ECONOMETRÍA I
Convocatoria Extraordinaria
22 de Septiembre de 2007

SOLUCIONES EXAMEN

Tipo 1

1.D	2.C	3.B	4.D	5.B
6.A	7.C	8.B	9.A	10.A
11.A	12.A	13.C	14.C	15.D
16.D	17.A	18.A	19.B	20.C
21.D	22.C	23.D	24.C	25.A
26.C	27.C	28.D	29.B	30.C
31.C	32.A	33.C	34.C	35.C
36.B	37.C	38.B	39.D	40.D

Tipo 2

1.A	2.B	3.C	4.A	5.C
6.D	7.B	8.C	9.D	10.D
11.D	12.D	13.B	14.B	15.A
16.A	17.D	18.D	19.C	20.B
21.A	22.B	23.A	24.B	25.D
26.B	27.B	28.A	29.C	30.B
31.B	32.D	33.B	34.B	35.B
36.C	37.B	38.C	39.A	40.A

Tipo 3

1.A	2.D	3.B	4.D	5.B
6.D	7.C	8.D	9.D	10.B
11.A	12.D	13.D	14.C	15.D
16.D	17.D	18.A	19.D	20.A
21.B	22.B	23.B	24.D	25.A
26.B	27.A	28.C	29.D	30.A
31.C	32.C	33.C	34.C	35.D
36.D	37.B	38.B	39.C	40.C

Tipo 4

1.C	2.A	3.D	4.C	5.D
6.B	7.A	8.D	9.B	10.B
11.B	12.B	13.A	14.A	15.C
16.C	17.B	18.B	19.D	20.A
21.C	22.A	23.C	24.A	25.B
26.A	27.A	28.C	29.D	30.A
31.A	32.B	33.A	34.A	35.A
36.D	37.A	38.D	39.C	40.C

UNIVERSIDAD CARLOS III DE MADRID
ECONOMETRÍA I
Curso 2006/07
EXAMEN FINAL (Convocatoria extraordinaria)

22 de Septiembre de 2007

GUIA DE SOLUCIONES

PROBLEMA 1

El presente problema se basa en la especificación más utilizada para la demanda de dinero de una economía. La ecuación (*), utilizando la notación del enunciado en logaritmos, puede escribirse como

$$(m_t - p_t) = \alpha_1 + \alpha_2 RS_t + \alpha_3 y_t + \zeta_t$$

Como $-v = (m - p) - y$, podemos escribir la ecuación de arriba de forma equivalente como (restando y_t en ambos lados de la ecuación).

$$-v_t = \alpha_1 + \alpha_2 RS_t + (\alpha_3 - 1)\, y_t + \zeta_t.$$

La especificación (*) es por tanto equivalente a la especificación (2), donde

$$\begin{aligned} \delta_1 &= \alpha_1, \\ \delta_2 &= \alpha_2, \\ \delta_3 &= \alpha_3 - 1. \end{aligned}$$

y por tanto el término de error ε_t verifica que $\varepsilon_t = \zeta_t$.

Por el contrario, la especificación (1) es una versión restringida del modelo, en la que

$$\begin{aligned} \beta_1 &= \alpha_1 \\ \beta_2 &= \alpha_2 \\ \alpha_3 &= 1 \end{aligned}$$

Nótese que la especificación (1) es menos general que la ecuación (*), ya que impone la restricción de que $\alpha_3 = 1$. Eso no implica que la ecuación sea independiente de y_t, puesto que está implícita en la definición de la variable dependiente $-v_t$: lo que impone es que para un tipo de interés a corto plazo dado, la elasticidad de la demanda de dinero con respecto a la renta es igual a la unidad.

El modelo se estima con datos de series temporales, que pueden presentar un problema potencial de autocorrelación sobre el que se pregunta en el examen.

Tanto en la especificación (1) como en la (2), la variable dependiente es el opuesto del logaritmo neperiano de la velocidad de circulación del dinero $(-v)$, lo que tiene que tenerse en cuenta a la hora de evaluar el signo de los efectos de las variables del lado derecho sobre la velocidad de circulación.

La discusión referente a las SALIDAS 1 a 6, a continuación, se basa en que la especificación (1) es correcta (en particular, que la restricción que impone respecto a la especificación general (*) se verifica).

