

Tarea 2

Fertilidad y participación laboral

Fecha de entrega: 19 de abril

Nota: Deberán subir a *Canvas* un archivo de texto con sus respuestas. Pueden utilizar el formato de su preferencia (e.g. \LaTeX , Word u hojas escritas a mano y escaneadas). Además, deberán subir otro archivo que genere todos sus resultados de las preguntas prácticas. Puede ser un R-script, Do-File o script de algún otro software similar.

A lo largo de las últimas 4 décadas se ha observado una tendencia a la alza de la participación laboral de las mujeres. Asimismo, las tasas de fertilidad se han reducido en gran medida. Estos fenómenos están mutuamente relacionados: (i) por un lado una mayor participación laboral de las mujeres incrementa el valor de su tiempo, lo cual desincentiva el embarazo y con ello la fertilidad y, por otro lado, (ii) un aumento en el control de fertilidad llevó a la creación de familias de menor tamaño (en términos de miembros del hogar) lo cual les permitió a las mujeres ingresar en la fuerza laboral con mayor facilidad. En esta tarea tendrán que investigar el efecto causal del tamaño de la familia (utilizado como indicador de fertilidad) sobre variables de participación laboral. Entre otras variables de participación laboral tendrán disponible una dummy de si la madre está empleada, número de semanas empleada y salario. Estas variables y otras se encuentran en la base de datos `pums80.dta` que está disponible en *Canvas*. El archivo `pums80.desc.pdf` contiene una descripción de la base de datos.

1. Llena las primeras dos columnas de la *Tabla 1* llevando a cabo las estimaciones necesarias. [Tip R: Para estimaciones con IV y controles puedes utilizar el comando `ivreg()` de la librería `AER` de la siguiente manera:
`ivreg(y~x+w_1+...+w_K|w_1+...+w_K+z_1+...+z_L, data)`, donde $y \sim x + w_1 + \dots + w_K$ es la ecuación de tu modelo, x es la variable explicativa de interés, w_k es el k -ésimo control y z_l es el l -ésimo instrumento que utilizarás. Tip Stata: Para IV necesitarás el comando `ivregress 2sls`]
2. Da una interpretación lo mas específica posible del coeficiente de *kidcount* para el caso de la estimación de la columnas (1) y (2). Indica claramente si hay significancia estadística en estas estimaciones y argumenta si el efecto es de tamaño considerable.
3. En el caso de la estimación de la columna (2) utilizamos como instrumento la variable *samesex*. Para que esta variable sea un instrumento adecuado debe cumplirse el supuesto de relevancia. Critica y/o da argumentos a favor de utilizar este supuesto. Deberás ser muy específico acerca de donde obtienes la evidencia para respaldar este supuesto (es decir, si debes estimar una regresión, repórtala con formato de ecuación).

4. Además de relevancia, el instrumento debe cumplir con otro supuesto. Indica cuál es este supuesto. Critica y/o da argumentos a favor de este supuesto. De ser posible presenta evidencia (aunque sea limitada) que ayude a respaldar o contradecir este supuesto.
5. En vez de utilizar el comando `ivreg()` (R) o `ivregress` (Stata), podríamos generar el mismo estimador puntual para el coeficiente de *kidcount* de la columna (2) utilizando dos regresiones: la primera etapa y la forma reducida. Lleva a cabo estas estimaciones, repórtalas con formato de ecuación e indica cómo se genera el valor estimado del coeficiente de *kidcount* partiendo de estas estimaciones. (Ojo: no es necesario derivar el error estándar, únicamente el valor estimado del parámetro de interés)
6. Una alternativa a utilizar *samesex* como instrumento consiste en utilizar como instrumentos las variables *twoboys* y *twogirls*. Estas variables las tendrán que generar y consisten en dummies que indican si los dos primeros hijos son niños (*twoboys*) o niñas (*twogirls*). Asimismo, la base de datos incluye como variable adicional *multi2nd*, que es una variable dummy que indica si el segundo parto fue múltiple (i.e. más de un bebé).
 - (a) Utilizando estas variables completa las estimaciones de la Tabla 1. En un párrafo breve describe los resultados que obtuviste. Reflexiona si los resultados muestran lo que esperabas y si deberían o no ser comparables entre las distintas columnas.
 - (b) Describe para qué tipo de individuos aplican los efectos mostrados en las columnas 2-5 de la Tabla 1. Es decir, deberás de describir en este contexto qué es un *complier* en cada caso.
 - (c) Estima en cada caso (columnas 2-5) el porcentaje de *compliers*, *never-takers* y *always-takers*. Explica por qué es necesario asumir que no existen los *defiers* para llevar a cabo tu cálculo anterior.
 - (d) ¿Es el resultado del inciso anterior para las columnas 3 y 4 consistente con el podcast de “*Do Baby Girls Cause Divorce?*”
7. A continuación utilizamos más de un instrumento para llevar a cabo nuestras estimaciones. En las primeras tres columnas de la *Tabla 2* utilizamos una estimación de *Mínimos Cuadrados en Dos Etapas (2SLS)*.
 - (a) Compara los resultados de la primera columna con los resultados de las columnas 3 y 4 de la *Tabla 1*. ¿Era de esperar este resultado?
 - (b) Compara los resultados de las columnas 2 y 3 de la *Tabla 2*. La única diferencia está en el uso de controles para llevar a cabo la estimación. ¿Era de esperar el resultado que obtuviste? Explica por qué si o por qué no.

