FINAL EXAM ECON 7801 SPRING 2001

ECONOMICS DEPARTMENT, UNIVERSITY OF UTAH

This is a closed book exam but you may bring one sheet with formulas with you; write your name on the formula sheet and submit it together with your exam. General formulas, please, not the answers to the questions in the class notes. (This same ruling will also apply to the field exam.)

The first two questions are repeats from the Midterm, and they are obligatory for everyone. Other than that you should pick and choose and do as many questions as you can in the 90 minutes time.

Problem 16. *2 points We are in the multiple regression model* $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ *with intercept, i.e.,* $\boldsymbol{X}$ *is such that there is a vector* $\boldsymbol{a}$ *with* $\boldsymbol{\iota} = \boldsymbol{X}\boldsymbol{a}$*. Define the row vector* $\bar{\boldsymbol{x}}^\top = \frac{1}{n}\boldsymbol{\iota}^\top\boldsymbol{X}$*, i.e., it has as its* j*th component the sample mean of the* j*th independent*

Date of exam Tuesday, May 1st, 9–10:30 am.

variable. Using the normal equations $\boldsymbol{X}^\top \boldsymbol{y} = \boldsymbol{X}^\top \boldsymbol{X}\hat{\boldsymbol{\beta}}$, show that $\bar{y} = \bar{\boldsymbol{x}}^\top \hat{\boldsymbol{\beta}}$ (i.e., the regression plane goes through the center of gravity of all data points).

Answer. Premultiply the normal equation by $\boldsymbol{a}^\top$ to get $\boldsymbol{\iota}^\top \boldsymbol{y} - \boldsymbol{\iota}^\top \boldsymbol{X}\hat{\boldsymbol{\beta}} = 0$. Premultiply by $1/n$ to get the result. □

Problem 17. *Assume $\hat{\hat{\boldsymbol{\beta}}}$ is the constrained least squares estimator subject to the constraint $\boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{o}$, and $\hat{\boldsymbol{\beta}}$ is the unconstrained least squares estimator.*

- **a.** *1 point With the usual notation $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}$ and $\hat{\hat{\boldsymbol{y}}} = \boldsymbol{X}\hat{\hat{\boldsymbol{\beta}}}$, show that*

$$\boldsymbol{y} = \hat{\hat{\boldsymbol{y}}} + (\hat{\boldsymbol{y}} - \hat{\hat{\boldsymbol{y}}}) + \hat{\boldsymbol{\varepsilon}} \tag{1}$$

Point out these vectors in the `reggeom` *simulation.*

Answer. In the `reggeom`-simulation, $\boldsymbol{y}$ is the purple line; $\boldsymbol{X}\hat{\hat{\boldsymbol{\beta}}}$ is the red line starting at the origin, one could also call it $\hat{\hat{\boldsymbol{y}}}$; $\boldsymbol{X}(\hat{\boldsymbol{\beta}} - \hat{\hat{\boldsymbol{\beta}}}) = \hat{\boldsymbol{y}} - \hat{\hat{\boldsymbol{y}}}$ is the light blue line, and $\hat{\boldsymbol{\varepsilon}}$ is the green line which does not start at the origin. In other words: if one projects $\boldsymbol{y}$ on a plane, and also on a line in that plane, and then connects the footpoints of these two projections, one obtains a zig-zag line with two right angles. □

- **b.** *4 points Show that in (1) the three vectors $\hat{\hat{\boldsymbol{y}}}$, $\hat{\boldsymbol{y}} - \hat{\hat{\boldsymbol{y}}}$, and $\hat{\boldsymbol{\varepsilon}}$ are orthogonal. You are allowed to use, without proof, the following formula for $\hat{\hat{\boldsymbol{\beta}}}$:*

$$\hat{\hat{\boldsymbol{\beta}}} = \hat{\boldsymbol{\beta}} - (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{R}^\top \big(\boldsymbol{R}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{R}^\top\big)^{-1}(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{u}). \tag{2}$$

