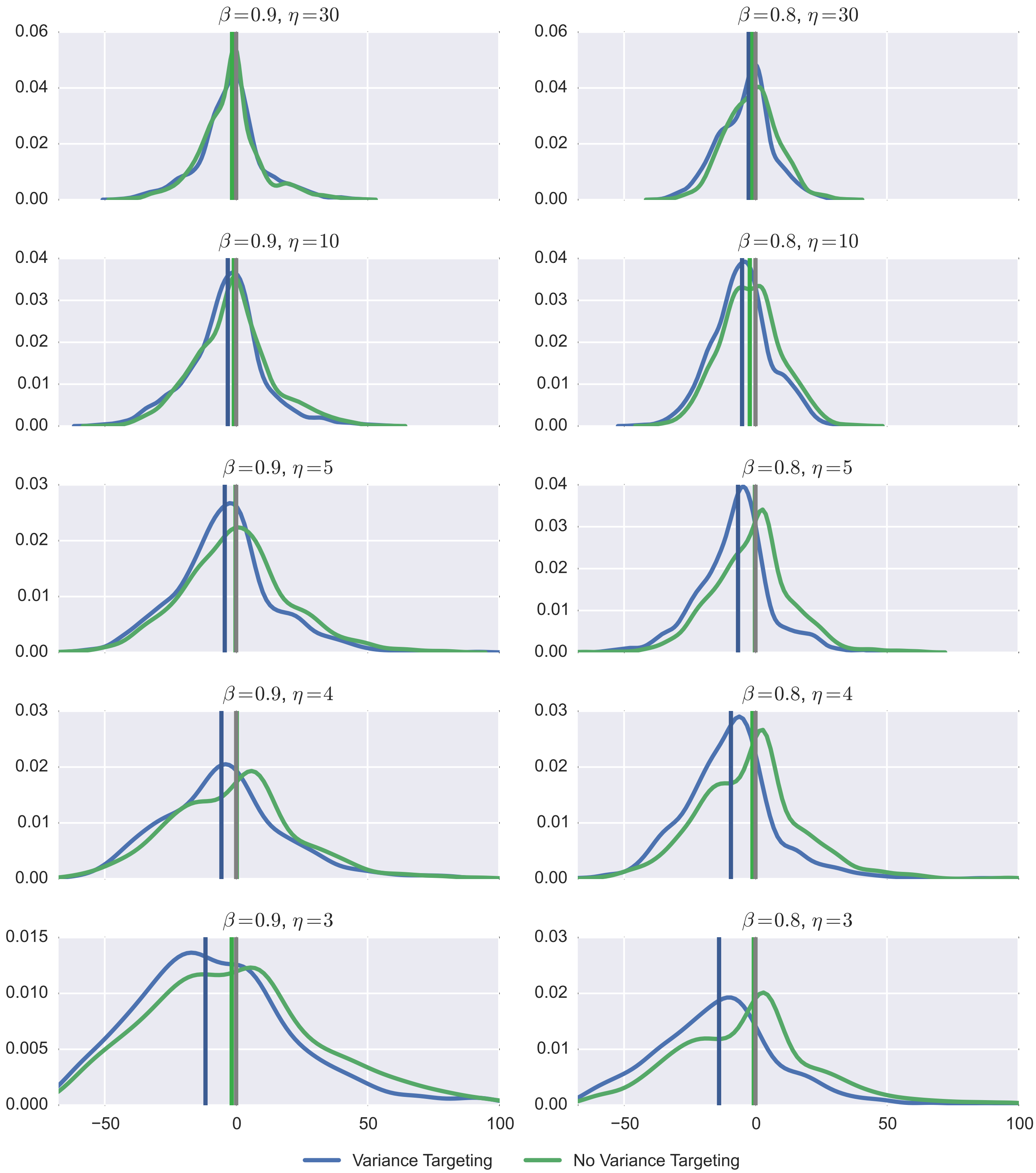

Figure 4: News impact parameter bias, B_α, for varying β and η.
Horizontal axis is in percentage of the true parameter value. Vertical solid lines are medians and zero (gray). Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

