UNIVERSIDAD TORCUATO DI TELLA
MAESTRÍA EN ECONOMÍA Y ECONOMETRÍA
2022

Datos de Panel
Problem Set 0
Repaso OLS, GLS & FGLS
Lectura y resumen de datos de panel en Stata

---

1. Considere el siguiente modelo de regresión:

$$\begin{aligned} ltotexp_i = & \beta_0 + \beta_1 suppins_i + \beta_2 phylim_i + \beta_3 actlim_i + \beta_4 totchr_i + \beta_5 age_i + \\ & \beta_6 female_i + \beta_7 income_i + u_i, \quad i = 1, \ldots, N \end{aligned}$$

   a) Use la base de datos “mus03data.dta”, la cual contiene datos de corte transversal de gastos médicos, para estimar la ecuación por OLS usando comandos de matrices en Stata. Adicionalmente, reporte los errores estándar usuales de OLS y los estadísticos t asociados.
   b) Utilice el comando regress para verificar los resultados obtenidos.
   c) Implemente un test de significatividad individual para totchr.
   d) Implemente un test de significatividad conjunta para todas las variables del modelo, excluyendo el intercepto.

2. En este ejercicio vamos a aprender cómo setear los datos como panel en Stata y cómo generar estadísticas descriptivas del panel. Adicionalmente, veremos cómo convertir los datos de *wide form* a *long form* y cómo generar un panel para simulaciones.

   a) Utilice la base mus08psidextract.dta y describa la base de datos de la manera usual y como un panel.
   b) Utilice la base pigweights.dta. Los datos se encuentran en formato *wide*. Utilice el comando reshape para llevarlos a formato *long*. Luego, describa la base de la misma forma que en el inciso (a).
   c) Genere un panel de 5000 observaciones con 10 períodos temporales y 500 unidades en el corte transversal. El panel debe estar en formato *long*. Genere observaciones de $x_{it} \sim \mathcal{N}(0,1)$, $u_{it} \sim \mathcal{N}(0,1)$ y además $y_{it} = 1 + x_{it} + u_{it}$. Estime por POLS.

3. Considere el siguiente modelo:

$$\begin{aligned} y_i &= \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i, \quad i = 1, \ldots, N \\ u_i &= \sqrt{\exp\left(-1 + 0{,}2 \cdot x_{2i}\right)} \cdot \varepsilon_i, \quad i = 1, \ldots, N \end{aligned}$$

   con $\beta_1 = 1$, $\beta_2 = 1$, $\beta_3 = 1$, $x_2 \sim \mathcal{N}(0,25)$, $x_3 \sim \mathcal{N}(0,25)$ y $\varepsilon \sim \mathcal{N}(0,25)$ Luego, el error u es heterocedástico con una varianza condicional igual a $25 \cdot \exp(-1 + 0{,}2 \cdot x2)$.

   a) Genere 1000 muestras de N=10 observaciones a partir del modelo presentado. Para cada muestra estime por OLS, GLS y FGLS los parámetros del modelo y realice un test de hipótesis para contrastar que $H_0 : \beta_3 = 1$. Reporte tamaño del test al 1 %. Adicionalmente, reporte la media, mediana y desvío estándar de las estimaciones de β_1, β_2 y β_3.
   b) Repita el punto anterior con N igual a 20, 30, 100, 200 y 500.
   c) Describa detalladamente las propiedades de muestra finita de FGLS de acuerdo a lo que observó de los puntos anteriores.

Universidad Torcuato Di Tella
Maestría en Economía y Econometría
2022

Datos de Panel
Problem Set 1
Modelo de Regresión Lineal

---

1. Utilice la base de datos provista *"cornwell.dta"*.

   a) A partir de los datos de los siete años, y utilizando los logaritmos de todas las variables, estime un modelo por POLS que relacione la tasa de crimen con *prbarr*, *prbconv*, *prbpris*, *avgsen* y *polpc* y que incluya un conjunto de *dummies* de año.

   b) Compute los errores estándar robustos a heteroscedasticidad arbitraria y a autocorrelación serial arbitraria.

   c) Implemente un contraste de Correlación Serial.

   d) Implemente un contraste de Heterocedasticidad.

