Title

To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Trends: Comment.

Journal

Journal of Applied Econometrics, 6(4)

Author

Leamer, Edward E

Publication Date

2023-02-17

Comment on `To Criticize the Critics'
Author(s): Edward E. Leamer
Source: *Journal of Applied Econometrics*, Vol. 6, No. 4 (Oct. - Dec., 1991), pp. 371-373
Published by: Wiley
Stable URL: https://www.jstor.org/stable/2096688
Accessed: 18-02-2023 00:08 UTC

JOURNAL OF APPLIED ECONOMETRICS, VOL. 6, 371–373 (1991)

COMMENT ON 'TO CRITICIZE THE CRITICS'

EDWARD E. LEAMER
Department of Economics, UCLA, Los Angeles, CA 90024, USA

1. INTRODUCTION

This is wonderful.

I have argued for many years that for most issues of inference in economics prior information matters in the sense that two economists can legitimately make substantially different inferences from the same data set. Faced with this dependence of the inferences on the prior information, our profession needs to take the following three steps:

1. Select a method for combining sample information with prior information.
2. Identify the range of alternative priors.
3. Characterize the mapping from the set of alternative priors into the corresponding inferences.

I see these three steps as the essential aspects of this paper by Peter Phillips.

The first step is to decide on a method of combining prior information with sample information. The obvious choice theoretically is the Bayesian model of inference. This would be mine, and it is also Phillips' choice. It is not the profession's choice, however. Most economists prefer to combine prior information with sample information through specification searches in which many different models are estimated with the same data set, and those few results that conform sufficiently with the priors are selected for reporting purposes. If you insist on this method of combination you will have difficulties with the next two steps: identifying the range of priors and characterizing the dependence of the inferences on the prior. However, the Bayesian model forces you to form a prior distribution, an exercise which can stretch your patience and tax your credulity in high dimensional settings. But, like Phillips, let us accept the Bayesian model for this discussion.

The next step is to identify the range of alternative prior distributions. For purposes of discussion, consider the simplest autoregressive model discussed by Phillips:

$$y_t = \rho y_{t-1} + \varepsilon_t; \quad t = 1, 2, \ldots, n; \ y_0 \text{ fixed.} \tag{1}$$

For this kind of model Sims has adopted the flat, improper prior, which is uniform with indefinite range. Phillips has argued instead for Jeffreys' improper prior depicted in his Figure 1. This prior rises continuously over the whole line. Which of these prior distributions do you like? Each departs substantially from my own prior, but my guess is that Sims' will give me a better idea of my posterior mean, and Phillips' will give a better idea of my posterior mode.

But my first comment is that I do not think it is wise to form a prior about ρ without first identifying y and the unit of time. Do you have the same ideas about ρ if y is nominal GNP, real wages, the saving rate, the Dow–Jones average, ... ? I don't. Do you have the same ideas

0883–7252/91/040371–03$05.00
Received April 1991

about ρ if the time unit is centuries, decades, years, days and seconds. I don't. Let's take y to be the US savings rate in the second half of the twentieth century and the time unit to be quarters of a year. For this variable I expect ρ to be close to one, but almost certainly less than one and certainly positive. Thus my prior takes on the value zero at $\rho = 0$ and rises smoothly up until $\rho = 1$, at which point it falls discontinuously and thereafter falls away smoothly for ever. (I am a bit disturbed by the fact that a savings rate necessarily lies between zero and one, a restriction that my model (1) chooses to ignore. I am wondering if I should adjust my prior to allow somehow for the approximate nature of the model. Of course, any model is only an approximation, valid for some ranges of the data but not for others. Let's keep in mind that (1) applies to savings rates that are adequately far from zero and one.) With this as my prior, what do I think of Sims' and Phillips' suggestions? Sims' flat prior does not favour values of ρ close to one, whereas Phillips' does, so I prefer Phillips'. But Phillips' prior favours values in excess of one. Mine is sharply against these values. Since Sims' prior is at least neutral with regard to values of ρ greater than or less than one, I can see why Sims' prior might be a better approximation of my own than is Phillips'. Actually, I suspect that the mode of my posterior distribution is best approximated if I use Phillips' prior, since the behaviour of his prior for values in excess of one is not likely to have much effect on the mode of the posterior near the sample estimate, which I strongly suspect will be less than one. But the posterior mean is likely to be dragged upwards by the enormous amount of prior probability that Phillips allocates to values of ρ in excess of one. Thus for approximating the mean of my posterior distribution, I suspect that Sims' distribution will do the better job.

