REPLY

Peter E. Kennedy

In the first footnote of my paper, I advised that it was not meant to be a contribution 'to the scholarly debate on econometric methodology, a realm in which sinning is not allowed', but rather that it offered 'commentary and advice relevant to the real world of applied econometrics, where sinning is ubiquitous'. One reason I adopted this theme is captured neatly by Stewart and Gill (1998, p. 382): 'The literature on methodology is very revealing, but also somewhat intimidating for the applied economist, who is left with the feeling that whatever he or she may do, there will inevitably be some departure from a currently fashionable model of research activity'. My paper tried to provide a diversion from the 'intimidating' econometric methodology literature, by offering rules for bounding the inevitable 'departures'. Two of the commentators recognized my paper's contribution, but one did not.

David Hendry has misunderstood the point of my 'sinning' paper. The main premise of my paper is that it is an unavoidable fact of life that applied econometricians will sin. Under this circumstance, the paper asks what rules should be taught to students to bound this sinning? As evidenced by the title and content of his reply, Hendry instead wants to stop the sinning altogether, and has used his reply to promote his own view of how sinning can be avoided and to advertise his own software. I suggest two possible explanations for Hendry's reaction.

First, Hendry's misunderstanding may be deliberate. Hendry may perceive himself as a Copernicus designated to pass truth on to the world, and to do so avails himself of every opportunity to deliver a sermon, regardless of the circumstances. As is typical of Hendry, and consistent with his respected status in the econometrics profession, his sermon is a good one. He lays out his methodological view clearly, supplies new references, raises provocative questions, criticizes sharply, and champions his cause with conviction. His main concern relates to rule 5 on simplicity and rule 7 on data mining, which contain comments on the general-to-specific (top down) versus specific-to-general (bottom up) approaches to specification. The essence of my recommendation is that researchers should follow a 'sinning combination' of top down and bottom up.[1] A general specification is tested down and then checked for adequacy. If, to use Hendry's words, new information has been learned, a different general model is specified and the process is repeated. It seems to me that this is what Hendry does

0950-0804/02/04 0615–06 JOURNAL OF ECONOMIC SURVEYS Vol. 16, No. 4

in his own work; otherwise, he would have to argue that his methodology is based on 'think of the correct answer at the start', so as to begin with the right general model.

A second view of Hendry's misunderstanding is that he has lost touch with reality and is better compared to Quixote than to Copernicus. Cloistered in Oxford, he is sure to encounter only students who are several standard errors above the norm, the very few who are capable of following his methodology without falling into sin. Evidence for this view comes from Hendry himself (Hendry, 1999, p. 359): 'It is surprising that so many of the published recommended practices were not reflected in the study by Siegert'. Only someone unaware of how applied econometrics is actually done by the vast majority of the economics profession would be surprised by such sinning. This view of Hendry's misunderstanding is consistent with the theme of my paper — instructors of econometrics do not have a good sense of how much sinning takes place out in the real world of econometric analysis, of the fact that it is impossible to prevent this sinning, and of what can be done to alleviate it.

In contrast to Hendry, both Jan Magnus and Ron Smith have recognized the message of my paper and have reacted in the spirit in which it was intended. Both add several additional rules/questions/examples/insights that offer readers more wisdom on doing applied econometrics, and so are welcome additions to my paper. Magnus's comment that it is somewhat dangerous to write down these rules struck a sensitive chord; I was well aware of this danger when I embarked on this paper, but decided that drawing the profession's attention to these rules was worth suffering the consequences of annoying econometric theorists. Smith has highlighted this by noting that it is a mistake to teach applied econometrics as if it were a continuation of theoretical econometrics. Indeed, a major inspiration for writing my paper was the fact that so many people told me that their 'applied econometrics' course was really an econometric theory course with some examples and an 'applied paper' requirement. Further, Smith has correctly identified the level at which instruction is delinquent by noting that 'A lot of learning applied econometrics is learning how to stop making very simple mistakes'.