   e) Asuma que se cumple el supuesto de exogeneidad estricta y que u_{it} sigue un proceso AR(1). Compute el estimador de FGLS siguiendo el enfoque de Prais-Winsten. Una descripción del procedimiento puede encontrarla en Wooldridge (2010), sección 7.8.6.
   Observación: GLS necesita exogeneidad estricta para conseguir estimadores consistentes.

   f) Compute los errores estándar robustos a heteroscedasticidad arbitraria y a autocorrelación serial arbitraria para el modelo con las variables transformadas del inciso previo.
   Sugerencia de Wooldridge. *"..If we have any doubts about the homoskedasticity assumption, or whether the AR(1) assumption sufficiently captures the serial dependence, we can just apply the usual fully robust variance matrix and associated statistics to pooled OLS on the transformed variables. This allows us to probably obtain an estimator more efficient than POLS (on the original data) but also guards against the rather simple structure we imposed on Ω. Of course, failure of strict exogeneity generally causes the Prais-Winsten estimator of β to be inconsistent."*

2. En este ejercicio examinará un modelo para el costo total de producción en la industria aeronáutica a modo de ilustrar una aplicación de un modelo heteroscedástico por grupos. Considere la siguiente función de costos:

$$\ln \text{cost}_{jt} = \beta_1 + \beta_2 \ln \text{output}_{jt} + \beta_3 \text{ load factor}_{jt} + \beta_4 \ln \text{fuel price}_{jt}$$
$$\delta_2 \text{ Firm}_2 + \delta_3 \text{ Firm}_3 + \delta_4 \text{ Firm}_4 + \delta_5 \text{ Firm}_5 + \delta_6 \text{ Firm}_6 + \varepsilon_{jt}$$

   a) Utilice la base de datos provista *"greene97.dta"*, la cual contiene datos para seis compañías aéreas observadas anualmente durante 15 años. Estime la ecuación por POLS.

   b) Ahora, asuma que dentro de cada compañía aérea se tiene que:

$$Var\left[\varepsilon_{jt} \mid \mathbf{x}_{jt}\right] = \sigma_j^2, \quad t = 1, \ldots, T$$

   Por lo tanto, si las varianzas fueran conocidas, el estimador de GLS sería:

$$\hat{\boldsymbol{\beta}} = \left[\sum_{j=1}^{N}\left(\frac{1}{\sigma_j^2}\right)\mathbf{X}_j'\mathbf{X}_j\right]^{-1}\left[\sum_{j=1}^{N}\left(\frac{1}{\sigma_j^2}\right)\mathbf{X}_j'\mathbf{y}_j\right]$$

   donde $\mathbf{X}_j$ es una matriz $T \times K$. Sin embargo, en este caso práctico las varianzas son desconocidas. Luego, se le solicita computar el estimador de FGLS a través de los siguientes métodos:

1) Estime el modelo calculando el estimador necesario para la varianza específica de la compañía aérea a partir de los residuos de OLS, es decir, $\hat{\sigma}_j^2 = \frac{\mathbf{e}_j'\mathbf{e}_j}{n_j}$.
2) Estimar el modelo tratándolo como una forma del modelo de heteroscedasticidad multiplicativa de Harvey (1976). Utilice el procedimiento en dos etapas.

c) Compare los resultados obtenidos en el inciso b).

3. Considere la siguiente ecuación de salarios:

$$y_{jt} = \beta_0 + \beta_1 x_{jt} + u_{jt}, \quad j = 1, 2, \dots N; t = 1, 2 \tag{1}$$

donde $\beta_0 = \beta_1 = 1$, $u_j \sim N(0, \Omega)$, $\Omega = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$ y $x_j \sim U[1, 20]$

Genere 1000 muestras de $N = 5$ observaciones de corte transversal a partir del modelo (1). Para cada muestra estime por FGLS los parámetros del modelo y realice un test de hipótesis para contrastar que $H_0 : \beta_1 = 1$. Reporte tamaño del test al 1 % y el poder del test cuando $\beta_1 = 0{,}8$. Luego, repita el procedimiento con $N = 500$. ¿Se aprecia algún cambio en el tamaño y/o en el poder del test ante el incremento de N?

