Queen's University
School of Graduate Studies and Research
Department of Economics

Economics 850 **Econometrics I** **Fall, 2023**

Professor James MacKinnon

Answers to Final Examination

December 18, 2023. Time: 3 hours

Notes: The examination is in two parts. Please answer the only question in Part I and three (3) questions from Part II. Tables with some critical values of the χ^2 and Student's t distributions appear at the end of the examination.

Part I. Please answer the following question, which is worth 28% of the final mark.

1. Consider the linear regression model

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + u_i,$$

where there are 87 observations, the regressors are assumed to be exogenous, and the four β_j parameters are unknown.

a) Write down $\hat{\beta}_4$, the OLS estimate of β_4, using a compact and readable notation. **Hint:** You may wish to define some vectors and matrices.

b) Explain how you would estimate the standard error of $\hat{\beta}_4$ and obtain a 95% confidence interval for β_4 under the assumption that the u_i are independently and identically distributed. Would this interval be exact without making additional assumptions?

c) Now suppose that $\mathrm{E}(u_i^2) = \exp(\gamma_0 + \gamma_1 x_{3i})$, with the γ_j unknown. Would the confidence interval from part b) be valid, even asymptotically? Are there any restrictions on the γ_j under which it would be (asymptotically) valid? Are there circumstances in which it would be asymptotically valid even without restrictions on the γ_i?

d) Briefly explain two different ways to obtain the standard error of $\hat{\beta}_4$ under the assumptions of part c). How would you use each of these standard errors to form a confidence interval? Would these intervals be asymptotically valid? Would they be valid for the actual sample in this case? Which of the two intervals would you expect to be longer?

ANSWER [5, 7, 8 and 8 marks for the four parts]

a) Define $\boldsymbol{x}_4$ as the vector with typical element x_{4i}, $\boldsymbol{y}$ as the vector with typical element y_i, and $\boldsymbol{X}_1$ as the 87×3 matrix with $[1 \;\; x_{2i} \;\; x_{31}]$ in the i^{th} row. Then the regression becomes

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \beta_4\boldsymbol{x}_4 + \boldsymbol{u}.$$

Using the FWL Theorem,

$$\hat{\beta}_4 = \frac{\boldsymbol{x}_4^\top \boldsymbol{M}_1 \boldsymbol{y}}{\boldsymbol{x}_4^\top \boldsymbol{M}_1 \boldsymbol{x}_4},$$

where $\boldsymbol{M}_1$ projects orthogonally off $\boldsymbol{X}_1$.

b) Under the IID assumption, the usual estimate of the standard error of $\hat{\beta}_4$ is

$$s_4 = \left(\frac{\boldsymbol{y}^\top \boldsymbol{M}_{\boldsymbol{X}} \boldsymbol{y}/83}{\boldsymbol{x}_4^\top \boldsymbol{M}_1 \boldsymbol{x}_4}\right)^{1/2}.$$

The confidence interval is

$$[\hat{\beta}_4 - c_{.975}s_4, \;\; \hat{\beta}_4 + c_{.975}s_4],$$

where $c_{.975}$ is the .975 quantile of the $t(83)$ distribution. This interval is would not be exact without the additional assumption that the u_i are normally distributed.

c) In general, the interval from b) will not be valid, even asymptotically, because there is now heteroskedasticity. There is also serial dependence, but not serial correlation, unless the regressors are serially independent. It will, of course, be valid if $\gamma_1 = 0$, because then the u_i are homoskedastic. It will also be valid if there is no correlation, asymptotically, between $\boldsymbol{M}_1\boldsymbol{x}_4$ and $\exp(\gamma_0 + \gamma_1 x_{3i})$.

d) The obvious methods are HC_1, HC_2, and HC_3. They just need to propose two of them. For reference, they all have the form

$$(\boldsymbol{X}^\top\boldsymbol{X})^{-1}[\text{middle matrix}](\boldsymbol{X}^\top\boldsymbol{X})^{-1}.$$

