Universidad Torcuato Di Tella
Maestría en Economía
Econometría de datos de panel
Trabajo Práctico 1 2023

Oscar Jaramillo

Propiedades de Muestra Finita en Paneles No Balanceados

Considere el siguiente modelo:

$$y^*_{jt} = \beta x_{jt} + c_j + u_{jt}, \quad j = 1, 2, ...N; \quad t = 1, 2, ...T. \tag{1}$$

$$s^*_{jt} = \gamma_1 \omega_{jt} + \gamma_2 z_{jt} + \alpha_j + \epsilon_{jt}$$

$$s_{jt} = 1, \quad \text{si } s^*_{jt} > 0, \quad \text{en cualquier otro caso}$$

$$y_{jt} = y^*_{jt} \times s_{jt}$$

Las dos variables de la ecuación de selección (ω y z) son independientes y normalmente distribuidas con media cero y varianza uno. La única variable de la ecuación de interés es $x_{jt} = \omega_{jt}$, por lo tanto tenemos una viariable que está excluida de la ecuación de interés. El error idiosincrático de la ecuación de selección, ϵ_{jt}, sigue una distribución normal con media cero y varianza uno. El error idiosincrático de la ecuación de interés (1) se define como: $u_{jt} = 0{,}6\epsilon_{jt} + 0{,}8\psi_1$ siendo una variable independiente y normal estándar. Con esta especificación la correlación etre los errores idiosincráticos de las ecuaciones de interés y de selección es 0.6. Además defina tres variables aleatorias: ψ_2, ψ_3 y ψ_4. Estas variables son específicas de corte transversal. Son todas independientes y normal estándar. Con esta especificación general defina tres modelos:

Modelo A: Este modelo asume que los dos términos de efectos no observables, c_j y α_j están correlacionados. Más especificamente:

$$\alpha_j = \psi_2 + \psi_4 \tag{2}$$

$$c_j = \psi_3 + \psi_4 \tag{3}$$

Modelo B: Este modelos asume que las variables observables y no observables están correlacionadas, pero que no hay correlación entre los efectos no observables entre ecuaciones.

$$\alpha_j = \psi_2 + \frac{z_{j1} + z_{j2}}{2} \tag{4}$$

$$c_j = \psi_3 + \frac{x_{j1} + x_{j2}}{2} \tag{5}$$

Modelos C: Este es el modelo más general porque permite ambos tipos de correlación entre los dos componentes no obsesrvables y entre estos componentes y las variables observables.

$$\alpha_j = \psi_2 + \frac{z_{j1} + z_{j2}}{2} + \psi_4 \tag{6}$$

$$c_j = \psi_3 + \frac{x_{j1} + x_{j2}}{2} + \psi 4 \quad (7)$$

Todos los parámetros son iguales a uno ($\beta = \gamma_1 = \gamma_2 = 1$). $N = 20, 40, 100$ y $T = 2, 10$

1. Caso 1: Para cada modelo y combinación de $N = 20, 40, 100$ y $T = 2$ realice un experimento de Monte Carlo de $S = 1000$ simulaciones y reporte el sesgo medio, el sesgo mediano, el error estándar, el RMSE y la desviación media absoluta de la estimación de Wooldridge para los tres parámetros del modelo (β, γ_1, y γ_2). Explique sus resultados:

   La primera parte del código en el archivo Jaramillo1263_1.1.do realiza un experimento de Montecarlo para los tres modelos y resume los resultados. Los datos son generados con las especificaciones pedidas y para la estimación de Wooldridge se utilizó el enfoque de Mundlak (modelo B y modelo C), para el modelo A no se incluyó la variable $\bar{z}_j$ como parte de la estimación debido a que la heterogeneidad no observada no está relacionada con las variables independientes.

   En cuanto a los modelos B y C donde se permite que la heterogeneidad no observada esté relacionada con z_{jt}, se cumplen todos los supuestos necesarios para estimar el modelo bajo una corrección por truncamiento incidental como en Wooldridge 1995. Los datos son los siguientes:

   Modelo A

   Los resultados para el modelo A son:

Cuadro 1: Sesgo modelo A

Sesgo del estimador		Tamaño de la muestra		
		20	40	100
β	Sesgo medio	-0.0195	0.01142	0.00832
	Sesgo mediano	-0.0435	-0.0097	0.00903
	Error estándar	0.81143	0.43865	0.25202
γ_1	Sesgo medio	-0.3211	-0.3885	-0.4091
	Sesgo mediano	-0.3664	-0.4009	-0.4116
	Error estándar	0.33079	0.19620	0.11260
γ_2	Sesgo medio	-0.3165	-0.3766	-0.4073
	Sesgo mediano	-0.3652	-0.3975	-0.4182
	Error estándar	0.33791	0.20531	0.11549