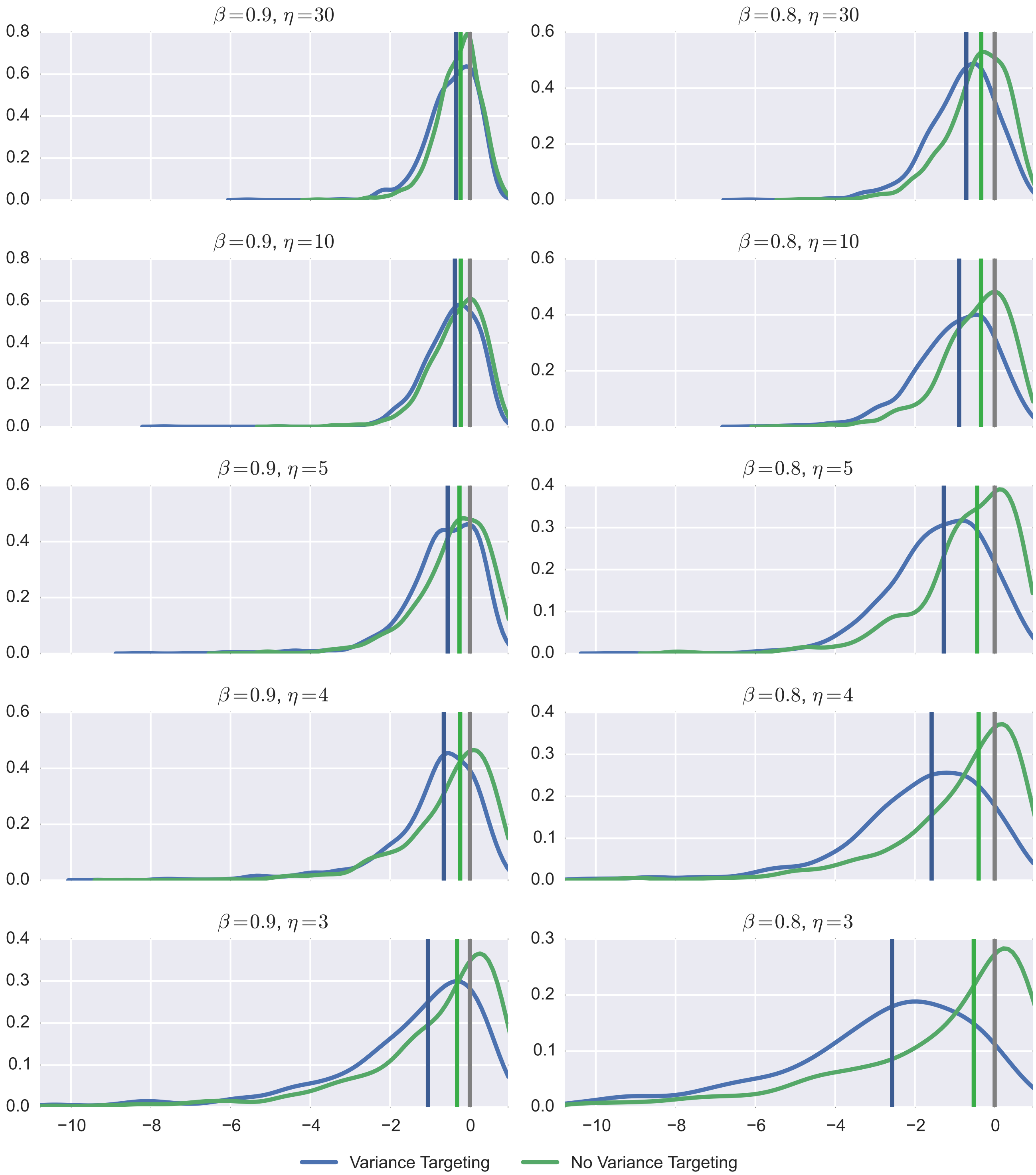

Figure 5: Persistence bias, $B_{\alpha+\beta}$, for varying β and η. % of the true parameter value. Horizontal axis is in percentage of the true parameter value. Vertical solid lines are medians and zero (gray). Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