Las SALIDAS 1 y 2 (en las que la variable dependiente es $-v$ y la única variable explicativa es RS) estiman el modelo restringido (1) por MCO y por MC2E. En este segundo caso, se tiene en cuenta la posible endogeneidad del tipo de interés a corto plazo RS, para el que se utilizan como instrumentos el tipo de interés contemporáneo a largo plazo, RL, y el tipo de interés a corto plazo observado hace dos años, $RS(-2)$. La SALIDA 6 estima igualmente el modelo (1) por MCO pero incluye errores estándar robustos a autocorrelación (de manera que los coeficientes estimados son idénticos a los de la SALIDA 1 pero los errores estándar de dichos coeficientes pueden diferir).

La SALIDA 3 presenta la ecuación auxiliar de forma reducida de la variable potencialmente endógena, que consiste en la proyección lineal de RS sobre RL y $RS(-2)$. Dicha salida nos permite evaluar si se cumple una de las dos condiciones requeridas para que los instrumentos sean válidos, que estén correlacionados con la variable potencialmente endógena que se debe instrumentar. Puede verse que tanto RL como $RS(-2)$ son individual y conjuntamente significativos en dicha ecuación auxiliar, lo que implica que ambos instrumentos tienen capacidad para predecir RS y cumplen por tanto la condición de correlación con dicha variable. Recordemos que la segunda condición para que un instrumento sea válido, que no esté correlacionado con el término de error del modelo a estimar u_t (es decir, que sean exógenas respecto al término de error del modelo de interés) es generalmente no contrastable (en todo caso, el ejercicio no aporta información que permita llevar a cabo un contraste de restricciones de sobreidentificación que permitiera contrastar parcialmente esta segunda condición).

La SALIDA 4 es una estimación MCO del modelo (1) ampliado, incluyendo los residuos de la ecuación auxiliar de forma reducida anterior (SALIDA 3) como variable adicional. El contraste de Hausman puede llevarse a cabo contrastando la significación de dicha variable adicional. Bajo la hipótesis de exogeneidad de RS, la variable $res3$ debería tener un coeficiente estadísticamente igual a 0. En vista de la SALIDA 4, $res3$ no es estadísticamente significativa, por lo que no rechazamos la exogeneidad de RS.

En consecuencia, dado que bajo la hipótesis de exogeneidad tanto el estimador MCO como el de MC2E de la especificación (1) son consistentes (en el caso de MC2E, siempre que los instrumentos sean válidos) debemos considerar los resultados de la SALIDA 1 frente a los de la SALIDA 2, dado que el estimador MCO es más eficiente bajo la hipótesis nula de exogeneidad.

Si nos concentramos por tanto en la SALIDA 1, vemos que el tipo de interés a corto plazo tiene un efecto **positivo** sobre la velocidad de circulación del dinero (es decir, que la velocidad de circulación del dinero aumenta al aumentar el tipo de interés a corto plazo). Por otro lado, como la variable dependiente es el logaritmo de la velocidad de circulación (con signo menos), la constante cambiada de signo se interpretaría como el valor medio del logaritmo de la velocidad de circulación cuando el tipo de interés a corto plazo es igual a cero. En cuanto a la interpretación del efecto del tipo de interés a corto plazo, RS, podemos concluir, dado que dicho tipo está medido en tanto por uno, que si éste aumenta en 0.1 puntos porcentuales, la velocidad de circulación aumentaría en un 0.7% (el efecto sobre la velocidad de circulación se mide en tanto por ciento al estar transformada en logaritmos). Nótese que la especificación de la SALIDA 1 implica que el efecto del tipo de interés a corto sobre el logaritmo de la velocidad de circulación, v, es constante e igual a β_2 (de manera que el efecto sobre la velocidad de circulación V no es constante).

Una de las posible hipótesis que se plantean en el examen es $H_0 : \beta_2 = -10$ (es decir: que ante una aumento del tipo de interés a corto plazo en 0.1 puntos porcentuales, la velocidad de circulación aumenta en un 1%. Suponiendo que el modelo (1) cumple los supuestos del modelo de regresión clásico, el estadístico de contraste, utilizando la SALIDA 1, sería aproximadamente

$$t = \frac{b_2 - (-10)}{se(b_2)} = \frac{-7 - (-10)}{0.63} \simeq 4.8,$$

que se distribuye asintóticamente como una $N(0, 1)$ bajo H_0. El p-valor del estadístico es muy inferior al 1%, de manera que a los niveles de significación habituales rechazamos la hipótesis nula.