   Vemos que el sesgo medio y mediano del estimador de β del modelo A tiende a cero a medida que aumenta el tamaño de la muestra, de igual manera la desviación entre las estimaciones del experimento disminuyen por lo que podemos decir que dado el modelo A que no permite una relación entre la heterogeneidad no observada y las variables independientes del modelo, el estimador de β será insesgado en muestras grandes. Y ya que el error estándar también disminuye existe evidencia que en muestras grandes, la estimación de Wooldridge del modelo A permite encontrar una estimación insesgada y de mínima varianza uniforme para β.

   En cuanto al sesgo de las estimaciones de γ_1 y γ_2 parece aumentar en promedio a medida que la muestra se hace mas grande, sin embargo la desviación se hace más pequeña lo que nos lleva a pensar que puede estar acercándose a algún tipo de asíntota o límite diferente de cero, esto quiere decir que la estimación de Wooldridge para los parámetros γ_1 y γ_2 no será insesgada y por lo tanto no puede cumplir el teorema de Markov. Esto se puede deber a la naturaleza de la heterogeneidad no observada que en el modelo A no se le permite estar relacionada con ninguna variable independiente, por lo que en lugar de la corrección de Wooldridge podría ser preferible una estimación de efectos fijos corrigiendo el truncamiento incidental.

Cuadro 2: RMSE modelo A

Ecuaciones estimadas	Tamaño de la muestra		
	20	40	100
Selección	0.99375	0.94571	0.92722
Interés	1.80527	1.10830	1.45620

Si comparamos la raíz de los errores promedio al cuadrado (RMSE) de las dos ecuaciones, vemos que para la ecuación de interés el RMSE disminuye a medida que aumenta el tamaño de la muestra por lo que en muestras grandes la corrección de Wooldridge con el modelo A estaría ajustando mejor la estimación de la ecuación de selección, el RMSE para la ecuación de selección representa el nivel de ajuste de la estimación de γ_1 y γ_2 así que en promedio las predicciones de la ecuación de selección estimada se desvían una unidad de la estimación real, y al hacerlo por ecuación estamos considerando el impacto conjunto de la estimación de γ_1 y γ_2.

En cuanto a la ecuación de interés, el RMSE alcanza un mínimo cuando el tamaño de la muestra es 40 y vuelve a subir cuando la muestra es 100, por lo que en muestras más grandes las predicciones del modelo A parecen tender a errarle a los datos reales, esto también se puede deber a que el RMSE es sensible a valores atípicos y como vimos en el cuadro anterior la estimación sesgada de γ_1 y γ_2 así como el truncamiento incidental podrían estar generando valores atípicos que afectan mucho las predicciones del modelo.

Cuadro 3: DMA modelo A

Estimaciones	Tamaño de la muestra		
	20	40	100
β	0.40946	0.25539	0.16813
γ_1	0.19882	0.12606	0.07465
γ_2	0.18152	0.12956	0.07221

La desviación media absoluta del modelo A para la estimación de β, γ_1 y γ_2 disminuye a medida que aumenta el tamaño de la muestra y para una muestra de tamaño 100 presenta valores relativamente bajos por lo que sirve como contraste del cuadro anterior y nos habla de la precisión del modelo, podemos decir que en muestras grandes la corrección del sesgo de truncamiento incidental de Wooldridge producirá estimadores cercanos a los reales.

Modelo B

Los resultados para el modelo B son:

Cuadro 4: Sesgo modelo B

Sesgo del estimador		Tamaño de la muestra		
		20	40	100
β	Sesgo medio	-0.0370	-0.0015	-0.0099
	Sesgo mediano	-0.4282	-0.0024	-0.0117
	Error estándar	0.56722	0.35118	0.23161
γ_1	Sesgo medio	18.8586	-0.2239	-0.2701
	Sesgo mediano	-0.1900	-0.2477	-0.2779
	Error estándar	348.872	0.24664	0.13503
γ_2	Sesgo medio	7.12403	-0.2308	-0.2630
	Sesgo mediano	-0.1737	-0.2514	-0.2698
	Error estándar	190.410	0.33374	0.17745

Los resultados del experimento para el modelo B son similares para β, un aumento del tamaño de la muestra disminuye el sesgo acercándose cada vez más a cero, igual para el error estándar, por lo que la estimación de β en muestras grandes podría cumplir el teorema de

Markov y ser insesgada.

