August 8, 2023

This document describes the multi-pronged approach to model efficacy. The three approaches include

1. Cross-validation, for estimation of out-of-sample efficacy
2. Bayes factor, for estimation of model fit relative to a baseline (in progress)
3. A sanity check that uses extensive margin tests for conditional efficacy and meaningful probability statements on exposures.

# 1 Cross-validation

The standard OOS $R^2$ as described in Campbell-Thomson 2007 trivially applies to a leave-one-out (LOO) cross-validation setting:

$$\begin{aligned} R^2_{ct} =& 1 - \frac{TE\left(r_t, \hat{r}_{-t}\right)}{Var\left(r_t\right)} \\ =& 1 - \frac{E\left[\left(r_t - \hat{r}_{-t}\right)^2\right]}{E\left[\left(r_t - \overline{r}_{-t}\right)^2\right]} \\ \implies \hat{R}^2_{ct} =& 1 - \frac{\sum_{t\in 1:T}\left(r_t - \hat{r}_{-t}\right)^2}{\sum_{t\in 1:T}\left(r_t - \overline{r}_{-t}\right)^2} \end{aligned}$$

In the above, $r_t$ corresponds to the reported returns and $\hat{r}_{-t}$ denotes the model-predicted returns using data up to but excluding time $t$. Similarly $\overline{r}$ represents the predictions under a "null" or baseline model. For the purposes of this discussion, the baseline model is a Bayesian estimation of a running average of returns. Consistent with the time-series setting, both future data and the hold-out date are excluded from the available data when fitting the models. Simply plugging in expectations gives a naive estimate using the posterior in a Bayesian setting:

$$\begin{aligned} R^2_{nbcv} =& 1 - \frac{E\left[\left(r_t - \int_{\Theta}\hat{r}_{-t}p\left(\hat{r}_{-t}|F_{-t}\right)d\Theta_{-t}\right)^2\right]}{E\left[\left(r_t - \int_{\Theta^0}\overline{r}_{-t}p\left(\overline{r}_{-t}|F_{-t}\right)d\Theta^0_{-t}\right)^2\right]} \\ \hat{R}^2_{nbcv} =& 1 - \frac{\sum_{t\in 1:T}\left(r_t - \frac{1}{J}\sum_{j\in 1:J}\hat{r}_{-t}\right)^2}{\sum_{t\in 1:T}\left(r_t - \frac{1}{J}\sum_{j\in 1:J}\overline{r}_{-tj}\right)^2} \end{aligned}$$

The above naive application however does not account for parameter uncertainty and only minimally exploits the distributional information provided by the model. Consider instead integrating the squared deviation across the posterior distribution at each point in time, and then averaging as above:

$$
\begin{aligned}
R^2_{bcv} =& 1-\frac{E\left[E_t\left(r_t-\hat{r}_{-t}\right)^2|f_t\right]}{E\left[E_t\left(r_t-\overline{r}_{-t}\right)^2\right]}\\
=& 1-\frac{E\left[\int_{\Theta}\left(r_t-\hat{r}_{-t}\right)^2 p\left(\hat{r}_{-t}|F_{-t}\right)d\Theta_{-t}\right]}{E\left[\int_{\Theta^0}\left(r_t-\overline{r}_{-t}\right)^2 p\left(\overline{r}_{-t}|F_{-t}\right)d\Theta^0_{-t}\right]}\\
\hat{R}^2_{bcv} =& 1-\frac{\sum_{t\in 1:T}\left[\sum_{j\in 1:J}\left(r_t-\hat{r}_{-tj}\right)^2\right]}{\sum_{t\in 1:T}\left[\sum_{j\in 1:J}\left(r_t-\overline{r}_{-tj}\right)^2\right]}
\end{aligned}
$$

Here, $\overline{r}_{-tj}$ is the prediction of a baseline model with parameters $\Theta^0_j$ while $\hat{r}_{-tj}$ is the prediction of a focal model of parameters $\Theta_j$. Based on the literature reviewed, the above adaptation of the Campbell-Thomson $R^2$ metric to a Bayesian setting is novel. It intuitively compares the OOS efficacy of the two prediction methodologies. In G3, the baseline model is a normal-gamma running average. Note that the metric ranges such that $R^2_{bcv} \in (-\infty, 1]$. This is in contrast to standard definitions of $R^2$ in an Ordinary Least Squares (OLS) setting but consistent with the standard Campbell-Thomson metric.

