REFEREE REPORT ON
OPTIMAL NON-NEGATIVE FORECAST RECONCILIATION

1. The abstract is too long and does not get to the point in a concise way, please reformulate.
2. The matrix $\boldsymbol{C}$ after eq (2) is not defined
3. Non-negative least square is not new can you provide references since the beginning?
4. Assuming that $\boldsymbol{\Lambda}_h$ is positive definite does not seem so straightforward, that means that forecast errors are linearly independent, but if variables are highly collinear this might not happen. I think it all depends on how you compute $\hat{\boldsymbol{y}}_t(h)$. If forecast capture all comovements than I can believe the assumption. Some more explanation should be given. I see you compute univariate forecasts (via ARMA) but why not using multivariate methods in high-dimensions, e.g. factor models (see Stock and Watson, 2002), this would control better $\boldsymbol{\Lambda}_h$.
5. "large size of the structures that typically arise in forecast reconciliation" what are the orders of magnitude? Your application has 555 series which can be handled by mulitvariate forecasting methods.
6. When you define $\boldsymbol{\lambda}^*$ that is equal to the derivative of q wrt $\boldsymbol{b}$ while you?re writing the opposite fix the notation.
7. Why in (9) you use the norm to define a quadratic loss? Keep the same notation as before.
8. Given the computational cost of running a constrained minimization with respect to the unconstrained case, what would happen if once we compute (1) we just throw away the negative forecasts and compute $\tilde{\boldsymbol{y}}$ with the remaining ones? Are we really after all m forecasts? Can you provide more motivation? And anyway what happens if you just impose the negative forecast to be zero. Can you repeat Table 7 in this case? Still the numbers there do not seem dramatically large.
9. What is an ETS model?
10. You say you log-transform data to compute forecasts and then you back-transform but doesn?t this amount to taking an exponential and therefore you must get positive values? This part should be clarified.

References

James H Stock & Mark W Watson (2002) Forecasting Using Principal Components From a Large Number of Predictors, Journal of the American Statistical Association, 97:460, 1167-1179