Sin embargo si nos fijamos en el sesgo medio y mediano de los estimadores γ_1 y γ_2 parecen alcanzar un mínimo en una muestra de tamaño 40, lo que parece señalar que la corrección de Wooldridge en este modelo arroja estimadores sesgados que no podrían cumplir el teorema de Markov, nótese también que para una muestra de tamaño 20, el sesgo medio y mediano de γ_1 y γ_2 así como el error estándar arrojan valores relativamente altos lo que parece indicar que en muestras pequeñas las estimaciones son "bastante"sesgadas.

Cuadro 5: RMSE modelo B

Ecuaciones estimadas	Tamaño de la muestra		
	20	40	100
Selección	33.4404	1.14395	1.07211
Interés	1.17416	1.28365	1.34874

Al considerar la raíz de los errores promedio al cuadrado del modelo B donde la heterogeneidad no observada si tiene permitido estar relacionada con las variables independientes, vemos que al igual que con el modelo A, a medida que aumenta el tamaño de la muestra disminuye el RMSE y en particular es menor que 1 para la estimación de β.

Para la ecuación de interés en este modelo, al igual que con las estimaciones de γ_1 y γ_2 el RMSE parece tener un mínimo con una muestra de tamaño 40 lo que señala que la capacidad predictiva del modelo en muestras pequeñas sería mejor que en muestras grandes, lo que coincide con las propiedades de muestra finita que vimos para el estimador de β.

Cuadro 6: DMA modelo B

Estimaciones	Tamaño de la muestra		
	20	40	100
β	0.35908	0.22742	0.15401
γ_1	0.25279	0.16250	0.09119
γ_2	0.31650	0.20802	0.12405

La desviación media absoluta del modelo B, al igual que en el modelo A es decreciente mientras aumenta el tamaño de la muestra, va de la mano con el sesgo de γ_1 y γ_2 y la insesgadez de β en muestras grandes, si lo relacionamos con el RMSE podemos decir que en muestras "pequeñas"las predicciones de ambas ecuaciones tienen más oportunidades de alcanzar el valor real si bien habrá sesgo, especialmente para la ecuación de selección.

Modelo C

Los resultados para el modelo C son:

Cuadro 7: Sesgo modelo C

Sesgo del estimador		Tamaño de la muestra		
		20	40	100
β	Sesgo medio	-0.0462	0.00662	-0.0069
	Sesgo mediano	-0.0915	-0.0082	-0.0085
	Error estándar	0.72647	0.42781	0.26798
γ_1	Sesgo medio	1.59946	-0.3731	-0.4016
	Sesgo mediano	-0.3276	-0.3959	-0.4028
	Error estándar	42.6974	0.21258	0.12194
γ_2	Sesgo medio	2.14184	-0.3612	-0.4027
	Sesgo mediano	-0.3178	-0.3846	-0.4121
	Error estándar	59.5408	0.29014	0.15925

El modelo C es un modelo más general que los dos anteriores ya que permite dos diferentes tipos de correlación entre la heterogeniedad no observada de la ecuación de selección y de la ecuación de interés, a través de z_j como en el modelo B y a través de ψ_4 como en el modelo A, por lo que podríamos esperar resultados similares. En el cuadro 7 se puedo ver que el sesgo para β disminuye a medida que aumenta el tamaño de la muestra y de igual manera los errores estándar lo que nos lleva a pensar en una estimación insesgada que cumple el teorema que Markov y es mínimo en la clase de estimadores lineales insesgados.

Al igual que en el modelo B vemos que las estimaciones de γ_1 y γ_2 en el modelo C tienen un mínimo donde el tamaño de la muestra es 40, lo que además también tiende a confirmar el resultado anterior, es en muestras finitas donde la estimación de Wooldridge permite acercarnos de mejor manera a los valores reales, teniendo estimaciones sesgadas en muestras grandes para γ_1 y γ_2.

Cuadro 8: RMSE modelo C

Ecuaciones estimadas	Tamaño de la muestra		
	20	40	100
Selección	5.51767	1.02228	0.97339
Interés	1.36413	1.49763	1.57099

Al igual que en el modelo B y en el model A, vemos que el RMSE del modelo C disminuye para la ecuación de selección, y tiene un mínimo en una muestra de tamaño 40 para la ecuación de interés, esto va conforme a las propiedades ya enunciadas, este modelo más general mantiene las propiedades ya vistas en los modelos B y A, donde el sesgo de γ_1 y γ_2 disminuye en muestras finitas y el sesgo del estimador de β tiende a cero, se debe utilizar con mucho cuidado las predicicones de los tres modelos. Se puede ver también que para una muestra de tamaño 20 se encuentran errores muy atípicos que son recogidos y .$^{\text{ex}}$agerados"por el RMSE en una muestra de tamaño 20, esta particularidad disminuye a medida que se aumenta el tamaño de la muestra debido a las propiedades de muestra finita de los estimadores.