## 1.1 Importance Sampling

Bayesian cross-validation is computationally demanding as it necessitates re-fitting the model for each fold. If performance is a consideration, importance weighting provides a partial solution. Applying importance sampling involves weighting each prediction relative to its probability density:

$$
\hat{R}^2_{cv} = 1-\frac{E\left[\sum_{j\in 1:J}\left(r_t-\hat{r}_{-tj}\right)^2 w_{jt}\right]}{E\left[\sum_{j\in 1:J}\left(r_t-\overline{r}_{-tj}\right)^2 w^0_{jt}\right]}
$$

Adjusting Vehtari 2016/2022 to a time-series setting by including only dates before the hold-out date:

$$
\begin{aligned}
\tilde{w}_{jt} \equiv& \frac{p\left(\hat{r}_{tj}|F_{-t}\right)}{p\left(\hat{r}_{tj}|F\right)}\\
w_{jt} =& \frac{\tilde{w}_{jt}}{\sum_j \tilde{w}_{jt}}
\end{aligned}
$$

Unfortunately, the above estimator on its own can lead to very poor results when a necessarily small number of samples from low density regions of the full-sample dominate the distribution. The issue is related to the thick-tail requirement in other importance sampling settings, though as discussed Vehtari et al 2022 the problem can arise even when the numerator has thick tails. While the following techniques mitigate the issue, each involve a trade-off between the number of times refitting the model versus the quality of the estimation.

Brukner-Gabry-Vehtari 2020 describes the basic approach. First, the stability of the estimation weights increases by estimating the model for a sequence of periods before the hold-out date as opposed to the full sample. This leads to unnormalized weights as follows. For all $t$ and any $s \leq t$:

$$
\tilde{w}_{jt} \equiv \frac{p\left(\hat{r}_{tj}|F_{-t}\right)}{p\left(\hat{r}_{tj}|F_{-s}\right)}
$$

Despite this mitigation, the stability of the weights still depends on the degree to which the model fitted up to $s$ matches the distribution of the returns at $t$. Periodic refitting combined with modifications to the tails of the distribution helps further improve the stability of the estimates. The three weighted approaches include using the original importance weights as described above, truncation as described in Ionides 2008, and Pareto Smoothed Importance Sampling (PSIS) as described in Vehtari 2016,2022.

#### 1.1.1 Ionides truncation

Ionides 2008 proposes truncating the weights for the $J$ posterior draws at a multiple of the average weight

$$\begin{aligned}\tilde{z}^I_{jt} &= \min\left(\tilde{w}_{jt}, \bar{\tilde{w}}_{jt} \times \sqrt{J}\right) \\ w^I_{jt} &= \frac{\tilde{z}^I_{jt}}{\sum_j \tilde{z}^I_{jt}}\end{aligned}$$

Vehtari 2022 discusses that while effective at reducing the variance of the estimates, the Ionides 2008 technique can bias the true posterior distribution by more than is necessary.

#### 1.1.2 Vehtari-Simpson-Gelman-Yao-Gabry 2022 PSIS

Start by sorting the raw importance ratios $\tilde{w}$ in order of magnitude. Then calculate the cutoff $M$ as follows:

$$M = \min\left(0.2J, 3\sqrt{J}\right)$$

For all $j > M$, fit a Generalized Pareto distribution $\left(GPD\left(\tilde{w_m}; \mu, \sigma, k\right)\right)$. To do this, set $\mu = \tilde{w}_M$. For simplicity relative to Vehtari, estimate $\sigma$ and $k$ using maximum likelihood:

$$\begin{aligned}\sigma, k &= arg \min_{\sigma,k} ll\left(\sigma, k\right) \\ &s.t. \\ ll\left(\sigma, k\right) &= \sum_{m \in (M+1):J} \log GPD\left(\tilde{w_m}; \mu, \sigma, k\right)\end{aligned}$$

Then define the unnormalized smoothed weights as:

$$\tilde{z}^{PSIS}_{jt} = \begin{cases} \tilde{w}_{jt} & j \leq M \\ \min\left[F^{-1}_{GPD}\left(\frac{\tilde{w}_{jt} - 0.5}{M}\right), \tilde{w}_{jt}\right] & j > M \end{cases}$$

Finally, normalize the weights in the usual manner:

$$w^{PSIS}_{jt} = \frac{\tilde{z}^{PSIS}_{jt}}{\sum_j \tilde{z}^{PSIS}_{jt}}$$

As discussed in Vehtari et al 2022 and Brukner et al 2020, an advantage of this approach is $k$ serves as a intrinsic diagnostic. Following Vehtari, an adaptive algorithm accepts the PSIS estimate if $k \leq 0.7$ and re-estimates the cross-validation if $k > 0.7$. Anecdotally, the algorithm speeds up the performance by about 4x.

