Journal of Financial Econometrics, 2020, Vol. 18, No. 4, 715–720
doi: 10.1093/jjfinec/nbaa003
Advance Access Publication Date: 13 June 2020
Commentary

Comment on: Pseudo-True SDFs in Conditional Asset Pricing Models

Lars Peter Hansen
University of Chicago

Address correspondence to Lars Peter Hansen, Departments of Economics and Statistics and the Booth School of Business, University of Chicago, Chicago, IL 60637, or e-mail: lhansen@uchicago.edu.

Received March 6, 2020; editorial decision February 18, 2020; accepted February 19, 2020

The Antoine, Proulx and Renault (APR) paper assembles a rich set of results on empirical methods for asset pricing by exploring the impact misspecification and conditioning information have on stochastic discount factor (SDF) models. The SDF formulation of an asset pricing model with conditioning information follows directly from Hansen and Richard (1987). It investigates the impact of using mean square pricing errors as a criterion for model-fitting following Hansen and Jagannathan (HJ) (1997). By design, the HJ approach gives an alternative to generalized method of moments (GMM) estimation for SDF models that is meant to tolerate misspecification and provide more meaningful model comparisons.

Papers by Nagel and Singleton (2011); Fang, Ren, and Yuan (2011); Gagliardini and Ronchetti (2019) and APR explore (a) how to best incorporate conditioning information in ways that are tractable and (b) how to supply inferential procedures that support these methods. These and related contributions are part of an under-appreciated but, to my mind, an important research agenda. It is indeed valuable to see the role of conditioning information in econometric implementation treated in a systematic and formal way rather than the common *ad hoc* implementation that has occurred in many applied papers.

The APR paper covers a lot of ground, and I will only comment on part of it. My discussion will revisit a so-called population analysis that underlies the HJ analysis and will provide a complementary perspective. I will not, however, direct any of my comments to their discussion of inferential methods. The more limited scope of my comment is not intended to suggest that the contribution to statistical inference is less important. It happens that I have more to add about the population analysis. The remainder of my comment will be organized around four questions that I pose. My hope is that these questions will help to frame future research in this area. I give my own perspective on all four questions and why they interest me.

1 Conditional Misspecification

I start by reviewing the bound construction and derivation. In so doing, I follow the derivation in Hansen, Heaton, and Luttmer (1995) which then allows for extensions that include some forms of transaction costs and the inclusion of positivity restrictions. While this derivation is not “conditional” by imitating constructions in Hansen and Richard (1987) and following the lead of Gagliardini and Ronchetti (2019), the conditional extension is straightforward.

Let $\{m(\theta) : \theta \in \Theta\}$ denote a parameterized family of possibly misspecified SDFs. For a given θ, find a conditional (on a sigma algebra $\Im$) least squares approximation to the family of correctly specified discount factors by solving:

Problem 1.1

$$\min_M \frac{1}{2}\mathbb{E}([m(\theta) - M]^2|\Im)$$

subject to

$$\mathbb{E}(MR|\Im) = 1_n$$

where R is an n-dimensional vector of one period gross returns and 1_n is an n-dimensional vector of ones.

The constraint included in Problem 1.1 restricts the correctly specified SDFs M to satisfy the pricing equation for returns. I have included a one-half scaling of the objective only for notational convenience.

Problem 1.1 has some interesting generalizations that replace the pricing equalities by pricing inequalities due to transaction costs. Mathematically, this extension is tractable and substantively interesting, and it could be developed along the lines of Hansen, Heaton, and Luttmer (1995). But for APR to explore them, they would lose some of the quasi-analytical simplicity they exploit in some of their very interesting characterizations. In my discussion, I will follow APR by abstracting from pricing inequalities.

Through standard duality calculations applied the minimization problem, we are led to the first-order conditions:

$$-m(\theta) + M + R'\lambda = 0,$$

where λ is a conditional vector of Lagrange multipliers measurable with respect to $\Im$. Thus, the difference between $m(\theta)$ and the closest M in a least squares sense is

$$m(\theta) - M = R'\lambda \tag{1}$$

and the implied pricing error vector for the parametric model $m(\theta)$ of the SDF satisfies

$$\mathbb{E}[m(\theta)R|\Im] - 1_n = \mathbb{E}(RR'|\Im)\lambda. \tag{2}$$

Armed with this calculation, it follows directly that the minimized objective of Problem 1.1 is

$$\mathsf{L}(\theta|\Im) = \frac{1}{2}(\mathbb{E}[m(\theta)R - 1_n|\Im])'[\mathbb{E}(RR'|\Im)]^{-1}(\mathbb{E}[m(\theta)R - 1_n|\Im]),$$

which is a quadratic form in the vector of pricing error for the misspecified model. The APR paper and the earlier Gagliardini and Ronchetti (2019) paper propose to use this for evaluating asset pricing models. The important novelty in this work is that it confronts explicitly the conditioning information captured by $\mathfrak{I}$. Conditioning information comes into play in both the construction of the matrix: $\mathbb{E}(RR'|\mathfrak{I})$ and in the construction of the pricing error vector (2).