Answer. One has to verify that the scalar products of the three vectors on the right hand side of (1) are zero. $\hat{\hat{\boldsymbol{y}}}^\top \hat{\boldsymbol{\varepsilon}} = \hat{\hat{\boldsymbol{\beta}}}^\top \boldsymbol{X}^\top \hat{\boldsymbol{\varepsilon}} = 0$ and $(\hat{\boldsymbol{y}} - \hat{\hat{\boldsymbol{y}}})^\top \hat{\boldsymbol{\varepsilon}} = (\hat{\boldsymbol{\beta}} - \hat{\hat{\boldsymbol{\beta}}})^\top \boldsymbol{X}^\top \hat{\boldsymbol{\varepsilon}} = 0$ follow from $\boldsymbol{X}^\top \hat{\boldsymbol{\varepsilon}} = \boldsymbol{o}$; geometrically on can simply say that $\hat{\boldsymbol{y}}$ and $\hat{\hat{\boldsymbol{y}}}$ are in the space spanned by the columns of $\boldsymbol{X}$, and $\hat{\boldsymbol{\varepsilon}}$ is orthogonal to that space. Finally, using (2) for $\hat{\boldsymbol{\beta}} - \hat{\hat{\boldsymbol{\beta}}}$,

$$
\begin{aligned}
\hat{\hat{\boldsymbol{y}}}^\top (\hat{\boldsymbol{y}} - \hat{\hat{\boldsymbol{y}}}) &= \hat{\hat{\boldsymbol{\beta}}}^\top \boldsymbol{X}^\top \boldsymbol{X} (\hat{\boldsymbol{\beta}} - \hat{\hat{\boldsymbol{\beta}}}) = \\
&= \hat{\hat{\boldsymbol{\beta}}}^\top \boldsymbol{X}^\top \boldsymbol{X} (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{R}^\top \big(\boldsymbol{R} (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{R}^\top \big)^{-1} \boldsymbol{R} \hat{\boldsymbol{\beta}} = \\
&= \hat{\hat{\boldsymbol{\beta}}}^\top \boldsymbol{R}^\top \big(\boldsymbol{R} (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{R}^\top \big)^{-1} \boldsymbol{R} \hat{\boldsymbol{\beta}} = 0
\end{aligned}
$$

because $\hat{\hat{\boldsymbol{\beta}}}$ satisfies the constraint $\boldsymbol{R} \hat{\hat{\boldsymbol{\beta}}} = \boldsymbol{o}$, hence $\hat{\hat{\boldsymbol{\beta}}}^\top \boldsymbol{R}^\top = \boldsymbol{o}^\top$. □

Problem 18. *3 points Explain the meanings of all the terms in the following identity*

$$
\hat{y}_i = (1 - h_{ii}) \hat{y}_i(i) + h_{ii} y_i \tag{3}
$$

and use that equation to explain why h_{ii} is called the "leverage" of the ith observation. Is every observation with high leverage also "influential" (in the sense that its removal would greatly change the regression estimates)?

Answer. $\hat{y}_i$ is the fitted value for the ith observation, i.e., it is the BLUE of η_i, of the expected value of the ith observation. It is a weighted average of two quantities: the actual observation y_i (which has η_i as expected value), and $\hat{y}_i(i)$, which is the BLUE of η_i based on all the other observations except the ith. The weight of the ith observation in this weighted average is called the "leverage" of the ith observation. The sum of all leverages is always k, the number of parameters in the regression. If the leverage of one individual point is much greater than k/n, then this point has much more influence on its own fitted value than one should expect just based on the number of observations,

Leverage is not the same as influence; if an observation has high leverage, but by accident the observed value y_i is very close to $\hat{y}_i(i)$, then removal of this observation will not change the regression results much. Leverage is potential influence. Leverage does not depend on any of the observations, one only needs the $\boldsymbol{X}$ matrix to compute it. □

Problem 19. *Prove the following facts about the diagonal elements of the so-called "hat matrix" $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top$, which has its name because $\boldsymbol{H}\boldsymbol{y} = \hat{\boldsymbol{y}}$, i.e., it puts the hat on $\boldsymbol{y}$.*

- **a.** *1 point $\boldsymbol{H}$ is a projection matrix, i.e., it is symmetric and idempotent.*

Answer. Symmetry follows from the laws for the transposes of products: $\boldsymbol{H}^\top = (\boldsymbol{A}\boldsymbol{B}\boldsymbol{C})^\top = \boldsymbol{C}^\top\boldsymbol{B}^\top\boldsymbol{A}^\top = \boldsymbol{H}$ where $\boldsymbol{A} = \boldsymbol{X}$, $\boldsymbol{B} = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}$ which is symmetric, and $\boldsymbol{C} = \boldsymbol{X}^\top$. Idempotency $\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top = \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top$. □

- **b.** *1 point Prove that a symmetric idempotent matrix is nonnegative definite.*

