Bayesian Estimation of Alternative Asset Returns

Arnav Sheth

December 4, 2023

Contents

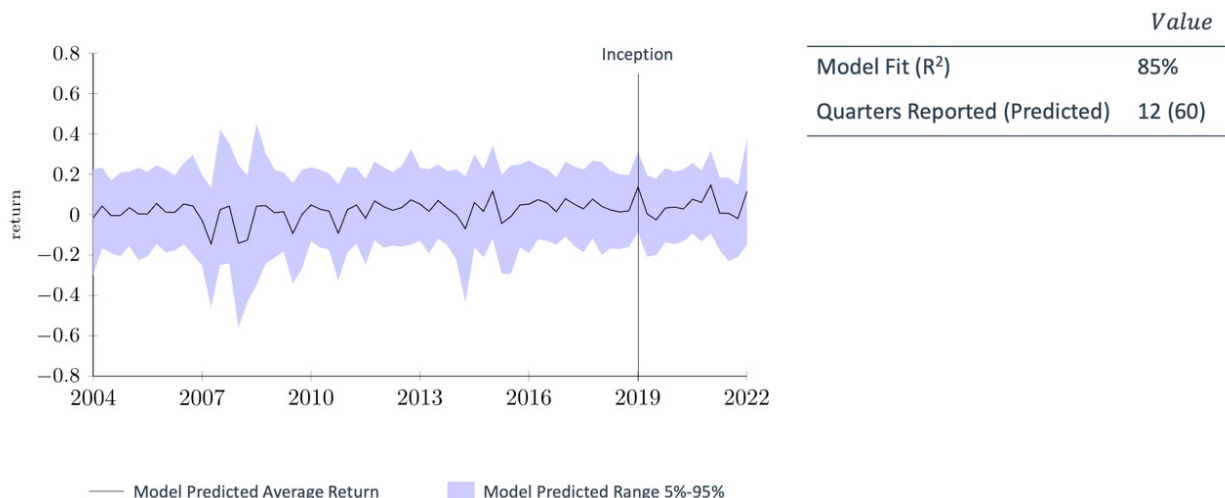|	Value
Model Fit (R^2)	85%
Quarters Reported (Predicted)	12 (60)

Figure 1: An example of a fund's returns as generated by GMAM 3.0. The fund was incepted in 2019, and backcasted to 2004 using a hierarchical Bayesian technique. De-smoothed returns were generated, and re-smoothed again to reflect what the actual fund returns would look like. This helps investors to pair the fund with a portfolio of publicly available securities.

1 Introduction

According to Bain & Company[1], private equity (PE) buyout deals for 2022 alone were valued at $654 billion globally. This was the second-highest valuation since 2008, and brought the value of the total deal volume to $2.4 trillion[2]. Additionally, total private markets AUM reached $11.7 trillion as of June 2022. This makes PE a major asset class, with institutions like colleges, foundations, pension funds, and more investing in it. Furthermore, SEC rules defining what an accredited investor is, have recently been broadened to include "individual investors that have the knowledge and expertise to participate in [financial] markets." One no longer needs to pass the income requirement. All of this has hastened the market expansion and liquidity of alternative investments and made it a viable option for accredited investors and private wealth funds[3].

A significant limitation for the analysis for PE and other alternative asset returns is the lack of availability of performance metrics based on actual transactions since alts are largely exempt from public disclosure requirements. This limitation impedes the process of optimal wealth allocation since this process requires data on the risk, return, and covariance of asset classes. In liquid markets, these can be calculated via historical return data. So how does one optimally combine such assets with publicly available investment opportunities?

---

[1]Source:https://www.bain.com/insights/topics/global-private-equity-report/

[2]Source:https://www.mckinsey.com/industries/private-equity-and-principal-investors/our-insights/mckinseys-private-markets-annual-review/

[3]Source:https://icapital.com/insights/practice-management/untapped-potential-alternative-investments-and-the-wealth-management-channel/

iCapital Network's attempts to solve this problem and others associated with returns on illiquid, alternative investments via our Generalized Multi-Asset Model version 3.0 (henceforth, GMAM 3). GMAM 3 is a returns-generating model that uses hierarchical Bayesian modeling techniques to generate economically significant returns for any single fund that trades on iCapital's marketplace. This model helps to educate investors on whether or not a particular fund is suitable for them, based on their investment needs and risk preferences.

There are several unique aspects of the model, and this technical manual explains the motivation behind the choices made as well as their economic rationale.

1.1 Why Bayesian?

Bayesian techniques offer a more robust and flexible approach to capture the complex dynamics of private equity returns due to their ability to:

- incorporate prior knowledge,
- handle limited data,
- account for iliquidity, and
- update estimates as new information becomes available.

They allow one to incorporate prior beliefs or information about specific funds into the estimation of returns. This prior information is leveraged to handle limited data. These techniques also allow for the incorporation of uncertainty into the return estimates by providing probability distributions rather than point estimates. They provide a framework that aligns well with the unique characteristics of private equity investments, allowing for more accurate and informed returns estimation.

For example, it well-known in the academic finance literature that for private equity (henceforth, PE) returns there is a dearth of performance metrics based on actual transactions (Kaplan and Schoar [2005]). Also, the available time series data often relies on non-market valuations or multiyear internal rates of return (IRR) that are segmented by the vintage years of the funds which further complicates the calculation. For investors seeking to optimally allocate their wealth across public securities and private equity, this is a significant barrier. For instance, even a simple Markowitz-style (Markowitz [1952]) mean-variance optimization requires a sufficiently long history of returns which is unavailable in the case of PE.

As mentioned above, one commonly used metric in the PE world to gauge a fund's performance is the IRR. It provides a convenient way by which to compare and benchmark various funds. Furthermore, it is an easy way to account for the timing and magnitude of cash flows since PE investments are typically illiquid and longer investment horizons. And finally, it helps investors

assess the risk of a fund.

While these are good reasons to use the IRR, there are numerous measurement concerns related to it. One implicit assumption when calculating IRRs is that the cash proceeds will be reinvested to achieve a particular return. This assumption does not coincide with the reality of cash flow distributions by PE. For example, if a PE fund reports a 50% IRR and has returned cash early in its life, the assumption is that the cash will be reinvested at 50%. However, it is unlikely that the fund will find such an investment opportunity every time cash is distributed.

Furthermore, IRR can distort management incentives, upwardly bias performance measures, and misrepresent volatility estimates (Phalippou [2008]). As stated in Sorensen and Jagannathan [2013], "The IRR may not exist, and it may not be unique." Even though the Modified IRR (MIRR) largely tackles some well-known pitfalls of IRR, its practical implementation in a private partnership is not obvious.

GMAM 3, on the other hand, uses time-weighted returns.

$$y_t = \frac{NAV_t + Distributions_t}{NAV_{t-1} + Contributions_t} \tag{1}$$

Though these have their own caveats[4], they also have their strengths. For instance, this method of calculation accounts for the contributions and distributions, as well as their timing. Also, it allows for a more standardized comparison of performance across different investment vehicles and asset classes. This applies especially to investors who wish to combine PE with publicly available investments. And finally, it is a widely understood and commonly used method in the broader investment industry. Its use in measuring PE performance can facilitate understanding and communication about fund performance, especially for stakeholders more familiar with traditional investment vehicles.

One of the strengths of the model's Bayesian estimation process provides measures of uncertainty around the estimated returns, and this allows the user to decide for themselves the degree to which they can rely on these estimated, time-weighed returns.

2   The Model

At its core, the model predicts fund returns from factor returns. The essence of the GMAM 3 returns-generating process can be modeled as:

$$x_{i,t} = f'_t \beta_{i,t} + r^f_t + \varepsilon_{i,t} \tag{2}$$

---

[4]They may not fully capture the complexities of private equity investments, such as the strategic timing of cash flows and the long-term, illiquid nature of these assets.

where f_t is a length K vector of factor returns including an intercept dummy, $\beta_{i,t}$ is a $K \times 1$ matrix of exposures, $x_{i,t}$ is the predicted return, r_t^f is the risk-free rate of return, and $\varepsilon_{i,t}$ is the error. The subscripts here are for an individual fund (i) for a given period (t). Henceforth, assume that all measures are for an individual fund, and so the i-subscript will be dropped.

We use a superset of thirteen factors: Alt Commodities, Alt Hedge Fund Crowding, Alt Oil, Alt Trend, Emerging Markets, Equity Market, Equity Momentum, Equity Quality, Equity SmallCap, Equity Value, Fixed Credit, Fixed Duration, and US Dollar. Detailed information about factor construction is available upon request.

While this may seem like a conventional way to estimate returns, there are two important things to be noted about the equation above:

1. x_t is NOT the final return provided to the user since these returns are de-smoothed. Alternative investment returns are both, serially correlated and lagged. The returns generated are the true economic returns. To mimic the underlying characteristics of the reported returns to a greater degree[5], we perform a re-smoothing of these returns. Section 2.1 below discusses the economic rationale and mathematical process in greater detail. The Appendix has more details on this as well.

2. We start with a superset of thirteen economic factors as listed above. Rather than include them all in the regression, we use the stochastic search variable selection (SSVS) technique (George and McCulloch [1993]) to select a subset of regressors that are economically meaningful to a particular fund. SSVS is a type of Bayesian linear regression which is useful when the number of predictors is larger than the number of observations – a problem frequently encountered in alternative investment data. More on this can be found in Section 2.3, and the Appendix.

It suffices to say here, that GMAM 3 uses Markov Chain Monte Carlo (MCMC) to compute estimates of β as well as the distributional parameters of ε_t. The final product displays only point estimates – which are calculated as averages across draws in the MCMC process. For MCMC, each draw is a distribution and not a point estimate (more details on this in the Appendix). So, the expected, predicted, de-smoothed returns estimated from N simulations are given by:

$$\overline{x}_{n,t} = \frac{1}{N} \sum_{n \in 1:N} \left(f_t' \beta_{n,t} + r_t^f + \varepsilon_{n,t} \right) \tag{3}$$

Since this is the result of a hierarchical Bayesian regression, the true distribution of the error terms is hard to characterize precisely – it is a mixture distribution. However, we can approximately

---

[5]This re-smoothing, while replicating the essential attributes of the reported returns more closely, does come at the cost of not reflecting the true economic changes in value over time.

say that[6]:

$$\varepsilon_{n,t} \sim Tdist\left(0, \sigma_{\varepsilon_n}, \nu_n\right)$$

Here, $\sigma_{\varepsilon_n}^2$ is derived from model parameters sampled at each iteration and ν_n is a t-distribution degree of freedom parameter likewise sampled. In other words, the error terms approximate a t-distribution to a great degree.
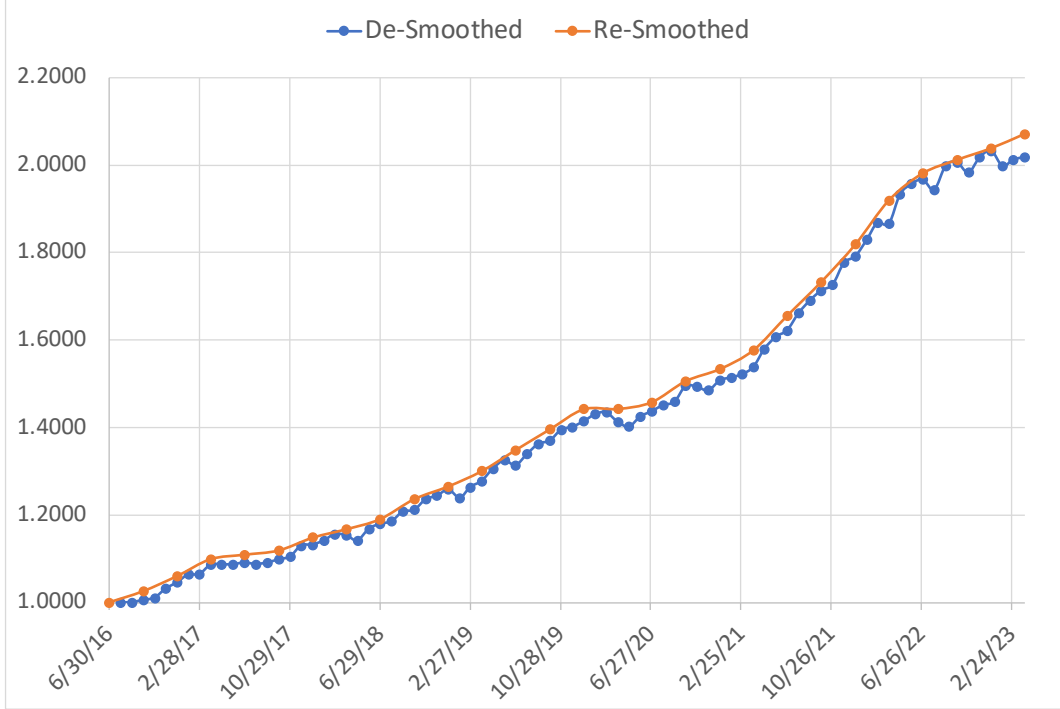
2.1   Re-Smoothing

Figure 2: An example of a fund's de-smoothed (x), and re-smoothed (y) returns as generated by GMAM 3. The reported re-smoothed returns are quarterly, and the de-smoothed returns are monthly. Shown here is the cumulative return of $1 invested in each of the set of returns.