## 1.2 Drawing from the mean model:

The baseline model's posterior is proportional to:

$$p^0(\mu,\tau|y) \propto \prod_{s=1}^{S} N\left(y_s;\mu,\tau^{-1}\right) \times N\left(\mu;\mu_0,\frac{1}{\tau_{\mu 0}\tau}\right) \times Gamma\left(\tau;\alpha_0,\zeta_0\right)$$

Integrating out $\tau$ in a semi-conditional approach leads to independent draws on each iteration. Alternatively, implementation can utilize the normal-gamma distribution as described in the appendix. Start by compounding the normals:

$$\begin{aligned}
\log p^0(\mu,\tau|y) =& \frac{\tau}{2}\left(-\sum_{s=1}^{S}(\mu-y_s)^2 - \tau_{\mu 0}(\mu-\mu_0)^2\right) + c_1^{\mu} \\
=& -\frac{\tau}{2}\left((S+\tau_{\mu 0})\mu^2 - 2\mu\left(\sum_{s=1}^{S} y_s + \tau_{\mu 0}\mu_0\right)\right) + c_2^{\mu} \\
=& -\frac{\tau(S+\tau_{\mu 0})}{2}\left(\mu^2 - 2\mu\left(\sum_{s=1}^{S} y_s + \tau_{\mu 0}\mu_0\right)(S+\tau_{\mu 0})^{-1}\right) + c_2^{\mu} \\
=& -\frac{\tau_\mu}{2}(\mu-\mu_\mu)^2 + c_3^{\mu} \\
=& \log N\left(\mu;\mu_\mu,\tau_\mu^{-1}\right) + c_4^{\mu} \\
s.t.& \\
\tau_\mu =& \tau(S+\tau_{\mu 0}) \\
\mu_\mu =& \tau\tau_\mu^{-1}\left(\sum_{s=1}^{S} y_s + \tau_{\mu 0}\mu_0\right) \\
c_1^{\mu} \equiv& \frac{S+1}{2}\left(\log(\tau) - \log(2\pi)\right) + \frac{1}{2}\log(\tau_{\mu 0}) + \log Gamma(\tau;\alpha_0,\zeta_0) + c_1^{ev0} \\
c_2^{\mu} \equiv& c_1^{\mu} - \frac{\tau}{2}\sum_{s=1}^{S} y_s^2 - \frac{\tau\tau_{\mu 0}}{2}\mu_0^2 \\
c_3^{\mu} \equiv& c_2^{\mu} + \frac{\tau_\mu}{2}\mu_\mu^2 \\
c_4^{\mu} \equiv& c_3^{\mu} - \frac{1}{2}\log\left(\frac{\tau_\mu}{2\pi}\right)
\end{aligned}$$

Marginalize out $\mu$ to allow for independent draws:

$$
\begin{aligned}
\log p^0(\tau|y) =& \frac{S+1}{2}\log(\tau) + (\alpha_0 - 1)\log\tau - \zeta_0\tau - \frac{\tau}{2}\sum_{s=1}^{S} y_s^2 - \frac{\tau\tau_{\mu 0}}{2}\mu_0^2 + \frac{\tau_\mu}{2}\mu_\mu^2 - \frac{1}{2}\log(\tau_\mu) + c_1^\tau \\
=& \frac{S+1}{2}\log(\tau) + (\alpha_0 - 1)\log\tau - \frac{1}{2}\log(\tau) \\
& - \left(\frac{1}{2}\sum_{s=1}^{S} y_s^2 + \frac{\tau_{\mu 0}}{2}\mu_0^2 - \tau\left(\frac{S+\tau_{\mu 0}}{2}\right)\mu_\mu^2 + \zeta_0\right)\tau + c_1^\tau \\
=& (\alpha - 1)\log\tau - \zeta\tau + c_1^\tau \\
=& \log\left(Gamma\left(\tau;\alpha,\zeta\right)\right) + c_2^\tau \\
s.t. & \\
\alpha \equiv& \frac{S}{2} + \alpha_0 \\
\zeta \equiv& \frac{1}{2}\sum_{s=1}^{S} y_s^2 + \frac{\tau_{\mu 0}}{2}\mu_0^2 - \frac{\left(\sum_{s=1}^{S} y_s + \tau_{\mu 0}\mu_0\right)^2}{2\left(S+\tau_{\mu 0}\right)} + \zeta_0 \\
c_1^\tau \equiv& \frac{1}{2}\log(\tau_{\mu 0}) + \alpha_0\log\zeta_0 - \log\Gamma(\alpha_0) - \frac{S}{2}\log(2\pi) + c_1^{ev0} \\
c_2^\tau \equiv& c_1^\tau - \frac{1}{2}\log(S+\tau_{\mu 0}) \\
c_3^\tau \equiv& c_1^\tau - \alpha\log\zeta + \log\Gamma(\alpha)
\end{aligned}
$$