Answer. If $\boldsymbol{H}$ is symmetric and idempotent, then for arbitrary $\boldsymbol{g}$, $\boldsymbol{g}^\top\boldsymbol{H}\boldsymbol{g} = \boldsymbol{g}^\top\boldsymbol{H}^\top\boldsymbol{H}\boldsymbol{g} = \|\boldsymbol{H}\boldsymbol{g}\|^2 \geq 0$. But $\boldsymbol{g}^\top\boldsymbol{H}\boldsymbol{g} \geq 0$ for all $\boldsymbol{g}$ is the criterion which makes $\boldsymbol{H}$ nonnegative definite. □

- **c.** *2 points Show that*

$$0 \leq h_{ii} \leq 1 \tag{4}$$

Answer. If $\boldsymbol{e}_i$ is the vector with a 1 on the ith place and zeros everywhere else, then $\boldsymbol{e}_i^\top \boldsymbol{H}\boldsymbol{e}_i = h_{ii}$. From $\boldsymbol{H}$ nonnegative definite follows therefore that $h_{ii} \geq 0$. $h_{ii} \leq 1$ follows because $\boldsymbol{I} - \boldsymbol{H}$ is symmetric and idempotent (and therefore nonnegative definite) as well: it is the projection on the orthogonal complement. □

• **d.** *2 points Show: the average value of the* h_{ii} *is* $\sum h_{ii}/n = k/n$, *where* k *is the number of columns of* $\boldsymbol{X}$. *(Hint: for this you must compute the trace* $\operatorname{tr}\boldsymbol{H}$.*)*

Answer. The average can be written as

$$\frac{1}{n}\operatorname{tr}(\boldsymbol{H}) = \frac{1}{n}\operatorname{tr}(\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top) = \frac{1}{n}\operatorname{tr}(\boldsymbol{X}^\top\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}) = \frac{1}{n}\operatorname{tr}(\boldsymbol{I}_k) = \frac{k}{n}.$$

Here we used $\operatorname{tr}\boldsymbol{B}\boldsymbol{C} = \operatorname{tr}\boldsymbol{C}\boldsymbol{B}$. □

• **e.** *1 point Show that* $\frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top$ *is a projection matrix. Here* $\boldsymbol{\iota}$ *is the n-vector of ones.*

• **f.** *2 points Show: If the regression has a constant term, then* $\boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top$ *is a projection matrix.*

Answer. If $\boldsymbol{\iota}$, the vector of ones, is one of the columns of $\boldsymbol{X}$ (or a linear combination of these columns), this means there is a vector $\boldsymbol{a}$ with $\boldsymbol{\iota} = \boldsymbol{X}\boldsymbol{a}$. From this follows $\boldsymbol{H}\boldsymbol{\iota}\boldsymbol{\iota}^\top = \boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{X}\boldsymbol{a}\boldsymbol{\iota}^\top$ $\boldsymbol{X}\boldsymbol{a}\boldsymbol{\iota}^\top = \boldsymbol{\iota}\boldsymbol{\iota}^\top$. One can use this to show that $\boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top$ is idempotent: $(\boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top)(\boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top) = \boldsymbol{H}\boldsymbol{H} - \boldsymbol{H}\frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top\boldsymbol{H} + \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top\frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top = \boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top + \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top = \boldsymbol{H} - \frac{1}{n}\boldsymbol{\iota}\boldsymbol{\iota}^\top$. □

• **g.** *1 point Show: If the regression has a constant term, then one can sharpen inequality (4) to* $1/n \leq h_{ii} \leq 1$.

Answer. $\boldsymbol{H} - \boldsymbol{\iota}\boldsymbol{\iota}^\top/n$ is a projection matrix, therefore nonnegative definite, therefore its diagonal elements $h_{ii} - 1/n$ are nonnegative. □

Problem 20. *3 points What are the main concepts used in modern "Regression Diagnostics"? Can it be characterized to be a careful look at the residuals, or does it have elements which cannot be inferred from the residuals alone?*

Answer. Leverage (sometimes it is called "potential") is something which cannot be inferred from the residuals, it does not depend on $\boldsymbol{y}$ at all. □

Problem 21. *2 points In the regression model with random regressors $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, you only know that $\operatorname{plim} \frac{1}{n}\boldsymbol{X}^\top\boldsymbol{X} = \boldsymbol{Q}$ is a nonsingular matrix, and $\operatorname{plim} \frac{1}{n}\boldsymbol{X}^\top\boldsymbol{\varepsilon} = \boldsymbol{o}$. Using these two conditions, show that the OLS estimate is consistent.*