Continuando con el análisis de la SALIDA 1, vemos que el R^2 es igual a 0.56, de manera que, ateniéndonos a la definición del R^2, si el modelo (1) cumple los supuestos del modelo de regresión clásico, podemos concluir que el tipo de interés a corto plazo explica el 56% de la varianza del logaritmo de la velocidad de circulación. En cuanto a la estimación de la varianza condicional (a RS) de v, no es sino el cuadrado de la desviación típica de los residuos, es decir, $(0.109142)^2 \simeq 0.012$. Nótese que la varianza condicional de v y de $-v$ son la misma cosa.

Además, en la SALIDA 1 encontramos que los coeficientes estimados de la constante y de la pendiente son estadísticamente distintos de cero de forma individual. Como uno de los coeficientes es el de la constante, y la constante es una variable con correlación cero con cualquier otra (en particular, con RS) entonces podemos cocnluir que el hecho de que individualmente son estadísticamente distintos de cero implica que también son estadísticamente distintos de cero conjuntamente. Naturalmente, esta propiedad no tiene por qué cumplirse cuando se trata de la significación estadística de los coeficientes de dos variables explicativas, ninguna de las cuales es constante, que están generalmente correlacionadas.

Por último, la SALIDA 1 presenta, al tratarse de datos de series temporales, el estadístico de Durbin-Watson. Recordemos que el estadístico de Durbin-Watson es aproximadamente

$$DW \simeq 2\left(1 - \widehat{\rho}\right),$$

donde $\widehat{\rho}$ es el coeficiente de autocorrelación muestral de primer orden de los residuos del modelo, es decir, la estimación basada en la muestra de $cov(e_t, e_{t-1})$. Recordemos que este estadístico sólo es informativo de si hay o no autocorrelación de primer orden, pero no aporta evidencia sobre si hay autocorrelación de orden superior. Si nos basamos en este estadístico, si no existe autocorrelación de primer orden, el valor del estadístico DW debería estar cercano a 2. Dado que el valor está muy alejado (mucho menor que 2), hay evidencia de autocorrelación (positiva) de primer orden. Utilizando la expresión de arriba, el coeficiente de autocorrelación muestral de primer orden está en torno a 0.84.

Sin embargo, la evidencia basada en el estadístico DW no nos dice nada acerca de la posibilidad de que exista autocorrelación de orden superior. La SALIDA 5 permite contrastar la hipótesis nula de no correlación de orden igual o inferior a 2. En concreto, consiste en una regresión de los residuos contemporáneos sobre los residuos desfasados uno y dos períodos. En vista de que el coeficiente de $e(-1)$ es estadísticamente significativo pero el coeficiente de $e(-2)$ es estad´siticamente igual a cero, podemos concluir que existe evidencia de autocorrelación de primer orden pero no de segundo orden. (Nótese que el coeficiente estimado de $e(-1)$ es numéricamente similar al coeficiente de autocorrelación muestral calculado a partir del estadístico DW). No obstante, los resultados de la SALIDA 5 no nos permiten concluir que no existe autocorrelación de orden tres o superior, dado que no se han incluido retardos de los residuos de orden tres o superior.

La existencia de autocorrelación de primer orden no afecta a la consistencia del estimador MCO de los coeficientes del modelo, pero en dicha situación la los errores estándar convencionales de dichos coeficientes no estiman consistentemente las desviaciones típicas de los coeficientes estimados. Ello implica que no es posible hacer inferencia válida a partir de los resultados de la SALIDA 1. Como puede verse, los errores estándar de la SALIDA 6 son sustancialmente mayores que los de la SALIDA 1, lo que es coherente con la evidencia hallada de autocorrelación de primer orden. Esto quiere decir que cualquier contraste apropiado sobre el modelo (1) debe utilizar los resultados de la SALIDA 7. En todo caso, la variable RS mantiene su significación estadística cuando se utilizan errores estándar apropiados, robustos a autocorrelación.