Cuadro 9: DMA modelo C

Estimaciones	Tamaño de la muestra		
	20	40	100
β	0.44150	0.28388	0.18039
γ_1	0.24514	0.13148	0.08152
γ_2	0.27675	0.17742	0.10089

Finalmente el DMA del modelo C también disminuye al igual que en los otros modelos, en general, recoge los efectos de los residuos que vimos especialmente en el modelo B, ya que este modelo es una generalización del modelo B, de igual manera se debe tener cuidado al utilizar las predicciones de este modelo, sin embargo podemos confiar en la insesgadez y la varianza mínima en la clase de estimadores insesgados del parámetro β, lo que quiere decir que la estimación de Wooldridge si es capaz de brindar estimadores insesgados corrigiendo el truncamiento incidental, si bien la estimación no es capaz de replicar la ausencia de datos, lo cual por supuesto no es necesario dado que esa información no es necesariamente de interés.

En cuanto a los tres modelos, el modelo B y el modelo C son muy cercanos, el modelo C implica una generalización del modelo B, ambos modelos permiten que la heterogeneidad no observada se relaciona con las variables independientes, lo que no sucede en el modelo A, es justo esta diferencia la que puede evitar encontrar un mínimo en el sesgo de las estimaciones de γ_1 y γ_2 debido a los supuestos necesarios bajo los cuales la coreción de Wooldridge arroja estimadores insesgados. En el caso del modelo A donde la heterogeneidad no observada no se le permite relacionarse con las variables independientes puede ser mejor utilizar una estimación de FE con una transformación Within corrigiendo por truncamiento incidental a la Heckman, o con menor probabilidad de éxito realizar una estimación de FE aumentando los promedios de las variables a la Mundlak o las variables dicotómicas perdidas a la Chamberlain.

En general las predicciones de los tres modelos deben ser tomados con cuidado sin embargo, los modelos B y particularmente el C pueden ser usados como predicciones insesgadas de β con una muy buena capacidad explicativa como nos lo resumen los resultados de RMSE y DMA, si existiera evidencia para utilizar el modelo B debería descartarse el modelo C, sin embargo en ausencia de esa evidencia utilizar el modelo general puede ser mejor ya que permite dos tipos diferentes de correlación entre la heterogeneidad no observada de la ecuación de interés y de la ecuación de selescción. Predecir γ_1 y γ_2 no sería necesario en un contexto real, pero en el caso que se deseara hacerlo la estimación probit de Wooldridge solamente permite encontrar un sesgo mínimo pero producirá estimadores sesgados en muestras finitas, debido a la especificación de los valores no observados. En un experimento donde los datos estén truncados de forma no aleatoria esta estimación podría ser más compleja y podría requerir algún otro tipo de estimación.

2. Caso 2: Para cada modelo y combinación de $N = 20, 40, 100$ y $T = 10$ considere lo siguiente.

   El código para esta segunda parte se encuentra en Jaramillo1263_1,2 si bien ambos códigos podían estar unidos, se separaron para mejorar la presentación y para computar de mejor forma los resultados.

   a) Tome la primera generación de los datos ($S = 1$) como si fuera una muestra obtenida de la realidad.

   b) Estime el modelo de Wooldridge y guarde los coeficientes estimados (llámelos $\beta^{(b)}$, $\gamma_1^{(b)}$, y $\gamma_2^{(b)}$) como así también los residuos de la ecuación de selección, $\epsilon_{jt}^{(b)}$.

   c) Realice un procedimiento de bootstrapping para construir intervalos de 95 % de confiabiabilidad para los tres estimadores de la siguiente manera: (i) utilice un procedimiento de muestreo aleatorio simple con reemplazo para obtener una nueva muestra de errores de la ecuación de selescción (Nota: observer que debe generar para cada j una nueva muestra de residuos $\epsilon_{jt}^{(b)}$ de dimensión T). Con estos nuevos residuos, con $\gamma_1^{(b)}$, y $\gamma_2^{(b)}$ y con los valores originales de ω, z (es decir, con los valores generados para $S = 1$) y α_j construya una nueva variables dependeinte del modelo de selección. (ii) Usando esta nueva variables dependiente del modelo de selección, los valores de $\beta^{(b)}$ y x_{jt} y c_j construya una nueva variable dependiente de la ecuación de interés. (iii) COn la nuevas variables dependientes de ambas ecuaciones construidas, estime por Wolldridge los tres parámetros del modelo. (iv) Vuelva al paso (i) y repita el procedimiento. Repit (iv) para construir $B = 1000$ muestras de bootstrapping y en cada caso guarde los coeficientes estimados por Wooldridge de β, γ_1 y γ_2. Reporte, para $B = 1000$, intervalos de bootstrpping de 95 % de confiabilidad. Compare estos resultados con intervalos de 95 % de significatividad construidos a partir de la matriz de varianzas y covarianzas asintótica para el modelo estimado con los datos de $S = 1$. ¿Qué conclusiones puede sacar?

