Author Manuscript

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/1475-4932.12406](https://doi.org/10.1111/1475-4932.12406)

This article is protected by copyright. All rights reserved

Introduction to the Theory of Econometrics, by Jan R. Magnus (VU University Press, Amsterdam, 2017), pp. xvi+122.

This book is a write-up of the notes used by the author “to teach a course in econometric theory for second-year undergraduate students,” students whose first year had included “some calculus, matrix algebra, and statistics, but no econometrics.” The first thing one notices about this book is its size, or rather its lack thereof. This slender volume weighs in at a mere 122 pages. If one ignores the two appendices (one on matrix algebra and one on statistics), the index, and the list of further readings, then the main body of the text is comprised of a mere 93 pages of text.¹ Not surprisingly, almost every page is dense with ideas. That said, the style of writing is invariably elegant and lucid. There is also a liberal dose of whimsy throughout the text which helps to keep the presentation fresh and inviting. It is a book that wants to be read rather than one that needs to be worked through and there are rewards aplenty for any who take the time to do so.

To my mind this is a book of two halves. The first half, comprised of the first three chapters, is relatively straight-forward, covering fundamental concepts of the linear multiple regression model (LMRM) $y = X\beta + u$, where y is the data to be modelled, X a matrix of observations on the explanatory variables, β a vector of unknown coefficients, and u a vector of disturbances. The second half of the book deals with more advanced topics, still focussed on the LMRM, and the material is noticeably more challenging than that of the earlier chapters.

The first chapter, entitled Approximation, introduces the method of least squares, which yields an approximation to the unknown β . Wonderfully, this is done in the complete absence of any of the probabilistic complications that so clutter, typically erroneously at that, many treatments of what is a purely numerical procedure. This chapter also includes a brief discussion of least absolute deviations, the primary purpose of which is to illustrate the possibility of other criteria being adopted when ‘fitting’ a curve to data. Notions of goodness of fit and the consequences of over- and under-fitting models are also introduced here.

Chapter 2 explores Estimation. Here we encounter the usual ideal conditions for least squares ‘estimation’ and explore their implications for the sampling properties of $\hat{\beta}$, the least squares approximation to β . There is also discussion of prediction, an exploration of the consequences of imposing linear restrictions during estimation, the use of mean squared errors as a basis for comparison of estimators, and a discussion of the relative importance of significance (precision of estimation) versus importance of parameter estimates. What is done is done with style and grace but the treatment is disappointingly conventional. The statistic is taken as given and then we make some assumptions, realistic or otherwise, under which the statistic $\hat{\beta}$ is shown to have certain properties.

Far and away the most interesting aspect of this chapter is the discussion of s^2 , the unbiased estimator of the disturbance variance, σ^2 . This discussion is noteworthy as much for what it doesn’t say as for what it does. Most importantly, there

¹Strictly, it is even less than this small number, because chapters open on right, or odd numbered pages, which occasionally results in a preceding blank page. Moreover, each chapter is prefaced by a full-page cartoon.

is no suggestion that s^2 has anything to do with the least squares ‘estimation’ of β which, of course, it doesn’t as least squares is completely silent on any notion of variance. Rather, with a majestic sleight of hand, it is argued that variances are quadratic concepts and so it makes sense to base estimation of σ^2 on a quadratic function of the data. In particular, s^2 results when one chooses a quadratic function of the form $s^2 = y' Ay$, for some matrix A , where A is chosen such that s^2 is unbiased for σ^2 . Implicitly, we have constructed the estimator to have certain desired sampling properties. This is an important principle. By extension, this principle suggests that we should make assumptions reflecting the data generating process and then try to construct estimators with particular sampling properties, *given these assumptions*. It is perhaps surprising that the analysis of $\hat{\beta}$ was not in a similar vein given that, to the best of my knowledge, the place where this principle is most clearly espoused is one of the author’s earlier works (Magnus and Neudecker, 1988). In any event, even if this principle of estimation was not championed in this text, the discussion does illustrate just how arbitrary a thing s^2 really is. Of course, in an ideal world the argument might have been made that its most truly desirable property is that, under ideal conditions, t-statistics based on s^2 have t distributions, although I happily concede that to be, perhaps, a little obscure for second-year undergraduates, the target audience.

From Estimation, the text moves on to Inference. We see the usual treatments of confidence intervals, t and F tests, and prediction intervals. The chapter ends with a discussion of pre-testing and model averaging. These latter two are big topics to go to the trouble of introducing without really using and so it is a little hard to discern the reasoning behind their inclusion. Perhaps, it is nothing more profound than the belief that it is better that these concepts be introduced than not.

