Journal of Financial Econometrics, 2016, Vol. 14, No. 2, 248–252
doi: 10.1093/jjfinec/nbv006
Advance Access Publication Date: 27 May 2015
Commentary

OXFORD

Comments on: Reflections on the Probability Space Induced by Moment Conditions with Implications for Bayesian Inference

Dante Amengual and Enrique Sentana

CEMFI

Address correspondence to Enrique Sentana, CEMFI, Casado del Alisal 5, 28014, Madrid, Spain, or e-mail: sentana@cemfi.es

Received February 17, 2015; revised February 17, 2015; accepted April 22, 2015

The traditional Bayesian approach to inference is based on the combination of a fully specified density for the data conditional on the model parameters (the likelihood) with prior views on those parameters. Aside from other methodological considerations, the advantage of using prior information may be particularly important in low-frequency macro/finance time series contexts in which the number of observations is insufficient to precisely pin down the values of the unknown model parameters.

Nevertheless, a potential drawback of the traditional Bayesian approach is that it is a full information procedure, which requires the correct specification of features of the distribution of the observed variables that the researcher might not be particularly interested in. In fact, many researchers prefer to use limited-information frequentist procedures, often with semi-parametric components, because under certain regularity conditions they reduce the potential for inconsistent estimation resulting from incorrect distributional assumptions. Whether those regularity conditions hold in any particular application (see Sims 2007), or whether the finite-sample performance of the limited-information, semi-parametric procedures agrees with the usual first-order asymptotic approximations even when they hold (see e.g., Cattaneo and Jansson 2014), is a different matter.

Although the methodological debate might never get settled, the classical approach clearly dominates its Bayesian counterpart in terms of both the number and computational simplicity of limited-information procedures available to empirical researchers. For that reason, any attempt to provide Bayesian counterparts to such frequentist procedures is extremely welcome. This is particularly true of the proposal considered in the paper by Gallant (2015), which discusses conditions under which one can give a proper Bayesian interpretation to the Laplace estimators proposed by Chernozhukov and Hong (2003).

Recent econometric practice, especially in macroeconomics, provides an alternative reason for being interested in Bayesian versions of limited-information procedures. As is well known, the availability of theoretically sound fast simulation procedures has freed up Bayesian statistics and econometrics from the restrictions imposed by the need to have closed-form solutions for posterior distributions. In fact, nowadays it is often easier to generate a huge number of draws from the posterior distribution of many complicated models for which the likelihood is not easily computable than to obtain the maximum likelihood estimators. However, sometimes it is unclear whether the seemingly sharp results obtained are due to the specification of a full parametric model or the use of a prior density, as the prior sensitivity analyses usually reported tend to be fairly local in scope. In this context, the ability to compare full and limited-information procedures with a common prior is also especially relevant (see Gómez-Jareño (2004) for an application to stochastic volatility models).

Finally, there is a third reason for being interested in limited-information Bayesian methods. Certain increasingly popular estimators, such as the continuously updated Generalized Methods of Moments (GMM) estimator (CU-GMM) of Hansen, Heaton, and Yaron (1996), sometimes give rise to implausible results. For example, in the context of consumption-based asset pricing models similar to one considered by Gallant in his paper, the simulations in Hansen, Heaton, and Yaron (1996) indicate that CU-GMM occasionally generates extreme estimators that lead to large pricing errors with even larger variances. While Peñaranda and Sentana (2015) suggest to bound the values that the prices of risk parameters can take by imposing good deal restrictions (see Cochrane and Saa-Requejo 2000), the possibility of imposing more flexible priors is certainly worth it.

As potential users of the methodology analysed in Gallant's (2015) paper, we have not focused our comments and questions on purely theoretical considerations, but rather on its appeal for practitioners.

1 Other Bayesian GMM Approaches

There have been several approaches to introduce limited-information Bayesian methods in econometric practice, the earliest precedent being the old simultaneous equations literature on limited-information methods with a Gaussian likelihood (see Zellner 1971, Section 9.5, or the survey by Dreze and Richard 1983).