   Vamos a resumir los resultados de todo el proceso para cada uno de los parámetros, para cada modelo y cada tamaño de muestra, los intervalos de confianza de bootstrap y los intervalos del modelo real $S = 1$ de la primera realización de los datos:

   Los resultados para γ_1:

Cuadro 10: Intervalos de confianza γ_1, modelo A

Tamaño de la muestra		Modelo A, $\gamma_1 = 0{,}59465$	
		Límite inferior	Límite superior
20	Real (S=1)	0.33022	0.78978
	Bootstrap	0.15177	0.43775
40	Real (S=1)	0.36235	0.65838
	Bootstrap	0.24249	0.44032
100	Real (S=1)	0.49723	0.69206
	Bootstrap	0.33556	0.45820

Cuadro 11: Intervalos de confianza γ_1, modelo B

Tamaño de la muestra		Modelo B, $\gamma_1 = 0{,}75294$	
		Límite inferior	Límite superior
20	Real (S=1)	0.44220	0.96098
	Bootstrap	0.38329	0.84996
40	Real (S=1)	0.49169	0.83063
	Bootstrap	0.36725	0.64601
100	Real (S=1)	0.64215	0.86373
	Bootstrap	0.49826	0.67204

Cuadro 12: Intervalos de confianza γ_1, modelo C

Tamaño de la muestra		Modelo C, $\gamma_1 = 0{,}576881$	
		Límite inferior	Límite superior
20	Real (S=1)	0.33778	0.81403
	Bootstrap	0.33141	0.76148
40	Real (S=1)	0.35952	0.66104
	Bootstrap	0.30073	0.56778
100	Real (S=1)	0.47892	0.67428
	Bootstrap	0.42858	0.57807

En los cuadros 10, 11 y 12, vemos los resultados de los intervalos de confianza al 95 % para γ_1 según el modelo real que es la primera realización de los datos aleatorios, a esta primera estimación se remuestreo a la bootstrap los errores para calcular los nuevos estimadores para la ecuación de selección y para la ecuación de interés, de igual manera realizando la corrección de Wooldridge para el truncamiento incidental. Este procedimiento se repite de forma general para los tres modelos y para cada parámetro, considerando por supuesto las definiciones del ejercicio y añadiendo el promedio de la variable independiente para el modelo B y C según el enfoque de Mundlak cuando la heterogeneidad no observada puede depender de las variables independientes.

En el caso de γ_1 vemos que el intervalo de confianza que proviene de Bootstrap es más pequeño que el real en los tres modelos, a medida que aumenta el tamaño de la muestra los límites inferior y superior del modelo A y del modelo B aumentan la distancia entre los límites se reduce, lo que va de la mano con las conclusiones anteriores.

Para γ_1 los límtes superiores del modelo C del modelo real y de botstrap parecen encontrar un mínimo con un tamaño de muestra 40, a partir de aquí aumentan en el modelo C, esto también sucede con menor fuerza en el modelo B, esto se puede deber a la inclusión de una variable en las regresiones que es el promedio de las variables anteriores siguiendo el enfoque de Mundlak, también va de la mano con las conclusiones del caso 1 de las propiedades de muestra finita de los estimadores.

Se ve también que el intervalo de confianza de bootstrap no solo es más chico para los tres modelos sino que tiene problemas para atrapar al parámetro γ_1 esto se puede deber a que como vimos en el experimento de Montecarlo, hay razones para creer que la estimación de Wooldridge de γ_1 y γ_2 es sesgada para 1, Bootstrap asume que la muestra es insesgada, esta puede ser la razón por la que bootrap no puede atrapar al parámetro real, además a medida que aumenta el tamaño de la muestra, el intervalo de bootstrap llega a capturar al parámetro al menos en el modelo C (más general) lo cual es evidencia a favor del sesgo de las estimaciones.