My prior is different if the variable y is stock prices divided by the CPI. The only way for this simple autoregressive model to account for the positive expected real return on stocks is to have a value of ρ in excess of one. If the time period is years, I am thinking of a rate of return of, say, 3 per cent per annum. My prior for ρ thus peaks around $1\cdot 03$ and falls continuously and rather steeply on either side, steeper to the left than to the right. Again, and for essentially the same reasons, I suspect that Sims' prior will do better for my mean, and Philips' will do better for my mode.

Note that I have not bought Phillips' argument in favour of the Jeffreys' prior. This prior is only a convention that solves the following communication problem: suppose I use the parameter ρ and you use the parameter $\theta = \rho^3$. How can we be sure we have the same posterior distributions after transforming to a common parameter? One answer is to use Jeffreys' method for forming a prior, you for θ and me for ρ. Another way to achieve this invariance is for us to do a little communicating. Let us make sure that your prior for θ is equivalent to my prior for ρ. Then we will have the same posterior distributions for whatever parameter we choose. Expressed this way, the Jeffreys' prior seems rather odd, doesn't it? It is a solution to a rather unusual problem; and the solution has some rather disturbing properties, in this case the dependence of the prior on the sample size.

Note also that I refuse to use the words 'objective' and 'ignorance'. It seems to me clear that both Sims' prior and Phillips' prior embody information about the probable values of the parameters. In my opinion this traditional Bayesian model with a fully defined prior distribution is incapable of characterizing a state of ignorance.[1] The word 'ignorance' is best interpreted in terms of the range of equally good approximations to your current state of mind. If there is one and only one prior distribution that you are willing to maintain, then you are not in a state of ignorance. Confusion and ignorance occur when there is a large range of

[1] For further discussion see, for example, Leamer (1978, 1982, 1986, 1987, 1989, and Leamer and Leonard, 1983).

alternative distributions that are adequate representations of your state of mind. Thus ignorance arises in the third step of the programme that I have listed above: mapping the alternative priors into their corresponding inferences.

Now let's move on to this third step which characterizes the sensitivity of the inferences to the choice of prior. I may be putting words into Sims' mouth, but I suspect that he uses the flat prior not because it is a perfect description of his prior state of mind but rather because he imagines that a more careful definition of his prior wouldn't matter enough to make the effort worthwhile. That would clearly have been my opinion. In this context the sample information seems to me likely to be so substantial compared with whatever prior information I may have, that I might as well use the flat prior. The surprise in Phillips' paper is that the prior matters. More accurately, his prior leads to inferences that are substantially different from Sims'. The fact that there is some prior that leads to a substantially different posterior can be no surprise—this is always true. The surprise is that there is a 'reasonable' prior that leads to a substantially different inference. But wait a moment. Is Phillips' prior reasonable? I am not so sure. It favours high values of ρ by a very large amount. I wonder if the prior were defined over a finite interval, say $0 \leqslant \rho \leqslant 1 \cdot 5$, what would the cumulative distribution look like. I suspect that the prior probability of $\rho \leqslant 1$ would be very small. If it were 10^{-10}, Phillips' finding of sensitivity would not be disturbing to me, since I would think his distribution too unusual to be taken seriously. Nonetheless, the effect that this paper has had on me is to make me much more concerned about the choice of priors in time-series settings.

Last, with great enthusiasm I offer a hearty welcome to Peter Phillips into the Bayesian religion. It is with near ecstasy that I embrace him as a convert into the tiny sect which emphasizes the sacrament of sensitivity analysis. Others, of course, are welcome.

REFERENCES

Leamer, E. E. (1978), *Specification Searches*, John Wiley and Sons, New York.

Leamer, E. E. (1982), 'Sets of posterior means with bounded variance priors', *Econometrica*, **50**, 725–736.

Leamer, E. E., and H. B. Leonard (1983), 'Reporting the fragility of regression estimates', *Review of Economics and Statistics*, **LXV**, 306–317.

Leamer, E. E. (1986), 'Bid–ask spreads for subjective probabilities', in P. K. Goel and A. Zellner (eds), *Bayesian Inference and Decision Techniques, Essays in Honor of Bruno de Finetti*, North-Holland, Amsterdam, pp. 217–232.

Leamer, E. E. (1987), 'Econometric metaphors', in T. Bewley (ed.), *Advances in Econometrics, Fifth World Congress*, Cambridge University Press, Cambridge, pp. 1–29.

Leamer, E. E. (1989), 'Planning, criticism and revision', *Journal of Applied Econometrics*, **4**, S5–S27, (*supplement*).