The middle matrices are

$$\text{HC}_1\text{:} \quad \frac{87}{83}\sum_{i=1}^{87} \hat{u}_i^2 \boldsymbol{X}_i \boldsymbol{X}_i^\top$$

$$\text{HC}_2\text{:} \quad \sum_{i=1}^{87} \frac{\hat{u}_i^2}{1-h_i} \boldsymbol{X}_i \boldsymbol{X}_i^\top$$

$$\text{HC}_3\text{:} \quad \sum_{i=1}^{87} \frac{\hat{u}_i^2}{(1-h_i)^2} \boldsymbol{X}_i \boldsymbol{X}_i^\top$$

None of these is valid for the actual sample of 87 observations. They are all valid asymptotically. We can be confident that the HC_1 interval is the shortest and the HC_3 interval the longest.

Part II. Please answer three (3) of the following five (5) questions. Each question has four (4) parts and is worth 24% of the final mark.

2. Consider the linear regression model

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i,$$

which is to be estimated using a sample of 73 observations. The regressors x_{2i} and x_{3i} are assumed to be exogenous. You are interested in the parameter $\gamma \equiv \beta_2/\beta_3$.

a) Explain how you would obtain an estimate $\hat\gamma$ and an asymptotically valid standard error $s(\hat\gamma)$ analytically (i.e., without doing any simulations) under the assumption that $\mathrm{E}(u_i u_j) = 0$ for $i \neq j$ and $\mathrm{E}(u_i^2 \mid \boldsymbol{X}_i) = \sigma_i^2$. Do not waste time doing the algebra.

b) Explain how you would perform a bootstrap test of the hypothesis that $\gamma = 3$ under the assumptions of part a). Would your bootstrap DGP incorporate any restrictions? Explain precisely how it would work.

c) Explain how you would obtain a bootstrap standard error $s^*(\hat\gamma)$ under the assumptions of part a) and how you could use that standard error to form a 95% confidence interval. Would your bootstrap DGP incorporate any restrictions? Explain precisely how it would work.

d) Explain how you would construct a 95% studentized bootstrap confidence interval for γ under the assumptions of part a). Would your bootstrap DGP incorporate any restrictions? Explain precisely how it would work.

ANSWER [6 marks for each part]

a) We need to use the delta method here. Since $\gamma = \beta_2/\beta_3$, the delta-method standard error is

$$s(\hat\gamma) = \left(\begin{bmatrix} 1/\hat\beta_2 & -\hat\beta_2/\hat\beta_3^2 \end{bmatrix} \begin{bmatrix} \hat V_{22} & \hat V_{23} \\ \hat V_{23} & \hat V_{33} \end{bmatrix} \begin{bmatrix} 1/\hat\beta_2 \\ -\hat\beta_2/\hat\beta_3^2 \end{bmatrix} \right)^{-1/2}.$$

Here the $\hat V_{jk}$ are elements of any heteroskedasticity-robust variance matrix estimator. Students were told not to waste time on the algebra.

b) Estimate the model under the assumption that $\gamma = 3$. This may be done by regressing y_i on $3x_{2i} + x_{3i}$, which will yield estimates $\tilde\beta_1$, $\tilde\beta_3$, and $\tilde\beta_2 = 3\tilde\beta_3$. You can obtain wild bootstrap samples from

$$y_{ib}^* = \tilde\beta_1 + \tilde\beta_2 x_{2i} + \tilde\beta_3 x_{3i} + v_b^* \tilde u_i,$$

where the $\tilde{u}_i$ are the restricted residuals and the v_b^* are Rademacher random variates. You could also transform the $\tilde{u}_i$ as in HC_2 or HC_3 before doing this, but that is not needed.