Los resultados para γ_2:

Cuadro 13: Intervalos de confianza γ_2, modelo A

Tamaño de la muestra		Modelo A, $\gamma_2 = 0{,}56703$	
		Límite inferior	Límite superior
20	Real (S=1)	0.37808	0.82654
	Bootstrap	0.26309	0.52973
40	Real (S=1)	0.48188	0.79871
	Bootstrap	0.40361	0.59106
100	Real (S=1)	0.46644	0.66762
	Bootstrap	0.34523	0.46921

Cuadro 14: Intervalos de confianza γ_2, modelo B

Tamaño de la muestra		Modelo B, $\gamma_2 = 0{,}76609$	
		Límite inferior	Límite superior
20	Real (S=1)	0.52067	1.04379
	Bootstrap	0.49738	0.95288
40	Real (S=1)	0.68237	1.08085
	Bootstrap	0.59857	0.89678
100	Real (S=1)	0.64503	0.88715
	Bootstrap	0.51391	0.69592

Cuadro 15: Intervalos de confianza γ_2, modelo C

Tamaño de la muestra		Modelo C, $\gamma_2 = 0{,}58319$	
		Límite inferior	Límite superior
20	Real (S=1)	0.34321	0.81544
	Bootstrap	0.38265	0.78811
40	Real (S=1)	0.52184	0.86355
	Bootstrap	0.52852	0.82111
100	Real (S=1)	0.47659	0.68978
	Bootstrap	0.43288	0.59632

En el caso de γ_2 y aumentando el análisis para γ_1 vemos que de igual manera el intervalo de bootstrap es más pequeño que el intervalo real, y que en una muestra de tamaño 40 encuentra un mínimo, en el caso de γ_2 se aprecia con mayor claridad, con una muestra de 100, el intervalo de bootstrap no es capáz de atrapar el valor verdadero del parámetro, por lo que podriamos decir que la estimación de Wooldridge de los parámetros de la ecuación de selección se 'comporta mejor' en muestras pequeñas, igual hay que considerar que en Montecarlo, tanto γ_1 como γ_2 parecen ser sesgados y que bootstrap asume que la muestra inicial no tiene estos problemas, por lo que se debe tener cuidado al utilizar estos resultados.

Los resultados para β:

Cuadro 16: Intervalos de confianza β, modelo A

Tamaño de la muestra		Modelo A, $\beta = 1{,}02615$	
		Límite inferior	Límite superior
20	Real (S=1)	0.79143	1.64668
	Bootstrap	1.13857	1.78828
40	Real (S=1)	0.66316	1.28698
	Bootstrap	0.78306	1.23224
100	Real (S=1)	0.81767	1.23462
	Bootstrap	0.90754	1.24837

Cuadro 17: Intervalos de confianza β, modelo B

Tamaño de la muestra		Modelo B, $\beta = 0{,}91801$	
		Límite inferior	Límite superior
20	Real (S=1)	0.58789	1.26299
	Bootstrap	0.69029	1.21641
40	Real (S=1)	0.57151	1.01846
	Bootstrap	0.57257	0.96386
100	Real (S=1)	0.76676	1.06944
	Bootstrap	0.73488	1.00378

Cuadro 18: Intervalos de confianza β, modelo C

Tamaño de la muestra		Modelo C, $\beta = 0{,}96321$	
		Límite inferior	Límite superior
20	Real (S=1)	0.76903	1.59468
	Bootstrap	0.89452	1.57793
40	Real (S=1)	0.54314	1.17906
	Bootstrap	0.67551	1.12253
100	Real (S=1)	0.77718	1.14923
	Bootstrap	0.80375	1.10970

En el caso de β vemos que los intervalos de confianza de bootrstrap tienen un mejor rendimiento son más pequeños que los intervalos reales de igual manera pero capturan al valor real, la amplitud del intervalo se reduce a medida que qumenta el tamaño de la muestra pero siguen capturando al valor real, esto se debe a que la estimación de Wooldridge de β es insesgada y ya que bootrap asume que la muestra inicial es representativa de la población subyacente, ya que remuestreamos los errores de la estimación y boostrap parece tener buenos resultados incluso más precisos que el intervalo real en el caso de una muestra de tamaño 40 y 100, esto es evidencia a favor de que los residuos de la estimación inicial que se uso de muestra representa los residuos del modelo por lo que es evidencia a favor del teormea de Markov, la estimación de β es 'buena'.