It is well-known in the finance practitioner literature that returns from alternative investment such as PE, hedge funds, or real estate funds, are highly serially correlated (see, for example, Getmansky et al. [2004]). In other words, past values are correlated with present values. The serial correlation mentioned here occurs because of the lack of liquidity in the fund itself or some of the assets held within it. For instance, these illiquid assets could be ones that don't trade frequently, making it challenging to always determine their market prices. When funds contain such illiquid assets, their reported returns may seem steadier than their actual economic returns (returns that consider all available market information about those securities). This can lead to a downward bias

---

[6]Strictly speaking, this approximates a t-distribution since the final distribution is the result of several layers of Bayesian analysis. More information about this can be found in the Appendix.

5

in estimating return variance and result in positive serial return correlation.

It is also well-known that returns from real estate funds, are lagged (Geltner [1993]). For instance, for many funds, returns are compiled and reported quarterly. But valuations of many properties included in the funds are effectively updated only annually. Each quarter some properties have their valuations updated, and others do not. Those properties that don't get a new valuation within a quarter carry over their last known value into the current quarter.

Finally, Financial Accounting Standard 157, released by the FASB in 2006 during the run-up to the financial crisis of that time, and now called Accounting Standards Code Topic 820, required companies to mark their assets to market. This was a radical change from historic cost accounting and required LPs to periodically mark the assets to market, even though markets in many portfolio companies are illiquid. This backwards appraisal may also result in a managerial bias towards smoothing asset values.

The latent returns $x_{i,t}$, henceforth just x for notational simplicity, generated above are not smoothed. They are the true economic returns. Since GMAM 3 attempts to mimic true reported returns, we must re-smooth the latent returns to reflect the reported returns so that investors can appropriately compare their investments in PE with publicly available investments. In GMAM 3, this is done using a moving average (MA) process with smoothing parameters estimated using a Bayesian linear regression. Note that the MA is defined in an econometric sense, and not in a literal sense, although it may be interpreted as such. More specifically, the final returns presented to the user are re-smoothed using the following process:

$$y = \Phi x \tag{4}$$

where y are the reported returns and Φ is a matrix consisting of smoothing coefficients, ϕ and $\tilde{\phi}$. More specifically, when both x and y have the same reporting frequency,

$$\Phi = \begin{bmatrix} \phi_1 & \cdots & \phi_P & \tilde{\phi}_{P+1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \phi_1 & \cdots & \phi_P & \tilde{\phi}_{P+1} & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \phi_1 & \cdots & \phi_P & \tilde{\phi}_{P+1} & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \phi_1 & \cdots & \phi_P & \tilde{\phi}_{P+1} \end{bmatrix} \tag{5}$$

ϕ is a vector of P unrestricted coefficients, and $\tilde{\phi}$ is a vector of $P+\Delta t$ restricted and unrestricted coefficients. P and Δt represent terms that map the observation period of the reported, smoothed

returns y to the latent, de-smoothed returns, x. In particular,

$$y_s = \left(\tilde{\phi}_{1:(P+\Delta t)}\right)' x_{(t[s]-P-\Delta t+1):t[s]} + \varepsilon_t^y \tag{6}$$

$$= \left(\tilde{\phi}_{(P+1):(P+\Delta t)}\right)' x_{(t[s]-\Delta t+1):t[s]} + \phi' x_{(t[s]-P):(t[s]-\Delta t)} + \varepsilon_t^y \tag{7}$$

The reported returns are y_s s.t. $s \in 1 : S$, and the latent returns are x_t s.t. $t \in 1 : T$. In equation (5) above, $S = T$, and the reported returns and de-smoothed returns have the same frequency. However, S may not always equal T, such as when the reported returns are quarterly, and the de-smoothed returns are monthly[7]. To resolve this, the following formula maps t to s:

$$t(s) = P + (s \times \Delta t) \tag{8}$$

For example, if the reported returns y are quarterly, and the latent returns x are monthly, $P = 3$, and $\Delta t = 3$ (since each quarter consists of three months). And so, for the first quarter, $s = 1$, and $t = 6$. When $\Delta t > 1$, the $\tilde{\phi}_j$, for $j \in 1 : \Delta t$ must add up to the same value, $\forall j$. If not, months that fall earlier in the quarter will have a different long-run impact on NAV than months later in the quarter. In the current version of GMAM 3 in production, they add up to 3 when reported returns are quarterly, and they add up to 1 when reported returns are monthly[8].

The economic assumption behind this process is that the observed fund returns y are a weighted average of the fund's economic returns x over the most recent $(P + \Delta t)$ periods, including the current period. The econometric implications of this are that under the given assumption, the observed fund returns follow a MA process of order $P + \Delta t$.

This restriction is similar to Getmansky et al. [2004] where the observed return (R_t^0) for some period t, is a weighted average of the "true" returns (R_t^C) over the most recent $k + 1$ periods: $R_t^0 = \theta_0 R_t^C + \cdots + \theta_k R_{(t-k)}^C$, with $\sum_{i=0}^k \theta_i = 1$ to ensure that all information is eventually incorporated into observed returns, and $\theta_i \in [0, 1]$ for $i = 1, \ldots, k$. In our case the observed returns are the reported returns, y, the "true" returns are the latent returns generated using the factors, x, and the θ_i terms from Getmansky et al. [2004] are denoted as ϕ_i in our estimation process. They are generated using a multivariate normal distribution in the hierarchical Bayesian model, as follows:

$$p(\phi|rest) \sim MN\left(\phi_0, \frac{1}{\tau_y \tau_\phi} M_0^{-1}\right) \tag{9}$$

where τ_y is a global precision parameter estimated for independent measurement using a

---

[7]Our factor returns are almost always monthly, as are the de-smoothed returns.

[8]Note that the vector $\phi \in \tilde{\phi}$. For quarterly reported data, say $\phi = [a, b, c]'$ where $a, b, c \in \mathbb{R}$. Then, $\tilde{\phi} = [a, b, c, 1 - a, 1 - b, 1 - c]'$ and the sum of all terms in $\tilde{\phi}$ is 3. For monthly reported data, say $\phi = [a, b, c, d, e]'$ where $a, b, c, d, e \in \mathbb{R}$. Then $\tilde{\phi} = [a, b, c, d, e, 1 - a - b - c - d - e]'$ and the sum of all terms in $\tilde{\phi}$ here, is 1. This process is also described in more detail in the Appendix.

Gamma-distributed prior: $p(\tau_y) \sim Gamma\left(\alpha_{y_0}, \zeta_{y_0}\right)$ with given shape hyperparameter,α_{y_0}, and given inverse scale hyperparameter, ζ_{y_0}; τ_ϕ is a global precision multiplier parameter, distributed as $p(\tau_\phi) \sim Gamma\left(\alpha_{\phi_0}, \zeta_{\phi_0}\right)$ with given shape and inverse scale hyperparameters; M_0 is a $P \times P$ matrix consisting of precision hyperparameters for ϕ_0; and ϕ_0 is a hyperparameter (i.e., fed into the model).

And the complete posterior distribution of y is given by:

$$p\left(y \mid rest\right) \sim MN\left(\Phi x, \frac{1}{\tau_y}I\right)) \tag{10}$$

with Φ, x, and τ_y defined as above; and I is an identity matrix used for computational purposes.

2.2   A Brief Digression on Priors

Since we are beginning to describe our posterior distributions with their respective priors, it might make sense to motivate some understanding about prior distributions in the context of a hierarchical Bayesian model.

Priors, while influencing the results with fewer datapoints and noisier data, do not ultimately matter much as data increases. Priors are, of course, essential in Bayesian analysis. However, their influence generally decreases as the amount of data increases, especially if the data is strongly informative. In theory, Bayesian models exhibit a property known as 'consistency.' This means that as the sample size grows to infinity, the posterior distribution of the parameter of interest will converge to the "true" parameter value[9], assuming the model is correctly specified. In such a scenario, the choice of prior becomes less critical as the sample size increases.

That said, the use of the gamma distribution with an inverse scale parametrization above is convenient for several reasons:

1. Conjugacy: conjugate priors are advantageous for computational and analytical simplicity.

2. Positive Values: the gamma distribution is defined for positive values only.

3. Informative and Non-Informative Settings: the gamma distribution can be used to encode both informative and non-informative priors. By adjusting its parameters, you can represent strong prior beliefs or specify a prior that has minimal influence on the posterior, allowing the data to drive the inference.

4. Regularization and Stability: In Bayesian hierarchical models, the use of gamma priors can help in regularizing the estimates, particularly in complex models or in the presence of limited data. This can prevent overfitting and improve the model's stability and predictive performance.

---

[9]In the frequentist sense of the word.

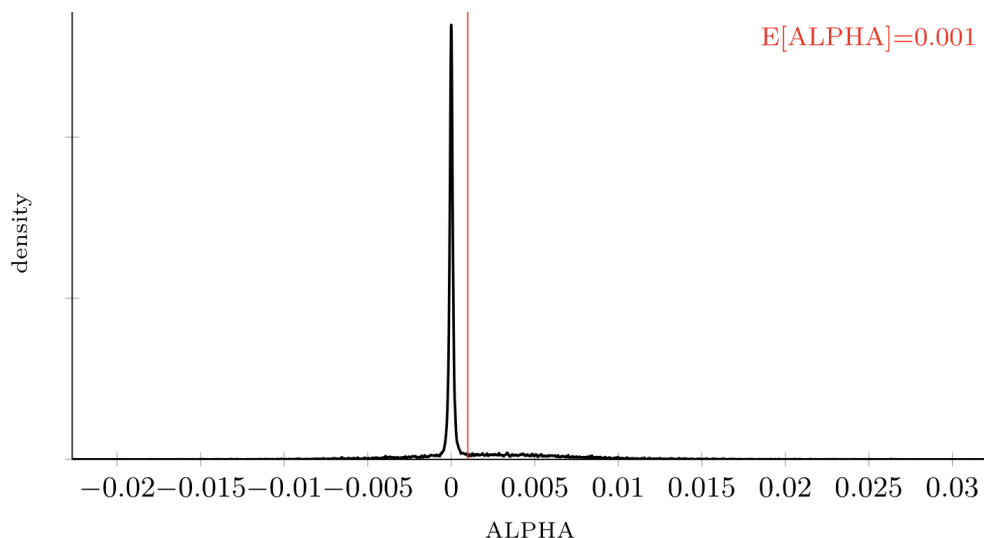## 2.3 Stochastic Search Variable Selection (SSVS)

Figure 3: The probability mass function of a fund's intercept term (β_0) from the SSVS (spike and slab) regression. We see a spike here because for this particular fund the probability mass is concentrated around zero with a small positive tilt. Additionally, there is low evidence of an intercept term. We have 14 months of history for this fund.

As mentioned above, the regression performed to estimate x is using a SSVS technique which defines prior regression coefficient variances consistent with inclusion or effective exclusion. In other words, it is a predictor selection technique. Such techniques commonly focus on which predictors to retain, though they also aim for improved predictive performance through developing an encompassing model, or model simplification without adversely affecting predictive accuracy (Piironen and Vehtari [2017]). Formal model choice is simplified for normal linear regression, as marginal likelihoods may be obtained analytically, but for many predictors, comparison of the many possible models becomes infeasible.

An imperfect analogy is stepwise regression using AIC or BIC as a selection criterion. However, when the number of regressors is large (as mentioned above, we have 13), the computational requirements for these procedures can be prohibitive. A common workaround is to use a heuristic method to restrict attention to a potential subset of regressors, i.e., include or exclude variables, based on R^2 considerations (say).

SSVS uses a Bayesian approach with a normal mixture model for regression analysis. In this method, selector variables are employed to pinpoint which subsets of predictors are worth considering. It works by identifying those predictors that have a higher chance of being relevant, based

on their posterior probability. To do this, SSVS uses a technique called Gibbs sampling[10]. This approach helps to sample from the distribution of all possible subsets of predictors. The subsets that show up more often in these samples are considered promising because they have a higher probability of being relevant.

We use the SSVS technique of George and McCulloch [1993], also known as a **spike and slab regression**. The term was coined by Mitchell and Beauchamp [1988] and referred to the prior for the regression coefficients used in their Bayesian hierarchy. This prior was chosen such that the regression parameters were mutually independent with a two-point mixture distribution made up of a uniform flat distribution (the slab) and a degenerate distribution at zero (the spike).