Finalmente, el problema presenta las SALIDAS 7, 8 y 9, correspondientes al modelo (2), que es un modelo más general que el modelo (1) (al no restringir el coeficiente de y a la unidad) y equivalente a la especificación (*). La interpretación del coeficiente de RS es parecida a la del modelo anterior, excepto por el hecho de que ahora dicho coeficiente refleja el efecto de RS sobre la velocidad de circulación para una renta dada (es decir, el efecto ceteris paribus). En este sentido, condicional a RS y y, el efecto ceteris paribus del tipo de interés a corto sobre la velocidad de circulación sigue sin ser constante (lo que es constante es el efecto sobre el logaritmo de la velocidad de circulación). La magnitud del efecto es parecida a la obtenida en el modelo (1). En cuanto al efecto de la renta, recordemos que $\delta_3 = \alpha_3 - 1$, donde, al estar la demanda de dinero y la renta en logaritmos, α_3 es la elasticidad de la demanda de dinero con respecto a la renta, que es constante. El modelo (2) permite que esa elasticidad sea distinta de la unidad. Este modelo nos permite contrastar la validez de la hipótesis de elasticidad unitaria de la demanda de dinero con respecto a la renta, $H_0 : \alpha_3 = 1$, o, de forma equivalente, $H_0 : \delta_3 = 0$. De no rechazarse dicha hipótesis, optaríamos por el modelo (1), y basaríamos nuestra inferencia en la SALIDA 6.

En este caso, tenemos tres estimaciones MCO del mismo modelo (por tanto, los coeficientes estimados son los mismos) con estimaciones alternativas de los errores estándar: no robustos a autocorrelación y heterocedasticidad (SALIDA 9), robustos a heterocedasticidad (SALIDA 8) y robustos a autocorrelación de orden igual o inferior a tres (SALIDA 7). Dada la información, la decisión de qué salida utilizar debe basarse en la comparación de los errores estándar. En ausencia de heterocedasticidad y de autocorrelación, los errores estándar robustos y no robustos deberían ser numéricamente muy parecidos. A la vista de los resultados de las SALIDAS 8 y 9, no hay evidencia de heterocedasticidad condicional, dado el gran parecido en los errores estándar de ambas. Sin embargo, los errores estándar robustos a autocorrelación son sustancialmente mayores que los respectivos errores estándar convencionales. Este hecho apunta un problema de autocorrelación condicional que invalida la inferencia de la SALIDA 9 (MCO con errores estándar no robustos), debiendo basarla en la SALIDA 7 (MCO con errores estándar robustos a heterocedasticidad). La elección de la SALIDA correcta tiene consecuencias críticas de cara al contraste de la hipótesis $H_0 : \alpha_3 = 1$, o $H_0 : \delta_3 = 0$. El estimador MCO de δ_3 es $d_3 = 0.043$ con un error estándar de 0.037. Por tanto, el estadístico t es

$$t = \frac{0.043}{0.037} \simeq 1.15$$

con un p-valor asociado de un 25%, por lo que no se rechaza la hipótesis de elasticidad unitaria de la demanda de dinero respecto a la renta a niveles de significación habituales.

En consecuencia, dados los resultados, el modelo de nuestra elección sería el modelo (1), con la restricción arriba mencionada, y nuestra inferencia se basaría en la SALIDA 6. Nótese que ello no implica que la renta no sea relevante para la demanda de dinero, porque el modelo (1) se puede escribir en la forma (*) con $\ln\left(\frac{M}{P}\right) = m - p$ como variable dependiente y $\ln Y = y$ como variable explicativa con coeficiente restringido a la unidad (lo que lleva a redefinir la variable dependiente como $-v = m - p - y$. Podemos concluir por tanto que la demanda de dinero depende positivamente de la renta (con elasticidad unitaria) y del tipo de interés a corto plazo.

Conviene aclarar que, como el modelo (2) es más general que el (1), ambas especificaciones son válidas y si se cumplen los supuestos habituales, los estimadores MCO de los coeficientes en ambos modelos son consistentes. Sin embargo, a la hora de escoger uno de los dos modelos, optaríamos por el (1) dada la ganancia de eficiencia que supone estimar un modelo que incorpora una restricción ($\delta_3 = 0$) que es cierta (como puede verse en los menores errores estándar de la constante y de la pendiente de RS en la SALIDA 6 frente a los de la SALIDA 7). Eso implica que nuestras estimaciones preferidas implican que la elasticidad de la demanda de dinero respecto a la renta es igual a 1 y que el efecto estimado de un incremento de 1 punto porcentual en el tipo de interés a corto aumenta en promedio la velocidad de circulación en un 7%.