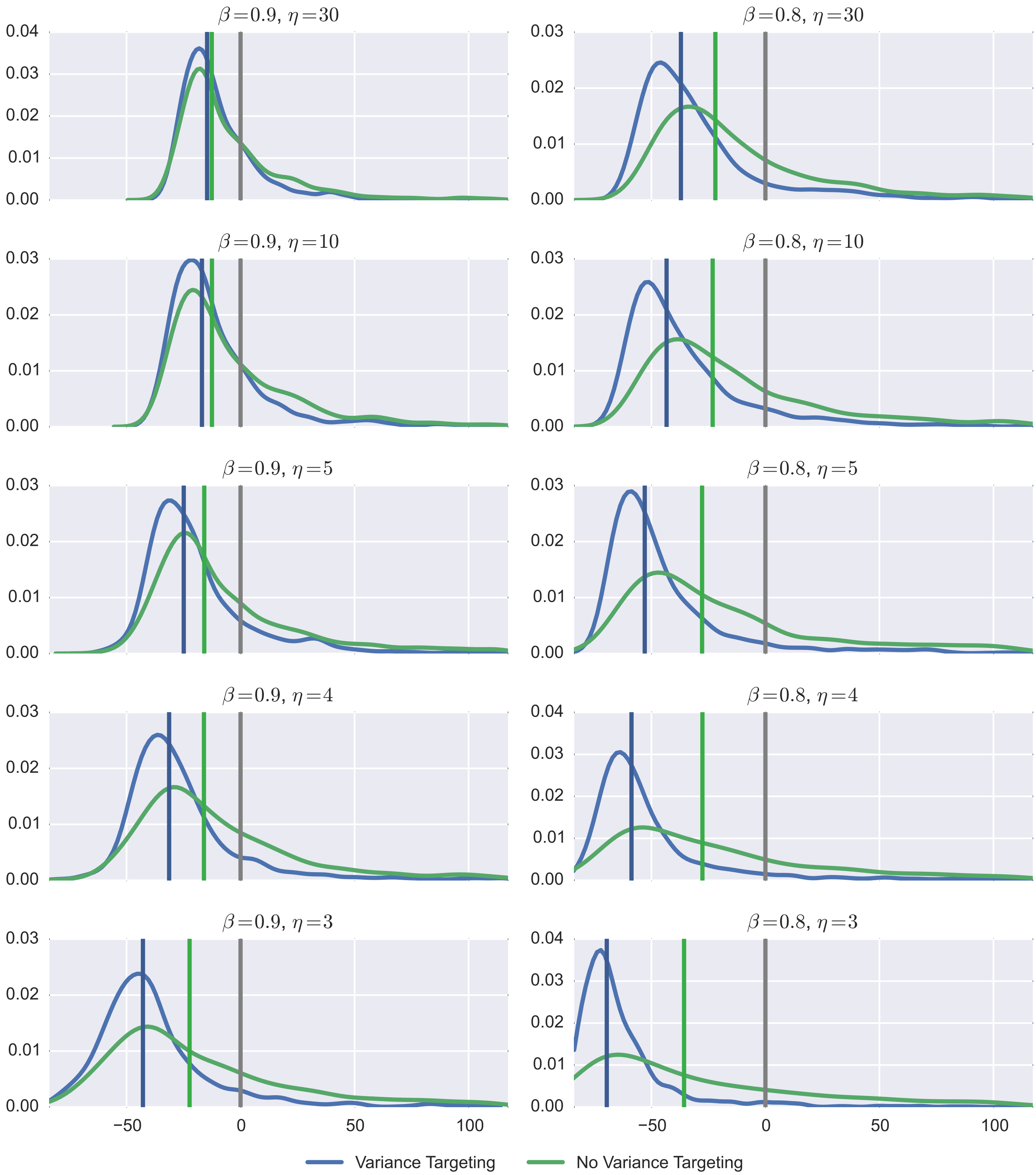

Figure 6: Unconditional variance bias, B_{σ^2}, for varying β and η.
Horizontal axis is in percentage of the true parameter value. Vertical solid lines are medians and zero (gray). Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

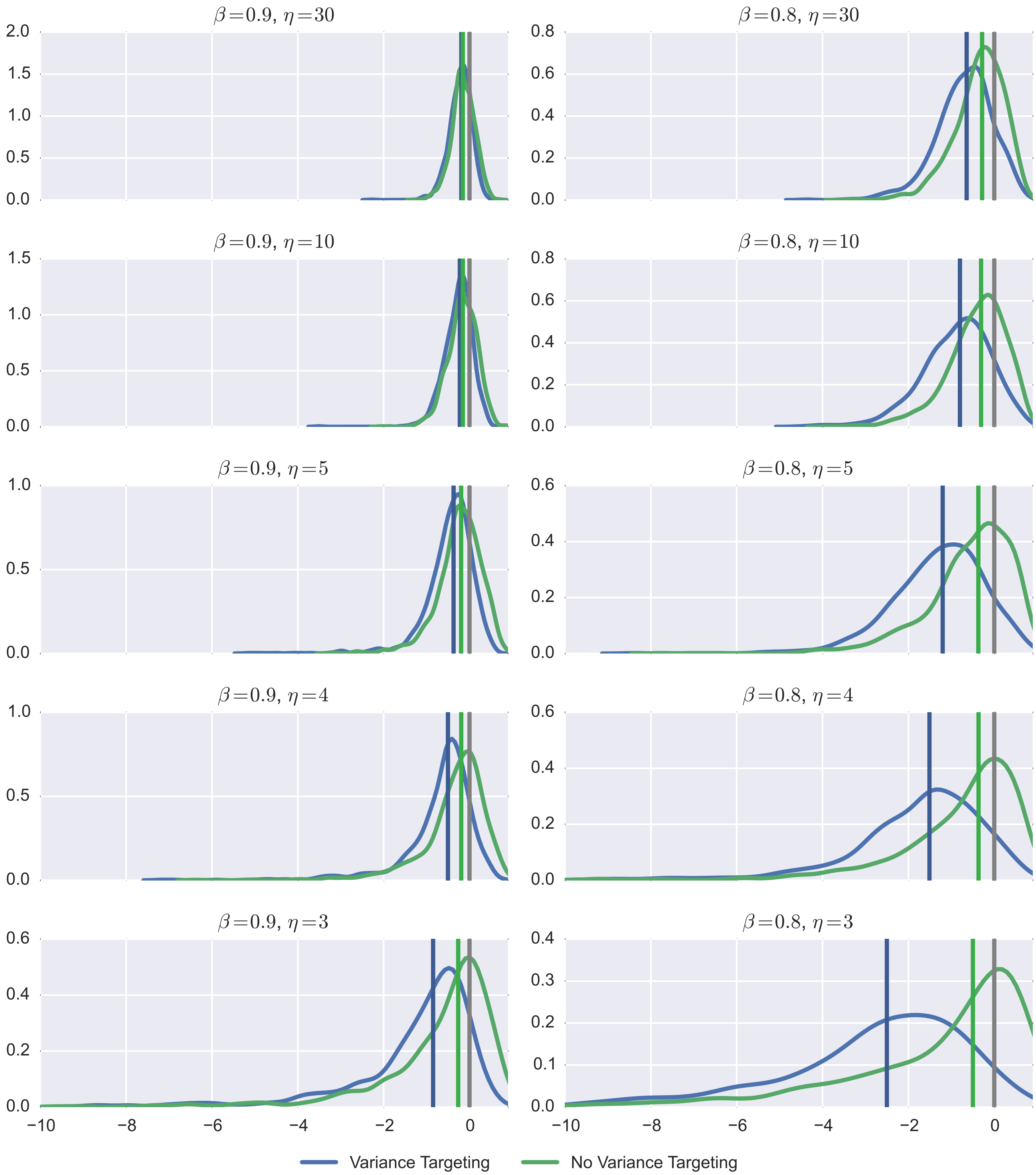

Figure 7: Log of one period ahead forecast, B_{+1}, for varying β and η.
Values on horizontal axis are multiplied by 100. Vertical solid lines are medians and zero (gray). Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