Ishwaran and Rao [2005] characterize a spike and slab model as being any model with a Bayesian hierarchy specified as follows:

$$p(y \mid X, \beta, \sigma^2) \sim N(X\beta, \sigma^2 I_n)$$
$$p(\beta \mid \gamma) \sim N(0, \Gamma)$$
$$\gamma := (\gamma_1, \ldots, \gamma_p)^T \sim \pi(\cdot)$$
$$p(\sigma^2) \sim \mu(\cdot)$$

Here, $\beta = (\beta_1, \ldots, \beta_p)^T$ is the regression vector, and $\Gamma = diag(\gamma_k)_{1 \leq k \leq p}$ is its $p \times p$ hypervariance matrix. The prior π for the hypervariance γ plays a critical role in how effective the technique is for variable selection. A successful and popular choice of π are priors that make use of mixture distributions involving a spike near zero. In George and McCulloch [1993], the prior for γ_k was assumed to have a two-component distribution of the form:

$$p\left(\gamma_k \mid \tau_k, c_k, \overline{\omega}_k\right) \overset{\text{ind}}{\sim} (1 - \overline{\omega}_k \delta_{\tau_k}(\cdot) + \overline{\omega}_k \delta_{c_k \tau_k}(\cdot), \quad k = 1, \ldots p.$$

The value for $\tau_k > 0$ (the spike) is chosen to be some small value, where "small" is typically based on the data at hand, while $c_k > 0$, also data-specific, is chosen so that $c_k \tau_k$ (the slab) is sufficiently large. Selecting the two hyperparameters in this way allows γ_k to be small or large, and this in turn enables the posterior of β_k to shrink towards zero or be some nonzero value. The values $(\overline{\omega}_k)_{k=1}^p$ are complexity parameters that influence the likelihood of a coefficient being shrunk towards zero. In principle, each variable can have a unique complexity value, but a common practice is to set $\overline{\omega}_k = 1/2, \forall k$, in which case the prior will be referred to as an indifference prior.

---

[10]Gibbs sampling is a statistical technique used for generating sequences of samples from the probability distribution of multiple variables. It's a kind of Markov Chain Monte Carlo (MCMC) method. It is particularly useful in scenarios where directly sampling from the joint distribution is difficult, but sampling from the conditional distribution of each variable is feasible.
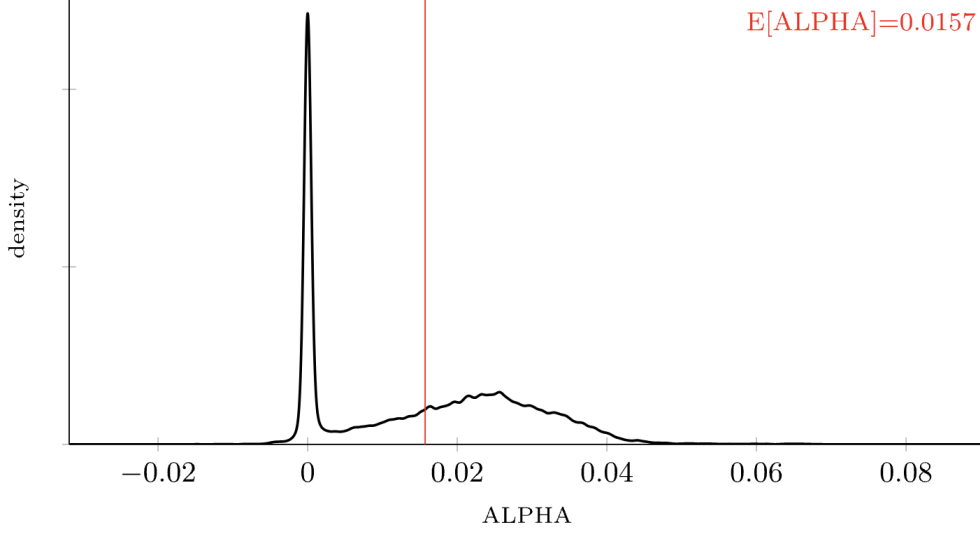
Figure 4: The probability mass function of a fund's intercept term (β_0) from the SSVS (spike and slab) regression. Though we see a spike here, the bulk is in the long right tail representing substantial upside and a significantly greater weight of evidence for a positive alpha. This fund has over two years of history.

For our model, we use a Bernoulli or binomial prior on γ_k, the selector variable:

$$p\left(\gamma_k\right) \sim Bern\left(\omega\right) \tag{11}$$

Here, $p\left(\omega\right) \sim Beta\left(\kappa_0, \delta_0\right)$ which is the probability of factor selection with hyperparameters κ_0 and δ_0 specified as given. γ_k for $k \in 1:K$ is a $K \times 1$ vector and selector variable used to determine which of the thirteen factors are economically significant in a probabilistic sense. Initially, each of the factors have an equally likely chance of being included. Conditional on a factor being in the regression, we specify a prior distribution for the regression coefficient associated with that variable. In our case, this is multivariate normal:

$$p\left(\beta|rest\right) \sim MN\left(\beta_0 + D^{-1}\beta_0^{\Delta}, \frac{1}{\tau_x\tau_y\tau_\beta}\left[DA_0D\right]^{-1}\right) \tag{12}$$

$$d_k = \sqrt{\gamma_k + (1 - \gamma_k)\frac{1}{v^2})} \tag{13}$$

Here, β_0 is the prior mean of β conditional on exclusion (i.e., when $\gamma_k = 0$); β_0^{Δ} is the shift in the prior mean of β conditional on inclusion (i.e., when $\gamma_k = 1$); d_k for $k \in 1:K$ is a $K \times 1$ vector – a function of γ which adjusts β for sparsity; D is a $K \times K$ matrix with $D = diag(d_k)$; and ν is the variance of the spike distribution as a fraction of the slab variance. And finally, τ_x, τ_y, and τ_β are all gamma-distributed with shape and inverse-scale hyperparameters as shown earlier.

11
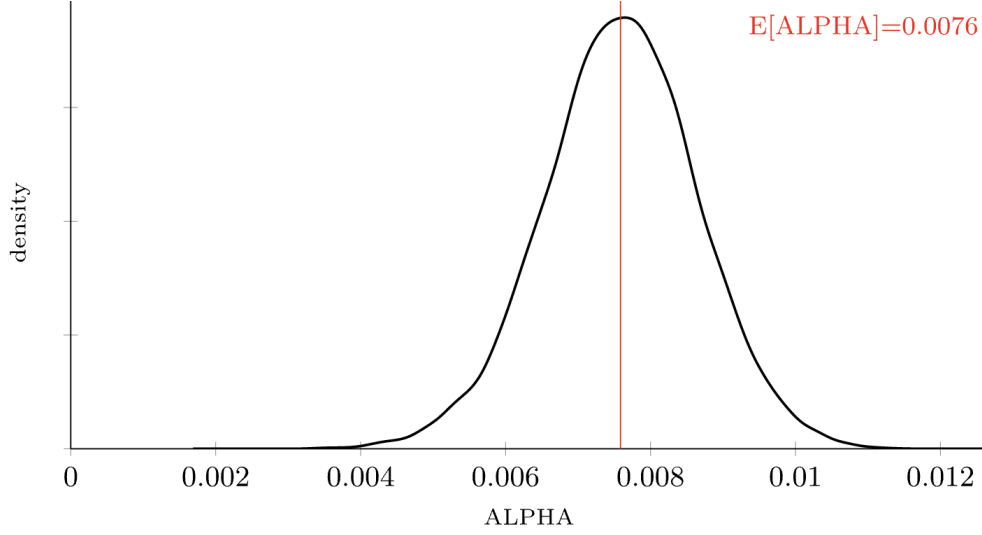
Figure 5: The probability mass function of a fund index intercept term (β_0) from the SSVS (spike and slab) regression. This index has almost twenty years of history and clearly shows positive alpha, though its magnitude is still uncertain.

3   Conclusion

We use a hierarchical Bayesian model to generate fund returns when data is sparse and noisy. The data generation process includes de-smoothed returns, generated using a spike and slab regression, and re-smoothed returns used to mimic reported returns. Our technique provides a systematic and rigorous way by which fund performance can be quantified.

References

[1] David Geltner. Estimating market values from appraised values without assuming an efficient market. *The Journal of Real Estate Research*, 8(3):324–345, 1993.

[2] Edward I. George and Robert E. McCulloch. Variable selection via gibbs sampling. *Journal of the American Statistical Association*, 88(423):881–889, 1993.

[3] Mila Getmansky, Andrew W. Lo, and Igor Makarov. An econometric model of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economics*, 74(3):529–609, 2004.

[4] Hemant Ishwaran and J. Sunil Rao. Spike and slab variable selection: frequentist and bayesian strategies. *The Annals of Statistics*, 33(2):730–733, 2005.

[5] Steven N Kaplan and Antoinette Schoar. Private equity performance: Returns, persistence, and capital flows. *The Journal of Finance*, 60(4):1791–1823, 2005.

[6] Harry M. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

[7] T.J. Mitchell and J.J. Beauchamp. Bayesian variable selection in linear regression. *Journal of the American Statistical Association*, 83:1023–1036, 1988.

[8] Ludovic Phalippou. The hazards of using irr to measure performance: The case of private equity. 2008. URL `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1111796`.

[9] Juho Piironen and Aki Vehtari. Comparison of bayesian predictive methods for model selection. *Statistics and Computing*, 27:711–735, 2017.

[10] Morten Sorensen and Ravi Jagannathan. The public market equivalent and private equity performance, 2013. URL `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2347972`.

4 Appendix

The complete generative process is given by:

$$p\left(y|rest\right) \sim MN\left(X_L R\phi + x_S, \frac{1}{\tau_y}I\right) \quad \text{equivalently, } MN\left(\Phi x, \frac{1}{\tau_y}I\right) \tag{10}$$

$$p\left(x|rest\right) \sim MN\left(F\beta + r, \frac{1}{\tau_x \tau_y}\Psi^{-1}\right) \tag{14}$$

$$p\left(\phi|rest\right) \sim MN\left(\phi_0, \frac{1}{\tau_y \tau_\phi}M_0^{-1}\right) \tag{9}$$

$$p\left(\beta|rest\right) \sim MN\left(\beta_0 + D^{-1}\beta_0^\Delta, \frac{1}{\tau_x \tau_y \tau_\beta}[DA_0D]^{-1}\right) \tag{12}$$

$$d_k = (\gamma_k + (1 - \gamma_k)\frac{1}{v^2})^{0.5} \tag{13}$$

$$p\left(\gamma_k\right) \sim Bern\left(\omega\right) \tag{11}$$

$$p\left(\omega\right) \sim Beta\left(\kappa_0, \delta_0\right) \tag{15}$$

$$p\left(\tau_y\right) \sim Gamma\left(\alpha_{y0}, \zeta_{y0}\right) \tag{16}$$

$$p\left(\tau_x\right) \sim Gamma\left(\alpha_{x0}, \zeta_{x0}\right) \tag{17}$$

$$p\left(\psi_t\right) \sim Gamma\left(\nu/2, \nu/2\right) \tag{18}$$

$$p\left(\nu\right) \sim Gamma\left(\alpha_{\nu0}, \zeta_{\nu0}\right) \tag{19}$$

$$p\left(\tau_\phi\right) \sim Gamma\left(\alpha_{\phi0}, \zeta_{\phi0}\right) \tag{20}$$

$$p\left(\tau_\beta\right) \sim Gamma\left(\alpha_{\beta0}, \zeta_{\beta0}\right) \tag{21}$$

Note that here y can be written in two forms using matrix notation. When x and y have the same frequency:

$$y = \Phi x = X_L R\phi + x_S \tag{4}$$

Where

$$X_L R\phi + x_S = \begin{bmatrix} x_1 & x_2 & \cdots & x_{P-1} & x_P & x_{P+1} \\ x_2 & x_3 & \cdots & x_P & x_{P+1} & x_{P+2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ x_{T-P-1} & x_{T-P} & \cdots & x_{T-3} & x_{T-2} & x_{T-1} \\ x_{T-P} & x_{T-P+1} & \cdots & x_{T-2} & x_{T-1} & x_T \end{bmatrix} [R] \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{P-1} \\ \phi_P \end{bmatrix} + \begin{bmatrix} x_{P+1} \\ x_{P+2} \\ \vdots \\ x_{T-1} \\ x_T \end{bmatrix} \tag{22}$$

And

$$R = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -1 & -1 & \cdots & -1 & -1 \end{bmatrix}$$

To generalize to other frequencies, recall that $t[s] \equiv P + s * \Delta t$ and define:

$$\Phi_{sj} \equiv \begin{cases} \phi_{P-(t[s]-\Delta t-j)} & 1 \leq P - (t[s] - \Delta t - j) \leq P \\ 1 - \left(\iota^{\phi}_{\Delta t - (t[s]-j)}\right)' \phi & t[s] - \Delta t < j \leq t[s] \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

$$X_{Lsj} \equiv x_{t[s]-(P+\Delta t-j)}$$

$$\iota^{\phi}_{pl} \equiv \begin{cases} 1 & p+l \mod \Delta t = 0 \\ 0 & \text{otherwise} \end{cases}$$

The above matrix formulation (22) can be generalized by only including rows of X_L where $t \in \{t[1...S]\}$ and adjusting the restriction matrix. The general version is then:

$$y = \Phi x = X_L R \phi + x_S, \quad \text{s.t. } x_S = X_L \iota_{\Delta t} \tag{4}$$

where $\iota_{\Delta t}$ is a vector of P zeros followed by Δt ones, making x_S the sum of the last Δt columns of X_L.