Universidad Torcuato Di Tella
Maestría en Economía y Econometría
2022

Datos de Panel
Problem Set 2
Modelos de Datos de Panel Lineales

---

1. Utilice nuevamente la base de datos *"cornwell.dta"* provista para el Problem Set 1. Considere el siguiente modelo de regresión:

$$\begin{aligned} \ln crmrte_{it} =& \beta_0 + \beta_1 \ln prbarr_{it} + \beta_2 \ln prbconv_{it} + \beta_3 \ln prbpris_{it} + \beta_4 \ln avgsen_{it} \\ &+ \beta_5 \ln polpc_{it} + \sum_{\tau=82}^{87} \beta_\tau \cdot I\{t=\tau\} + \mu_i + \varepsilon_{it} \end{aligned}$$

   a) Utilizando el comando *egen* de STATA, construya las medias individuales de las variables del modelo.
   b) Aplique la transformación *within* al modelo. Luego, estime el modelo transformado por POLS.
   c) Comente sobre la validez de los errores estándar del inciso previo.
   d) Utilice el comando *xtreg* para estimar nuevamente el modelo usando efectos fijos.
   e) Estime el modelo usando diferencias finitas de primer orden.

2. Utilice la base de datos provista *"murder.dta"*. La base de datos es una muestra longitudinal de estados de EE.UU., para los años 1987, 1990 y 1993.

   a) Estime por OLS el efecto de las ejecuciones (x) sobre la tasa de homicidios (*murder rates*, m) controlando por desempleo (u) y año:

   $$m_{i,t} = \alpha + \beta_x x_{i,t} + \beta_u u_{i,t} + \beta_{90} d_{90,t} + \beta_{93} d_{93,t} + \nu_{i,t}$$

   Note que se omitió la dummy temporal para el año 1987. Interprete los resultados.
   b) ¿Por qué podría ser importante tener en consideración los efectos temporales agregados en el modelo?
   c) Ahora, considere la siguiente modificación en el modelo:

   $$m_{i,t} = \alpha + \beta_x x_{i,t} + \beta_u u_{i,t} + \beta_{90} d_{90,t} + \beta_{93} d_{93,t} + c_i + e_{i,t}$$

   donde c_i es un efecto individual por estado. Estime la ecuación usando efectos fijos.
   d) Repita la estimación del inciso previo usando diferencias finitas de primer orden.
   e) Brinde un ejemplo bajo el cual la variable de ejecuciones no sería estrictamente exógena (condicional en c_i). ***Observación.*** Para obtener estimaciones consistentes, el modelo de efectos fijos asume exogeneidad estricta de las variables explicativas condicionadas en c_i.
   f) Repita la estimación del inciso c) usando el estimador de GLS para diferencias finitas de primer orden. Compruebe que los coeficientes estimados son iguales a los obtenidos por FE.
   g) Reestimar el modelo del inciso c) usando efectos aleatorios. Implementar el test de Hausman. ¿Cuál es el mejor estimador?

3. Considere el siguiente modelo:

$$y_{it} = x_{it}\beta + \mu_i + \nu_{it}, \quad i = 1, 2, \ldots, N, \; t = 1, 2, \ldots, T$$

donde $x_{it} \overset{iid}{\sim} .\mathcal{N}(0,1)$, $u_i \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\mu^2\right)$, $\nu_{it} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\nu^2\right)$ y $\mu_i \perp \nu_{it}$ para todo i, t. Suponga que $\beta = \sigma_\mu^2 = \sigma_\nu^2 = 1$ y $T = 10$. La idea es realizar experimentos de Monte Carlo para evaluar la eficiencia de distintos estimadores de β.

   a) Caso 1: $N = 5$. Realice un experimento de Monte Carlo con 1000 simulaciones. Reporte media, desvío estándar y RMSE de la estimación de β usando: POLS, RE y FE.

   b) Repita el punto anterior con $N = 10, 30, 50, 100$ y 500.

   c) Comente los resultados obtenidos y su conclusión de qué estimador debiera utilizarse en la práctica.

4. *Basado en el Ejercicio 10.18 de Wooldridge (2010).* Utilice la base de datos *wagepan.dta* para responder las preguntas a continuación.

   a) Utilizando *lwage* como variable dependiente, estimar un modelo que contenga un intercepto y las variables *dummy* de año $d81$ a $d87$. Estime el modelo por POLS, RE, FE y FD. ¿Qué puede concluir acerca de los coeficientes de las variables *dummy*?

   b) Añada las variables constantes en el tiempo *educ*, *black* e *hisp* al modelo, y estímelo por POLS y RE. ¿Cómo se comparan los coeficientes? ¿Qué ocurre si se estima la ecuación por FE?

   c) ¿Son iguales los errores estándar de POLS y RE del inciso b)? ¿Cuáles son probablemente más fiables?

   d) Obtenga los errores estándar robustos para POLS. ¿Prefiere estos o los errores estándar habituales de RE?

   e) Obtenga los errores estándar robustos de RE. ¿Cómo se comparan con los errores estándar robustos de POLS, y por qué?

