# Online Supplementary Appendix

## Proof of Proposition 1

Conditional on , the normalisation term can be treated as a constant.

.

Since , we recognize that .

To compute the incremental weights , we use the Bayes formula:

 .

## Proof of Proposition 2

To prove Proposition 2, we use a statistical property of the multivariate truncated normal distribution. Though the marginal distributions do not have analytic forms in general, an exception is the case in which a variable block is unrestricted. The results are summarized in the following lemma:

**Lemma**: suppose that . Then the marginal distribution for remains truncated normal: , while the marginal distribution for is the extended skewed normal with the density , where , .

**Proof:** The normalisation term for the joint distribution of and satisfies:

,

where , .

Thus the marginal densities for and are given by

,

.

Now we start to prove Proposition 2. We will show that the optimal importance function follows a truncated normal distribution, while the incremental weight is an extended skewed normal density. By the preceding lemma, we only need to show and  are multivariate truncated normal distributions.

.

We have dropped the normalisation term , as we are using the symbol . The term denotes the conditional distribution in the context of a Gaussian linear state space model:

,

where , , .

In a Gaussian linear model, all the marginal and conditional distributions are normal. So is a multivariate normal density. It follows that is a multivariate truncated normal distribution, in which is free while is truncated to the region .

Similarly, , where is the counterpart of in the Gaussian linear state space model. Therefore, is a multivariate truncated normal density. By the lemma, remains multivariate truncated normal. Apply the lemma again, is an extended skewed normal density.

Now we show how to use the Kalman filter to partial out so as to obtain the filtering distributions.

First, conditioning on , the predictive distribution for are computed by the Kalman filter based on the GLSM. The predictive moments and represent the mean and variance for the Gaussian density .

Second, consider the joint distribution , which is a multivariate truncated normal distribution:

,

Third, we partial out and obtain , which is

.

Conditioning on , we obtain , which follows . As for the incremental weights , we apply the lemma and obtain Eq (8).

## Proof of Proposition 3

For a linear state space model with deterministic initial states, the filtered states are linear with respect to initial states. See De Jong (1991). The augmented Kalman filter satisfies:

.

By the law of iterated expectations, we have

.

For the initial state smoothing, we use the fact that the prediction errors of the observations are also linear with respect to the initial states.

.

Recall that is represented by particles with weights . By the Bayes formula, the smoothing distribution can be represented by the same particles with the updated weights such that

 .

The likelihood function has the following decomposition form:

.

Note that is a consistent estimator for .

 .

Hence Eq (13) is a consistent estimator for .

## Proof of Proposition 4

Conditional on the period particles and observations , the weights and have been fully determined. Leaving the conditioning variables implicit in our notation, we have , , , .

## Proof of Proposition 5

Invert the state equation: , hence .

Since , we have .

Also, , hence the conditional distribution

.

## Proof of Proposition 6

Note that is the state filtering distribution corresponding to a single-period PCSSM with the deterministic initial state . Applying Proposition 4, we obtain .

The incremental weights can be decomposed as

.

Since , we have

.

Recall that , hence

.

## Proof of Proposition 7

Conditioning on , the normalisation term is a constant for both ICSSM and PCSSM. Therefore, the proof for Proposition 2 applies, and the optimal importance function is truncated normal: . As for the incremental importance weights, it has the decomposition form:

where and equals the truncation probability , as .

## Temporal and Cross-sectional Rao-Blackwellisation

Temporal Rao-Blackwellisation can be carefully interpreted as a special case of cross-sectional Rao-Blackwellisation, if we allow time-varying state dimensions, linear algebra with empty matrices, and judicious state classification in the last unconstrained period. In Section 4, we partition the state vector into the constrained and unconstrained , and then simulate by the particle filter. To apply that algorithm under time-varying constraints, we put for , . For period , all the state variables are classified as the unconstrained states. There are no new particles to generate. However, the incremental importance weights must be computed in order to update the unnormalised weights. This is exactly the smoothing procedure given by Eq (12). Most importantly, we give a special treatment by artificially labelling the period- states as constrained states with infinity bounds, because cross-sectional Rao-Blackwellisation requires that the normalisation term cannot be a function of the past unconstrained states. We generate period- particles from with the importance weights given by Eq (12). If we resample particles and reverse the order of sampling and resample under the optimal importance function, this is exactly the sampling procedure given by Eq (14).

## Filtering and Estimating AFNS model with ZLB constraint

The original AFNS model has three state variables: . We reparametrize them as , where . The PCSSM Rao-Blackwellised filter only generates particles for , while and are marginalized. Also, we augment the state vector by the static variables , which are recursively estimated by the Rao-Blackwellised particle filter. There are no particles for , since their analytic filtering distributions are available conditional on . In summary, the enlarged discrete-time transition equation is given by

,

where , , , , , , , , , , and , , are independent standard normal noises.

There is a difference between ICSSM and PCSSM in this application. The unconstrained state variables and have an impact on the ICSSSM normalisation term, and they are classified as constrained states with infinity bounds in order to apply Rao-Blackwellisation. In contrast, the PCSSM normalisation term is a constant, and all the unconstrained states, including , can be marginalized by cross-sectional Rao-Blackwellisation. In addition, temporal Rao-Blackwellisation is employed, as the ZLB constraint is imposed after December 2008.

The unknown parameters include and the noise variance for observation equations. Given a set of parameter values, the Rao-Blackwellised particle filter delivers the likelihood function by the average unnormalised weights. Under flat priors, the posterior density is largely determined by the likelihood function. Parameters are estimated by a random-walk Metropolis sampler. Our proposal distribution is a multivariate t distribution with the covariance matrix obtained from the unconstrained-model MLE covariance matrix. A proposal draw is accepted with the probability equal to the ratio of the proposal and current densities, or one, whichever is smaller. We generate 50000 draws with 10000 burn-in samples, and the trace plot suggests that the chain converges and mixes well. The Metropolis acceptance ratio is around 0.25. We tried other proposal distributions, including multivariate normal and t with the baseline covariance matrix multiplied by a constant between 0.5 and 2. The estimation results are similar to those reported in Table 2.