Answer. $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{y} = \boldsymbol{\beta} + (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\varepsilon}$ due to (**??**), and

$$\operatorname{plim}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\varepsilon} = \operatorname{plim}(\frac{\boldsymbol{X}^\top\boldsymbol{X}}{n})^{-1}\frac{\boldsymbol{X}^\top\boldsymbol{\varepsilon}}{n} = \boldsymbol{Q}\boldsymbol{o} = \boldsymbol{o}.$$

□

Problem 22. *3 points We are in the simple regression $y_t = \alpha + \beta x_t + \varepsilon_t$. If one draws, for every value of x, a 95% confidence interval for $\alpha + \beta x$, one gets a "confidence band" around the fitted line, as shown in Figure 1. Is the probability that this confidence band covers the true regression line* over its whole length *equal to 95%, greater than 95%, or smaller than 95%? Give a good verbal reasoning for*

FIGURE 1. Confidence Band for Regression Line

your answer. You should make sure that your explanation is consistent with the fact that the confidence interval is random and the true regression line is fixed.

Problem 23. *3 points We are in the model* $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ *and we have a matrix* $\boldsymbol{W}$ *of "instrumental variables" which satisfies the following three conditions:* $\operatorname{plim} \frac{1}{n}\boldsymbol{W}^\top\boldsymbol{\varepsilon} = \boldsymbol{o}$, $\operatorname{plim} \frac{1}{n}\boldsymbol{W}^\top\boldsymbol{W} = \boldsymbol{Q}$ *exists, is nonrandom and positive definite, and*

$\operatorname{plim} \frac{1}{n}\boldsymbol{W}^\top\boldsymbol{X} = \boldsymbol{D}$ *exists, is nonrandom and has full column rank. Show that the instrumental variables estimator*

$$\tilde{\boldsymbol{\beta}} = \Big(\boldsymbol{X}^\top\boldsymbol{W}(\boldsymbol{W}^\top\boldsymbol{W})^{-1}\boldsymbol{W}^\top\boldsymbol{X}\Big)^{-1}\boldsymbol{X}^\top\boldsymbol{W}(\boldsymbol{W}^\top\boldsymbol{W})^{-1}\boldsymbol{W}^\top\boldsymbol{y} \tag{5}$$

is consistent. Hint: Write $\tilde{\boldsymbol{\beta}}_n - \boldsymbol{\beta} = \boldsymbol{B}_n \cdot \frac{1}{n}\boldsymbol{W}^\top\boldsymbol{\varepsilon}$ *and show that the sequence of matrices* $\boldsymbol{B}_n$ *has a* plim.

Answer. Write it as

$$\begin{aligned}
\tilde{\boldsymbol{\beta}}_n &= \Big(\boldsymbol{X}^\top\boldsymbol{W}(\boldsymbol{W}^\top\boldsymbol{W})^{-1}\boldsymbol{W}^\top\boldsymbol{X}\Big)^{-1}\boldsymbol{X}^\top\boldsymbol{W}(\boldsymbol{W}^\top\boldsymbol{W})^{-1}\boldsymbol{W}^\top(\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon})\\
&= \boldsymbol{\beta} + \Big(\boldsymbol{X}^\top\boldsymbol{W}(\boldsymbol{W}^\top\boldsymbol{W})^{-1}\boldsymbol{W}^\top\boldsymbol{X}\Big)^{-1}\boldsymbol{X}^\top\boldsymbol{W}(\boldsymbol{W}^\top\boldsymbol{W})^{-1}\boldsymbol{W}^\top\boldsymbol{\varepsilon}\\
&= \boldsymbol{\beta} + \Big((\frac{1}{n}\boldsymbol{X}^\top\boldsymbol{W})(\frac{1}{n}\boldsymbol{W}^\top\boldsymbol{W})^{-1}(\frac{1}{n}\boldsymbol{W}^\top\boldsymbol{X})\Big)^{-1}(\frac{1}{n}\boldsymbol{X}^\top\boldsymbol{W})(\frac{1}{n}\boldsymbol{W}^\top\boldsymbol{W})^{-1}\frac{1}{n}\boldsymbol{W}^\top\boldsymbol{\varepsilon},
\end{aligned}$$

i.e., the $\boldsymbol{B}_n$ and $\boldsymbol{B}$ of the hint are as follows:

$$\begin{aligned}
\boldsymbol{B}_n &= \Big((\frac{1}{n}\boldsymbol{X}^\top\boldsymbol{W})(\frac{1}{n}\boldsymbol{W}^\top\boldsymbol{W})^{-1}(\frac{1}{n}\boldsymbol{W}^\top\boldsymbol{X})\Big)^{-1}(\frac{1}{n}\boldsymbol{X}^\top\boldsymbol{W})(\frac{1}{n}\boldsymbol{W}^\top\boldsymbol{W})^{-1}\\
\boldsymbol{B} &= \operatorname{plim}\boldsymbol{B}_n = (\boldsymbol{D}^\top\boldsymbol{Q}^{-1}\boldsymbol{D})^{-1}\boldsymbol{D}^\top\boldsymbol{Q}^{-1}
\end{aligned}$$