Prior to exploring parameter choice, let me add two twists on the previous discussion. First, from the pricing-error Equation (2), the vector of conditional multipliers is

$$\lambda^* = [\mathbb{E}(RR\prime|\mathfrak{I})]^{-1}[\mathbb{E}[m(\theta)R|\mathfrak{I}] - 1_n].$$

From the first-order condition (1), the closest discount factor to $m(\theta)$ that prices the assets is

$$M^* = m(\theta) - R'\lambda^*.$$

This leads me to my first query:

Why not use the misspecified stochastic discount factor $m(\theta)$ to select the valid SDF M^* that is closest to $m(\theta)$ and prices the vector returns correctly?

The better the model, the smaller the adjustment. While we may choose to call a minimizing choice θ over the parameter space Θ the pseudo-true parameter, why not refer the corresponding M^* as the implied pseudo-true SDF?

Depending on how we use the conditioning information in practice, the $m(\theta)$ adjustment, M^*, may end up a nonparametric component as the formula for M^* involves conditional first and second moments. Nonparametric estimators have slower rates of convergence and this may make the practical construction of M^* less reliable. Therefore, depending upon the application there are perhaps good reasons to commit to the parameterized family *a priori* of misspecified SDFs rather than to the corrected versions. I am not fully convinced by this argument, however, and I suggest that the M^* construction should be of interest in its own right.

My second query has to do with the positivity of the SDF. For some problems, the parameterized family of SDFs is positive, such as in the exponentially conditionally affine models explored by the APR paper and for models that link back directly to the intertemporal marginal rate of substitution of the marginal investors. In such cases, it might also be of interest to impose $M > 0$ with probability 1 as a conditional counterpart to the parallel analysis in HJ (1997).

Why not impose that $M > 0$ with probability one in the constraint set for minimization Problem 1.1?

To do so requires that we entertain the possibility that the infimum in the objective is not attained, or equivalently that we replace the strict inequality in the constraint with a weaker one whereby $M \geq 0$. An SDF $M \geq 0$ can be approximated arbitrarily well by a sequence of strictly positive SDFs provided that there is at least one SDF that satisfies the pricing relation. As is well known, the existence of a strictly positive SDF follows from the absence of arbitrage within the observable set of asset returns R.

Under the restriction that $M \geq 0$, it may be shown that the corrected discount factor is

$$M^* = [m(\theta) - R'\lambda^*]_+,$$

where the notation $[\cdot]_+$ means that whenever the random variable argument is negative, it gets replaced by zero. The vector λ^* solves a maximization problem with first-order conditions that are equivalent to the pricing restrictions. APR do not explore this case, presumably because they want to preserve the quasi-analytic characterization that underlies some of their revealing representation results. Nevertheless, I see the restriction $M > 0$ as a worthy one to explore in more depth when using conditioning information.

2 Conditional Parameter Estimation

APR make continued references to the "pseudo-true parameter." As they note, it could be constructed in two ways. The first way is the θ that solves:

Problem 2.1

$$\min_{\theta \in \Theta} \mathbb{L}(\theta|\mathfrak{I}).$$

Solving this problem in population leads to a pseudo-true parameter value θ^* that depends on conditioning information. In other words, the resulting parameter θ would vary over time as a function of this information. This may seem uninteresting for some parameterized models, but it is germane for conditional factors or their exponential counterparts.

Why not let the "parameter" θ be a scaling of the underlying factors for conditional factor models or their exponential counterparts so that:

$$m(\theta) = \theta \cdot F, \text{ or}$$
$$m(\theta) = \exp(\theta \cdot F)$$

and solve conditional minimization Problem 2.1?

By solving Problem 2.1, the θ vector inherits the conditional dependence built in the specification-error criterion $\mathbb{L}(\theta|\mathfrak{I})$. Implementing Problem 2.1 would thus let the factor coefficients depend on this same conditioning information. Instead of embracing this approach, APR focus on specifications with invariant parameters and more limited conditioning information dependence for the factor coefficients. This leaves it to the applied researcher to deal with these two forms conditioning in potentially distinct manners. While the APR paper approach follows some of the prior literature, it would be a welcome addition to have a better understanding of the advantages and disadvantages to treating the impact of conditioning information in these different ways.

One advantage to adopting the approach mentioned here is that the interpretation of parameter minimization by APR extends nicely to the presence of conditioning information. To see this, assuming a smooth parameterization with an interior minimum, the first-order conditions with respect to θ can be stated conditionally as:

$$\left(\mathbb{E}\left[\frac{\partial m}{\partial \theta}|_{\theta=\theta_o} R'|\mathfrak{I}\right]\right)[\mathbb{E}(RR'|\mathfrak{I})]^{-1}\mathbb{E}[m(\theta_o)R - 1_n|\mathfrak{I}] = 0.$$

These first-order conditions imply that the payoff vector

$$\left(\mathbb{E}\left[\frac{\partial m}{\partial \theta}|_{\theta=\theta_o} R' \mid \mathfrak{I}\right]\right)[\mathbb{E}(RR' \mid \mathfrak{I})]^{-1}R$$

is "priced correctly" by the chosen misspecified discount factor $m(\theta_o)$ conditioned in $\mathfrak{I}$. This avoids the weaker "on average" conclusion that APR note in their paper when θ is not chosen conditionally.