Universidad Torcuato Di Tella
Master en Economía y Econometría
2021

Datos de Panel
Problem Set 3: Modelos de Panel Dinámicos

1. Considere la base de datos *mod_abdata.dta* que fue utilizada por Arellano y Bond en su famoso paper de 1991. Se trata de un panel de 140 empresas británicas encuestadas anualmente entre 1976 y 1984. El panel original no es balanceado, pero la versión para este ejercicio se trata de un panel balanceado de empresas con observaciones para exactamente 6 años entre 1977 y 1982. La variable que identifica la empresa es *id* y la variable que identifica el tiempo es *year*. La variable n es el empleo de la empresa. Luego, considere un modelo muy simplificado del siguiente tipo:

$$\begin{aligned} \ln n_{it} &= \rho \ln n_{it-1} + \varepsilon_{it} \\ \varepsilon_{it} &= c_i + \nu_{it} \\ E[c_i] &= E[\nu_{it}] = E[c_i \nu_{it}] = 0 \end{aligned}$$

donde n_{it} es el empleo de la empresa i en el año t.

a) Estime el modelo por OLS. ¿Qué sesgo esperaría encontrar y por qué?

b) Estime el modelo usando efectos fijos (FE). ¿Permite la transformación *within* eliminar el sesgo de paneles dinámicos?

c) Considere una transformación de diferencias finitas de primer orden del modelo. ¿Continúa siendo la variable dependiente rezagada potencialmente endógena?

d) Implemente el estimador de Anderson-Hsiao a partir del comando *ivregress* en Stata.

e) Ahora, obtenga la estimación GMM de ρ utilizando todos los instrumentos posibles en niveles para el modelo en primeras diferencias. Para ello utilice el comando *xtabond2*.

f) Obtenga la estimación de GMM de ρ utilizando todos los instrumentos posibles en niveles para el modelo en primeras diferencias e Δy_{it-1} como instrumento para el modelo en niveles.

g) Repita las estimaciones de los incisos e) y f) incluyendo efectos fijos de tiempo.

2. En este ejercicio se ilustrará el hecho de que los estimadores de Arellano-Bond y de Blundell-Bond pueden extenderse en forma directa a modelos que incluyan regresores estríctamente exógenos y regresores secuencialmente exógenos.

   En su paper original, Arellano y Bond modelaron el empleo de las empresas (n) utilizando un modelo de ajuste parcial para reflejar los costos de contratación y despido, incluyendo dos rezagos de la variable empleo. Otras variables incluidas fueron el nivel salarial actual y el rezagado (w), el stock de capital actual, rezagado una y dos veces (k) y la producción agregada actual, rezagada una y dos veces en el sector de la empresa (ys). Todas las variables se expresan en logaritmos. También se incluye un conjunto de variables *dummy* de tiempo.

   a) Estime el modelo por OLS. Compute los errores estándar robustos a heterocedasticidad y correlación serial.

   b) Estime el modelo por FE. Compute los errores estándar robustos a heterocedasticidad y correlación serial.

   c) Implemente el estimador de Anderson-Hsiao usando n_{it-2} como instrumento.

   d) Estime la ecuación de empleo usando el estimador de Arellano-Bond. Asuma que la única endogeneidad presente es en el rezago de la variable dependiente.

   e) Ahora, considere como hicieron Blundell y Bond (1998) que los salarios y el stock de capital no deben tomarse como estrictamente exógenos en este contexto (como se hizo en los modelos anteriores). Reestime el modelo usando el estimador de A-B y considerando a los salarios y el stock de capital como regresores secuencialmente exógenos.

   f) Adicionalmente, Blundell y Bond (1998) eliminan de su modelo los rezagos más largos (de dos períodos) del empleo y el capital, y prescinden del nivel de producto agregado sectorial. Considerando esta cuestión, compute el estimador de Blundell-Bond.