4.1 Complete Posterior Distribution

The complete posterior distribution is given by:

$$\begin{aligned} p(\Theta|y, F) &\propto p(y|x, \gamma, \omega, \beta, \phi, \tau_x, \tau_y, \tau_\phi, \tau_\beta, \psi, \nu, F) \times p(x|\beta, \phi, \tau_x, \tau_y, \psi, F) \\ &\times p(\phi|\tau_y, \tau_\phi) \times p(\beta|\gamma, \tau_x, \tau_y, \tau_\beta) \times p(\gamma|\omega) \times p(\omega) \\ &\times p(\psi|\nu) \times p(\nu) \times p(\tau_x) \times p(\tau_y) \times p(\tau_\phi) \times p(\tau_\beta) \\ &= MN\left(y; \Phi x, \frac{1}{\tau_y} I\right) \times MN\left(x; F\beta + r, \frac{1}{\tau_x} \Psi^{-1}\right) \\ &\times MN\left(\phi; \phi_0, \frac{1}{\tau_y \tau_\phi} M_0^{-1}\right) \times MN\left(\beta; \beta_0 + D^{-1}\beta_0^{\Delta}, \frac{1}{\tau_x \tau_\beta}[DA_0 D]^{-1}\right) \\ &\times \prod_{k=1}^{K} Bern(\gamma_k; \omega) \times Beta(\omega; \kappa_0, \delta_0) \\ &\times \prod_{t=1}^{T} Gamma(\psi; \nu/2, \nu/2) \times Gamma(\nu; \alpha_{\nu 0}, \zeta_{\nu 0}) \\ &\times Gamma(\tau_x; \alpha_{x0}, \zeta_{x0}) \times Gamma(\tau_y; \alpha_{y0}, \zeta_{y0}) \\ &\times Gamma(\tau_\phi; \alpha_{\phi 0}; \zeta_{\phi 0}) \times Gamma(\tau_\beta; \alpha_{\beta 0}, \zeta_{\beta 0}) \end{aligned} \tag{24}$$

4.2 Posterior of ϕ

$$\log p(\phi|rest) = \log MN\left(\phi; \mu_\phi, \Lambda_\phi^{-1}\right) + c_4^{\phi} \tag{25}$$

where

$$\Lambda_\phi = \tau_y \left(\tilde{X}_L' \tilde{X}_L + \tau_\phi M_0 \right)$$

$$\mu_\phi = \tau_y \Lambda_\phi^{-1} \left(\tilde{X}_L' \tilde{y} + \tau_\phi M_0' \phi_0 \right)$$

$$c_1^\phi = \frac{S+P}{2} \log \left(\frac{\tau_y}{2\pi} \right) + \frac{P}{2} \log \tau_\phi + \frac{1}{2} \log det \left(M_0 \right)$$

$$+ \log \left[MN \left(x; F\beta + r, \frac{T}{\tau_x \tau_y} \Psi^{-1} \right) \times MN \left(\beta; \beta_0 + D^{-1}\beta_0^\Delta, \frac{1}{\tau_x \tau_y \tau_\beta} \left[DA_0 D \right]^{-1} \right) \right.$$

$$\times \prod_{k=1}^{K} Bern \left(\gamma_k; \omega \right) \times Beta \left(\omega; \kappa_0, \delta_0 \right) \times \prod_{t=1}^{T} Gamma \left(\psi; \nu/2, \nu/2 \right)$$

$$\times Gamma \left(\nu; \alpha_{\nu 0}, \zeta_{\nu 0} \right) \times Gamma \left(\tau_x; \alpha_{x0}, \zeta_{x0} \right) \times Gamma \left(\tau_y; \alpha_{y0}, \zeta_{y0} \right)$$

$$\left. \times Gamma \left(\tau_\beta; \alpha_{\beta 0}, \zeta_{\beta 0} \right) \times Gamma \left(\tau_\phi; \alpha_{\phi 0}, \zeta_{\phi 0} \right) \right] + c^{ev}$$

$$c_2^\phi = c_1^\phi - \frac{\tau_y}{2} \left[\tau_\phi \phi_0' M_0 \phi + \tilde{y}' \tilde{y} \right]$$

$$c_3^\phi = c_2^\phi + \frac{1}{2} \mu_\phi' \Lambda_\phi \mu_\phi$$

$$c_4^\phi = c_3^\phi + \frac{P}{2} \log 2\pi - \frac{1}{2} \log det \Lambda_\phi$$

$$c_{ev} = -\log p \left(y_t \right)$$

The last constant makes the full posterior into a valid probability distribution.

Note that it may make sense to impose additional restrictions not easily captured by R. For instance, the prior distribution can be truncated to force $\phi \in [0,1]$. With this restriction, the above analysis applies over the interval, with the support of the posterior and prior distribution truncated and re-normalized as appropriate. Then the approximate posterior of ϕ is:

$$\log q \left(\phi \right) = \log MN \left(\phi; \overline{\mu}_\phi, \overline{\Lambda}_\phi^{-1} \right) + \overline{c}_4^\phi \tag{26}$$

where

$$\overline{\Lambda}_\phi = E[\tau_y]\left(E\left[\tilde{X}'_L\tilde{X}_L\right] + E\left[\tau_\phi\right]E\left[M_0\right]\right)$$

$$\overline{\mu}_\phi = E[\tau_y]\Lambda_\phi^{-1}\left(E\left[\tilde{X}'_L\tilde{y}\right] + E\left[\tau_\phi\right]E\left[M_0\right]\phi_0\right)$$

$$\overline{c}_1^\phi = \frac{S+P}{2}E\left[\log\left(\frac{\tau_y}{2\pi}\right)\right] + 0.5\log det\left(M_0\right) + \frac{P}{2}\log\tau_\phi$$

$$+ E_{-\phi}\log\left[MN\left(x; F\beta + r, \frac{1}{\tau_x\tau_y}\Psi^{-1}\right) \times MN\left(\beta;\beta_0 + D^{-1}\beta_0^\Delta, \frac{1}{\tau_x\tau_y\tau_\beta}\left[DA_0D\right]^{-1}\right)\right.$$

$$\times \prod_{k=1}^{K} Bern\left(\gamma_k;\omega\right) \times Beta\left(\omega;\kappa_0,\delta_0\right) \times \prod_{t=1}^{T} Gamma\left(\psi;\nu/2,\nu/2\right)$$

$$\times Gamma\left(\nu;\alpha_{\nu 0},\zeta_{\nu 0}\right) \times Gamma\left(\tau_x;\alpha_{x0},\zeta_{x0}\right) \times Gamma\left(\tau_y;\alpha_{y0},\zeta_{y0}\right)$$

$$\left.\times Gamma\left(\tau_\beta;\alpha_{\beta 0},\zeta_{\beta 0}\right) \times Gamma\left(\tau_\phi;\alpha_{\phi 0},\zeta_{\phi 0}\right)\right] + c^{ev}$$

$$\overline{c}_2^\phi = \overline{c}_1^\phi - \frac{E[\tau_y]}{2}\left(E\left[\tilde{y}'\tilde{y}\right] + E\left[\tau_\phi\right]\phi_0'E\left[M_0\right]\phi_0\right)$$

$$\overline{c}_3^\phi = \overline{c}_2^\phi + \frac{1}{2}\left(\overline{\mu}'_\phi\overline{\Lambda}_\phi\overline{\mu}_\phi\right)$$

$$\overline{c}_4^\phi = \overline{c}_3^\phi + \frac{P}{2}\log 2\pi - \frac{1}{2}\log\left(det\left(\overline{\Lambda}_\phi\right)\right)$$

4.3 Posterior of x

$$\log p\left(x|rest\right) = \log MN\left(x;\mu_x,\Lambda_x^{-1}\right) + c_4^x \tag{27}$$

where

$$\Lambda_x = \tau_y \left(\Phi' \Phi + \tau_x \Psi \right)$$

$$\mu_x = \tau_y \Lambda_x^{-1} \left(\Phi' y + \tau_x \Psi \left(r + F\beta \right) \right)$$

$$c_1^x = \frac{S+T}{2} \log \left(\frac{\tau_y}{2\pi} \right) + \frac{T}{2} \log \tau_x + \frac{1}{2} \log Det \left(\Psi \right)$$

$$+ \log \left[MN \left(\phi; \phi_0, \frac{1}{\tau_y \tau_\phi} M_0^{-1} \right) \times MN \left(\beta; \beta_0 + D^{-1} \beta_0^\Delta, \frac{1}{\tau_x \tau_y \tau_\beta} \left[DA_0 D \right]^{-1} \right) \right.$$

$$\times \prod_{k=1}^{K} Bern \left(\gamma_k; \omega \right) \times Beta \left(\omega; \kappa_0, \delta_0 \right) \times \prod_{t=1}^{T} Gamma \left(\psi; \nu/2, \nu/2 \right) \times Gamma \left(\nu; \alpha_{\nu 0}, \zeta_{\nu 0} \right)$$

$$\times Gamma \left(\tau_x; \alpha_{x0}, \zeta_{x0} \right) \times Gamma \left(\tau_y; \alpha_{y0}, \zeta_{y0} \right)$$

$$\times Gamma \left(\tau_x; \alpha_{x0}, \zeta_{x0} \right) \times Gamma \left(\tau_y; \alpha_{y0}, \zeta_{y0} \right)$$

$$\left. \times Gamma \left(\tau_\beta; \alpha_{\beta 0}, \zeta_{\beta 0} \right) \times Gamma \left(\tau_\phi; \alpha_{\phi 0}, \zeta_{\phi 0} \right) \right] + c^{ev}$$

$$c_2^x = c_1^x - \frac{\tau_y \tau_x}{2} \left(r + F\beta \right)' \Psi \left(r + F\beta \right) - \frac{\tau_y}{2} y' y$$

$$c_3^x = c_2^x + \frac{\mu_x' \Lambda_x \mu_x}{2}$$

$$c_4^x = c_3^x + \frac{T}{2} \log 2\pi - \log det \Lambda_x$$

The approximate posterior is:

$$\log q \left(x \right) = \log \left(N \left(x; \overline{\mu}_x, \overline{\Lambda}_x^{-1} \right) \right) + \overline{c}_4^x \tag{28}$$

where

$$\overline{\Lambda}_x = E\left[\tau_y\right]\left(E\left[\Phi'\Phi\right] + E\left[\tau_x\right]E\left[\Psi\right]\right)$$

$$\overline{\mu}_x = E\left[\tau_y\right]\Lambda_x^{-1}\left(E\left[\Phi\right]'y + E\left[\tau_x\right]E\left[\Psi\right]\left(F\overline{\mu}_\beta + r\right)\right)$$

$$\overline{c}_1^x = E_{-x}\left[\frac{S+T}{2}\log\frac{\tau_y}{2\pi} + \frac{T}{2}\log\tau_x + \frac{1}{2}\log det\Psi\right.$$

$$+ \log\left(MN\left(\phi;\phi_0,\frac{1}{\tau_y\tau_\phi}M_0^{-1}\right) \times MN\left(\beta;\beta_0 + D^{-1}\beta_0^\Delta, \frac{1}{\tau_x\tau_y\tau_\beta}\left[DA_0D\right]^{-1}\right)\right.$$

$$\times \prod_{k=1}^{K}Bern\left(\gamma_k;\omega\right) \times Beta\left(\omega;\kappa_0,\delta_0\right) \times \prod_{t=1}^{T}Gamma\left(\psi;\nu/2,\nu/2\right) \times Gamma\left(\nu;\alpha_{\nu0},\zeta_{\nu0}\right)$$

$$\times Gamma\left(\tau_x;\alpha_{x0},\zeta_{x0}\right) \times Gamma\left(\tau_y;\alpha_{y0},\zeta_{y0}\right)$$

$$\left.\left.\times Gamma\left(\tau_\beta;\alpha_{\beta0},\zeta_{\beta0}\right) \times Gamma\left(\tau_\phi;\alpha_{\phi0},\zeta_{\phi0}\right)\right)\right] + c^{ev}$$

$$\overline{c}_2^x = \overline{c}_1^x - E_{-x}\left[\frac{\tau_y}{2}\left(\tau_x\left(\beta'F' + r'\right)\Psi\left(F\beta + r\right) + y'y\right)\right]$$

$$\overline{c}_3^x = \overline{c}_2^x + \frac{\overline{\mu}_x'\overline{\Lambda}_x\overline{\mu}_x}{2}$$

$$\overline{c}_4^x = \overline{c}_3^x + \frac{T}{2}\log\left(2\pi\right) - \frac{1}{2}\log\left(det\overline{\Lambda}_x\right)$$