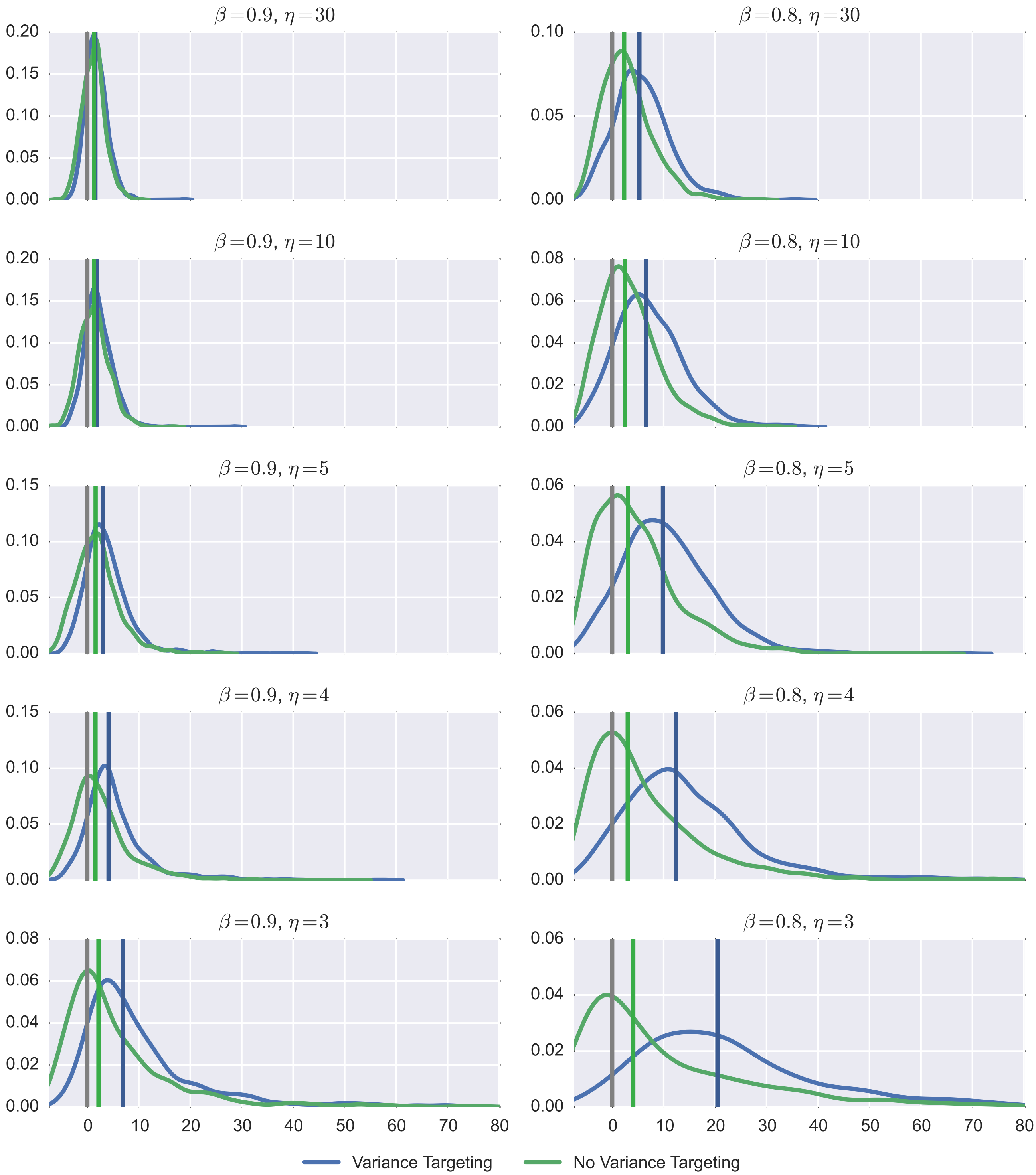

Figure 8: One period in advance Value-at-Risk bias, $B_{5\%}$, for varying β and η. Values on horizontal axis are multiplied by 100. Vertical solid lines are medians and zero (gray). Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