3. Cuando hay muchos instrumentos, surgen dos problemas principales:

- Sobreestimación (*overfitting*) de la variable endógena
- Mala estimación de la matriz de pesos W

En estos casos, se proponen las siguientes soluciones:

a) Probar diferentes especificaciones de IV recortando el número de rezagos en la matriz de instrumentos $\mathbf{Z}$.

b) Colapsar/combinar instrumentos. Se modifica la matriz de instrumentos para el individuo i:

$$\mathbf{Z}_i = \begin{bmatrix} y_{i1} & 0 & 0 & \cdots \\ y_{i1} & y_{i2} & 0 & \cdots \\ y_{i3} & y_{i2} & y_{i1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Si el modelo funciona debería dar resultados similares con distintos instrumentos.

Retome el ejercicio 2.e) para ver una aplicación de esta cuestión. Estime el modelo de empleo restringiendo el máximo rezago a 3 y 4 períodos. Por último, estime el modelo colapsando instrumentos. Analice si los resultados obtenidos son robustos.

4. Considere nuevamente el modelo del primer ejercicio. Obtenga el estimador LSDVC propuesto por Kiviet (1995) a partir del comando *xtlsdvc*. Luego, estime la matriz de varianzas y covarianzas de los coeficientes de Kiviet siguiendo el procedimiento explicado en clase.

Universidad Torcuato Di Tella
Maestría en Economía y Econometría
2022

Datos de Panel
Problem Set 4
Modelos Lineales en Paneles Desbalanceados

---

1. Utilice la base de datos *"keane.dta"* la cual contiene el historial de empleo y escolaridad de una muestra de hombres para los años 1981 a 1987. Luego, considere la siguiente ecuación de salarios:

$$\ln(wage_{it}) = \beta_0 + \beta_1 exper_{it} + \beta_2 educ_{it} + c_i + u_{it}, \quad t = 1, 2, \ldots, T \tag{1}$$

   donde $\ln(wage_{it})$ es el logaritmo del salario por hora, $exper_{it}$ son los años de experiencia en el mercado laboral y $educ_{it}$ son los años de escolaridad. Responda las siguientes preguntas:

   a) Estime la ecuación usando efectos fijos. ¿Cuál es el sesgo potencial en este contexto?

   b) Implemente el contraste de sesgo de selección propuesto por Wooldridge (1995) bajo el enfoque de Mundlak (1978).

   c) Implemente el contraste de sesgo de selección propuesto por Wooldridge (1995) bajo el enfoque de Chamberlain (1980).

2. Considerando nuevamente la ecuación de salarios del ejercicio previo, realice los siguientes procedimientos:

   a) Estime el modelo por Wooldridge (1995) bajo el enfoque de Chamberlain (1980).

   b) Estime el modelo por Wooldridge (1995) bajo el enfoque de Mundlak (1978).

   c) Comente sobre los errores estándar de las estimaciones anteriores.

   d) Estime los errores estándar vía bootstrapping.

   e) Estime los errores estándar analíticos (varianza asintótica).

Universidad Torcuato Di Tella
Maestría en Economía y Econometría
2022

Datos de Panel
Problem Set 5
Modelos de Variable Dependiente Discreta

---

1. El archivo *wagepan.dta* contiene los datos utilizados por Vella y Verbeek (1998). Estos datos contienen información para 545 hombres que trabajaron cada año de 1980 a 1987. Utilice los datos para analizar el impacto de la escolaridad ($educ_{it}$) en la probabilidad de estar afiliado a un sindicato ($union_{it}$). Las variables se describen en el conjunto de datos. Observe que la educación no cambia con el tiempo.

   a) Use Pooled OLS para estimar el modelo:

   $$P\left(union_{it}=1 \mid educ_{it}\right)=\beta_0+\beta_1 educ_{it} \tag{1}$$

   De acuerdo a los resultados obtenidos, ¿tiene impacto un año más de escolaridad sobre la probabilidad de estar afiliado a un sindicato?

   b) Use Pooled Probit para estimar el modelo:

   $$P\left(union_{it}=1 \mid educ_{it}\right)=\Phi\left(\beta_0+\beta_1 educ_{it}\right) \tag{2}$$

   Comente sobre el impacto de un año más de educación en la probabilidad de estar afiliado a un sindicato.

   c) Use Pooled Logit para estimar el modelo:

   $$P\left(union_{it}=1 \mid educ_{it}\right)=\Lambda\left(\beta_0+\beta_1 educ_{it}\right)=\frac{e^{\beta_0+\beta_1 educ_{it}}}{1+e^{\beta_0+\beta_1 educ_{it}}} \tag{3}$$