4.4 Posterior of τ_y

• Let $\tilde{\beta} \equiv \beta - \beta_0$. Then the conditional posterior is:

$$\log p\left(\tau_y|rest\right) = \log Gamma\left(\tau_y;\alpha_y,\zeta_y\right) + c_2^{\tau_y} \tag{29}$$

where

$$\alpha_y = \frac{S+T+P+K}{2} + \alpha_{y0}$$

$$\zeta_y = \zeta_{y0} + \frac{1}{2}\left(\tilde{y} - \tilde{X}_L\phi\right)'\left(\tilde{y} - \tilde{X}_L\phi\right) + \frac{\tau_\phi}{2}\left(\phi - \phi_0\right)' M_0 \left(\phi - \phi_0\right)$$
$$+ \frac{\tau_x}{2}\left((x-r) - F\beta\right)' \Psi \left((x-r) - F\beta\right)$$
$$+ \frac{\tau_x \tau_\beta}{2}\left(\tilde{\beta} - D^{-1}\beta_0^\Delta,\right)' DA_0D\left(\tilde{\beta} - D^{-1}\beta_0^\Delta\right)$$

$$c_1^{\tau_y} = +\alpha_{y0}\log\zeta_{y0} - \log\Gamma\left(\alpha_0\right) - \frac{S+T+K+P}{2}\log 2\pi + \frac{T+K}{2}\log\tau_x$$
$$+ \frac{1}{2}\log det M_0 + \frac{1}{2}\log det\Psi + \frac{1}{2}\log det\left(DA_0D\right) + \frac{P}{2}\log\tau_\phi + \frac{K}{2}\log\tau_\beta$$
$$+ \log\left(\prod_{k=1}^{K} Bern\left(\gamma_k;\omega\right) \times Beta\left(\omega;\kappa_0,\delta_0\right) \times \prod_{t=1}^{T} Gamma\left(\psi;\nu/2,\nu/2\right)\right.$$
$$\times Gamma\left(\nu;\alpha_{\nu0},\zeta_{\nu0}\right) \times Gamma\left(\tau_x;\alpha_{x0},\zeta_{x0}\right)$$
$$\left. \times Gamma\left(\tau_\beta;\alpha_{\beta0},\zeta_{\beta0}\right) \times Gamma\left(\tau_\phi;\alpha_{\phi0},\zeta_{\phi0}\right)\right) + c^{ev}$$

$$c_2^{\tau_y} = c_1^{\tau_y} + \log\Gamma\left(\alpha_y\right) - \alpha_y\log\zeta_y$$

And the approximate unconditional posterior for τ_y is:

$$\log q\left(\tau_y\right) = \log\left(Gamma\left(\tau_y;\overline{\alpha}_y,\overline{\zeta}_y\right)\right) + \overline{c}_2^{\tau_y} \tag{30}$$

where

$$\overline{\alpha}_y = \frac{S + T + P + K}{2} + \alpha_{y0}$$

$$\overline{\zeta}_y = \frac{1}{2} \left(E\left[\tilde{y}'\tilde{y}\right] + E\left[\phi'\tilde{X}_L'\tilde{X}_L\phi\right] - E\left[\tilde{y}'\tilde{X}_L\right]\mu_\phi - \mu_\phi'E\left[\tilde{X}_L'\tilde{y}\right] \right.$$

$$+ E\left[g_y^{-1}\right]\left(E\left[\phi'M_0\phi\right] + \phi_0 M_0 \phi_0 - \phi_0'M_0\mu_\phi - \mu_\phi'M_0\phi_0\right)$$

$$+ E\left[\tau_x\right]\left(E\left[x'\Psi x\right] + E\left[\left(\beta'F' + r'\right)\Psi\left(F\beta + r\right)\right] - \mu_x'E\left[\Psi\right]\left(F\mu_\beta + r\right) - \left(\mu_\beta'F' + r'\right)E\left[\Psi\right]\mu_x\right)$$

$$\left. + E\left[\tau_x\right]E\left[\tau_\beta\right]\left(E\left[\tilde{\beta}'DA_0D\tilde{\beta}\right] + \left(\beta_0^\Delta\right)'A_0\beta_0^\Delta - \left(\beta_0^\Delta\right)'A_0E\left[D\right]E\left[\tilde{\beta}\right] - E\left[\tilde{\beta}\right]'E\left[D\right]A_0\beta_0^\Delta\right)\right) + \zeta_{y0}$$

$$\overline{c}_1^{\tau_y} = E_{-\tau_y}\left[-\frac{S + T + P + K}{2}\log 2\pi + \frac{1}{2}\log det\left(\Psi\right) + \frac{1}{2}\log det\left(M_0\right) + \frac{1}{2}\log det\left(DA_0D\right) \right.$$

$$+ \frac{T + K}{2}\log \tau_x + \alpha_{y0}\log \zeta_{y0} - \log\Gamma\left(\alpha_{y0}\right) + \frac{P}{2}\log \tau_\phi + \frac{K}{2}\log \tau_\beta$$

$$+ \log\left(\prod_{k=1}^{K} Bern\left(\gamma_k; \omega\right) \times Beta\left(\omega; \kappa_0, \delta_0\right)\right.$$

$$\times \prod_{t=1}^{T} Gamma\left(\psi; \nu/2, \nu/2\right) \times Unif\left(\nu; \nu_0^-, \nu_0^+\right) \times Gamma\left(\tau_x; \alpha_{x0}, \zeta_{x0}\right)$$

$$\left.\left. \times Gamma\left(\tau_\beta; \tau_{\beta0}, \zeta_{\beta0}\right) \times Gamma\left(\tau_\phi; \tau_{\phi0}, \zeta_{\phi0}\right)\right)\right] + c^{ev}$$

$$\overline{c}_2^{\tau_y} = \overline{c}_1^{\tau_y} - \overline{\alpha}_y\log\overline{\zeta}_y + \log\Gamma\left(\overline{\alpha}_y\right)$$

4.5 Posterior for τ_x

$$\log p\left(\tau_x|rest\right) = \log Gamma\left(\tau_x; \alpha_x, \zeta_x\right) + c_2^{\tau_x} \tag{31}$$

where

$$\alpha_x = \frac{T+K}{2} + \alpha_{0x}$$

$$\zeta_x = \frac{\tau_y}{2}\left((x-r) - F\beta\right)' \Psi \left((x-r) - F\beta\right) + \frac{\tau_y \tau_\beta}{2}\left(\tilde{\beta} - D^{-1}\beta_0^\Delta\right)' DA_0 D\left(\tilde{\beta} - D^{-1}\beta_0^\Delta\right) + \zeta_{0x}$$

$$c_1 = \frac{T+K}{2}\log\frac{\tau_y}{2\pi} + \frac{1}{2}\log \det DA_0 D + \frac{1}{2}\log \det \Psi + \frac{K}{2}\log\tau_\beta$$
$$\quad + \alpha_{y0}\log\zeta_{y0} - \log\Gamma\left(\alpha_{y0}\right)$$
$$\quad + \log\left(MN\left(\phi;\phi_0,\frac{1}{\tau_y\tau_\phi}M_0^{-1}\right) \times MN\left(y;\Phi x,\frac{1}{\tau_y}I\right)\right.$$
$$\quad \times \prod_{k=1}^{K} Bern\left(\gamma_k;\omega\right) \times Beta\left(\omega;\kappa_0,\delta_0\right)$$
$$\quad \times \prod_{t=1}^{T} Gamma\left(\psi;\nu/2,\nu/2\right) \times Gamma\left(\nu;\alpha_{\nu0},\zeta_{\nu0}\right) \times Gamma\left(\tau_y;\alpha_{y0},\zeta_{y0}\right)$$
$$\quad \left. \times Gamma\left(\tau_\beta;\alpha_{\beta0},\zeta_{\beta0}\right) \times Gamma\left(\tau_\phi;\alpha_{\phi0},\zeta_{\phi0}\right)\right) + c^{ev}$$

$$c_2 = c_1 + \log\Gamma\left(\alpha_x\right) - \alpha_x\log\zeta_x$$

The unconditional posterior is approximated as:

$$\log q\left(\tau_x\right) = \log Gamma\left(\tau_x;\overline{\alpha}_x,\overline{\zeta}_x\right) + \overline{c}_2^{\tau_x} \tag{32}$$

where

$$\overline{\alpha}_x = \frac{T+K}{2} + \alpha_{x0}$$

$$\overline{\zeta}_x = \zeta_{x0} + \frac{E\left[\tau_y\right]}{2}\left(E\left[x'\Psi x\right] + E\left[(\beta'F' + r')\Psi(F\beta + r)\right] - \mu_x'E\left[\Psi\right](F\mu_\beta + r) - \left(\mu_\beta'F' + r'\right)E\left[\Psi\right]\mu_x\right)$$

$$+ \frac{E\left[\tau_y\right]E\left[\tau_\beta\right]}{2}\left(E\left[\tilde{\beta}'DA_0D\tilde{\beta}\right] + (\beta_0^\Delta)'A_0\beta_0^\Delta - (\beta_0^\Delta)'A_0E\left[D\right]E\left[\tilde{\beta}\right] - E\left[\tilde{\beta}\right]'E\left[D\right]A_0\beta_0^\Delta\right)$$

$$\overline{c}_1^{\tau_x} = E_{-\tau_x}\left[-\frac{T+K}{2}\log 2\pi + \frac{1}{2}\log det\left(\Psi\right) + \frac{1}{2}\log det\left(DA_0D\right)\right.$$

$$+ \frac{T+K}{2}\log \tau_y + \alpha_{x0}\log \zeta_{x0} - \log \Gamma\left(\alpha_{x0}\right) + \frac{K}{2}\log \tau_\beta$$

$$+ \log\left(MN\left(y;\Phi x, \frac{1}{\tau_y}I\right) \times MN\left(\phi;\phi_0, \frac{1}{\tau_y\tau_\phi}M_0^{-1}\right)\right.$$

$$\times \prod_{k=1}^{K} Bern\left(\gamma_k;\omega\right) \times Beta\left(\omega;\kappa_0,\delta_0\right)$$

$$\times \prod_{t=1}^{T} Gamma\left(\psi;\nu/2,\nu/2\right) \times Gamma\left(\nu;\alpha_{\nu0},\zeta_{\nu0}\right) \times Gamma\left(\tau_y;\alpha_{y0},\zeta_{y0}\right)$$

$$\left.\left.\times Gamma\left(\tau_\beta;\alpha_{\beta0},\zeta_{\beta0}\right) \times Gamma\left(\tau_\phi;\alpha_{\phi0},\zeta_{\phi0}\right)\right)\right] + c^{ev}$$

$$\overline{c}_2^{\tau_x} = \overline{c}_1^{\tau_x} - \overline{\alpha}_x\log \overline{\zeta}_x + \log \Gamma\left(\overline{\alpha}_x\right)$$

4.6   Posterior for τ_ϕ

$$\log p\left(\tau_\phi|rest\right) = \log Gamma\left(\tau_\phi;\alpha_\phi,\zeta_\phi\right) + c_2^{\tau\phi} \tag{33}$$

where

$$\alpha_\phi = \alpha_{\phi0} + \frac{P}{2}$$

$$\zeta_\phi = \zeta_{\phi0} + \frac{\tau_y}{2}\left(\phi - \phi_0\right)'M_0\left(\phi - \phi_0\right)$$

$$c_1 = \alpha_{\phi0}\log \zeta_{\phi0} - \log \Gamma\left(\alpha_{\phi0}\right) + \frac{1}{2}\log Det\left(M_0\right) + \frac{P}{2}\log\left(\frac{\tau_y}{2}\pi\right)$$

$$+ \log\left[MN\left(y;\Phi x, \frac{1}{\tau_y}I\right) \times MN\left(x;F\beta + r, \frac{1}{\tau_x\tau_y}\Psi^{-1}\right) \times MN\left(\beta;\beta_0 + D^{-1}\beta_0^\Delta, \frac{1}{\tau_x\tau_y\tau_\beta}\left[DA_0D\right]^{-1}\right)\right.$$

$$\times \prod_{k=1}^{K} Bern\left(\gamma_k;\omega\right) \times Beta\left(\omega;\kappa_0,\delta_0\right) \times \prod_{t=1}^{T} Gamma\left(\psi;\nu/2,\nu/2\right) \times Gamma\left(\tau_\beta;\tau_{\beta0},\zeta_{\beta0}\right)$$

$$\left.\times Gamma\left(\nu;\alpha_{\nu0},\zeta_{\nu0}\right) \times Gamma\left(\tau_x;\alpha_{x0},\zeta_{x0}\right) \times Gamma\left(\tau_y;\alpha_{y0},\zeta_{y0}\right)\right] + c^{ev}$$

$$c_2 = c_1 - \alpha_\phi\log \zeta_\phi + \log \Gamma\left(\alpha_\phi\right)$$

Approximate unconditional posterior:

$$\log q\left(\tau_\phi\right) = \log InvGamma\left(\tau_\phi; \overline{\alpha}_\phi, \overline{\zeta}_\phi\right) + \overline{c}_2^{\tau\phi} \tag{34}$$

where

$$\overline{\alpha}_\phi = \alpha_{\phi 0} + \frac{P}{2}$$

$$\overline{\zeta}_\phi = \zeta_{\phi 0} + E\left[\phi' M_0 \phi\right] + \phi_0 M_0 \phi_0 - \phi_0' M_0 \mu_\phi - \mu_\phi' M_0 \phi_0$$

$$\overline{c}_1^{\tau\phi} = \frac{P}{2} E\left[\log\left(\frac{\tau_y}{2\pi}\right)\right] + \frac{1}{2}\log det\left(M_0\right) + \alpha_{\phi 0}\log\zeta_{\phi 0} - \log\Gamma\left(\alpha_{\phi 0}\right)$$

$$+ E_{-gy}\log\left[MN\left(y; \Phi x, \frac{1}{\tau_y}I\right) \times MN\left(x; F\beta, \frac{1}{\tau_x\tau_y}\Psi^{-1}\right) \times MN\left(\beta; \beta_0 + D^{-1}\beta_0^\Delta, \frac{1}{\tau_x\tau_y\tau_\beta}\left[DA_0D\right]^{-1}\right)\right.$$

$$\times \prod_{k=1}^{K} Bern\left(\gamma_k; \omega\right) \times Beta\left(\omega; \kappa_0, \delta_0\right) \times \prod_{t=1}^{T} Gamma\left(\psi; \nu/2, \nu/2\right) \times Gamma\left(\tau_\beta; \tau_{\beta 0}, \zeta_{\beta 0}\right)$$

$$\left. \times Gamma\left(\nu; \alpha_{\nu 0}, \zeta_{\nu 0}\right) \times Gamma\left(\tau_x; \alpha_{x0}, \zeta_{x0}\right) \times Gamma\left(\tau_y; \alpha_{y0}, \zeta_{y0}\right)\right] + c^{ev}$$

$$\overline{c}_2^{\tau\phi} = c_1^{\tau\phi} - \overline{\alpha}_\phi\log\overline{\zeta}_\phi + \log\Gamma\left(\overline{\alpha}_\phi\right)$$