Note that since $c_1^{ev0} = -P(D|M)$:

$$
\begin{aligned}
P(D|M) =& \frac{1}{2}\log(\tau_{\mu 0}) + \alpha_0\log\zeta_0 - \log\Gamma(\alpha_0) - \frac{S}{2}\log(2\pi) \\
& - \frac{1}{2}\log(S+\tau_{\mu 0}) - \alpha\log\zeta + \log\Gamma(\alpha)
\end{aligned}
$$

# 2 Bayes Factor

# 3 Extensive margin tests

This approach identifies assets for which the model is particularly ineffective. Model analysis for such assets should not be shown on the Architect platform.

## 3.1 Motivation and Discussion

- The tests start from the principle of permissivness. The tests err on the side of providing information on risks that might not exist, on the presumption that such errors are less costly than failing to inform about an existing risk.
- The model currently provides the following important information to clients and advisors.
  - Return predictions (via the back-cast)
  - Factor exposures

- Determining if the model provides "valuable enough" inference about an asset entails assessing an asset's goodness of fit and the significance of exposures. If the model's inferences on an asset fail both the goodness of fit test and the exposure test, the model is unlikely to provide enough valuable insights to be worth showing on the platform. However due to the permissive principle above, passing these tests is only weak evidence that the model's inferences are generally reliable. Rather, passing one of these tests indicates that it is more likely than not that some information that the model is providing is of value.

## 3.2 Methodology

To show model results, an asset must pass at least one of the following two tests:

1. Goodness of fit test
    - Define the model's capacity to provide useful predictions as its efficacy relative to the mean model.
        - Measuring efficacy in terms of tracking error, the goal is to assess if

$$\sum_{s\in 1:S} \left(y_s - \Phi\left(F\beta + r\right)\right)^2 \geq \sum_{s\in 1:S} \left(y_s - \overline{y}\right)^2$$

          where $y$ is the data, $\Phi$ is the transformation matrix that forms quarterly predicted returns from monthly desmoothed returns, $\beta$ is the exposure vector (including the intercept), $\overline{y}$ is the average reported return, $F$ is the matrix of factor returns over the backcast, and $r$ is the risk free rate over the back-cast.
        - Consider a two part goodness of fit test utilizing the following test statistic. While motivated by the above tracking error, the calculation for the test statistic coincides with that of $R^2_{ESS} \in (-\infty, 1]$ calculated using the residual methodology:

$$R^2_{ESS} = 1 - \frac{\sum_{s\in 1:S} \left(y_s - \Phi\left(F\beta + r\right)\right)^2}{\sum_{s\in 1:S} \left(y_s - \overline{y}\right)^2}$$

        - As an initial permissive gate, insist that the actual values used on the platform are additive over the mean model. In other words, for all assets for which Architect displays analytics, impose the requirement

$$freqR^2 = 1 - \frac{\sum_{s\in 1:S} \left(y_s - \mathbb{E}\left[\Phi\right]\left(F\mathbb{E}\left[\beta\right] + r\right)\right)^2}{\sum_{s\in 1:S} \left(y_s - \overline{y}\right)^2}$$

          where all expectations are taken with respect to the posterior distribution $p\left(\Theta|D\right)$. This metric does not have an obvious Bayesian interpretation, but serves as a starting point and is representative of the analytics shown on the MVP platform. For this reason, any asset for which $freqR^2 < 0$ automatically fails the goodness of fit test.
        - A more rigorous test incorporates the full distribution of outcomes predicted by the model. From a Bayesian perspective, the this distribution of outcomes is implied by the data and the priors. The test statistic specifically computes the probability that a particular draw from the

posterior is valuable, with value defined as performance better than the mean model:

$$p\left(R_{ESS}^2 > 0\right) = \int_\Theta \iota \left[1 - \frac{\sum_{s\in 1:S}\left(y_s - \Phi\left(F\beta + r\right)\right)^2}{\sum_{s\in 1:S}\left(y_s - \overline{y}\right)^2} > 0\right] p\left(\Theta|D\right) d\Theta$$

where $\iota$ is an indicator function. High values indicate that the model is likely to impart useful information over a large portion of the parameter distribution. An asset's analysis is considered valuable enough to show on the platform if $p\left(R_{ESS}^2 > 0\right) > c_{ESS}$, where $c_{ESS}$ is the minimum cutoff probability (tentatively 80%). As this test is generally much more restrictive than the previously described test, any asset that passes this test and fails the previous should be flagged for further analysis.