A much closer precedent is Kwan (1999), who makes use of the uniform asymptotic normality (or "Hajek regularity") of certain classical estimators to provide a Bayesian interpretation for them. Effectively, his results rely on an asymptotic pivotal argument which allows to reverse the sampling distribution of $\hat{\theta}|\theta$ into the posterior $\theta|\hat{\theta}$. Given that Kwan (1999) explicitly mentions that similar pivotal arguments have been used by other authors in different contexts (Fraser 1968, 1972; Boos and Monahan 1986; Florens, Mouchart, and Rolin 1990, Chapter 8, Examples 4 and 9; or Doksum and Lo 1990), it would be useful to understand the relationship between Assumption 1 in Gallant's (2015) paper and the assumptions those authors make.

More recently, Müller (2013) explains how an asymptotically valid posterior density can be constructed from a GMM estimator and its second moment. He also shows that, under some regularity conditions, the asymptotic distribution of the limited-information posterior is the mirror image of the corresponding result in the classical distribution theory.

Again, a discussion of the relationship between those regularity conditions and Assumption 1 would be most welcome.

Although the Laplace estimation methodology of Chernozhukov and Hong (2003) is rather general, in practice it is often applied by exponentiating the continuously updated version of the GMM criterion function. The numerical equivalence of CU-GMM and Euclidean empirical likelihood estimators (see Antoine, Bonnal, and Renault 2007) immediately begs the question of the relationship between Gallant's (2015) proposal and earlier attempts to provide Bayesian interpretations to single-step GMM procedures of the Generalized Empirical Likelihood variety. Although the proliferation of nuisance parameters associated with the probabilities of the different observations makes a straightforward application of Bayesian techniques impractical, Inoue (2001), Lazar (2003), Schennach (2005), and Ragusa (2007) contain careful theoretical analyses of this idea. In that regard, Schennach (2005) explicitly relates her approach to the limited-information Bayesian approach in Kim (2002), who generalized the maximum entropy approach of Golan, Judge, and Miller (1996) and Zellner (1996, 1997) to GMM.

From a practical perspective, it would be interesting to understand the relationship between the Bayesian version of the Laplace estimators of Chernozhukov and Hong (2003) considered in this paper with those procedures.

A final approach that also enjoys some popularity is the nonparametric Bayesian methods considered by Ferguson (1973) and Rubin (1981). Chamberlain and Imbens (2003) contain a couple of interesting economic applications. A clarification of the relationship between those nonparametric bootstrap methods and Assumption 1 in Gallant's (2015) paper is particularly important given that for tractability reasons those methods often rely on the assumption of multinomial distributions for the observed variables.

2 Other Comments

2.1 Finite-Sample Performance of the Limited-Information Procedure

Obviously, a price to pay for using limited-information procedures is that we no longer have finite-sample inferences and the asymptotic approximation of the limited-information estimator may be poor even in large samples.

2.2 Primitive Conditions and Asymptotic Approximations

Given that the asymptotic Gaussian approximation to the orthonormalized moment conditions is usually based on a central limit theorem for the sample mean of the moments and a uniform law of large numbers for a Heteroskedasticity and Autocorrelation Consistent (HAC) estimator of their limiting variance, it would be useful to think of more primitive assumptions for the data generation process that would justify Assumption 1 in Gallant (2015).

2.3 Conflict Between Bayesian and Classical Approaches

Although this is well beyond the scope of the present article, there are some well-known examples in which classical and Bayesian results may lead to different conclusions. Unit roots or weak instruments are obvious examples. It would be useful to understand whether those issues affect the results of the paper by Gallant (2015).

2.4 Reparameterizations

Another topic that is often sidestepped in theoretical discussions but which is of much practical importance, are parameterization issues. Again, the paper by Gallant (2015) is silent about this, but there are well-known examples of moment condition models which are identified in terms of a particular parameterization but whose Jacobian becomes severely ill conditioned when written in terms of alternative parameterizations (see Hillier (1990) and Peñaranda and Sentana (2015) for some examples in the context of IV and linear factor pricing models, respectively).

2.5 Moment Choice

In some cases, a researcher has a potentially very large number of moments at her disposal. Florens and Simoni (2015) have recently studied the use of Bayesian methods to impose priors on the validity and importance of those moments.

References

Antoine, A., H. Bonnal, and E. Renault. 2007. On the Efficient Use of the Informational Content of Estimating Equations: Implied Probabilities and Euclidean Empirical Likelihood. *Journal of Econometrics* 138: 461–487.

Boos, D. D., and J. F. Monahan.1986. Bootstrap Methods Using Prior Information. *Biometrika* 73: 77–83.

