Journal of Econometrics 177 (2013) 75–84

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Finite-sample exact tests for linear regressions with bounded dependent variables

Olivier Gossner[a,b,*], Karl H. Schlag[c]

[a] Paris School of Economics (CNRS, UMR 8545), Paris, France
[b] London School of Economics, London, United Kingdom
[c] Department of Economics, University of Vienna, Austria

ARTICLE INFO

Article history:
Received 20 June 2012
Received in revised form
26 April 2013
Accepted 29 June 2013
Available online 13 July 2013

JEL classification:
C12

Keywords:
Nonparametric linear regression
Exact test
Heteroskedasticity

ABSTRACT

We introduce tests for finite-sample linear regressions with heteroskedastic errors. The tests are exact, i.e., they have guaranteed type I error probabilities when bounds are known on the range of the dependent variable, without any assumptions about the noise structure. We provide upper bounds on probability of type II errors, and apply the tests to empirical data.

1. Introduction

The fundamental goal of hypothesis testing, as set by Neyman and Pearson (1930), is the minimization of both type I and type II error probabilities. To cite Neyman and Pearson (1930, p. 100): (1) *we must be able to reduce the chance of rejecting a true hypothesis to as low a value as desired*; (2) *the test must be so devised that it will reject the hypothesis tested when it is likely to be false*.

In the usual model of linear regressions, how well are these goals achieved? When error terms are normally distributed and homoskedastic, the classical test[1] has a type I error probability equal to the nominal level of the test. But error terms in real data almost never have a precisely normal distribution, let alone a homoskedastic one. For any given heteroskedastic noise structure, White (1980)'s robust test guarantees a type I error probability that approaches the nominal level when the sample size goes to infinity. But without restrictions on the (unknown) noise structure, and for any sample size, the probability of a type I error resulting from the use of White (1980)'s test can be as large as 1. In fact, this is a consequence of a general impossibility result due to Bahadur and Savage (1956) and Dufour (2003) that shows that no meaningful test can be constructed in which the probability of a type I error is guaranteed to be less than 1.

The use of statistical tools in situations where the underlying distributional assumptions are not satisfied can have catastrophic consequences. Practitioners can be led to greatly underestimate the probability of certain outcomes, and remain unprepared to those outcomes while thinking they are safe. This is what must have happened to David Viniar, CFO of Goldman Sachs, who declared in August 2007 about the financial crisis: "We were seeing things that were 25-standard deviation moves, several days in a row" (quoted in Larsen, 2007). Since the probability of a 25-standard deviation event under the normal distribution is less than 1 over 10^{137}, we can safely conclude that the distributional assumptions used by Viniar and his colleagues were not satisfied.

In this paper our message is a positive one. We identify an important class of statistical problems where the negative conclusions from Bahadur and Savage (1956) and Dufour (2003) do not apply, and we introduce tests with guaranteed upper bounds on type I and type II errors for this class of problems. The tests are exact in the sense that they guarantee a type I error probability

* Correspondence to: Paris School of Economics, 48 Boulevard Jourdan, 75014 Paris, France.
E-mail addresses: ogossner@gmail.com (O. Gossner), karl.schlag@univie.ac.at (K.H. Schlag).

[1] In the remaining of the paper, "classical test" refers to the classical OLS t-test assuming normal and homoskedastic errors.

0304-4076/$ – see front matter
http://dx.doi.org/10.1016/j.jeconom.2013.06.003

below the nominal level independently of the error structure.[2] We also implement our these tests in practical numerical examples.

The class of problems we consider is the class in which a bound on the dependent variable is known. This condition is satisfied in a large range of applications. For instance, it is warranted by the very nature of the endogenous variable (e.g., proportions, success or failure, test scores) in 43 of the 75 papers using linear models published in 2011 in the *American Economic Review*. It should be noted that even under the boundedness condition, the existence of exact tests was previously an open problem. Previous exact tests were derived by Schlag (2006, 2008a) for the mean of a random variable and for the slope of a simple linear regression. Exact tests for linear regressions under the alternative assumption that error terms have median zero are developed by Dufour and Hallin (1993), Boldin et al. (1997), Chernozhukov et al. (2009), Coudin and Dufour (2009), Dufour and Taamouti (2010).

We refer to our two tests as the nonstandardized and the Bernoulli tests. We briefly summarize their constructions here. Each test relies on a linear combination of the dependent variables (such as in the OLS method) that is an unbiased estimator of the coefficient to be tested.

The nonstandardized test relies on inequalities due to Cantelli (1910), Hoeffding (1963), and Bhattacharyya (1987), as well as on the Berry–Esseen inequality (Berry, 1941; Esseen, 1942; Shevtsova, 2010), to bound the tail probabilities of the unbiased estimator. One challenge in the construction is to apply the Berry–Esseen inequality even though there is no lower bound on the variance of any of the error terms.

The Bernoulli test generalizes the methodology introduced by Schlag (2006, 2008b) for mean tests. Each term of the linear combination that constitutes the unbiased estimator is probabilistically transformed into a Bernoulli random variable. We then design a test for the mean of the family obtained using Hoeffding (1956)'s bound on the sum of independent Bernoulli random variables. This defines a randomized test, on which we then rely to construct a deterministic test.