I must thank Smith for identifying a major shortcoming in my paper, namely that I failed to explain adequately what I meant by my recommendation that assignments be designed to 'force students to fend for themselves in real-world contexts, with vague general instructions, rather than specific step-by-step directions telling them what to do'. Although I followed that recommendation with a warning that 'These assignments are not easy to design — they must be focussed on specific issues that should become apparent to students as they investigate the data, and they must have model answers showing the steps that students should have taken and the results that they should have uncovered', I should have been more specific about how that should be done to avoid Smith's experience that the students hate such assignments.[2]

Smith is correct that in general students hate assignments that are completely vague. Give them thirty years of data on money, income, and the interest rate and ask them to run with them and they will surely hate you, particularly if it is early

in the course when they have not yet become accustomed to dealing with the difficulties of real world data. But this is not what I had in mind. My concern was with the type of assignment found in the otherwise-excellent applied econometrics book by Berndt (1991), where students are led through the steps of investigating a specific issue, rather than being told to 'investigate such and such an issue'. The result is that students learn how to do some mechanical things, but miss the learning that comes from thinking out for themselves the steps that need to be taken to analyze some issue.

I also agree with Smith that independent empirical projects are good learning experiences, but to me this should be the capstone assignment for a course, not the only assignment. In the earlier stages of my applied econometrics course I ask students to work on short assignments that are vague in the sense that the students are not provided step-by-step instructions, but sharply defined in the sense that they are given some data and asked to address a specific, straightforward question. The intention is to produce an assignment that is a good learning experience, capable of being completed in just a few hours. In addressing this question they could discover one or more 'problems' along the way, such as a 'wrong' sign, or data peculiarities. Recognizing and dealing with these problems can provide an extra learning experience, beyond the mechanics of doing the estimation. Three 'elementary' examples may make this clearer.

Example 1. For their group undergraduate project, three students in Professor K's class went to three different shopping malls, on the same day, to collect data from retail outlets. They each returned with eleven observations on a variable y and a variable x, the details of which are not important. They plan to lump these data together to produce a sample of size 33 and regress y on x as the first stage of their project. They asked Professor K if this was OK and he mumbled something about so long as you test for pooling. They were too embarrassed to ask what pooling meant, and have come to you for help. They have already put their data together, in (data source). This file consists of two columns of data, with no headings, each with 33 observations. The first column contains the x observations, and the second column the y observations. The first 11 observations were obtained by the first student, the next 11 by the second student, and the last 11 by the third student.

(a) Explain what is meant by testing for pooling, and how to test the hypothesis that these data can be pooled. Report the value of the test statistic, its degrees of freedom and its p value. On the basis of what you have seen so far, do you advise them to conclude that these data can be pooled? Explain why or why not.
(b) One of these students is a pain in the ass, and claims that it is surely the case that the intercepts differ from mall to mall because one mall is from a wealthy area of town, one from a poor area, and the third from a middle-income area. He asks you to explain how to test if the slopes are the same, supposing all the intercepts are *not* equal. Explain how to do this test, report the test statistic, its degrees of freedom, and its p value.

In this example students are asked to extend the traditional Chow test in two ways — to more than two regimes, and to a subset of the coefficients. They are not led through the steps of doing this, but the task is straightforward. The data are from Anscombe (1973) and demonstrate a case in which failure to look at the data will lead one to false conclusions based on formal tests. More advanced students could be expected to worry about error variance equivalence.

Example 2. A US national newsmagazine has reported that if academic performance (as measured by SAT scores) is regressed on expenditures on education, one discovers that the more money the government spends, the less students learn! The national teacher's union is furious about this and has hired Professor K to check this out — the union claims that 'there must be something wrong with this regression — they must have made a mistake'. Professor K instructed his research assistant, Robin, to find some data to allow him to investigate if US states that spend more on schooling produce students who score higher on the SAT (scholastic aptitude test), a test that many students take to enhance their chances of admission to the college of their choice.

Robin was very diligent, locating data in the 1997 Digest of Education Statistics, an annual publication of the U.S. Department of Education. She pulled out from its website a chunk of data that she thinks should contain what Professor K needs (expenditure per pupil and total SAT score) and has put it into (data source). Unfortunately, Robin became homesick for her boyfriend on the other side of the world and disappeared without a trace, leaving behind only the following cryptic explanation of the 1994 data on the 50 states (data description omitted). You have been hired as Robin's replacement and asked to supply a brief writeup of your investigation into this issue, so that Professor K can charge the teacher's union a large fee without doing any work. Produce this writeup.