4.7 Posterior for τ_β

$$\log p\left(\tau_\beta|rest\right) = \log Gamma\left(\tau_\beta; \alpha_\beta, \zeta_\beta\right) + c_2^{\tau\beta} \tag{35}$$

where

$$\alpha_\beta \equiv \alpha_{\beta 0} + \frac{K}{2}$$

$$\zeta_\beta \equiv \zeta_{\beta 0} + \frac{\tau_x\tau_y}{2}\left(\tilde{\beta} - D^{-1}\beta_0^\Delta\right)' DA_0D\left(\tilde{\beta} - D^{-1}\beta_0^\Delta\right)$$

$$c_1^{\tau\beta} \equiv \alpha_{\beta 0}\log\zeta_{\beta 0} - \log\Gamma\left(\alpha_{\beta 0}\right) + \frac{K}{2}\log\frac{\tau_x\tau_y}{2\pi} + \frac{1}{2}\log detDA_0D$$

$$+ \log\left(MN\left(\phi; \phi_0, \frac{1}{\tau_y\tau_\phi}M_0^{-1}\right) \times MN\left(y; \Phi x, \frac{1}{\tau_y}I\right)\right.$$

$$\times \prod_{k=1}^{K} Bern\left(\gamma_k; \omega\right) \times Beta\left(\omega; \kappa_0, \delta_0\right)$$

$$\times \prod_{t=1}^{T} Gamma\left(\psi; \nu/2, \nu/2\right) \times Gamma\left(\nu; \alpha_{\nu 0}, \zeta_{\nu 0}\right) \times Gamma\left(\tau_\phi; \tau_{\phi 0}, \zeta_{\phi 0}\right)$$

$$\left. \times Gamma\left(\tau_y; \alpha_{y0}, \zeta_{y0}\right) \times Gamma\left(\tau_x; \alpha_{x0}, \zeta_{x0}\right)\right) + c^{ev}$$

$$c_2^{\tau\beta} \equiv c_1^{\tau\beta} - \alpha_\beta\log\zeta_\beta + \log\Gamma\left(\alpha_\beta\right)$$

Approximate unconditional posterior:

$$\log q\left(\tau_\beta\right) = \log Gamma\left(\tau_\beta; \overline{\alpha}_\beta, \overline{\zeta}_\beta\right) + \overline{c}_2^{\tau\beta} \tag{36}$$

where

$$\overline{\alpha}_\beta = \alpha_{\beta 0} + \frac{K}{2}$$

$$\overline{\zeta}_\beta = \zeta_{\beta 0} + \frac{E\left[\tau_x\right]E\left[\tau_y\right]}{2}\left(E\left[\tilde{\beta}' D A_0 D \tilde{\beta}\right] + (\beta_0^\Delta)' A_0 \beta_0^\Delta - (\beta_0^\Delta)' A_0 E\left[D\right]E\left[\tilde{\beta}\right] - E\left[\tilde{\beta}\right]' E\left[D\right] A_0 \beta_0^\Delta\right)$$

$$\overline{c}_1^{\tau\beta} \equiv E_{-\tau\beta}\left[\alpha_{\beta 0}\log\zeta_{\beta 0} - \log\Gamma\left(\alpha_{x0}^g\right) + \frac{K}{2}\log\frac{\tau_x\tau_y}{2\pi} + \frac{1}{2}\log det D A_0 D\right.$$

$$+ \log\left(MN\left(\phi;\phi_0,\frac{1}{\tau_y\tau_\phi}M_0^{-1}\right)\times MN\left(y;\Phi x,\frac{1}{\tau_y}I\right)\right.$$

$$\times \prod_{k=1}^{K} Bern\left(\gamma_k;\omega\right)\times Beta\left(\omega;\kappa_0,\delta_0\right)$$

$$\times \prod_{t=1}^{T} Gamma\left(\psi;\nu/2,\nu/2\right)\times Gamma\left(\nu;\alpha_{\nu 0},\zeta_{\nu 0}\right)\times Gamma\left(\tau_\phi;\tau_{\phi 0},\zeta_{\phi 0}\right)$$

$$\left.\left.\times Gamma\left(\tau_y;\alpha_{y0},\zeta_{y0}\right)\times Gamma\left(\tau_x;\alpha_{x0},\zeta_{x0}\right)\right)\right] + c^{ev}$$

$$\overline{c}_2^{\tau\beta} \equiv -\overline{\alpha}_\beta\log\overline{\zeta}_\beta + \log\Gamma\left(\overline{\alpha}_\beta\right)$$

4.8 Posterior for β

Note that setting the multi-variate distribution to $MN\left(\beta;\beta_0 + D^{-1}\beta_0^\Delta,\frac{1}{\tau_x\tau_y\tau_\beta}\left[DA_0D\right]^{-1}\right)$ greatly improves tractability, particularly for the approximate unconditional posterior approximation.

First, define $\tilde{\beta}_0^\Delta \equiv \beta_0 - D^{-1}\beta_0$. The the conditional posterior $p\left(\beta|rest\right)$ is:

$$\log p\left(\beta|rest\right) = \log MN\left(\beta;\mu_\beta,\Lambda_\beta\right) + c_4^\beta \tag{37}$$

where

$$\Lambda_\beta \equiv \tau_x \tau_y \left[F'\Psi F + \tau_\beta D A_0 D \right]$$

$$\mu_\beta \equiv \tau_x \tau_y \Lambda_\beta^{-1} \left[F'\Psi (x - r) + \tau_\beta D A_0 \left(D\beta_0 + \beta_0^\Delta \right) \right]$$

$$c_1^\beta \equiv \frac{T + K}{2} \log \frac{\tau_x \tau_y}{2\pi} + \frac{1}{2} \log det D A_0 D + \frac{1}{2} \log det \Psi + \frac{K}{2} \log \tau_\beta$$

$$- \frac{T}{2} \log T + \log \left(MN \left(\phi; \phi_0, \frac{1}{\tau_y \tau_\phi} M_0^{-1} \right) \times MN \left(y; \Phi x, \frac{1}{\tau_y} I \right) \right.$$

$$\times \prod_{k=1}^{K} Bern \left(\gamma_k; \omega \right) \times Beta \left(\omega; \kappa_0, \delta_0 \right)$$

$$\times \prod_{t=1}^{T} Gamma \left(\psi; \nu/2, \nu/2 \right) \times Gamma \left(\nu; \alpha_{\nu 0}, \zeta_{\nu 0} \right)$$

$$\times Gamma \left(\tau_y; \alpha_{y0}, \zeta_{y0} \right) \times Gamma \left(\tau_x; \alpha_{x0}, \zeta_{x0} \right)$$

$$\left. \times Gamma \left(\tau_\beta; \alpha_{\beta 0}, \zeta_{\beta 0} \right) \times Gamma \left(\tau_\phi; \alpha_{\phi 0}, \zeta_{\phi 0} \right) \right) + c^{ev}$$

$$c_2^\beta = c_1^\beta - \frac{\tau_x \tau_y}{2} \left((x - r)' \Psi (x - r) + \tau_\beta \left(D\beta_0 + \beta_0^\Delta \right)' A_0 \left(D\beta_0 + \beta_0^\Delta \right) \right)$$

$$c_3^\beta = c_2^\beta + \frac{\mu_\beta' \Lambda_\beta \mu_\beta}{2}$$

$$c_4^\beta = c_3^\beta - \frac{1}{2} \log \Lambda_\beta + \frac{K}{2} \log 2\pi$$

And the approximate unconditional posterior $q(\beta)$:

$$\log q(\beta) = \log \left[N \left(\beta; \overline{\mu}_\beta, \overline{\Lambda}_\beta \right) \right] + \overline{c}_4^\beta \tag{38}$$

where

$$\Lambda_\beta \equiv E\left[\tau_x\right]E\left[\tau_y\right]\left(F'E\left[\Psi\right]F + E\left[\tau_\beta\right]E\left[DA_0D\right]\right)$$

$$\mu_\beta \equiv E\left[\tau_x\right]E\left[\tau_y\right]\overline{\Lambda}_\beta^{-1}\left(F'E\left[\Psi\right]\left(\overline{\mu}_x - r\right) + E\left[\tau_\beta\right]\left(E\left[DA_0D\right]\beta_0 + E\left[D\right]A_0\beta_0^\Delta\right)\right)$$

$$\overline{c}_1^\beta \equiv E_{-\beta}\left[\frac{T+K}{2}\log\left(\frac{\tau_x\tau_y}{2\pi}\right) + \frac{1}{2}\log det\left(\Psi\right) + \frac{1}{2}\log det\left(DA_0D\right) + \frac{K}{2}\log\tau_\beta\right.$$

$$+ \log\left(MN\left(y;\Phi x,\frac{1}{\tau_y}I\right)\times MN\left(\phi;\phi_0,\frac{1}{\tau_y\tau_\phi}M_0^{-1}\right)\right)$$

$$\times\prod_{k=1}^{K}Bern\left(\gamma_k;\omega\right)\times Beta\left(\omega;\kappa_0,\delta_0\right)$$

$$\times\prod_{t=1}^{T}Gamma\left(\psi;\nu/2,\nu/2\right)\times Gamma\left(\nu;\alpha_{\nu0},\zeta_{\nu0}\right)$$

$$\times Gamma\left(\tau_x;\alpha_{x0},\zeta_{x0}\right)\times Gamma\left(\tau_y;\alpha_{y0},\zeta_{y0}\right)$$

$$\left.\times Gamma\left(\tau_\beta;\alpha_{\beta0},\zeta_{\beta0}\right)\times Gamma\left(\tau_\phi;\alpha_{\phi0},\zeta_{\phi0}\right)\right)\right] + c^{ev}$$

$$\overline{c}_2^\beta \equiv \overline{c}_1^\beta - E_{-\beta}\left[\frac{\tau_y\tau_x}{2}\left(\tau_\beta\left(D\beta_0 + \beta_0^\Delta\right)'DA_0D\left(D\beta_0 + \beta_0^\Delta\right) + \left(x-r\right)'\Psi\left(x-r\right)\right)\right]$$

$$\overline{c}_3^\beta \equiv \overline{c}_2^\beta + \frac{1}{2}\overline{\mu}_\beta'\overline{\Lambda}_\beta\overline{\mu}_\beta$$

$$\overline{c}_4^\beta \equiv \overline{c}_3^\beta + \frac{K}{2}\log\left(2\pi\right) - \frac{1}{2}\log det\left(\overline{\Lambda}_\beta\right)$$

4.9 Posterior for γ

What follows is the conditional posterior for γ_k in the scenario where each γ_k is conditionally independent of the other values of γ (denoted as γ_{-k}). In other words, $p\left(\gamma_k|\gamma_{-k},rest\right) = p\left(\gamma_k|rest\right)$. Note that this does not imply unconditionally that $\gamma_k \perp \gamma_{-k}$ as other variables (e.g. β) influence both γ_k and γ_{-k}.