Then generate B bootstrap samples, and use each of them to compute a bootstrap test statistic for the hypothesis that $\gamma = 3$. The actual and bootstrap test statistics are

$$\tau = \frac{\hat{\gamma} - 3}{s(\hat{\gamma})} \quad \text{and} \quad \tau_b^* = \frac{\hat{\gamma}_b^* - 3}{s(\hat{\gamma}_b^*)}.$$

Compute a bootstrap P value in the usual way. Choose B appropriately.

c) This time use an unrestricted bootstrap DGP. Do not bother to compute the $s(\hat{\gamma}_b^*)$. Use the $\hat{\gamma}_b$ to compute

$$\text{se}^*(\hat{\gamma}) = \left(\frac{1}{B-1} \sum_{b=1}^{B} (\hat{\gamma}_b^* - \bar{\gamma}^*)^2 \right)^{1/2}.$$

The bootstrap confidence interval is then

$$[\hat{\gamma} - c_{.975}(70)\text{se}^*(\hat{\gamma}), \;\; [\hat{\gamma} + c_{.975}(70)\text{se}^*(\hat{\gamma})],$$

where $c_{.975}(70)$ is the .975 quantile of the $t(70)$ distribution.

d) The studentized bootstrap interval is

$$[\hat{\gamma} - c^*_{.975} s(\hat{\gamma}), \;\; [\hat{\gamma} - c^*_{.025} s(\hat{\gamma})],$$

where the $c^*(\alpha)$ are quantiles of the bootstrap distribution of the τ_b^*, and $s(\hat{\gamma})$ is the delta-method standard error from part a).

3. Consider the linear regression model

$$y_i = \boldsymbol{X}_i \boldsymbol{\beta} + u_i,$$

where the u_i are independently but not identically distributed. Specifically, it is assumed that $\text{E}(u_i \,|\, \boldsymbol{X}_i) = 0$, and $\text{E}(u_i^2 \,|\, \boldsymbol{X}_i) = \sigma_i^2$.

a) Write the OLS estimator $\hat{\boldsymbol{\beta}}$ as a function of $\boldsymbol{\beta}_0$, the true value of $\boldsymbol{\beta}$, $\boldsymbol{X}$, and $\boldsymbol{u}$, where $\boldsymbol{X}$ has typical row $\boldsymbol{X}_i$, and $\boldsymbol{u}$ has typical element u_i. The expression you obtain should have two terms, only one of which is stochastic. Of what order in the sample size N is the stochastic term under standard assumptions? Explain.

b) Under the assumptions made so far, will $\hat{\boldsymbol{\beta}}$ be unbiased? Will it be consistent? What can you say about its asymptotic distribution? Do you need to make any additional assumptions to answer these questions? Explain briefly.

c) Suppose now that the u_i are not assumed to be independent. Of course, the $\boldsymbol{X}_i$ were never assumed to be independent. Do your answers to part b) change in any way as a result of these new assumptions?

d) Consider the special case in which

$$u_i = v + w_i, \quad v \sim \mathrm{N}(0, \omega^2), \;\; \mathrm{E}(w_i w_j) = 0 \;\text{ for all } i, j.$$

What can you say about the stochastic term in the expression for $\hat{\boldsymbol{\beta}}$ that you obtained in part a)? Is $\hat{\boldsymbol{\beta}}$ consistent in this case? Explain.

ANSWER [5, 7, 6, 6 marks]

a) Evidently

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}\boldsymbol{\beta}_0 + (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{u} \\ &= \boldsymbol{\beta}_0 + (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{u}\end{aligned}$$

Under standard assumptions, $(\boldsymbol{X}^\top\boldsymbol{X})^{-1} = O(1/N)$ and $\boldsymbol{X}^\top\boldsymbol{u} = O_p(N^{1/2})$. Therefore, the stochastic term is $O_p(N^{-1/2})$.

b) Under the assumptions made so far, $\hat{\boldsymbol{\beta}}$ will not be unbiased, because the regressors are assumed to be predetermined rather than exogenous. If we make the additional assumption that $\boldsymbol{S}_{\boldsymbol{X}^\top\boldsymbol{X}} = \operatorname{plim}(\boldsymbol{X}^\top\boldsymbol{X}/N)$ is a positive definite matrix, and put enough limits on the σ_i^2 so that $\operatorname{plim} N^{-1}\boldsymbol{X}^\top\boldsymbol{u} = \boldsymbol{0}$, then it will be consistent.