3 Invariant Parameters

I now follow the featured case by Gagliardini and Ronchetti (2019) and APR in which there is a finite-dimensional parameter vector to be estimated that does not depend on the conditioning information. We can treat this case in one of two ways. Following the insights in Hansen and Singleton (1982) and Hansen and Richard (1987) and using the abstract definition of conditional expectations, we could revert to the unconditional HJ (1997) approach and introduce conditioning information through the back door by scaling returns by random variables that are $\mathfrak{I}$ measurable. The second way, the one used by Gagliardini and Ronchetti (2019) and APR, identifies a pseudo-true parameter vector by solving:

Problem 3.1

$$\delta^2 = \min_{\theta\in\Theta} \mathbb{E}[\mathbb{L}(\theta|\mathfrak{I})],$$

The more explicit approach can reveal better how best to include conditioning information in estimation and inference.

The first-order conditions from Problem 3.1 provide the so-called "estimating equations" for the pseudo-true parameter. When the returns used to construct the pricing errors are one-period, there is a revealing contrast between this approach to parametric estimation and the efficient GMM approach of Hansen (1985) and Nagel and Singleton (2011). This insight is evident from the APR paper and its precursor, Gagliardini and Ronchetti (2019). The conditional HJ approach uses the inverse of the conditional second moment matrix:

$$\mathbb{E}(RR'|\mathfrak{I})$$

whereas an efficient (continuously update) GMM approach could be implemented but instead uses the conditional second moment matrix of the pricing-error vector:

$$\mathbb{E}\Big([m(\theta)R - 1_n|\mathfrak{I}][m(\theta)R - 1_n|\mathfrak{I}]'|\mathfrak{I}\Big).$$

This latter matrix is parameter-dependent. This comparison of conditional second moment matrices provides a nice extension of an insight made by HJ (1997) and others contrasting the two approaches in previous research that did not include an explicit treatment of conditioning information.

Given that the parameterized family of models is misspecified, I believe it to be fruitful to do more than approximate the minimizer in Problem 3.1. In addition, I suggest the study of sets of parameters that satisfy some misspecification threshold. More precisely, I welcome the constructing of potentially "small" misspecification sets of the form:

$$\{\theta \in \Theta : \mathbb{E}[\mathbb{L}(\theta|\mathfrak{I})] \leq \delta^2\}$$

for alternative $\delta > {}_{\delta}$. For a given δ, the resulting set is a "misspecified" counterpart to a confidence set often used in applied research. This leads me to my fourth question:

> *Why the focus on inference about a pseudo-true parameter vector instead of on the more general construction of a set of parameters that satisfies a pre-specified misspecification bound?*

Chen, Hansen, and Hansen (2020) proposed a counterpart to this strategy, but one based on a different measure of model misspecification.[1]

One mark of a good paper is that it opens the door to subsequent new research. The APR paper most definitely lives up to this standard of success. The four questions I pose are evidence of this success.

References

Chen, X., L. P. Hansen, and P. Hansen. 2020. Robust Identification of Investor Beliefs. Unpublished manuscript.

Fang, Y., Y. Ren, and Y. Yuan. 2011. Nonparametric Estimation and Testing of Stochastic Discount Factor. *Finance Research Letters* 8: 196–205.

Gagliardini, P., and D. Ronchetti. 2019. Comparing Asset Pricing Models by the Conditional Hansen–Jagannathan Distance*. *Journal of Financial Econometrics* published online.

Hansen, L. P. 1985. A Method for Calculating Bounds on the Asymptotic Covariance Matrices of Generalized Method of Moments Estimators. *Journal of Econometrics* 30: 203–238.

Hansen, L. P., and R. Jagannathan. 1997. Assessing Specification Errors in Stochastic Discount Factor Models. *The Journal of Finance* 52: 557–590.

Hansen, L. P., and S. F. Richard. 1987. The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models. *Econometrica* 50: 587–614.

Hansen, L. P., and K. J. Singleton. 1982. Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models. *Econometrica* 50: 1269–1286.

Hansen, L. P., J. Heaton, and E. G. Z. Luttmer. 1995. Econometric Evaluation of Asset Pricing Models. *Review of Financial Studies* 8: 237–274.

Nagel, S., and K. J. Singleton. 2011. Estimation and Evaluation of Conditional Asset Pricing Models. *The Journal of Finance* 66: 873–909.

1 Chen, Hansen, and Hansen (2020) used potential belief distortions to dictate how severe the SDF misspecification.