Practically this implies A_0 is diagonal, such that $a_0 \equiv diag\left(A_0\right)$. Also recall $d_k^2 \equiv \gamma_k + \frac{1-\gamma_k}{v^2}$. As the only discrete distribution, the derivation for $p\left(\gamma\right)$ proceeds somewhat differently than others. The distribution for p_k with a conditionally independent prior for β is given by $\frac{\tilde{p}(\gamma_k=1)}{\tilde{p}_k(\gamma_k=0)+\tilde{p}_k(\gamma_k=1)}$.

$$\log p\left(\gamma_k\right) = \gamma_k\log p_{\gamma k} + \left(1-\gamma_k\right)\log\left(1-p_{\gamma k}\right) + c_3^{\gamma_k} \tag{39}$$

where

$$p_{\gamma k} = \frac{\tilde{p}_{\gamma k}|_1}{\tilde{p}_{\gamma k}|_0 + \tilde{p}_{\gamma k}|_1}$$

$$\tilde{p}_{\gamma k}|_1 \equiv \exp\left(-\frac{\tau_x \tau_y \tau_\beta a_{0k}}{2}\left(\tilde{\beta}_k^2 - 2\beta_{0k}^\Delta \tilde{\beta}_k\right)\right)\omega$$

$$\tilde{p}_{\gamma k}|_0 \equiv \exp\left(-\frac{\tau_x \tau_y \tau_\beta a_{0k}}{2}\left(\frac{\tilde{\beta}_k^2}{v^2} - \frac{2\beta_{0k}^\Delta \tilde{\beta}_k}{v}\right)\right)\frac{1-\omega}{v}$$

$$c_1^{\gamma k} \equiv \frac{1}{2}\log\frac{a_{0k}\tau_x \tau_y \tau_\beta}{2\pi}$$

$$+ \log\left(MN\left(\phi; \phi_0, \frac{1}{\tau_y \tau_\phi}M_0^{-1}\right) \times MN\left(y; \Phi x, \frac{1}{\tau_y}I\right)\right.$$

$$\times \prod_{j=1, j\neq k}^{K}\left(N\left(\beta_j; \frac{\beta_0^\Delta}{d_j} + \beta_0, \frac{1}{\tau_x \tau_y \tau_\beta a_{0j}d_j^2}\right) \times Bern\left(\gamma_j; \omega\right)\right) \times Beta\left(\omega; \kappa_0, \delta_0\right)$$

$$\times \prod_{t=1}^{T}Gamma\left(\psi; \nu/2, \nu/2\right) \times Gamma\left(\nu; \alpha_{\nu 0}, \zeta_{\nu 0}\right)$$

$$\times Gamma\left(\tau_y; \alpha_{y0}, \zeta_{y0}\right) \times Gamma\left(\tau_x; \alpha_{x0}, \zeta_{x0}\right)$$

$$\left.\times Gamma\left(\tau_\beta; \alpha_{\beta 0}, \zeta_{\beta 0}\right) \times Gamma\left(\tau_\phi; \alpha_{\phi 0}, \zeta_{\phi 0}\right)\right) + c^{ev}$$

$$c_2^{\gamma k} \equiv c_1^{\gamma k} - \frac{\tau_x \tau_y \tau_\beta a_{0k}\left(\beta_{0k}^\Delta\right)^2}{2}$$

$$c_3^{\gamma k} \equiv c_2^{\gamma k} + \log\left(\tilde{p}_{\gamma k}|_1 + \tilde{p}_{\gamma k}|_0\right)$$

Note that the normalization is accounted for in $c_3^{\gamma k}$. The normalization is fully revealed as the true probabilities must add to one. Similarly, the approximate distribution for any γ_k is given by $\frac{\tilde{q}_k(\gamma_k=1)}{\tilde{q}_k(\gamma_k=0)+\tilde{q}_k(\gamma_k=1)}$.

The approximate unconditional posterior is then:

$$\log q\left(\gamma_k\right) = \gamma_k \log \overline{p}_{\gamma k} + \left(1 - \gamma_k\right)\log\left(1 - \overline{p}_{\gamma k}\right) + \overline{c}_3^{\gamma k} \tag{40}$$

28

where

$$\overline{p}_{\gamma k} \equiv \frac{\tilde{q}\left(\gamma_k\right)|_1}{\tilde{q}\left(\gamma_k\right)|_1 + \tilde{q}\left(\gamma_k\right)|_0}$$

$$\tilde{q}\left(\gamma_k\right)|_1 = \exp\left(-\frac{E\left[\tau_x\right] E\left[\tau_y\right] E\left[\tau_\beta\right] a_{0k}}{2}\left(E\left[\tilde{\beta}_k^2\right] - 2E\left[\tilde{\beta}_k\right]\beta_{0k}^\Delta\right) + E\left[\log\left(\omega\right)\right]\right)$$

$$\tilde{q}\left(\gamma_k\right)|_0 = \exp\left(-\log\left(v\right) - \frac{E\left[\tau_x\right] E\left[\tau_y\right] E\left[\tau_\beta\right] a_{0k}}{2}\left(\frac{E\left[\tilde{\beta}_k^2\right]}{v^2} - \frac{2E\left[\tilde{\beta}_k\right]\beta_{0k}^\Delta}{v}\right) + E\left[\log\left(1-\omega\right)\right]\right)$$

$$\overline{c}_1^{\gamma k} \equiv E_{-\gamma_k}\left[\frac{1}{2}\log\left(\frac{\tau_x \tau_y \tau_\beta a_{0k}}{2\pi}\right)\right.$$

$$+ \log\left(MN\left(y; \Phi x, \frac{1}{\tau_y}I\right) \times MN\left(\phi; \phi_0, \frac{1}{\tau_y \tau_\phi}M_0^{-1}\right) \times MN\left(x; F\beta + r, \frac{1}{\tau_x \tau_y}\Psi^{-1}\right)\right.$$

$$\prod_{j=1, j\neq k}^{K}\left(N\left(\beta_j; \frac{\beta_{0j}^\Delta}{d_j} + \beta_{0j}, \frac{1}{\tau_x \tau_y \tau_\beta d_j^2 a_{0j}}\right) \times Bern\left(\gamma_j; \omega\right)\right) \times Beta\left(\omega; \kappa_0, \delta_0\right)$$

$$\times \prod_{t=1}^{T} Gamma\left(\psi; \nu/2, \nu/2\right) \times Gamma\left(\nu; \alpha_{\nu 0}, \zeta_{\nu 0}\right)$$

$$\times Gamma\left(\tau_x; \alpha_{x0}, \zeta_{x0}\right) \times Gamma\left(\tau_y; \alpha_{y0}, \zeta_{y0}\right)$$

$$\left.\left.\times Gamma\left(\tau_\beta; \alpha_{\beta 0}, \zeta_{\beta 0}\right) \times Gamma\left(\tau_\phi; \alpha_{\phi 0}, \zeta_{\phi 0}\right)\right)\right] + c^{ev}$$

$$\overline{c}_2^{\gamma k} \equiv \overline{c}_1^{\gamma k} - \frac{1}{2}E\left[\tau_x\right] E\left[\tau_y\right] E\left[\tau_\beta\right] a_{0k}\left(\beta_{0k}^\Delta\right)^2$$

$$\overline{c}_3^{\gamma k} \equiv \overline{c}_2^{\gamma k} + \log\left(\tilde{q}\left(\gamma_k\right)|_1 + \tilde{q}\left(\gamma_k\right)|_0\right)$$

4.9.1 Posterior for γ (General Case)

To get the off-diagonal terms for A_0, we use the below generalization:

$$p\left(\gamma_k\right) = \gamma_k \log p_{\gamma k} + \left(1 - \gamma_k\right)\log\left(1 - p_{\gamma k}\right) + c_3^{\gamma k} \tag{41}$$

where

$$p_{\gamma k} \equiv \frac{\tilde{p}_{\gamma k}|_1}{\tilde{p}_{\gamma k}|_0 + \tilde{p}_{\gamma k}|_1}$$

$$\tilde{p}_{\gamma k}|_1 \equiv \exp\left(-\frac{\tau_x \tau_y \tau_\beta}{2}\left[\left(\tilde{\beta}' D A_0 D \tilde{\beta} - 2\tilde{\beta}' D A_0 \beta_0^\Delta\right)\right]_{d_k=1}\right)\omega$$

$$\tilde{p}_{\gamma k}|_0 \equiv \exp\left(-\frac{\tau_x \tau_y \tau_\beta}{2}\left[\left(\tilde{\beta}' D A_0 D \tilde{\beta} - 2\tilde{\beta}' D A_0 \beta_0^\Delta\right)\right]_{d_k=v^{-1}}\right)\frac{1-\omega}{v}$$

$$c_1^{\gamma k} \equiv \frac{K}{2}\log\frac{\tau_x \tau_y \tau_\beta}{2\pi} + \frac{1}{2}\log \det A_0 + \sum_{j=1,j\neq k}^{K}\log d_j$$

$$+ \log\left(MN\left(\phi; \phi_0, \frac{1}{\tau_y \tau_\phi}M_0^{-1}\right) \times MN\left(y; \Phi x, \frac{1}{\tau_y}I\right)\right.$$

$$\times \prod_{j=1,j\neq k}^{K} Bern\left(\gamma_j; \omega\right) \times Beta\left(\omega; \kappa_0, \delta_0\right)$$

$$\times \prod_{t=1}^{T} Gamma\left(\psi; \nu/2, \nu/2\right) \times Gamma\left(\nu; \alpha_{\nu 0}, \zeta_{\nu 0}\right)$$

$$\times Gamma\left(\tau_y; \alpha_{y0}, \zeta_{y0}\right) \times Gamma\left(\tau_x; \alpha_{x0}, \zeta_{x0}\right)$$

$$\left.\times Gamma\left(\tau_\beta; \alpha_{\beta 0}, \zeta_{\beta 0}\right) \times Gamma\left(\tau_\phi; \alpha_{\phi 0}, \zeta_{\phi 0}\right)\right) + c^{ev}$$

$$c_2^{\gamma k} \equiv c_1^{\gamma k} - \frac{\tau_x \tau_y}{2g_x}\left(\beta_0^\Delta\right)' A_0 \beta_0^\Delta$$

$$c_3^{\gamma k} \equiv c_2^{\gamma k} + \log\left(\tilde{p}_{\gamma k}|_1 + \tilde{p}_{\gamma k}|_0\right)$$

And the approximate unconditional posterior is:

$$\log q\left(\gamma_k\right) = \gamma_k \log \overline{p}_{\gamma k} + \left(1 - \gamma_{\gamma k}\right)\log\left(1 - \overline{p}_{\gamma k}\right) + \overline{c}_3^{\gamma k} \tag{42}$$

where

$$\bar{p}_{\gamma k} \equiv \frac{\tilde{q}\left(\gamma_k\right)|_1}{\tilde{q}\left(\gamma_k\right)|_1 + \tilde{q}\left(\gamma_k\right)|_0}$$

$$\tilde{q}\left(\gamma_k\right)|_1 = \exp\left(-\frac{E\left[\tau_x\right]E\left[\tau_y\right]E\left[\tau_\beta\right]}{2}\left(E_{-\gamma k}\left[\tilde{\beta}'DA_0D\tilde{\beta} - 2\tilde{\beta}'DA_0\beta_0^\Delta | d_k = 1\right]\right) + E\log\left(\omega\right)\right)$$

$$\tilde{q}\left(\gamma_k\right)|_0 = \exp\left(-\log\left(v\right) - \frac{E\left[\tau_x\right]E\left[\tau_y\right]E\left[\tau_\beta\right]}{2}\left(E_{-\gamma k}\left[\tilde{\beta}'DA_0D\tilde{\beta} - 2\tilde{\beta}'DA_0\beta_0^\Delta | d_k = v^{-1}\right]\right) + E\log\left(1-\omega\right)\right)$$

$$\bar{c}_1^{\gamma k} \equiv E_{-\gamma k}\left[\frac{K}{2}\log\left(\frac{\tau_x\tau_y\tau_\beta}{2\pi}\right) + \frac{1}{2}\log det\left(A_0\right) + \sum_{j=1,j\neq k}^{K}\log d_j\right.$$

$$+ \log\left(MN\left(y;\Phi x,\frac{1}{\tau_y}I\right) \times MN\left(\phi;\phi_0,\frac{1}{\tau_y\tau_\phi}M_0^{-1}\right) \times MN\left(x;F\beta+r,\frac{1}{\tau_x\tau_y}\Psi^{-1}\right)\right.$$

$$\prod_{j=1,j\neq k}^{K}\left(Bern\left(\gamma_j;\omega\right)\right) \times Beta\left(\omega;\kappa_0,\delta_0\right)$$

$$\times \prod_{t=1}^{T} Gamma\left(\psi;\nu/2,\nu/2\right) \times Gamma\left(\nu;\alpha_{\nu 0},\zeta_{\nu 0}\right)$$

$$\times Gamma\left(\tau_x;\alpha_{x0},\zeta_{x0}\right) \times Gamma\left(\tau_y;\alpha_{y0},\zeta_{y0}\right)$$

$$\left.\left.\times Gamma\left(\tau_\beta;\alpha_{\beta 0},\zeta_{\beta 0}\right) \times Gamma\left(\tau_\phi;\alpha_{\phi 0},\zeta_{\phi 0}\right)\right)\right] + c^{ev}$$

$$\bar{c}_2^{\gamma k} \equiv \bar{c}_1^{\gamma k} - \frac{1}{2}E_{-\gamma k}\left[\tau_x\tau_y\tau_\beta\left(\beta_0^\Delta\right)'A_0\beta_0^\Delta\right]$$

$$\bar{c}_3^{\gamma k} \equiv \bar{c}_2^{\gamma k} + \log\left(\tilde{q}\left(\gamma_k\right)|_1 + \tilde{q}\left(\gamma_k\right)|_0\right)$$