□

Problem 24. *4 points The model is*

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{E}; \qquad \operatorname{vec}(\boldsymbol{E}) \sim (\boldsymbol{o}, \boldsymbol{\Sigma} \otimes \boldsymbol{I}). \tag{6}$$

$\boldsymbol{X}$ is a known matrix of constants, and $\boldsymbol{B}$ and $\boldsymbol{\Sigma}$ are matrices of parameters to be estimated. Show that the BLUE in this model is

$$\hat{\boldsymbol{B}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y} \tag{7}$$

(i.e., the BLUE does not *depend on $\boldsymbol{\Sigma}$), and that*

$$\mathcal{V}[\operatorname{vec}(\hat{\boldsymbol{B}})] = \boldsymbol{\Sigma} \otimes (\boldsymbol{X}^\top \boldsymbol{X})^{-1}. \tag{8}$$

Here are some properties of the Kronecker product and the vectorization operator which may be useful for your proof:

$$(\boldsymbol{A} \otimes \boldsymbol{B})^\top = \boldsymbol{A}^\top \otimes \boldsymbol{B}^\top \tag{9}$$

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = \boldsymbol{A}\boldsymbol{C} \otimes \boldsymbol{B}\boldsymbol{D} \tag{10}$$

$$(\boldsymbol{A} \otimes \boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1} \tag{11}$$

$$\operatorname{vec}(\boldsymbol{A}\boldsymbol{B}\boldsymbol{C}) = (\boldsymbol{C}^\top \otimes \boldsymbol{A}) \operatorname{vec}(\boldsymbol{B}). \tag{12}$$

Answer. To show that $\hat{\boldsymbol{B}}$ is the BLUE, write the equation $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{E}$ in vectorized form, using (12), as

$$\operatorname{vec}(\boldsymbol{Y}) = (\boldsymbol{I} \otimes \boldsymbol{X}) \operatorname{vec}(\boldsymbol{B}) + \operatorname{vec}(\boldsymbol{E}) \tag{13}$$

Since $\mathcal{V}[\operatorname{vec}(\boldsymbol{E})] = \boldsymbol{\Sigma} \otimes \boldsymbol{I}$, the GLS estimate is, according to (**??**),

$$\operatorname{vec}(\hat{\boldsymbol{B}}) = \Big((\boldsymbol{I} \otimes \boldsymbol{X})^\top (\boldsymbol{\Sigma} \otimes \boldsymbol{I})^{-1} (\boldsymbol{I} \otimes \boldsymbol{X})\Big)^{-1} (\boldsymbol{I} \otimes \boldsymbol{X})^\top (\boldsymbol{\Sigma} \otimes \boldsymbol{I})^{-1} \operatorname{vec}(\boldsymbol{Y}) \tag{14}$$

$$= \Big((\boldsymbol{I} \otimes \boldsymbol{X}^\top)(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I})(\boldsymbol{I} \otimes \boldsymbol{X})\Big)^{-1} (\boldsymbol{I} \otimes \boldsymbol{X}^\top)(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}) \operatorname{vec}(\boldsymbol{Y}) \tag{15}$$

$$= \Big(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{X}^\top \boldsymbol{X}\Big)^{-1} (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{X}^\top) \operatorname{vec}(\boldsymbol{Y}) \tag{16}$$

$$= \Big(\boldsymbol{I} \otimes (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top\Big) \operatorname{vec}(\boldsymbol{Y}) \tag{17}$$

and applying (12) again, this is equivalent to

$$\hat{\boldsymbol{B}} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y}. \tag{18}$$

From this vectorization one can also derive the dispersion matrix $\mathcal{V}[\operatorname{vec}(\hat{\boldsymbol{B}})] = \boldsymbol{\Sigma} \otimes (\boldsymbol{X}^\top \boldsymbol{X})^{-1}$. In other words, $\mathcal{C}[\hat{\boldsymbol{\beta}}_i, \hat{\boldsymbol{\beta}}_j] = \sigma_{ij} (\boldsymbol{X}^\top \boldsymbol{X})^{-1}$, which can be estimated by $\hat{\sigma}_{ij} (\boldsymbol{X}^\top \boldsymbol{X})^{-1}$. □