PROBLEMA 2

(El enunciado de este problema es similar al de un problema planteado en el examen de Enero de 2007, con la única diferencia de que la variable explicativa RAZA_ACUSADO se define de manera opuesta, es decir, toma el valor 1 si el acusado es de raza negra y 0 en caso contrario).

Se trata de un modelo de elección discreta para caracterizar la probabilidad de ser condenado a muerte. La SALIDA 1 incluye una única variable explicativa, RAZA_ACUSADO, mientras que la SALIDA 2 es un modelo más general que incluye una variable explicativa adicional, RAZA_VICTIMA. Nótese que, al igual que la variable dependiente CONDENA, las variables explicativas son binarias.

Comparando ambas salidas, vemos que la omisión de la variable RAZA_VICTIMA afecta de forma crítica a la hora de evaluar el efecto de la raza del acusado sobre la probabilidad de condena. Así, mientras que en la SALIDA 1 dicha variable no es significativa a niveles de significación habituales (de manera que concluiriamos que la condena es independiente de la raza del acusado) en ls SALIDA 2 vemos que, dada la raza de la víctima, la probabilidad de condena es mayor si el acusado es negro.

El texto también indica que la proporción de acusados de raza blanca en la muestra es del 72% (lo que representa 489 individuos de la muestra) y por tanto la proporción de acusados de raza negra en la muestra es del 28% (lo que supone 190 individuos). En la SALIDA 1 podemos ver que la media de CONDENA es 0.102, es decir que el 10.2% de los individuos de la muestra son condenados a muerte.

Al ser la variable binaria RAZA_ACUSADO, tenemos que, a partir de la SALIDA 1 podemos estimar la proporción en la muestra de acusados de raza negra que son condenados a muerte como

$$\begin{aligned} E\,(\text{CONDENA} \mid \text{RAZA\_ACUSADO} = 1) &= \Lambda\,(-2.08 - 0.39) \\ &= \frac{1}{1 + e^{2.08+0.39}} = 0.0779. \end{aligned}$$

que en tanto por ciento es aproximadamente el 8%.

Con la información que tenemos, podemos calcular los estimadores de la constante y la pendiente de la proyección lineal o PLO de CONDENA sobre RAZA_ACUSADO (es decir, el modelo de probabilidad lineal). Recordemos que la estimación MCO de la constante y la pendiente de $PLO\,(Y|\,X) = \beta_1 + \beta_2 X$ son

$$\begin{aligned} b_0 &= \overline{Y} - b_1\overline{X}, \\ b_1 &= \frac{\sum_i Y_i X_i - n\overline{Y}\,\overline{X}}{\sum_i X_i^2 - n\overline{X}^2}, \end{aligned}$$

Donde los sumatorios se refieren a todas las observaciones (de 1 a 679).

En este caso, Y = CONDENA, X = RAZA_ACUSADO, de manera que

- $\sum_i Y_i X_i$ = No. de acusado de raza negra que son condenados = 190×0.08.
- $\overline{Y}$ = Media de CONDENA = 0.102.

- $\overline{X}$ = Media de RAZA_ACUSADO $= \frac{1}{679}\sum_i X_i = \frac{190}{679} = 0.28$.

- Como X_i es binaria, $\sum_i X_i^2 = \sum_i X_i = 190$ (número de acusados de raza negra).