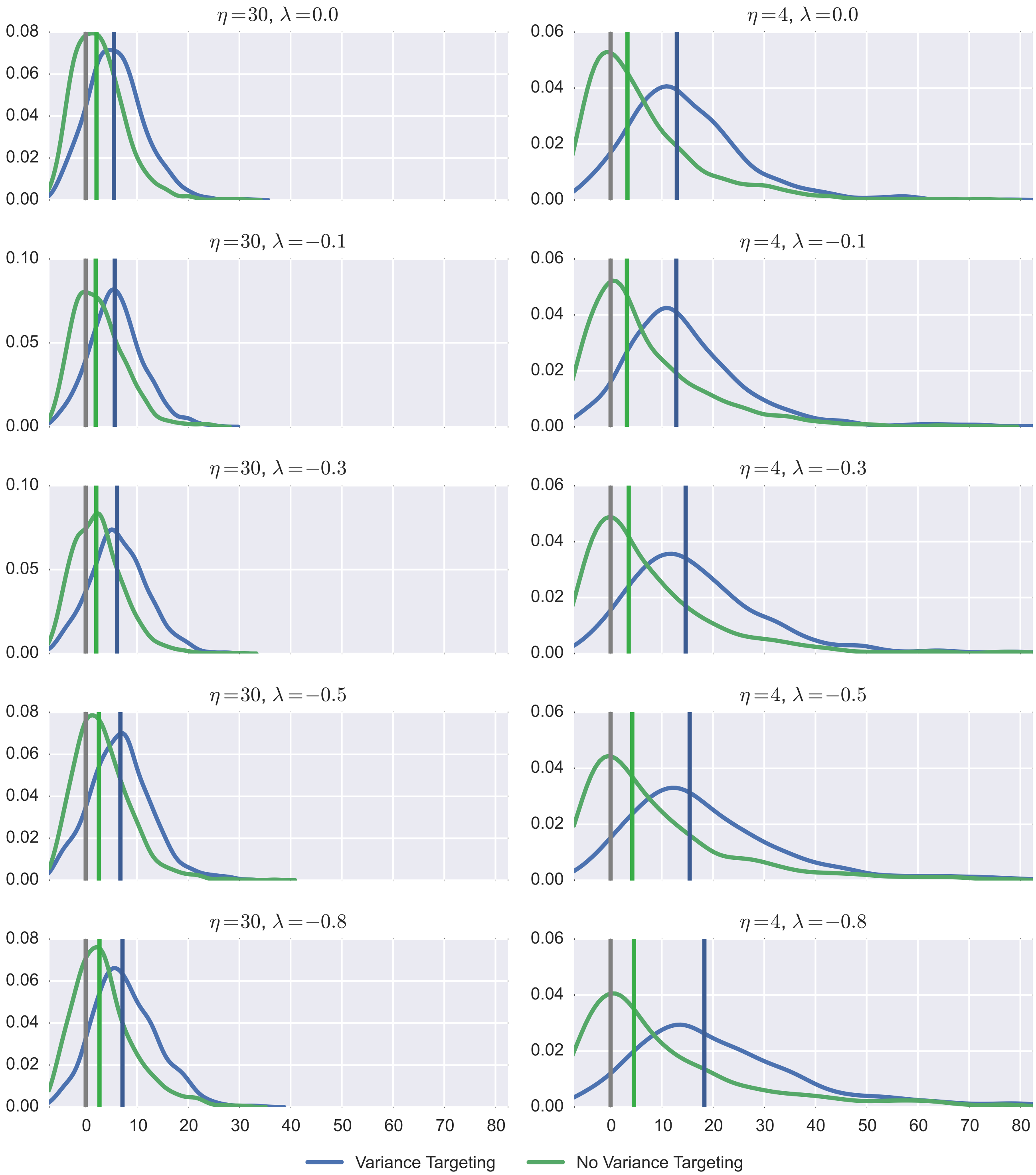

Figure 9: One period in advance Value-at-Risk bias, $B_{5\%}$, for varying η and λ. Values on horizontal axis are multiplied by 100. Vertical solid lines are medians and zero (gray). Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, feedback β is 0.8, and $\omega = 1 - \alpha - \beta$.