Problem 25. *5 points Define "seemingly unrelated equations" and discuss the estimation issues involved.*

Problem 26. [Gre97, p. 709 ff]. *Here is a demand and supply curve with $\boldsymbol{q}$ quantity, $\boldsymbol{p}$ price, $\boldsymbol{y}$ income, and $\boldsymbol{\iota}$ is the vector of ones. All vectors are t-vectors.*

$$\boldsymbol{q} = \alpha_0\boldsymbol{\iota} + \alpha_1\boldsymbol{p} + \alpha_2\boldsymbol{y} + \boldsymbol{\varepsilon}_d \qquad \boldsymbol{\varepsilon}_d \sim (\boldsymbol{o}, \sigma_d^2\boldsymbol{I}) \qquad \textit{(demand)} \tag{19}$$

$$\boldsymbol{q} = \beta_0\boldsymbol{\iota} + \beta_1\boldsymbol{p} + \boldsymbol{\varepsilon}_s \qquad \boldsymbol{\varepsilon}_s \sim (\boldsymbol{o}, \sigma_s^2\boldsymbol{I}) \qquad \textit{(supply)} \tag{20}$$

$\boldsymbol{\varepsilon}_d$ and $\boldsymbol{\varepsilon}_s$ are independent of $\boldsymbol{y}$, but amongst each other they are contemporaneously correlated, with their covariance constant over time:

$$\operatorname{cov}[\varepsilon_{dt}, \varepsilon_{su}] = \begin{cases} 0 & \textit{if } t \neq u \\ \sigma_{ds} & \textit{if } t = u \end{cases} \tag{21}$$

• **a.** *1 point Which variables are exogenous and which are endogenous?*

Answer. $\boldsymbol{p}$ and $\boldsymbol{q}$ are called *jointly dependent* or *endogenous*. $\boldsymbol{y}$ is determined outside the system or *exogenous*. □

• **b.** *2 points Assuming $\alpha_1 \neq \beta_1$, verify that the reduced-form equations for $\boldsymbol{p}$ and $\boldsymbol{q}$ are as follows:*

$$\boldsymbol{p} = \frac{\alpha_0 - \beta_0}{\beta_1 - \alpha_1}\boldsymbol{\iota} + \frac{\alpha_2}{\beta_1 - \alpha_1}\boldsymbol{y} + \frac{\boldsymbol{\varepsilon}_d - \boldsymbol{\varepsilon}_s}{\beta_1 - \alpha_1} \tag{22}$$

$$\boldsymbol{q} = \frac{\beta_1\alpha_0 - \beta_0\alpha_1}{\beta_1 - \alpha_1}\boldsymbol{\iota} + \frac{\beta_1\alpha_2}{\beta_1 - \alpha_1}\boldsymbol{y} + \frac{\beta_1\boldsymbol{\varepsilon}_d - \alpha_1\boldsymbol{\varepsilon}_s}{\beta_1 - \alpha_1} \tag{23}$$

Answer. One gets the reduced form equation for $\boldsymbol{p}$ by simply setting the righthand sides equal:

$$\beta_0\boldsymbol{\iota} + \beta_1\boldsymbol{p} + \boldsymbol{\varepsilon}_s = \alpha_0\boldsymbol{\iota} + \alpha_1\boldsymbol{p} + \alpha_2\boldsymbol{y} + \boldsymbol{\varepsilon}_d$$
$$(\beta_1 - \alpha_1)\boldsymbol{p} = (\alpha_0 - \beta_0)\boldsymbol{\iota} + \alpha_2\boldsymbol{y} + \boldsymbol{\varepsilon}_d - \boldsymbol{\varepsilon}_s,$$

hence (22). To get the reduced form equation for $\boldsymbol{q}$, plug that for $\boldsymbol{p}$ into the supply function (one might also plug it into the demand function but the math would be more complicated):

$$\boldsymbol{q} = \beta_0\boldsymbol{\iota} + \beta_1\boldsymbol{p} + \boldsymbol{\varepsilon}_s = \beta_0\boldsymbol{\iota} + \frac{\beta_1(\alpha_0 - \beta_0)}{\beta_1 - \alpha_1}\boldsymbol{\iota} + \frac{\beta_1\alpha_2}{\beta_1 - \alpha_1}\boldsymbol{y} + \frac{\beta_1(\boldsymbol{\varepsilon}_d - \boldsymbol{\varepsilon}_s)}{\beta_1 - \alpha_1} + \boldsymbol{\varepsilon}_s$$