4.10   Posterior for ω

$$\log p\left(\omega\right) = \log Beta\left(\kappa,\delta\right) + c_3^\omega \tag{43}$$

where

$$\kappa \equiv \kappa_0 + \sum_{k=1}^{K} \gamma_k$$

$$\delta \equiv \delta_0 + K - \sum_{k=1}^{K} \gamma_k$$

$$c_1^\omega \equiv -\log B\left(\kappa_0, \delta_0\right) + \log\left(MN\left(\phi; \phi_0, \frac{1}{\tau_y \tau_\phi} M_0^{-1}\right) \times MN\left(y; \Phi x, \frac{1}{\tau_y} I\right) \right.$$

$$\times MN\left(x; F\beta + r, \frac{1}{\tau_x \tau_y} \Psi^{-1}\right) \times MN\left(\beta; \beta_0 + D^{-1}\beta_0^\Delta, \frac{1}{\tau_x \tau_y \tau_\beta} [DA_0 D]^{-1}\right)$$

$$\times \prod_{t=1}^{T} Gamma\left(\psi; \nu/2, \nu/2\right) \times Gamma\left(\nu; \alpha_{\nu 0}, \zeta_{\nu 0}\right)$$

$$\times Gamma\left(\tau_y; \alpha_{y 0}, \zeta_{y 0}\right) \times Gamma\left(\tau_x; \alpha_{x 0}, \zeta_{x 0}\right)$$

$$\left. \times Gamma\left(\tau_\beta; \tau_{\beta 0}, \zeta_{\beta 0}\right) \times Gamma\left(\tau_\phi; \tau_{\phi 0}, \zeta_{\phi 0}\right) \right) + c^{ev}$$

$$c_2^\omega \equiv c_1^\omega + \log B\left(\kappa, \delta\right)$$

The approximate unconditional posterior is:

$$\log q\left(\omega\right) = \log\left(Beta\left(\overline{\kappa}, \overline{\delta}\right)\right) + \overline{c}_2^\omega \tag{44}$$

where

$$\overline{\kappa} \equiv \kappa_0 + \sum_{k} \overline{p}_{\gamma k}$$

$$\overline{\delta} \equiv \delta_0 + K - \sum_{k} \overline{p}_{\gamma k}$$

$$\overline{c}_1^\omega \equiv E_{-\omega}\left[-\log B\left(\kappa_0, \delta_0\right) + \log\left(MN\left(y; \Phi x, \frac{1}{\tau_y} I\right) \times MN\left(\phi; \phi_0, \frac{1}{\tau_y \tau_\phi} M_0^{-1}\right) \right.\right.$$

$$\times MN\left(x; F\beta + r, \frac{1}{\tau_x \tau_y} \Psi^{-1}\right) \times MN\left(\beta; \beta_0 + D^{-1}\beta_0^\Delta, \frac{1}{\tau_x \tau_y \tau_\beta} [DA_0 D]^{-1}\right)$$

$$\times \prod_{t=1}^{T} Gamma\left(\psi; \nu/2, \nu/2\right) \times Gamma\left(\nu; \alpha_{\nu 0}, \zeta_{\nu 0}\right)$$

$$\times Gamma\left(\tau_x; \alpha_{x 0}, \zeta_{x 0}\right) \times Gamma\left(\tau_y; \alpha_{y 0}, \zeta_{y 0}\right)$$

$$\left.\left. \times Gamma\left(\tau_\beta; \alpha_{\beta 0}, \zeta_{\beta 0}\right) \times Gamma\left(\tau_\phi; \alpha_{\phi 0}, \zeta_{\phi 0}\right) \right) \right] + c^{ev}$$

$$\overline{c}_2^\omega \equiv \overline{c}_1^\omega + \log B\left(\kappa, \delta\right)$$

4.11 Posterior for ψ_t

Conditional posterior:
$$\log p\left(\psi_t\right) = \log Gamma\left(\psi_t, \alpha_{\psi t}, \zeta_{\psi t}\right) + c_2^{\psi t} \tag{45}$$

where

$$\alpha_{\psi t} \equiv \frac{\nu+1}{2}$$

$$\zeta_{\psi t} \equiv \frac{\nu}{2} + \frac{\tau_x \tau_y}{2}\left((x_t - r_t) - f_t'\beta\right)^2$$

$$c_1^{\psi t} \equiv \frac{\nu}{2}\log\frac{\nu}{2} - \log\Gamma\left(\frac{\nu}{2}\right) + \frac{1}{2}\log\left(\frac{\tau_x \tau_y}{2\pi}\right)$$

$$+ \log\left(\prod_{j=1, j\neq t}^{T}\left[N\left(x_t; f_t'\beta + r, \frac{1}{\tau_x \tau_y \psi_t}\right) \times Gamma\left(\psi_j; \frac{\nu}{2}, \frac{\nu}{2}\right)\right]\right.$$

$$\times MN\left(\phi; \phi_0, \frac{1}{\tau_y \tau_\phi}M_0^{-1}\right) \times MN\left(y; \Phi x, \frac{1}{\tau_y}I\right)$$

$$\times MN\left(\beta; \beta_0 + D^{-1}\beta_0^{\Delta}, \frac{1}{\tau_x \tau_y \tau_\beta}[DA_0 D]^{-1}\right)$$

$$\times \prod_{k=1}^{K} Bern\left(\gamma_k; \omega\right) \times Beta\left(\omega; \kappa_0, \delta_0\right) \times Gamma\left(\nu; \alpha_{\nu 0}, \zeta_{\nu 0}\right)$$

$$\times Gamma\left(\tau_y; \alpha_{y0}, \zeta_{y0}\right) \times Gamma\left(\tau_x; \alpha_{x0}, \zeta_{x0}\right)$$

$$\left. \times Gamma\left(\tau_\beta; \alpha_{\beta 0}, \zeta_{\beta 0}\right) \times Gamma\left(\tau_\phi; \alpha_{\phi 0}, \zeta_{\phi 0}\right)\right) + c^{ev}$$

$$c_2^{\psi t} \equiv c_1^{\psi t} - \alpha_{\psi t}\log\zeta_{\psi t} + \log\Gamma\left(\alpha_{\psi t}\right)$$

Approximate unconditional posterior:
$$\log q\left(\psi_t\right) = \log\left(Gamma\left(\psi_t; \alpha_{\psi t}, \zeta_{\psi t}\right)\right) + \bar{c}_2^{\psi t} \tag{46}$$

where

$$\overline{\alpha}_{\psi t} \equiv \frac{E\left[\nu\right]}{2} + \frac{1}{2}$$

$$\overline{\zeta}_{\psi t} \equiv \frac{E\left[\tau_x\right] E\left[\tau_y\right]}{2} \left(E\left[x_t^2\right] - 2\mu_{xt}\left(f_t'\mu_\beta + r_t\right) + E\left[\left(\beta' f_t + r_t\right)\left(f_t'\beta + r_t\right)\right]\right) + \frac{E\left[\nu\right]}{2}$$

$$\overline{c}_1^{\psi t} \equiv E_{-\psi t} \left[\frac{1}{2} \log \frac{\tau_x \tau_y}{2\pi} + \frac{\nu}{2} \log \frac{\nu}{2} - \log \Gamma\left(\frac{\nu}{2}\right) \right.$$

$$+ \sum_{j=1, j \neq t}^{T} \log N\left(x_t; f_t'\beta + r, \frac{1}{\tau_x \tau_y \psi_t}\right) + \sum_{j=1, j \neq t}^{T} \log Gamma\left(\psi_j; \nu/2, \nu/2\right)$$

$$+ \log \left(MN\left(y; \Phi x, \frac{1}{\tau_y}I\right) \times MN\left(\phi; \phi_0, \frac{1}{\tau_y \tau_\phi}M_0^{-1}\right) \times MN\left(\beta; \beta_0 + D^{-1}\beta_0^\Delta, \frac{1}{\tau_x \tau_y \tau_\beta}\left[DA_0D\right]^{-1}\right) \right.$$

$$\times \prod_{k=1}^{K} Bern\left(\gamma_k; \omega\right) \times Beta\left(\omega; \kappa_0, \delta_0\right)$$

$$\times Unif\left(\nu; \nu_0^-, \nu_0^+\right) \times Gamma\left(\tau_x; \alpha_{y0}, \zeta_{y0}\right) \times Gamma\left(\tau_y; \alpha_{y0}, \zeta_{y0}\right)$$

$$\left. \left. \times Gamma\left(\tau_\beta; \alpha_{\beta0}, \zeta_{\beta0}\right) \times Gamma\left(\tau_\phi; \alpha_{\phi0}, \zeta_{\phi0}\right) \right) \right] + c^{ev}$$

$$\overline{c}_2^{\psi t} \equiv \overline{c}_1^{\psi t} + \log \Gamma\left(\alpha_{\psi t}\right) - \alpha_{\psi t} log\left(\zeta_{\psi t}\right)$$

4.12   Posterior for ν

Conditional:

$$p\left(\nu\right) = \left(\frac{\nu}{2}\right)^{\frac{T\nu}{2} + \alpha_{\nu0} - 1} \Gamma^{-T}\left(\frac{\nu}{2}\right) \exp\left(\frac{\nu\eta_1}{2}\right) \eta_2 \exp c_3^\nu \tag{47}$$

where

$$\eta_1 \equiv \sum_{t \in 1:T} \left(\log \psi_t - \psi_t\right) - 2\zeta_{\nu 0}$$

$$\eta_2 \equiv \left[\int_{\nu^-}^{\infty} \left(\frac{\nu}{2}\right)^{\frac{T\nu}{2} + \alpha_{\nu 0} - 1} \Gamma^{-T}\left(\frac{\nu}{2}\right) \exp\left(\frac{\nu \eta_1}{2}\right) d\nu\right]^{-1}$$

$$c_1^{\nu} \equiv \alpha_{\nu 0} \log \zeta_{\nu 0} - \log \Gamma\left(\alpha_{\nu 0}\right)$$

$$+ \log \left(MN\left(y; \Phi x, \frac{1}{\tau_y} I\right) \times MN\left(\phi; \phi_0, \frac{1}{\tau_y \tau_\phi} M_0^{-1}\right) \times MN\left(\beta; \beta_0 + D^{-1}\beta_0^\Delta, \frac{1}{\tau_x \tau_y \tau_\beta}\left[DA_0 D\right]^{-1}\right) \right.$$

$$\times MN\left(x; F\beta + r, \frac{1}{\tau_x \tau_y}\Psi^{-1}\right) \times \prod_{t=1}^{T} Gamma\left(\psi_t; \nu/2, \nu/2\right)$$

$$\times \prod_{k=1}^{K} Bern\left(\gamma_k; \omega\right) \times Beta\left(\omega; \kappa_0, \delta_0\right)$$

$$\times Gamma\left(\tau_x; \alpha_{y0}, \zeta_{y0}\right) \times Gamma\left(\tau_y; \alpha_{y0}, \zeta_{y0}\right)$$

$$\left. \times Gamma\left(\tau_\beta; \alpha_{\beta 0}, \zeta_{\beta 0}\right) \times Gamma\left(\tau_\phi; \alpha_{\phi 0}, \zeta_{\phi 0}\right) \right) \right] + c^{ev}$$

$$c_2^{\nu} \equiv c_1^{\nu} + \left(\alpha_{\nu 0} - 1\right) \log 2 - \sum_{t \in 1:T} \log \psi_t$$

$$c_3^{\nu} \equiv c_2^{\nu} - \log \eta_2$$

Unconditional:

$$q\left(\nu\right) = \left(\frac{\nu}{2}\right)^{\frac{T\nu}{2} + \alpha_{\nu 0} - 1} \Gamma^{-T}\left(\frac{\nu}{2}\right) \eta_2 \exp\left(\frac{\nu \eta_1}{2}\right) \exp c_3^{\nu} \tag{48}$$

where

$$c_1^\nu \equiv E_{-\nu}\left[\alpha_{\nu 0}\log\zeta_{\nu 0} - \log\Gamma\left(\alpha_{\nu 0}\right)\right.$$

$$+\log\left(MN\left(y;\Phi x,\frac{1}{\tau_y}I\right)\times MN\left(x;F\beta+r,\frac{1}{\tau_x\tau_y}\Psi^{-1}\right)\right.$$

$$\times MN\left(\phi;\phi_0,\frac{1}{\tau_y\tau_\phi}M_0^{-1}\right)\times MN\left(\beta;\beta_0+D^{-1}\beta_0^\Delta,\frac{1}{\tau_x\tau_y\tau_\beta}\left[DA_0D\right]^{-1}\right)$$

$$\times\prod_{k=1}^K Bern\left(\gamma_k;\omega\right)\times Beta\left(\omega;\kappa_0,\delta_0\right)$$

$$\times Gamma\left(\tau_x;\alpha_{x0},\zeta_{x0}\right)\times Gamma\left(\tau_y;\alpha_{y0},\zeta_{y0}\right)$$

$$\left.\left.\times Gamma\left(\tau_\beta;\alpha_{\beta 0},\zeta_{\beta 0}\right)\times Gamma\left(\tau_\phi;\alpha_{\phi 0},\zeta_{\phi 0}\right)\right)\right]$$

$$c_2^\nu \equiv c_1^\nu - \sum_t E\log\psi_t + \left(\alpha_{\nu 0}-1\right)\log\left(2\right)$$

$$c_3^\nu \equiv c_2^\nu - \log\eta_2$$

$$\overline{\eta}_1 = \sum_t E\left[\log\psi_t - \psi_t\right] - 2\zeta_{\nu 0}$$

$$\overline{\eta}_2 \equiv \left(\int_{\nu^-}^\infty\left[\left(\frac{\nu}{2}\right)^{\frac{T\nu}{2}+\alpha_{\nu 0}-1}\Gamma^{-T}\left(\frac{\nu}{2}\right)\exp\left(\frac{\nu\eta_1}{2}\right)\right]d\nu\right)^{-1}$$