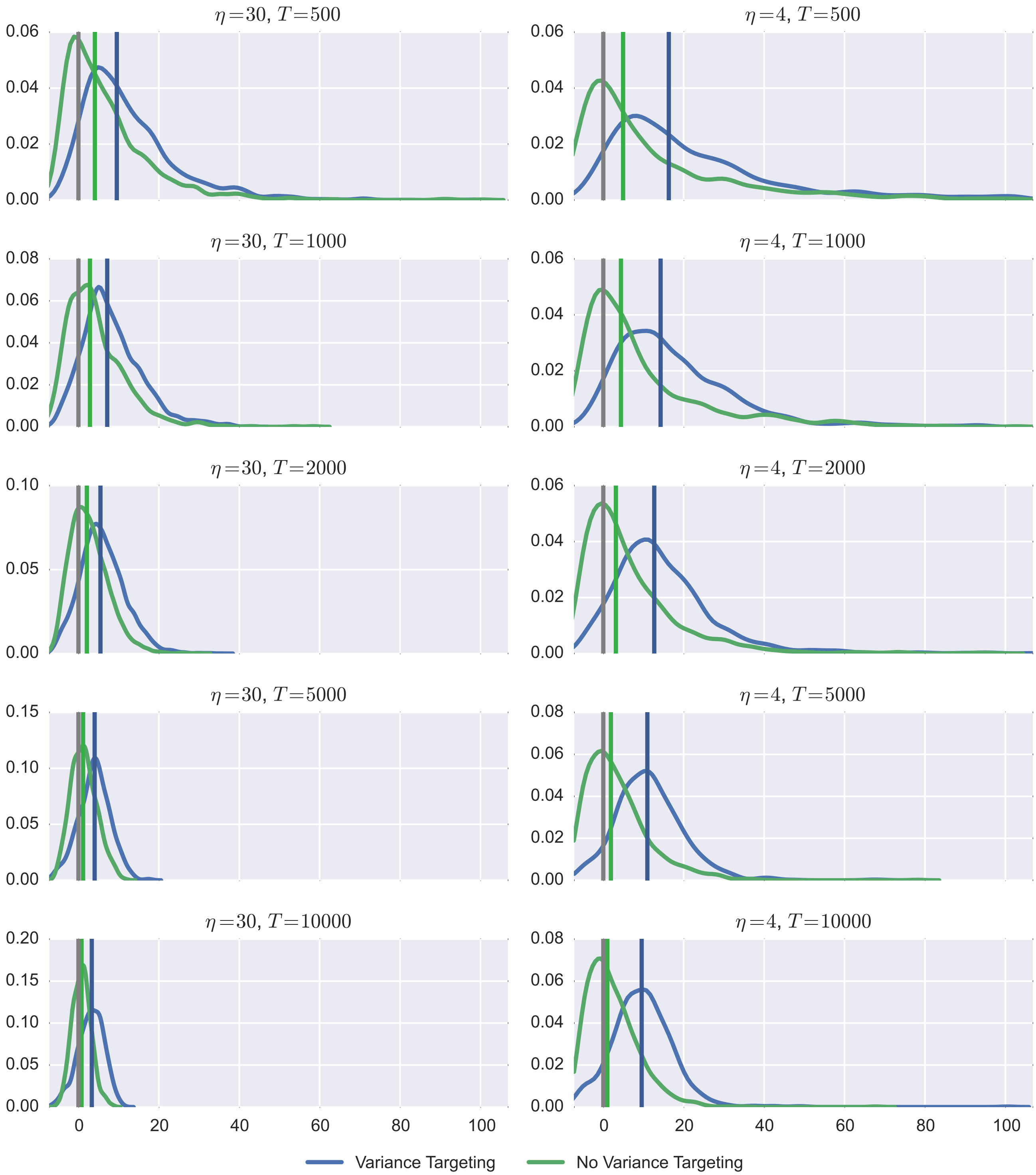

Figure 10: One period in advance Value-at-Risk bias, $B_{5\%}$, for varying η and T. Values on horizontal axis are multiplied by 100. Vertical solid lines are medians and zero (gray). True value of persistence $\alpha + \beta$ is 0.99, feedback β is 0.8, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

		No Variance Targeting					Variance Targeting				
β	η	5	25	50	75	95	5	25	50	75	95
0.9	30	-29.03	-3.10	6.34	30.97	71.08	-31.66	-7.34	15.28	43.72	94.54
	10	-34.46	-10.85	8.62	36.63	86.69	-33.83	-9.22	14.03	44.33	89.89
	5	-43.04	-16.13	10.61	40.73	108.89	-40.05	-12.65	16.86	46.19	106.00
	4	-47.65	-18.47	8.58	46.60	124.59	-42.63	-12.04	17.05	48.35	133.23
	3	-60.20	-23.12	13.96	60.10	173.06	-53.84	-17.53	18.50	70.16	175.32
0.8	30	-27.76	-9.88	6.70	24.54	54.98	-23.36	-7.55	7.16	26.03	58.22
	10	-27.98	-11.57	5.67	23.98	59.22	-26.80	-8.04	8.79	26.23	61.12
	5	-33.69	-10.81	7.08	25.18	68.33	-29.05	-7.20	8.03	25.87	55.99
	4	-39.76	-13.87	4.79	27.27	83.18	-38.52	-8.47	9.32	30.13	79.17
	3	-49.51	-16.61	7.72	35.98	109.70	-49.07	-9.92	12.88	42.72	110.69

Table 1: Intercept bias, B_ω, for varying β and η.
The values are in percentage of the true parameter value. Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

		No Variance Targeting					Variance Targeting				
β	η	5	25	50	75	95	5	25	50	75	95
0.9	30	-2.78	-0.12	0.01	0.18	1.47	-3.09	-0.19	-0.03	0.08	2.19
	10	-3.58	-0.21	0.00	0.25	2.42	-3.24	-0.18	-0.01	0.13	2.79
	5	-4.86	-1.86	-0.02	0.52	3.00	-4.60	-0.90	-0.01	0.18	3.21
	4	-5.32	-2.04	-0.04	0.96	3.48	-6.42	-1.90	-0.02	1.02	3.68
	3	-11.16	-3.24	-0.16	1.61	4.70	-10.00	-3.45	-0.01	1.93	4.97
0.8	30	-4.21	-1.84	-0.05	1.42	3.65	-4.61	-1.10	-0.00	1.18	3.94
	10	-5.16	-2.05	0.04	1.70	4.35	-5.23	-0.67	-0.00	1.44	4.55
	5	-6.66	-2.67	-0.02	0.95	5.34	-6.27	-0.17	-0.00	0.60	5.73
	4	-9.33	-2.97	-0.00	2.07	6.78	-9.08	-0.21	-0.00	1.49	6.76
	3	-17.26	-5.06	-0.04	2.65	8.44	-14.86	-2.84	-0.00	3.16	9.39

Table 2: Feedback bias, B_β, for varying β and η.
The values are in percentage of the true parameter value. Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