Cattaneo, M., and M., Jansson. 2014. Bootstrapping Kernel-Based Semiparametric Estimators. Mimeo, Berkeley.

Chamberlain, G., and G. W. Imbens. 2003. Nonparametric Applications of Bayesian Inference. *Journal of Business and Economic Statistics* 21: 12–18.

Chernozhukov, V., and H. Hong. 2003. An MCMC Approach to Classical Estimation. *Journal of Econometrics* 115: 293–346.

Cochrane, J. H., and J. Saa-Requejo. 2000. Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets. *Journal of Political Economy* 108: 79–119.

Doksum, K. A., and A. Y. Lo. 1990. Consistent and Robust Bayes Procedures for Location Based on Partial Information. *Annals of Statistics* 18: 443–453.

Dreze, J. H., and J. Richard. 1983. "Bayesian Analysis of Simultaneous Equation Systems." In Z. Griliches, and M. Intriligator (eds.), *Handbook of Econometrics*, vol. 1. Amsterdam: North-Holland.

Ferguson, T. 1973. A Bayesian Analysis of Some Nonparametric Problems. *Annals of Statistics* 1: 209–230.

Florens, J. P., M. Mouchart, and J. Rolin.1990. *Elements of Bayesian Statistics*. New York: Marcel Dekker.

Florens, J. P., and A. Simoni. 2015. *Gaussian Processes and Bayesian Moment Estimation*. Mimeo, CREST.

Fraser, D. A. S. 1968. *The Structure of Inference*. New York: Wiley.

Fraser, D. A. S. 1972. Bayes, Likelihood, and Structural. *Annals of Mathematical Statistics* 43: 777–790.

Gallant, A. R. 2015. Reflections on the Probability Space Induced by Moment Conditions with Implications for Bayesian Inference. *Journal of Financial Econometrics*, 14: 229–247.

Golan, A., G. G. Judge, and D. Miller. 1996. *Maximum Entropy Econometrics, Robust Estimation with Limited Data*. New York: Wiley.

Gomez-Jareño, S.. 2004. Laplace-Type Estimators for Stochastic Volalitility Models. Unpublished Master dissertation, CEMFI.

Hansen, L. P., J. Heaton, and A. Yaron. 1996. Finite Sample Properties of Some Alternative GMM Estimators. *Journal of Business and Economic Statistics* 14: 262–280.

Hillier, G. H. 1990. On the Normalization of Structural Equations: Properties of Direct Estimators. *Econometrica* 58: 1181–1194.

Inoue, A. 2001. *A Bayesian Method of Moments in Large Samples*. Mimeo, North Carolina State University.

Kim, J. Y. 2002. Limited information likelihood and Bayesian analysis, *Journal of Econometrics* 107, 175–193.

Kwan, Y. K. 1999. Asymptotic Bayesian Analysis Based on a Limited Information Estimator. *Journal of Econometrics* 88: 99–121.

Lazar, N. 2003. Bayesian Empirical Likelihood. *Biometrika* 90: 319–326.

Müller, U. K. 2013. Risk of Bayesian Inference in Misspecified Models, and the Sandwich Covariance Matrix. *Econometrica* 81: 1805–1849.

Peñaranda, F. and E. Sentana. 2015. A Unifying Approach to the Empirical Evaluation of Asset Pricing Models. *Review of Economics and Statistics* 97: 412–435.

Ragusa, G. 2007. Bayesian Likelihoods for Moment Condition Models. Mimeo, Irvine.

Rubin, D. 1981. Bayesian Bootstrap. *Annals of Statistics* 9: 130–134.

Schennach, S. M. 2005. Bayesian Exponentially Tilted Empirical Likelihood. *Biometrika* 92: 31–46.

Sims, C. A. 2007. Thinking about Instrumental Variables. Mimeo, Princeton.

Zellner, A. 1971. *An Introduction to Bayesian Inference in Econometrics*. New York: Wiley.

Zellner, A. 1996. “Bayesian Method of Moments (BMOM) Analysis of Mean and Regression Models.” In J. C. Lee, W. C. Johnson, and A. Zellner (eds.), *Modelling and Prediction Honoring Seymour Geisser*. New York: Springer, pp. 61–74.

Zellner, A. 1997. “The Bayesian Method of Moments (BMOM): Theory and Application.” In T. Fomby, and R. C. Hill (eds.), *Advances in Econometrics*. Cambdrige: Cambridge University Press.