Combining the first two and the last two terms gives (23). □

• **c.** *2 points Show that one will in general not get consistent estimates of the supply equation parameters if one regresses $\boldsymbol{q}$ on $\boldsymbol{p}$ (with an intercept).*

Answer. By (22) (the reduced form equation for $\boldsymbol{p}$), $\text{cov}[\varepsilon_{st}, p_t] = \text{cov}[\varepsilon_{st}, \frac{\varepsilon_{dt} - \varepsilon_{st}}{\beta_1 - \alpha_1}] = \frac{\sigma_{sd} - \sigma_s^2}{\beta_1 - \alpha_1}$. This is generally $\neq 0$, therefore inconsistency. □

• **d.** *2 points If one estimates the supply function by instrumental variables, using $\boldsymbol{y}$ as an instrument for $\boldsymbol{p}$ and $\boldsymbol{\iota}$ as instrument for itself, write down the formula for the resulting estimator $\tilde{\beta}_1$ of β_1 and show that it is consistent. You are allowed to use, without proof, the following equation for the simple instrumental variables estimator:*

$$\tilde{\beta} = \frac{\sum(w_t - \bar{w})(y_t - \bar{y})}{\sum(w_t - \bar{w})(x_t - \bar{x})} \tag{24}$$

Answer. $\tilde{\beta}_1 = \frac{\frac{1}{n}\sum(y_i - \bar{y})(q_i - \bar{q})}{\frac{1}{n}\sum(y_i - \bar{y})(p_i - \bar{p})}$. Its plim is $\frac{\text{cov}[\boldsymbol{y},\boldsymbol{q}]}{\text{cov}[\boldsymbol{y},\boldsymbol{p}]} = \frac{\beta_1\alpha_2 \text{ var}[\boldsymbol{y}]/(\beta_1 - \alpha_1)}{\alpha_2 \text{ var}[\boldsymbol{y}]/(\beta_1 - \alpha_1)} = \beta_1$. These covariances were derived from (22) and (23). □

• **e.** *2 points Show that the Indirect Least Squares estimator of β_1 is identical to the instrumental variables estimator.*

Answer. For indirect least squares one estimates the two reduced form equations by OLS:

$$\text{the slope parameter in (22), } \frac{\alpha_2}{\beta_1 - \alpha_1}, \text{ estimated by } \frac{\sum(y_i - \bar{y})(p_i - \bar{p})}{\sum(y_i - \bar{y})^2};$$

$$\text{the slope parameter in (23), } \frac{\beta_1\alpha_2}{\beta_1 - \alpha_1}, \text{ estimated by } \frac{\sum(y_i - \bar{y})(q_i - \bar{q})}{\sum(y_i - \bar{y})^2}$$

Divide to get

$$\beta_1 \text{ estimated by } \frac{\sum(y_i - \bar{y})(q_i - \bar{q})}{\sum(y_i - \bar{y})(p_i - \bar{p})}$$

which is the same $\tilde{\beta}_1$ as in part d. □

• **f.** *1 point Since the error terms in the reduced form equations are contemporaneously correlated, wouldn't one get more precise estimates if one estimates the reduced form equations as a seemingly unrelated system, instead of OLS?*

Answer. Not as long as one does not impose any constraints on the reduced form equations, since all regressors are the same. □

• **g.** *2 points We have shown above that the regression of* $\boldsymbol{q}$ *on* $\boldsymbol{p}$ *does not give a consistent estimator of* β_1*. However one does get a consistent estimator of* β_1 *if one regresses* $\boldsymbol{q}$ *on the predicted values of* $\boldsymbol{p}$ *from the reduced form equation. (This is 2SLS.) Show that this estimator is also the same as above.*

Answer. This gives $\tilde{\beta}_1 = \frac{\sum(q_i - \bar{q})(\hat{p}_i - \bar{p})}{\sum(\hat{p}_i - \bar{p})^2}$. Now use $\hat{p}_i - \bar{p} = \hat{\pi}_1(y_i - \bar{y})$ where $\hat{\pi}_1 = \sum \frac{(p_i - \bar{p})(y_i - \bar{y})}{\sum(y_i - \bar{y})^2}$. Therefore $\tilde{\beta} = \hat{\pi}_1 \frac{\sum(q_i - \bar{q})(y_i - \bar{y})}{\hat{\pi}_1^2 \sum(y_i - \bar{y})^2} = \frac{\sum(y_i - \bar{y})(q_i - \bar{q})}{\sum(y_i - \bar{y})(p_i - \bar{p})}$ again. □