If in addition we assume that a central limit theorem applies to the vector $N^{-1/2}\boldsymbol{X}^\top\boldsymbol{u}$, then we can show that

$$N^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \overset{a}{\sim} \mathrm{N}(0, \boldsymbol{V}^\infty),$$

where $\boldsymbol{V}^\infty$ denotes the asymptotic variance matrix of $N^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)$. This matrix will have the form

$$\boldsymbol{S}_{\boldsymbol{X}^\top\boldsymbol{X}}^{-1}\left(\operatorname{plim}\frac{1}{N}\boldsymbol{X}^\top\boldsymbol{\Omega}\boldsymbol{X}\right)\boldsymbol{S}_{\boldsymbol{X}^\top\boldsymbol{X}}^{-1},$$

where $\boldsymbol{\Omega}$ is a diagonal matrix with σ_i^2 as the i^{th} diagonal element.

c) Without imposing some restrictions on the amount of dependence, we can no longer claim that $\hat{\boldsymbol{\beta}}$ is either consistent or asymptotically normal. Moreover, the appropriate normalizing factor may no longer be $N^{1/2}$.

The asymptotic variance matrix, if it exists, has the same form as above, but $\boldsymbol{\Omega}$ is no longer a diagonal matrix. With too much dependence, no CLT may apply and the matrix in the middle may not be $O(1)$.

d) This term is no longer $O_p(N^{-12}$. We can write it as

$$v(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\iota} + (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{w},$$

where $\boldsymbol{\iota}$ is an N-vector with every element equal to 1, and $\boldsymbol{w}$ is an N-vector with typical element w_i. Under reasonable assumptions, the second term here will be $O_p(N^{-1/2})$, but the first term is simply the random variable v times $(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\iota} = O(1)$. So it is $O_p(1)$. Thus $\hat{\boldsymbol{\beta}}$ is not consistent in this case.

4. You are given a sample of 20,781 stroke patients treated at 122 hospitals located in 21 communities, along with data on the characteristics of the patients and the hospitals. Let y_{hi} denote the number of days in hospital for patient i in hospital h, and let $\boldsymbol{X}_{hi}$ be a vector of patient and hospital characteristics. Some hospitals claim to base their operations on "patient-centred care," and others do not. The variable z_{hi} is equal to 1 if hospital h made such a claim when patient i was treated, and equal to 0 otherwise.

Suppose you regress the y_{hi} on the z_{hi}, the $\boldsymbol{X}_{hi}$, and a vector of community fixed effects. The coefficient on z is β. The OLS estimate of β, say $\hat{\beta}$, is -0.734, which suggests that stroke patients are discharged about 3/4 of a day earlier in hospitals with patient-centred care.

a) You compute the standard error of $\hat{\beta}$ in several different ways. The HC_1, HC_2, and HC_3 standard errors are 0.112, 0.116, and 0.122, respectively. When you cluster by hospital, the CV_1 standard error is 0.226, the CV_2 standard error is 0.245, and the CV_3 standard error is 0.268. Which of these standard errors seems most believable? Explain why, and also explain precisely how to compute this standard error.

b) In view of the number of hospitals, it may seem odd that the various cluster-robust standard errors vary so much. What is the most likely explanation for this? What quantities could you compute and study in order to provide evidence for this explanation?

c) After you submit your paper on patient-centred care to a journal, a referee argues that you should have clustered by community instead of by hospital. You therefore compute three more cluster-robust standard errors, this time clustering by community. What would you do if these were a little smaller than the ones you obtained originally? What would you do if they were a little larger? What would you do if they were 30% to 40% larger? Would it make sense to use another method of inference to verify the results when you cluster by community?

d) Suppose the cluster-robust standard error that you prefer when you cluster by community is 0.366. Use this standard error and the table at end of the examination to obtain a 95% confidence interval for β. Does this interval suggest that the hypothesis $\beta = 0$ can be rejected at the .05 level?

ANSWER [6 marks for each part]

a) The HC standard errors should obviously not be believed, because they are much smaller than the CV ones. While a formal test exists, there is no way this could possibly have happened by accident.

The CV_3 (or cluster jackknife) standard error is the most believable. If all hospitals were about the same size, it might be a bit too large, and CV_2 might be better. But that is surely not the case here.