Hay que considerar que la corrección de Wooldridge tienen como objetivo principal estimar β considerando el sesgo por truncamiento incidental por lo que las estimaciones de γ_1 y γ_2 no tienen que ser necesariamente insesgadas y ni siquiera estar bien especificadas, simplemente reflejar el impacto que tiene el sesgo sobre la estimaciòn de β. La diferencia está en como se maneja la heterogeneidad no observada pero vemos que para el caso de β no existe demasiados cambios entre la capacidad de los intervalos de bootstrap entre los modelos, solo es necesario considerar que el modelo A podría arrojar mejores resultados con una estimación de efectos fijos por lo discutido en el experimento de Montecarlo.

Modelos de Panel no Lineales

El archivo KEANE.DTA contiene un subconjunto de los datos utilizados por Keane y Wolpin en formato Stata. Estos datos tienen información sobre empleo e historia escolar para una muestra de hombres para los años 1981-1987. El archivo KEANE.DES contiene una detallada descripción de los datos. Los dos archivos se adjuntan con el exámen.

Para responder esta pregunta considere solo los hombres negros.

Los códigos para esta sección se encuentran en el archivo Jaramillo1263_2.

(a) Use Pooled Probit para estimar el modelo:

$$P(employ_{jt} = 1 | employ_{jt-1}) = \Phi(\delta_0 + \rho employ_{jt-1})$$

Qué supuestos son necesarios para asegurar que los desvíos estándar y estadísticos t de pooled probit sean asintóticamente válidos?

Cuadro 19: Estimación probit (a)

Variable	**Coeficiente**	**Desvío estándar robusto**	**t**	**P>|t|**
$employ_{jt-1}$	1.38704	0.05021	27.63	0.000
Constante	-0.5400	0.03020	-17.88	0.000

En el cuadro 19 se puede ver los resultados de la estimación, se puede ver tanto para la variable independiente dinámica como para la constante las desviaciones estándar son estadísticamente relevantes.

Para asegurar que los desvíos estándar y los estadísticos t de pooled probit hay que considerar que en este caso estamos planteando on pooled probit dinámico, por lo que además de los supuestos de independencia que implica que las respuestas de un individuo no están correlacionadas con las respuestas de otros individuos, linealidad, ausencia de multicolinealidad y una muestra relativamente grande.

También debemos pedir homoscedasticidad para cada individuo, por lo que es necesario estimar el pooled probit en cluster o discriminando por individuos, además de homoscedasticadad requerimos independencia de los errores condicionales, es decir que los errores de un individuo no estén relacionado con los de otro individuo, aquí puede surgir además autorcorrelación serial entre individuos lo que requeriría otro ripo de estimación, para este caso solamente estimar como cluster es suficiente.

Debido a que es un modelo dinámico también es necesario hablar de completitud dinámica que permite modelar y estimar adecuadamente las relaciones dinámicas y las dependencias temporales de los datos, de no tener esto puede ser necesario otro tipo de estimaciones que corrijan el sesgo de los datos faltantes.

Por supuesto también los supuestos clásicos de distribución normal de los errores y en el caso de probit, los necesarios para estimar por máxima verosimilitud.

(b) Estime $P(employ_{jt} = 1|employ_{jt-1} = 1)$ y $P(employ_{jt} = 1|employ_{jt-1} = 0)$. Explique como obtendría errores estándar para estas estimaciones.

$$P(employ_{jt} = 1|employ_{jt-1} = 1) \rightarrow 0{,}80151$$
$$P(employ_{jt} = 1|employ_{jt-1} = 0) \rightarrow 0{,}29459$$

La probabilidad de que el hombre negro elija trabajar como white collar, blue collar o actividades de servicios dado que en el periodo anterior trabajaba en alguna de las mismas actividades es 0,80151, la probabilidad de que elija trabajar en estas actividades dado que en el periodo anterior no trabajaba de esta manera es 0,29459.

Para obtener los errores estándar de estas estimaciones sería necesario utilizar el método Delta sobre el cambio en la probabilidad esperada.

(c) Adicione al modelo en 1. el conjunto completo de variables binarias temporales y estime las probabilidades pedidas en (b). Para 1987. Hay diferencias importantes con las estimaicones de (b)?

Cuadro 20: Estimación probit (c)

Año	$employ_{jt-1} = 1$	$employ_{jt-1} = 0$
83	0.78158	0.29439
84	0.81252	0.33326
85	0.83145	0.36009
86	0.79566	0.31139
87	0.83249	0.36163

El cuadro 20 muestra las probabilidades de que la población de interés elija un empleo como white collar, blue collar o actividades de servicios dado que en el periodo anterior estaba trabajando en estas mismas actividades en la primera columna y dado que en el año anterior no trabajaba en estas actividades en la segunda columna, y para cada año diferente.