• **h.** *1 point So far we have only discussed estimators of the parameters in the supply function. How would you estimate the demand function?*

Answer. You can't. The supply function can be estimated because it stays put while the demand function shifts around, therefore the observed intersection points lie on the same supply function but different demand functions. The demand function itself cannot be estimated, it is underidentified in this system. □

Problem 27. *1 point If* $\mathcal{V}[\operatorname{vec} \boldsymbol{E}] = \boldsymbol{\Sigma} \otimes \boldsymbol{I}$*, this means (check the true answer or answers) that*

- *different rows of* $\boldsymbol{E}$ *are uncorrelated, and every row has the same covariance matrix, or*
- *different columns of* $\boldsymbol{E}$ *are uncorrelated, and every column has the same covariance matrix, or*

- *all ε_{ij} are uncorrelated.*

Answer. The first answer is right. □

Problem 28. *This example is adapted from* [JHG+88, (14.5.8) on p. 617]*:*

- **a.** *2 points Use the order condition to decide which of the following equations are exactly identified, overidentified, not identified.*

$$y_1 = -y_2\gamma_{21} - y_4\gamma_{41} + x_1\beta_{11} + x_4\beta_{41} + \varepsilon_1 \tag{25}$$

$$y_2 = -y_1\gamma_{12} + x_1\beta_{12} + x_2\beta_{22} + \varepsilon_2 \tag{26}$$

$$y_1 = -y_2\gamma_{23} - y_3\gamma_{33} - y_4\gamma_{43} + x_1\beta_{13} + x_4\beta_{43} + \varepsilon_3 \tag{27}$$

$$y_4 = x_1\beta_{14} + x_2\beta_{24} + x_3\beta_{34} + x_4\beta_{44} + \varepsilon_4 \tag{28}$$

Answer. (28) is exactly identified since there are no endogenous variable on the right hand side, but all exogenous variables are on the right hand side. (27) is not identified, it has 3 y's on the right hand side but only excludes two x's. (26) overfulfils the order condition, overidentified. (25) is exactly identified. □

• **b.** *1 point Write down the matrices* $\boldsymbol{\Gamma}$ *and* $\boldsymbol{B}$ *(indicating where there are zeros and ones) in the matrix representation of this system, which has the form*

$$(29)\quad \begin{bmatrix} \boldsymbol{y}_1 & \boldsymbol{y}_2 & \boldsymbol{y}_3 & \boldsymbol{y}_4 \end{bmatrix} \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} & \gamma_{24} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} & \gamma_{34} \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} \end{bmatrix} =$$

$$= \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 & \boldsymbol{x}_4 \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \\ \beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} \\ \beta_{31} & \beta_{32} & \beta_{33} & \beta_{34} \\ \beta_{41} & \beta_{42} & \beta_{43} & \beta_{44} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 & \boldsymbol{\varepsilon}_2 & \boldsymbol{\varepsilon}_3 & \boldsymbol{\varepsilon}_4 \end{bmatrix}$$

Answer.

$$\boldsymbol{\Gamma} = \begin{bmatrix} 1 & \gamma_{12} & 1 & 0 \\ \gamma_{21} & 1 & \gamma_{23} & 0 \\ 0 & 0 & \gamma_{33} & 0 \\ \gamma_{41} & 0 & \gamma_{43} & 1 \end{bmatrix} \qquad \boldsymbol{B} = \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \\ 0 & \beta_{22} & 0 & \beta_{24} \\ 0 & 0 & 0 & \beta_{34} \\ \beta_{41} & 0 & \beta_{43} & \beta_{44} \end{bmatrix}$$

□

Maximum number of points: 57.

References

[Gre97] William H. Greene, *Econometric analysis*, third ed., Prentice Hall, Upper Saddle River, NJ, 1997. 11

[JHG+88] George G. Judge, R. Carter Hill, William E. Griffiths, Helmut Lütkepohl, and Tsoung-Chao Lee, *Introduction to the theory and practice of econometrics*, second ed., Wiley, New York, 1988. 15

Economics Department, University of Utah, 1645 Campus Center Drive, Salt Lake City, UT 84112-9300, U.S.A

E-mail address: ehrbar@econ.utah.edu

URL: http://www.econ.utah.edu/ehrbar