   Comente sobre el impacto de un año más de educación en la probabilidad de estar afiliado a un sindicato. Compute el error estándar para esta estimación.

   d) Estime la siguiente extensión del modelo (2):

   $$P\left(union_{it}=1 \mid educ_{it}, c_i\right)=\Phi\left(\beta_0+\beta_1 educ_{it}+c_i\right) \tag{4}$$

   donde c_i son efectos no observables individuales. Use el modelo Probit de efectos aleatorios. ¿Cuál es el problema que surge al momento de estimar el efecto parcial de interés?

   e) Estime la siguiente extensión del modelo (3):

   $$P\left(union_{it}=1 \mid educ_{it}, c_i\right)=\Lambda\left(\beta_0+\beta_1 educ_{it}+c_i\right) \tag{5}$$

   donde c_i son efectos no observables individuales. Use el modelo Logit de efectos aleatorios. ¿Surge el mismo problema que en el inciso anterior al momento de estimar el efecto parcial de interés?

f) Compute el denominado estimador Logit de efectos fijos para el modelo (5). ¿Se puede computar el efecto de un año más de educación sobre la probabilidad de estar afiliado a un sindicato? Explique.

g) Considere la siguiente extensión del modelo (4):

$$\begin{aligned} &P\left(union_{it} = 1 \mid educ_{it}, black_{it}, married_{it}, c_i\right) = \\ &\Phi\left(\beta_0 + \beta_1 educ_{it} + \beta_2 black_{it} + \beta_3 married_{it} + c_i\right) \end{aligned} \tag{6}$$

donde $black_{it}$ es una variable binaria que toma valor 1 si la persona es afroamericana y $married_{it}$ es una variable binaria que toma valor 1 si la persona es casada. Asuma la siguiente versión de Mundlak (1978) del supuesto de Chamberlain (1980):

$$c_i \mid X_i \sim Normal\left(\psi + \xi \cdot \overline{married}_i, \sigma_a^2\right) \tag{7}$$

El modelo dado por (6) y (7) es un caso de lo que en la literatura se denomina modelo Probit de efectos aleatorios de Chamberlain. Al asumir solamente (6) y (7) se tiene que:

$$\begin{aligned} &P\left(union_{it} = 1 \mid educ_{it}, black_{it}, married_{it}\right) = \\ &\Phi\left[\left(\beta_0 + \beta_1 educ_{it} + \beta_2 black_{it} + \beta_3 married_{it} + \psi + \xi \overline{married}_i\right)\left(1 + \sigma_a^2\right)^{-1/2}\right] \equiv \\ &\Phi\left[\beta_{0,a} + \beta_{1,a} educ_{it} + \beta_{2,a} black_{it} + \beta_{3,a} married_{it} + \xi_a \overline{married}_i\right] \end{aligned}$$

Use Pooled Probit para estimar el modelo. Estime el efecto de la escolaridad sobre la probabilidad de estar sindicalizado para una persona afroamericana casada.

2. Considere los datos del ejercicio previo para analizar la probabilidad de estar afiliado a un sindicato según la situación de afiliación sindical del año previo.

a) Use Pooled Probit para estimar el modelo:

$$P\left(union_{it} = 1 \mid union_{it-1}\right) = \Phi\left(\psi + \rho \cdot union_{it-1}\right) \tag{8}$$

A continuación, obtenga una estimación para

$$P\left(union_{it} = 1 \mid union_{it-1} = 1\right)$$

y para

$$P\left(union_{it} = 1 \mid union_{it-1} = 0\right).$$

. Comente sobre el efecto marginal de estar afiliado a un sindicato en el año $t-1$ en la probabilidad de estar afiliado a un sindicato en el año t.

b) Adicione al modelo el conjunto completo de variables binarias temporales. Vuelva a estimar las probabilidades solicitadas para cada año de la muestra.

c) Estime un modelo de efectos no observables dinámico. Use el modelo Probit de efectos aleatorios incluyendo $union_{i,80}$ como una variable explicativa adicional. Luego, promedie las probabilidades estimadas a lo largo de $union_{i,80}$ para obtener la probabilidad promedio de estar afiliado a un sindicato en el año 1987 dado que estaba afiliado en el período anterior.