In this example there is a 'wrong' sign. The correct sign is easily found by putting in one of the additional explanatory variables available in the data. The key part of the assignment writeup is to explain why the wrong sign was obtained. This assignment is based on Guber (2000).

Example 3. Professor K is in jail. According to the police, Professor K and his mistress, Nancy, got drunk at a pub and the two of them, plus her unleashed dog Oliver (the very same dog as in Murray, 1994, *American Statistician* 48, pp. 37–39) staggered off into the snowy night. Some time thereafter, a crime (the details of which need not concern us) was committed, involving the infamous Oliver, aided and abetted (according to the police) by Professor K. Professor K claims that he and Oliver were not on good terms and so it is ridiculous for anyone to conclude that he (Professor K) was in any way responsible for the crime. In (data source) are the time paths of these three, measured by the police (from footprints in the snow) as deviations from the straight line route (across an open field) from the pub to Nancy's condo, for their first 400 steps, after which the data become unreliable because of the fight associated with the crime. With his one permitted phone call from jail, Professor K calls you, his research assistant, and asks you to analyze these data to determine what can be concluded from them regarding the relationship among him, Nancy and Oliver. He claims that he was just following

Nancy (for fear of getting lost) and was avoiding Oliver (because he and Oliver have never been friends — the key to Professor K's protestations of innocence to the police), and hopes that the data will support his story. Analyze these data to produce a clear, concise document that Professor K can use in court to defend himself.

This assignment is based on Smith and Harrison (1995); data can easily be generated by the instructor. In this example the student is not told that s/he should test for unit roots, then test for cointegration, then estimate the cointegrating relationships and interpret them. But the assignment is nonetheless straightforward.

Although two of the three examples above use fabricated data, my preference is to use actual data, but as I hope I have made clear above, fashion the assignment so that the student is faced with a specific task that can be accomplished in a reasonable length of time. Complementary to these types of assignments are 'referee-report' assignments that ask the student to prepare a referee report on a paper containing at least one substantive econometric error. Sometimes good examples can be found in published form, such as Datta, Iskandar-Datta, and Patel (1999). More often, though, examples come from requests the instructor has had to write referee reports on applied econometric work. In my own experience (influenced heavily by a dozen years editing the research section of the *Journal of Economic Education*), such examples abound, and indeed this is one of the main reasons why I wrote the 'sinning' paper in the first place — why do so many economics PhDs make such grievous errors in elementary applied econometric work?

Notes

1. This is not a novel suggestion. Hansen (1996, p. 1411), for example, writes: 'Indeed, it is easy to see that it is impossible to implement the general to specific approach fully. This would require an enormously complex exercise, with a complete model of the joint distribution of all variables, allowing for nonlinearities, heteroscedasticity, coefficient drift and non-Gaussian errors. It is clear that this would be too costly in terms of parameterization. The only practical solution is to mix-and-match the general to specific and specific to general methods'.

References

Anscombe, F. J. (1973) Graphs in statistical analysis. *American Statistician* 27, 1, 17–21.

Berndt, E. R. (1991) *The Practice of Econometrics: Classic and Contemporary*. Reading, MA: Addison-Wesley.

Datta, S., Iskandar-Datta, M. and Patel, A. (1999) Bank monitoring and the pricing of corporate public debt. *Journal of Financial Economics*, 51, 435–449.

Guber, D. C. (1999) Getting What You Pay For: The debate over equity in public school expenditures. *Journal of Statistics Education*, 7, 2. (www.amstat.org/publications/jse/secure/v7n2/datasets.guber.cfm).

Hansen, B. E. (1996) Methodology: alchemy or science? *Economic Journal*, 106, 1398–1413.

Hendry, D. F. (1999) An econometric analysis of US food expenditure 1931–1989. In J.

Magnus and M. Morgan (eds), *Methodology and Tacit Knowledge: Two Experiments in Econometrics* (pp. 341–61). Chichester, UK: Wiley.
Smith, A. and Harrison, R. (1995) A drunk, her dog and a boyfriend: an illustration of multiple cointegration and error correction. University of Canterbury Discussion Paper 9505.
Stewart, J. and L., Gill (1998) *Econometrics*, 2nd edition. London: Prentice-Hall Europe.