2. Exposure significance test

    - A model that fails to produce useful predictive returns may still provide useful information on exposures. For instance, the model might fail to capture the moving average components, or certain parameters may have very high plausible ranges given the data.
    - A second-pass test for significant exposures determines if the exposure information, in-of-itself, is valuable enough to warrant showing the analysis to clients. The test statistic checks if the data implies a high probability that the exposure is greater (less than) zero. Specifically, show the fund if, for any exposure $k$ excluding the intercept, $\max\left[p\left(\beta_k > 0\right), 1 - p\left(\beta_k > 0\right)\right] \geq c_{exp}$

    $$p\left(\beta_k > 0\right) \equiv \int \iota\left[\beta_k > 0\right] p\left(\Theta|D\right) d\Theta$$

    and $c_{exp}$ is a cutoff value (tentatively 80%).

- Due to the permissive and coarse nature of the tests, Architect should maintain the ability to override this criteria on an exception basis.

## 3.3 Sample Application

The methodology was applied within the research POC to the following assets.

| Research POC Label | Quarterly? | Rsq_ESS | p(Rsq_ESS) | Passes GOF? | max(p(beta<>0)) | Passes exp? | **Display?** |
|---|---|---|---|---|---|---|---|
| amgpantheonfundllc | No | 0.19 | 0.92 | yes | 0.98 | yes | **yes** |
| aresindustrialreitclasst | No | (0.28) | 0.00 | no | 0.71 | no | **no** |
| aresprivatemarketsclassi | No | 0.17 | 0.62 | no | 0.60 | no | **no** |
| aresrealestateincometrustinc | No | 0.02 | 0.24 | no | 0.74 | no | **no** |
| atlasenhancedfundlp | No | 0.29 | 1.00 | yes | 1.00 | yes | **yes** |
| blackstoneprivatecreditfund | No | 0.67 | 0.94 | yes | 0.93 | yes | **yes** |
| blackstonerealestateincometrustinc | No | 0.16 | 0.87 | yes | 0.84 | yes | **yes** |
| brookfieldreiticapitaloffshoreaccess... | No | 0.38 | 0.86 | yes | 0.96 | yes | **yes** |
| canyonbalancedhedgefocusfundltd | No | 0.72 | 1.00 | yes | 1.00 | yes | **yes** |
| carlyletacticalprivatecreditfund | No | 0.56 | 1.00 | yes | 0.94 | yes | **yes** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| icapitalcampbellabsolutereturnaccess... | No | 0.91 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalcoopersquareaccessfundlp | No | 0.85 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitaldoublelineopportunisticfundltd | No | 0.35 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalhgvoraaccessfundlp | No | 0.67 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalincomeopportunitiesfundlp | No | 0.42 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalkingstreetcapitalaccessfundlp | No | 0.36 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalkkrprivatemarketsfund | No | 0.26 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalmillenniumfundltd | No | 0.30 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalmultistrategyfundltd | No | 0.04 | 0.67 | no | 1.00 | yes | **yes** |
| icapitalnewalphaaccessfunduslp | No | 0.74 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitaloffshorestrategieshoneycomb... | No | 0.59 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalrenaissanceidgefundltd | No | 0.23 | 0.96 | yes | 0.97 | yes | **yes** |
| icapitalrenaissanceiefundltd | No | 0.37 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalsegpartnersfundlp | No | 0.65 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalshaccessfundlp | No | 0.73 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalsorobanopportunitiesfundlp | No | 0.80 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalthirdpointfundltd | No | 0.66 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalworldquantmillenniumsealslp | No | 0.42 | 0.75 | no | 0.60 | no | **no** |
| nuveenglobalcitiesreitinc | No | 0.44 | 0.95 | yes | 0.99 | yes | **yes** |
| owlrockcoreincomecorp | No | 0.61 | 0.89 | yes | 0.86 | yes | **yes** |
| ozdpiihedgefocusfundlp | No | 0.85 | 1.00 | yes | 0.99 | yes | **yes** |
| spfhedgefocusfundltd | No | 0.30 | 1.00 | yes | 0.99 | yes | **yes** |
| steelecreekcapitalcorporation | No | 0.19 | 0.83 | yes | 0.78 | no | **yes** |
| stepstoneprivatemarketssprim | No | (0.17) | 0.06 | no | 0.70 | no | **no** |
| wmasystematicequityalphalongshort... | No | 0.90 | 0.95 | yes | 0.68 | no | **yes** |
| worldquantmillenniumwmqsgeae... | No | 0.97 | 1.00 | yes | 1.00 | yes | **yes** |
| blackrockprivateinvestmentsfund | Yes | 0.65 | 0.43 | no | 0.87 | yes | **yes** |
| carlylecreditsolutionscars | Yes | 0.89 | 1.00 | yes | 0.91 | yes | **yes** |
| hlsfvicapitalaccessfundlp | Yes | 0.72 | 0.75 | no | 0.98 | yes | **yes** |
| hlsfvicapitaloffshoreaccessfundlp | Yes | 0.80 | 0.78 | no | 0.96 | yes | **yes** |
| icapitalbcpviiiaccessfundlp | Yes | 0.91 | 0.87 | yes | 0.92 | yes | **yes** |
| icapitalblackstonegrowthaccessfundlp | Yes | 0.99 | 1.00 | yes | 1.00 | yes | **yes** |
| icapitalpofivaccessfundlp | Yes | 0.55 | 0.88 | yes | 0.97 | yes | **yes** |
| icapitalspreviiaccessfundlp | Yes | 0.51 | 0.53 | no | 0.91 | yes | **yes** |
| icapitalvintageivaccessfund... | Yes | 0.96 | 0.98 | yes | 1.00 | yes | **yes** |
| icapitalvintagevaccessfunduslp | Yes | 0.96 | 0.96 | yes | 1.00 | yes | **yes** |