β	η	No Variance Targeting					Variance Targeting				
		5	25	50	75	95	5	25	50	75	95
	30	-21.30	-8.85	-1.54	2.52	19.93	-23.57	-8.51	-1.48	3.69	20.43
	10	-25.35	-10.99	-0.88	6.52	26.16	-28.52	-11.61	-3.18	3.43	20.39
0.9	5	-33.09	-12.97	-0.24	10.89	37.50	-35.53	-15.54	-4.34	4.51	29.10
	4	-39.11	-17.96	0.23	11.78	42.91	-41.50	-21.09	-5.61	7.05	38.27
	3	-55.33	-24.64	-1.70	20.94	83.92	-60.06	-30.26	-11.67	8.09	53.29
	30	-17.04	-8.04	-1.20	4.83	14.88	-20.82	-10.57	-2.60	1.92	13.12
	10	-20.28	-9.62	-2.18	4.86	18.25	-24.13	-12.17	-5.11	1.30	14.96
0.8	5	-25.31	-10.66	-0.22	6.03	23.96	-32.04	-15.82	-6.65	-1.05	17.22
	4	-34.20	-15.61	-1.16	6.07	31.69	-37.44	-20.56	-9.35	-1.26	21.27
	3	-49.33	-22.49	-0.65	9.45	57.96	-55.62	-30.34	-13.86	-1.01	29.82

Table 3: News impact parameter bias, B_α, for varying β and η.
The values are in percentage of the true parameter value. Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

β	η	No Variance Targeting					Variance Targeting				
		5	25	50	75	95	5	25	50	75	95
	30	-1.28	-0.61	-0.23	0.07	0.49	-1.58	-0.80	-0.35	0.03	0.41
	10	-1.58	-0.73	-0.23	0.17	0.57	-1.71	-0.87	-0.37	0.06	0.46
0.9	5	-2.18	-0.88	-0.26	0.25	0.78	-2.39	-1.12	-0.55	0.01	0.45
	4	-2.75	-1.07	-0.24	0.29	0.79	-3.48	-1.36	-0.65	-0.11	0.47
	3	-4.77	-1.58	-0.32	0.37	0.90	-6.32	-2.40	-1.05	-0.18	0.48
	30	-1.97	-0.86	-0.34	0.16	0.60	-2.41	-1.33	-0.71	-0.22	0.42
	10	-2.26	-0.99	-0.34	0.17	0.66	-2.87	-1.59	-0.89	-0.25	0.39
0.8	5	-3.06	-1.20	-0.44	0.22	0.78	-3.65	-2.13	-1.27	-0.49	0.35
	4	-4.32	-1.55	-0.40	0.28	0.84	-5.57	-2.77	-1.58	-0.65	0.34
	3	-7.75	-2.65	-0.52	0.44	0.90	-10.53	-4.46	-2.57	-1.23	0.15

Table 4: Persistence bias, $B_{\alpha+\beta}$, for varying β and η.
The values are in percentage of the true parameter value. Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

β	η	No Variance Targeting					Variance Targeting				
		5	25	50	75	95	5	25	50	75	95
	30	-27.46	-20.07	-12.62	3.92	47.97	-27.66	-20.98	-14.67	-3.21	25.03
	10	-31.02	-22.58	-12.54	9.65	70.89	-32.73	-25.15	-16.93	-3.99	32.34
0.9	5	-40.04	-27.26	-15.92	14.16	187.71	-43.25	-34.06	-24.91	-11.60	34.55
	4	-49.70	-32.35	-16.06	16.11	211.30	-51.59	-40.97	-31.32	-18.43	31.39
	3	-65.07	-44.13	-22.35	29.72	504.19	-70.23	-53.49	-42.81	-29.56	18.15
	30	-50.06	-37.25	-21.96	7.36	112.58	-56.14	-47.64	-37.06	-21.23	48.23
	10	-54.66	-40.54	-23.13	10.17	141.87	-61.05	-52.99	-43.34	-25.32	40.32
0.8	5	-63.48	-47.95	-27.78	20.48	244.89	-69.04	-61.60	-52.92	-37.01	45.17
	4	-69.55	-53.72	-27.62	26.84	333.32	-75.42	-67.01	-58.74	-45.20	25.61
	3	-81.09	-65.19	-35.68	46.58	678.36	-83.86	-76.23	-69.57	-58.41	-8.17

Table 5: Unconditional variance bias, B_{σ^2}, for varying β and η.
The values are in percentage of the true parameter value. Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

β	η	No Variance Targeting					Variance Targeting				
		5	25	50	75	95	5	25	50	75	95
	30	-0.61	-0.31	-0.16	0.04	0.27	-0.66	-0.37	-0.19	-0.03	0.16
	10	-0.73	-0.37	-0.16	0.08	0.35	-0.83	-0.46	-0.23	-0.04	0.21
0.9	5	-1.16	-0.48	-0.20	0.12	0.50	-1.35	-0.69	-0.37	-0.12	0.23
	4	-1.77	-0.61	-0.19	0.13	0.55	-2.24	-0.90	-0.50	-0.20	0.20
	3	-3.41	-1.08	-0.26	0.18	0.77	-4.75	-1.67	-0.85	-0.34	0.15
	30	-1.45	-0.67	-0.28	0.07	0.46	-1.86	-1.07	-0.64	-0.26	0.29
	10	-1.74	-0.77	-0.31	0.08	0.52	-2.31	-1.37	-0.80	-0.32	0.27
0.8	5	-2.58	-1.04	-0.37	0.15	0.61	-3.19	-1.90	-1.20	-0.57	0.29
	4	-3.69	-1.34	-0.37	0.18	0.66	-4.87	-2.50	-1.51	-0.76	0.18
	3	-7.07	-2.37	-0.50	0.27	0.80	-9.96	-4.18	-2.50	-1.29	-0.07