To compute the CV_3 variance matrix, we have to compute 122 sets of delete-one-hospital estimates, say $\hat{\boldsymbol{\beta}}^{(g)}$. This apparently requires 122 OLS regressions, but they can be done quite cheaply by using matrices of sums of squares and cross-products. For simplicity, consider the jackknife variance of $\hat{\beta}$ only. It is

$$\frac{G-1}{G}\sum_{g=1}^{G}(\hat{\beta}^{(g)} - \hat{\beta})^2.$$

The CV_3 standard error is the square root of this quantity.

b) The most obvious reason for the three CV standard errors to differ so much is that the sizes of the hospitals, and hence the number of observations per hospital, probably vary a lot. It is also possible that only a modest proportion of hospitals either claim, or do not claim, to provide patient-centred care, which would create some high-leverage observations.

To investigate these issues, you should compute the N_g and the partial leverages for all hospitals. If either or both of them vary a lot, that would explain the differences and also suggest that it is safest to use CV_3.

c) If the standard errors clustered by community do not differ much from those clustered by hospital, I would stick with the latter. It would be a bad idea to switch if the former were smaller, but probably not harmful to switch if they were a little larger. However, if they are 30% larger (or even 5% larger), I would switch to clustering by community.

Because there are only 21 communities, I would use the wild cluster bootstrap to verify the results from the CV_3 standard error. The best method to use is probably WCR-S, but there is no harm trying WCR-C as well.

d) The confidence interval based on the $t(20)$ distribution is

$$[-.734 - .366 * 2.086,\ \ -.734 + .366 * 2.086] = [-1.4975,\ \ 0.029476].$$

Since this interval includes zero, we cannot reject the null at the .05 level.

5. Consider the linear simultaneous equations model

$$\boldsymbol{y}_1 = \beta \boldsymbol{y}_2 + \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{u}, \tag{1}$$

$$\boldsymbol{y}_2 = \boldsymbol{W}\boldsymbol{\pi} + \boldsymbol{v} = \boldsymbol{W}_1\boldsymbol{\pi}_1 + \boldsymbol{Z}\boldsymbol{\pi}_2 + \boldsymbol{v}, \tag{2}$$

where $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are n–vectors of observations on endogenous variables, $\boldsymbol{Z}$ is an $n \times k$ matrix of observations on exogenous or predetermined variables, $\boldsymbol{W} \equiv [\boldsymbol{W}_1 \;\; \boldsymbol{Z}]$ is an $n \times l$ matrix of observations on exogenous or predetermined variables, and $\boldsymbol{u}$ and $\boldsymbol{v}$ are n–vectors of homoskedastic disturbances. Assume that $\mathrm{E}(u_t v_s) = 0$ for $t \neq s$ and $\mathrm{E}(u_t v_t) = \rho \sigma_u \sigma_v$, where u_t and v_s are elements of $\boldsymbol{u}$ and $\boldsymbol{v}$, and σ_u^2 and σ_v^2 are their variances. The rest of the notation should be obvious.

a) In general, will the OLS estimate of β be unbiased? Will it be consistent? Is there any special case in which it will be consistent? Explain carefully.

b) Explain how to compute the generalized IV estimator $\hat{\beta}_{\mathrm{IV}}$. What can you say about the mean squared error of this estimator relative to the mean squared error of the OLS estimator $\hat{\beta}_{\mathrm{OLS}}$? What are the key features of the data-generating process (DGP) that explain the relationship between the two MSE values? Explain.

c) How many overidentifying restrictions are there? Explain how to test them using an asymptotic test. What would you conclude if $n = 350$, $k = 12$, $l = 15$, and the test statistic (in χ^2 form) were 6.46? Note that there is a table of the χ^2 distribution at the end of the examination.

d) Suppose the F statistic for $\boldsymbol{\pi}_1 = \boldsymbol{0}$ in equation (2) were 2.24. Would you expect the standard error of $\hat{\beta}_{\mathrm{IV}}$ to be similar to the standard error of $\hat{\beta}_{\mathrm{OLS}}$ in this case? Would you expect the finite-sample distribution of $\hat{\beta}_{\mathrm{IV}}$ to be well approximated by its asymptotic distribution? Would your answers to either or both of these questions change if the value of this F statistic were 17.58?