Se puede ver que parece haber una mayor probabilidad en general de mantener el empleo que de conseguir un nuevo empleo en este tipo de actividades, las diferencias en las probabilidades de año a año deben estar relacionadas con las características propias del mercado laboral cada año, siento el año 83 el más difícil para mantenerse en el empleo y para encontrar un empleo de este tipo en la población de interés.

(d) Ahora estime un modelo de efectos no observables dinámico. En particular, adicione $employ_{j,81}$ como una variable explicativa adicional y use el modelo Probit de efectos aleatorios. Use el conjunto completo de variables binarias temporales.

Cuadro 21: Estimación probit (d)

Variables	**Coeficiente**	**desvío estándar**	**z**	**p**	95 % intervalo de confianza	
$employ_{jt-1}$	0.89149	0.06722	13.26	0.000	0.75974	1.02324
employ81	0.57189	0.08812	6.49	0.000	0.39918	0.74461
y83	0.43415	0.07946	5.46	0.000	0.27841	0.58981
y84	0.65075	0.08327	7.82	0.000	0.48755	0.81395
y85	0.79825	0.08784	9.09	0.000	0.62609	0.97042
y86	0.68967	0.08947	7.71	0.000	0.51431	0.86502
y87	0.84014	0.09046	9.29	0.000	0.66284	1.01745
constante	-1.0034	0.06526	-15.3	0.000	-1.1313	-0.8755
/lnsig2u	-1.1648	0.19666			-1.5502	-0.7795
sigma_u	0.55853	0.05491			0.46065	0.67721
rho	0.23778	0.03563			0.17505	031442

El cuadro 21 muestra las estimaciones para el modelo considerando todo el conjunto de las variables dicotómicas por año además de employ81 en un modelo probit de efectos aleatorios, como estamos considerando un modelo probit dinámico, algunas suposiciones sobre este modelo cambian y es necesario adicionar también variables.

(e) Hay evidencia de dependencia del estado anterior, condicional en c_j? Explique.

En el contexto del modelo probit dinámico con variables no observadas, la dependencia del estado anterior condicional en la heterogeneidad no observada se refiere al hecho de que el coeficiente de autocorrelación ρ en el modelo dinámico no es igual a cero después de controlar por el efecto individual α_j. Esto implica que el estado anterior tiene un impacto adicional en el estado contemporáneo, más allá de lo que se puede atribuir a la heterogeneidad no observada.

Para evaluar esta dependencia, se realiza un test de significatividad individual para el coeficiente ρ bajo la hipótesis nula $H_0 : \rho = 0$. Si el test indica que rho es significativamente diferente de cero, se concluye que el estado anterior tiene un efecto significativo en el estado contemporáneo, incluso después de controlar por la heterogeneidad no observada.

En la tabla presentada en el inciso anterior, se puede observar que la variable $employ_{jt-1}$ es estadísticamente significativa al 5 %. Esto indica que el estado anterior de la variable employ tiene un impacto estadísticamente significativo en el estado contemporáneo. Además, desde un punto de vista económico, el valor nulo del coeficiente se encuentra fuera del intervalo de confianza, lo que implica que su signo está determinado y está asociado con un aumento de alrededor de 0.891 en el índice del modelo probit.

Estos resultados proporcionan evidencia de la dependencia del estado anterior en el modelo probit dinámico de variables no observadas, lo que sugiere que el estado anterior tiene un efecto importante en el estado actual, incluso después de controlar por otros factores no observados.

(f) Promedie la probabilidades estimadas a lo largo de $employ_{j81}$ para obtener el efecto parcial promedio para 1987. Compare estas estimaciones con aquellas en (c).

$$\mathbf{E}_{employ_{j81}}[P(employ_{j87} = 1|employ_{j86} = 1, y87, employ_{j81}) \rightarrow 0{,}76766$$
$$\mathbf{E}_{employ_{j81}}[P(employ_{j87} = 1|employ_{j86} = 0, y87, employ_{j81}) \rightarrow 0{,}48730$$

Las probabilidades promedio bajo $employ_{j81}$ de la probabilidad de mantener el empleo en el año 87 dado que tiene empleo en el 86 es más baja que el calculo sin promediar en el ejercicio (C) esto quiere decir que es relativamente más coomplejo mantener un empleo en el tiempo, ya que además esta estimación nos habla de los efectos parciales que hay año a año, en el caso de encontrar empleo dado que en el 86 no tenía empleo en esta actividad específica, es más alto, por lo que dado que en el año anterior no tenía empleo y considerando los años anteriores la probabilidad de adquirir empleo es relativamente más grande considerando los efectos parciales.