Table 6: Log of one period ahead forecast, B_{+1}, for varying β and η.
Values are multiplied by 100. Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

β	η	No Variance Targeting					Variance Targeting				
		5	25	50	75	95	5	25	50	75	95
0.9	30	-2.24	-0.30	1.29	2.51	4.99	-1.30	0.23	1.56	3.03	5.39
	10	-2.90	-0.63	1.28	3.01	5.98	-1.69	0.34	1.85	3.76	6.84
	5	-4.15	-0.95	1.60	3.96	9.54	-1.87	1.00	3.05	5.66	11.04
	4	-4.56	-1.03	1.59	5.02	14.53	-1.67	1.68	4.12	7.42	18.35
	3	-6.33	-1.46	2.16	8.88	27.78	-1.27	2.81	6.96	13.65	38.62
0.8	30	-3.78	-0.54	2.33	5.53	11.87	-2.42	2.11	5.29	8.81	15.23
	10	-4.30	-0.68	2.53	6.32	14.23	-2.26	2.63	6.57	11.25	18.90
	5	-5.04	-1.23	3.05	8.53	21.06	-2.38	4.65	9.85	15.52	25.99
	4	-5.44	-1.46	3.01	10.96	30.07	-1.49	6.24	12.37	20.44	39.55
	3	-6.56	-2.18	4.09	19.34	57.10	0.61	10.58	20.45	34.02	79.94

Table 7: One period in advance Value-at-Risk bias, $B_{5\%}$, for varying β and η. Values are multiplied by 100. Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.

η	λ	No Variance Targeting					Variance Targeting				
		5	25	50	75	95	5	25	50	75	95
30	0.0	-4.00	-1.15	2.11	5.36	11.66	-2.59	2.10	5.50	9.27	15.76
	-0.1	-4.21	-1.08	1.95	5.52	11.58	-2.14	2.49	5.67	9.14	15.26
	-0.3	-4.44	-1.23	2.06	5.28	11.16	-2.61	2.74	6.11	10.02	16.42
	-0.5	-4.18	-0.50	2.57	6.34	13.39	-2.98	2.83	6.77	10.52	16.94
	-0.8	-4.58	-0.61	2.69	6.68	15.37	-1.31	3.49	7.18	11.66	19.16
4	0.0	-5.38	-1.53	3.30	10.81	30.89	-0.74	6.78	12.94	20.33	36.94
	-0.1	-5.85	-1.15	3.19	11.39	31.05	-0.52	6.98	12.84	20.30	38.18
	-0.3	-5.82	-1.43	3.54	12.46	36.16	-0.18	7.67	14.66	23.92	46.96
	-0.5	-5.97	-1.40	4.24	14.49	43.71	-0.38	8.25	15.44	26.21	53.68
	-0.8	-6.34	-1.18	4.54	17.19	55.07	0.30	10.21	18.31	30.74	63.22

Table 8: One period in advance Value-at-Risk bias, $B_{5\%}$, for varying η and λ. Values are multiplied by 100. Number of observations is 2000. True value of persistence $\alpha + \beta$ is 0.99, feedback β is 0.8, and $\omega = 1 - \alpha - \beta$.

		No Variance Targeting					Variance Targeting				
η	T	5	25	50	75	95	5	25	50	75	95
	500	-4.27	-0.75	4.09	11.04	28.57	-1.08	4.20	9.54	17.58	37.05
	1000	-4.41	-0.98	2.87	8.21	17.47	-1.84	3.28	7.16	12.25	21.55
30	2000	-3.86	-0.72	2.10	5.51	11.29	-2.51	2.21	5.48	9.35	15.94
	5000	-3.51	-1.05	1.16	3.46	7.20	-2.55	1.51	4.03	6.64	10.44
	10000	-2.83	-0.85	0.77	2.26	4.64	-3.02	1.00	3.32	5.36	8.02
	500	-6.23	-1.39	4.92	19.78	68.40	-0.03	7.39	16.29	32.07	84.10
	1000	-5.50	-0.99	4.39	13.95	47.10	0.27	7.11	14.25	25.21	47.05
4	2000	-5.40	-1.52	3.13	10.93	30.68	-1.02	6.62	12.68	20.52	38.37
	5000	-5.89	-2.28	1.89	7.00	19.34	-0.82	5.86	10.96	16.08	25.79
	10000	-5.53	-2.40	1.01	5.62	14.35	-2.44	4.99	9.54	14.12	21.50

Table 9: One period in advance Value-at-Risk bias, $B_{5\%}$, for varying η and T. Values are multiplied by 100. True value of persistence $\alpha + \beta$ is 0.99, feedback β is 0.8, skewness λ is 0, and $\omega = 1 - \alpha - \beta$.