ANSWER [5, 7, 5, and 6 marks]

a) In general the the OLS estimate of β is biased and inconsistent, because $\boldsymbol{y}_2$ is correlated with $\boldsymbol{u}$ through the correlation between $\boldsymbol{u}$ and $\boldsymbol{v}$. However, it would be consistent if the u_i and v_i were contemporaneously uncorrelated, that is, if $\rho = 0$.

b) The GIV estimator is

$$\hat{\beta}_{\mathrm{IV}} = \frac{\boldsymbol{y}_2{}^{\top}\boldsymbol{P}_{\boldsymbol{W}}\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{y}_1}{\boldsymbol{y}_2{}^{\top}\boldsymbol{P}_{\boldsymbol{W}}\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{P}_{\boldsymbol{W}}\boldsymbol{y}_2}.$$

This is what the FWL Theorem gives us when applied to the second-stage regression

$$\boldsymbol{y}_1 = \boldsymbol{P}_{\boldsymbol{W}}\boldsymbol{y}_2 + \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{u}.$$

Note that $\boldsymbol{P_W Z} = \boldsymbol{Z}$, so we can also write

$$\hat{\beta}_{\text{IV}} = \frac{\boldsymbol{y}_2{}^{\top}(\boldsymbol{P_W} - \boldsymbol{P_Z})\boldsymbol{y}_1}{\boldsymbol{y}_2{}^{\top}(\boldsymbol{P_W} - \boldsymbol{P_Z})\boldsymbol{y}_2}.$$

The IV estimator may or may not have a larger MSE than the OLS estimator. It will certainly have a larger variance, but it should have a smaller bias. Since MSE is the sum of the squared bias and the variance, it could go either way. If the reduced-form regression (2) fits well, then the IV variance should not be too much larger than the OLS variance. Thus the IV MSE is likely to be lower in that case, unless the bias of the OLS estimator is small. This bias depends in large part on ρ. The further ρ is from zero, the greater the bias.

c) There are $l - k - 1$ overidentifying restrictions. You can test them using a Sargan test. The test statistic is

$$\frac{\hat{\boldsymbol{u}}_1^{\top}\boldsymbol{M_W}\hat{\boldsymbol{u}}_1}{\hat{\sigma}_1^2}.$$

Here $\hat{\sigma}_1^2$ is the IV estimate of the variance of the first equation. This test statistic is asymptotically distributed as $\chi^2(l - k - 1)$.

The test statistic 6.56 lies between the .05 and .025 critical values of the $\chi^2(2)$ distribution, so we would reject the null if we believe that the asymptotic theory underlying this test is reliable.

d) This is quite a small F statistic, which suggests that the reduced-form equation is quite weak. Therefore, the IV standard error is probably a lot bigger than the OLS standard error. Moreover, because 2.24 is far short of the Stock-Yogo critical values, standard asymptotic theory probably works badly.

In contrast, if the F statistic were 17.58, the reduced-form regression would be quite strong. Thus the IV standard error might not be a lot larger than the OLS one, and asymptotic inference would probably be quite reliable.

Table 1. Upper-Tail Critical Values of the χ^2 Distribution

D.F. / Level	.10	.05	.025	.01
1	2.706	3.841	5.024	6.635
2	4.605	5.991	7.378	9.210
3	6.251	7.815	9.348	11.345
4	7.779	9.488	11.143	13.277
5	9.236	11.070	12.833	15.086
6	10.645	12.592	14.449	16.812

Table 2. Two-Tail Critical Values of the Student's t Distribution

D.F. / Level	.10	.05	.025	.01
10	1.812	2.228	2.634	3.169
15	1.753	2.131	2.490	2.947
16	1.746	2.120	2.473	2.921
17	1.740	2.110	2.458	2.898
18	1.734	2.101	2.445	2.878
19	1.729	2.093	2.433	2.861
20	1.725	2.086	2.423	2.845
21	1.721	2.080	2.414	2.831