We provide bounds on the probabilities of type II error of our tests. These bounds can be used to select – depending on the sample size and the realization of the exogenous variables – which of our tests is most appropriate. We also rely on these bounds to show that these tests have enough power for practical applications.

In two canonical numerical examples involving one covariate in addition to the constant, the bounds on the probability of type II errors show that the tests perform well even for small sample sizes (e.g., $n = 40$).

We implement our tests and compute confidence intervals using the empirical data from Duflo et al. (2011).[3] We compare the results relying on our test with the 95/ ones obtained using either the classical method and White's heteroskedastic robust method. The results show that, compared to the classical test or White's test, the losses of significance of our exact method are moderate, and the confidence intervals are in most cases augmented by a factor of no more than 50%.

The paper is organized as follows. Section 2 introduces the model. Sections 3 and 4 present the nonstandardized test and the Bernoulli test. In Section 5, we examine their efficiency using numerical examples. Section 7 shows an application of the tests to empirical data. The underlying data-generating process is discussed and extensions are discussed in Section 8. We conclude in Section 9. All proofs are presented in the Appendix.

2. Linear regression

We consider the standard linear regression model with random regressors, given by

$$Y_i = X_i\beta + \varepsilon_i, \quad i = 1, \ldots, n \tag{1}$$

where X_i is the i-th row of a random matrix $X \in \mathbb{R}^{n\times m}$ of independent variables, $\beta \in \mathbb{R}^m$ is the vector of unknown coefficients, and $\varepsilon \in \mathbb{R}^n$ is the random vector of errors. The *fixed regressor case* in which X is nonrandom and known ex-ante to the statistician is a special case. We assume (i) *strict exogeneity*: $\mathbf{E}(\varepsilon|X) = 0$ a.s., and (ii) almost surely no *multicolinearity*: X has rank m with probability 1. To keep the exposition simple, in most of the paper we also assume (iii) *conditional independence of errors*: $(\varepsilon_i)_i$ are independent conditional on X. Finally, we assume (iv) *bounded endogenous variable*[4]: there exist ω and ω' with $\omega < \omega'$ such that $P(Y_i \in [\omega, \omega']) = 1$ for $i = 1, \ldots, n$. In particular, (iv) implies that $X_i\beta \in [\omega, \omega']$ almost surely, and ensures the existence of all moments of ε_i for $i = 1, \ldots, n$. We assume that $\omega' = \omega + 1$; this is without loss of generality since we can reduce other cases to this one by dividing each side of Eq. (1) by $\omega' - \omega$. We relax (iii) and (iv) in Section 8. The assumptions (i)–(iv) are stronger than those of e.g. White (1980), but are sufficient to guarantee the existence of unbiased estimators (not just asymptotically so) of β. We do not make any further assumptions about the error terms, such as $\mathrm{Var}(\varepsilon_i) > 0$ or homoskedasticity.

We present two exact tests at the level of significance $\alpha > 0$ for the one-sided hypothesis $\mathrm{H}_0 : \beta_j \le \bar{\beta}_j$ against $\mathrm{H}_1 : \beta_j > \bar{\beta}_j$, where $\bar{\beta}_j \in \mathbb{R}$.[5] *Exact* means that the probability of a type I error of the test is proven to be at most α for any random vectors (X, ε) that satisfy (i)–(iv). In particular, bounds on the probabilities of type I errors are guaranteed for every given sample size.

Both tests have a type I error probability bounded by the nominal level conditional on the realization of X. This allows to combine both tests into another exact test according to the following procedure: given X, select the test for which our bounds guarantee a type II error probability below 0.5 for the largest range of the parameter to be tested. This is the procedure that we implement in our software and in the numerical applications.

3. The nonstandardized test

Assumption (ii) ensures the existence of $\tau_j \in \mathbb{R}^n$ such that $X'\tau_j = e_j$, where $e_{jj} = 1$ and $e_{jk} = 0$ for $k \ne j$. For such τ_j, $\hat{\beta}_j = \tau_j'Y$ is an unbiased estimate of β_j. One example of τ_j is the system of coefficients for which $\tau_j'Y$ is the OLS estimate of β_j. We present a test for a given such vector τ_j, and later discuss the choice of τ_j. We let $\|\ \|_\infty$ denote the supremum norm, and $\|\ \|$ denote the Euclidean norm, while Φ denotes the cumulative normal distribution.

Consider the functions defined for $\sigma > 0$, $t > 0$, and $\tau_j \in \mathbb{R}^n$:

$$\varphi_C(\sigma, t) = \frac{\sigma^2}{\sigma^2 + t^2}$$

---

[2] This is the same sense as in e.g. Yates (1934) and Dufour and Torres (2000).

[3] The software used to implement the tests is freely downloadable from the authors' webpages.

[4] Without any restriction on the support of Y, the possibility of very small or very large outcomes that occur with very small probability (fat tails) make it impossible to make any inference about $\mathbf{E}Y$ based on the observed values of Y, as shown by Bahadur and Savage (1956) when testing for means and by Dufour (2003) in linear regression analysis.

[5] Tests of $\mathrm{H}_0 : \beta_j \ge \bar{\beta}_j$, $\mathrm{H}_0 : \beta_j = \bar{\beta}_j$, and confidence intervals are derived easily, see Section 8.