Chapter 4 is concerned with Maximum Likelihood. Here the level of the material becomes noticeably more challenging. The discussion is kept within the context of the LMRM, and so the least squares solution, $\hat{\beta}$, is also the maximum likelihood estimator (mle) for β , although there is now an mle for σ^2 that is (i) an integral outcome of the estimation process, in stark contrast to s^2 , and (ii) different from s^2 . We learn about the usual properties of scores and Hessians, Fisher’s information matrix, and the Cramér-Rao lower bound (CRLB). There is also a discussion of Wald, Lagrange multiplier, and likelihood ratio tests of linear restrictions when the disturbance variance is both known and unknown (which lead variously to equality of the three statistics and a numerical inequality between them, respectively).

The penultimate chapter deals with Asymptotics. Skipping quickly over notions of limits, probability limits, Slutsky’s theorem, and the Markov and Chebyshev inequalities, we obtain simple proofs of convergence in probability which are then used to establish the consistency of $\hat{\beta}$ and σ^2 . Next, we are introduced to the notion of an asymptotic distribution and, in particular, the Lindeberg-Lévy central limit theorem. Again, the primary illustration is in respect of $\hat{\beta}$ and σ^2 . The asymptotic version of the CRLB is used to illustrate notions of asymptotic efficiency. Finally, there is a discussion of the asymptotic properties of the mle, although this last is much less satisfactory than the asymptotic analysis of $\hat{\beta}$ and σ^2 . The main problem is that Lindeberg-Lévy isn’t enough and that, in order to establish the asymptotic normality of the score and hence of the mle, we need a stronger central limit theorem

than we have, one due Lindeberg-Feller, and so the proof is left hanging. If it were me, I think I would either have omitted the asymptotic analysis of the mle or have at least have provided the requisite limit theorem as part of the statistical appendix. Very pleasing is the fact that the chapter ends with a sensible discussion of what asymptotics are about. Specifically, asymptotics are not a form of truth, as so many texts implicitly assert, rather they are a basis for approximation; approximations that sometimes work well and sometimes don't. Blind faith in asymptotics can be every bit as objectionable as the strong assumptions underlying exact results that so many happily rail against.

In the final chapter the author addresses the fact that, until now, the analysis has treated all regressors X as fixed. Relaxing that assumption leads to notions of Conditioning, which is the title of this final chapter. Here the Laws of Total Expectation and Variance are applied in the context of the linear model, leading to a regression function (or conditional expectation) which is shown to be the best predictor of y given X . The discussion culminates in a discussion of mean-independence and its place relative to uncorrelatedness (a weak concept) and independence (a strong concept).

When evaluating a text like this, one must assess it relative to its target audience, which in this case are mathematically well-prepared. There are clearly some intellectually challenging concepts covered here and so, without this preparation, I doubt whether many students would find the material introductory at all. That said, there is no consideration here of any form of non-linearity at all. Hence, excluded from consideration are dummy variables, binary choice models, all forms of dynamics, including autocorrelation and heteroskedasticity, etc. Consequently, this book is clearly not a substitute for a text like Stock and Watson (2010, $\approx$ 850 pages), Griffiths *et al.* (2011, $\approx$ 800 pages), Wooldridge (2013, $\approx$ 900 pages), or any of the other common choices in this market, which all provide a much more comprehensive coverage of material than does the book under review. It is this brevity of both pages and coverage, focussed on econometric theory rather than the more commonly encountered exhaustive coverage, that distinguishes this text in the market place.

Above I have attempted to outline those areas about which I had some reservation, and in truth my reservations are both few and minor, but have said less about why I like this text so much. It is beautifully written by a true master of his craft. Although terse, the discussion is always insightful and uncluttered by extraneous detail. My sense is that, for better or worse, there are fewer subjects these days than once there were where a book like this might be used as the primary text. However, as a supplement to many existing texts, one could do a lot worse. Indeed, this is one area where the text's brevity will prove to be a real bonus. Moreover, the clarity of presentation will make it a useful resource when preparing lecture notes, even if the primary text is something else. Highly recommended.

CHRISTOPHER L SKEELS
The University of Melbourne

REFERENCES

- Griffiths, W. E., Hill, R. C. and Lim, G. C. (2011), *Principles of Econometrics*, John Wiley & Sons, Inc., New York, 4th edn.
- Magnus, J. R. and Neudecker, H. (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons Ltd, Chichester.
- Stock, J. H. and Watson, M. W. (2010), *Introduction to Econometrics*, Addison-Wesley, Boston, 3rd edn.
- Wooldridge, J. M. (2013), *Introductory Econometrics: A Modern Approach*, South Western, Mason, Ohio, 5th edn.

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Skeels, CL

Title:

Introduction to the Theory of Econometrics

Date:

2018-06-01

Citation:

Skeels, C. L. (2018). Introduction to the Theory of Econometrics. *ECONOMIC RECORD*, 94 (305), pp.209-211. <https://doi.org/10.1111/1475-4932.12406>.

Persistent Link:

<http://hdl.handle.net/11343/284093>