Por tanto, la estimación de la constante y la pendiente del modelo de probabilidad lineal son iguales a

$$\begin{aligned} b_0 &= 0.102 - (-0.03)\, 0.28 \simeq 0.11, \\ b_1 &= \frac{190 \times 0.08 - 679 \times 0.102 \times 0.28}{190 - 679 \times 0.28^2} \simeq -0.03. \end{aligned}$$

Nótese que si el modelo de la SALIDA 1 se hubiese estimado como un modelo de probabilidad lineal, el modelo sería por construcción homocedástico, dado que al ser Y_i binomial, su varianza condicional sería igual a $P_i\,(1 - P_i)$, donde P_i depende del valor que tome la variable explicativa RAZA_ACUSADO, de manera que la varianza condicional no es constante. Otra propiedad del modelo de probabilidad lineal es que impone que el efecto de la variable explicativa sobre la probabilidad de condena es constante, dada la linealidad del modelo (a diferencia de un logit o un probit). Por último, aunque en general el modelo de probabilidad lineal no acota los valores predichos entre 0 y 1, en este caso sí lo hace, debido que el modelo tiene una única variable explicativa que es además dicotómica, de manera que las predicciones del modelo en este caso no pueden ser mayores que 1 o menores que 0. Puede comprobarse directamente viendo que las únicas posibilidades aquí son que RAZA_ACUSADO sea 0 (blanco) o 1 (negro). En el primer caso, la probabilidad predicha de condena es $0.11 - 0.03 \times 0 = 0.11$, mientras que en el segundo caso dicha probabilidad es $0.11 - 0.03 \times 1 = 0.08$. Si el modelo de probabilidad lineal incluyera dos o más variables explicativas, dicotómicas o no, no podría garantizarse que las probabilidades predichas para cualquier combinación posible de valores de dichas variables explicativas estén entre 0 y 1.

El hecho de que la única variable explicativa en el modelo de la SALIDA 1 sea dicotómica supone además que el efecto estimado de la raza del acusado sería **exactamente** el mismo si en lugar de un logit (distribución logística) hubiéramos utilizado un probit (distribución normal). Este resultado no se mantiene si el modelo tiene dos o más variables explicativas, dicotómicas o no.

Por otro lado, como se vio en teoría, el modelo logit puede estimarse de forma consistente (suponiendo que el modelo es correcto) tanto por mínimos cuadrados no lineales como por máxima verosimilitud. No obstante, bajo el supuesto de que la distribución es efectivamente logística, el estimador de máxima verosimilitud de dicho modelo sería más eficiente. Además, en el caso de la estimación del modelo logístico por mínimos cuadrados no lineales, deberíamos utilizar errores estándar de los coeficientes robustos a heterocedasticidad, porque los errores estándar convencionales serían inconsistentes.

A la hora de hacer inferencia o predecir probabilidades, debemos primero escoger el modelo más apropiado de entre los dos presentados. El criterio de selección a utilizar, dada la información disponible, es el logaritmo de la función de verosimilitud. Vemos que la mejora en el logaritmo de la función de verosimilitud es sustancial al pasar de la SALIDA 1 a la SALIDA 2. Formalmente, deberñiamos calcular un contraste LR de razón de verosimilitudes

$$LR = 2\left(l_S - l_R\right) \mathrel{\dot{\sim}} \chi_1^2,$$

donde l_S y l_R son los valores de los logaritmos de las funciones de verosimilitud del modelo no restringido (SALIDA 2, que no impone coeficiente cero a RAZA_VICTIMA) y del modelo restringido, respectivamente. Dicho contraste tiene en este caso un grado de libertad al existir una única restricción. El valor del estadístico es en este caso 20.36, por lo que a cualquier nivel de significación usual rechazamos la restricción que impone el modelo de la SALIDA 1 y optamos por tanto por la SALIDA 2.

Esto supone que:

- Si la víctima es blanca, es mucho más probable que el acusado sea condenado, con independencia de cuál sea la raza del acusado.
- La probabilidad de condena si la víctima es negra es mucho más pequeña que la probabilidad media de condena.
- El efecto relativo (en términos de cocientes de probabildiades) de la raza del acusado es mayor cuando la víctima es de raza negra.
- El efecto absoluto de la raza del acusado es mayor cuando la víctima es de raza blanca.

A partir de la SALIDA 2, podemos concluir, entre otras cosas, que la probabilidad predicha de condena para un acusado de raza negra cuando la víctima es blanca es aproximadamente

$$\begin{aligned} E(\text{CONDENA}|\text{RAZA\_ACUSADO} &= 1, \text{RAZA\_VICTIMA} = 0) = \Lambda\left(-2.04 + 0.83 \times 1 - 2.39 \times 0\right) \\ &= 0.2297 \simeq 0.23 \end{aligned}$$