# 4 Appendix

## 4.1 Normal-Gamma

The normal gamma distribution is sometimes provided in certain packages. It is equivalent to the semi-conditional approach described for the normal model, but may be easier to implement if sampling is available.

$$
\begin{aligned}
\log p\left(\mu,\tau|y\right) \propto & \left(\frac{S+1}{2}\right)\log\tau - \frac{\tau}{2}\left(\sum_{s=1}^{S}\left(\mu-y_s\right)^2 - \tau_{\mu 0}\left(\mu-\mu_0\right)^2\right) + \left(\alpha_0-1\right)\log\tau - \zeta_0\tau + c_1^{\mu\tau} \\
= & \left(\alpha_0+\frac{S-1}{2}\right)\log\tau - \left(\zeta_0+\frac{1}{2}\sum_{s=1}^{S}y_s^2+\frac{\tau_{\mu 0}}{2}\mu_0^2\right)\tau \\
& - \frac{\tau\left(S+\tau_{\mu 0}\right)}{2}\left(\mu^2-\frac{2\mu}{\left(S+\tau_{\mu 0}\right)}\left(\sum_{s=1}^{S}y_s+\tau_{\mu 0}\mu_0\right)\right)+c_1^{\mu\tau} \\
= & \alpha\log\tau-\zeta\tau-\frac{\tau_\mu\tau}{2}\left(\mu-\mu_\mu\right)^2+c_1^{\mu\tau} \\
= & \log NormalGamma\left(\mu,\tau;\mu,\tau_\mu,\alpha,\zeta\right)+c_3^{\mu\tau} \\
s.t. & \\
\alpha \equiv & \frac{S}{2}+\alpha_0 \\
\tau_\mu \equiv & S+\tau_{\mu 0} \\
\mu_\mu \equiv & \tau_\mu^{-1}\left(\sum_{s=1}^{S}y_s+\tau_{\mu 0}\mu_0\right) \\
\zeta \equiv & \zeta_0+\frac{1}{2}\sum_{s=1}^{S}y_s^2+\frac{\tau_{\mu 0}}{2}\mu_0^2-\frac{\tau_\mu}{2}\mu_\mu^2
\end{aligned}
$$

Note that this is effectively the same as approach #1, as drawing from the normal gamma implies drawing $\tau$ from $Gamma\left(\tau;\alpha,\zeta\right)$ and then drawing $\mu_\mu$ from $Normal\left(\mu;\mu_\mu,\frac{1}{\tau\lambda}\right)$.