8. En las columnas 4 y 5 de la *Tabla 2* llevamos a cabo una estimación de GMM tradicional, como la vista en clase.
 - (a) Reflexiona cuál de las dos estimaciones entre la columna 4 y 5 consideras que es mejor. Justifica tu razonamiento.
 - (b) Compara los resultados obtenidos entre las columnas 2 y 5. ¿Qué diferencias importantes notas? ¿Eran de esperar lo que obtuviste al comparar estas columnas? Justicia por qué sí o por qué no.
9. En las estimaciones previas se utilizan semanas trabajadas. En vez de utilizar *weeksm1* emplea *workedm* como variable dependiente. Utiliza los instrumentos y metodología que consideres más adecuados basado en los resultados de las *Tablas 1* y *2*. Reporta tus resultados en la columna (1) de la *Tabla 3*. Da una interpretación lo mas específica posible de tus resultados.
10. Además de efectos sobre la participación laboral, quisieras ver si existen efectos sobre los salarios. Decide qué especificación debes utilizar. Deberás reportar tus resultados en la columna (2) de la *Tabla 3*. Da una interpretación lo mas específica posible de tus resultados. Incluye una descripción acerca del tamaño de los efectos encontrados.
11. Supón que quieres dar evidencia para determinar si el efecto del tamaño de la familia sobre semanas trabajadas es mayor para las madres hispanas. Reporta tus resultados en la columna (3) de la *Tabla 3*. Deberás llenar las variables (tanto dependientes como regresores) que debes reportar. Emplea el método, instrumentos y controles que consideres más adecuados. Describe los resultados obtenidos.

Tabla 1
Un Instrumento

	(1)	(2)	(3)	(4)	(5)
Variable dependiente:	<i>weeksm1</i>	<i>weeksm1</i>	<i>weeksm1</i>	<i>weeksm1</i>	<i>weeksm1</i>
<i>kidcount</i>					
<i>N</i>					
<i>F</i> 1a etapa	—				
Método	MCO	IV	IV	IV	IV
Instrumentos	—	<i>samesex</i>	<i>twoboys</i>	<i>twogirls</i>	<i>multi2nd</i>

Todas las regresiones controlan por *agem1*, *agefstm*, *black*, *hispan*, *othrace*

Errores robustos entre paréntesis

***Significativo al 1% **Significativo al 5% *Significativo al 10%

Tabla 2
Varios Instrumentos

	(1)	(2)	(3)	(4)	(5)
Variable dependiente:	<i>weeksm1</i>	<i>weeksm1</i>	<i>weeksm1</i>	<i>weeksm1</i>	<i>weeksm1</i>
<i>kidcount</i>					
<i>N</i>					
<i>F</i> 1a etapa	—				
Método	2SLS	2SLS	2SLS	GMM	GMM
Instrumentos	<i>twoboys</i>	<i>samesex</i>	<i>samesex</i>	<i>twoboys</i>	<i>samesex</i>
	<i>twogirls</i>	<i>multi2nd</i>	<i>multi2nd</i>	<i>twogirls</i>	<i>multi2nd</i>
				<i>multi2nd</i>	
Controles	Si	Si	No	Si	Si

Todas las regresiones controlan por *agem1*, *agefstm*, *black*, *hispan*, *othrace*

Errores robustos entre paréntesis

***Significativo al 1% **Significativo al 5% *Significativo al 10%

Tabla 3
Mas especificaciones

	(1)	(2)	(3)	(4)
Variable dependiente:	<i>workedm</i>			
<i>kidcount</i>				
<i>N</i>				
<i>F</i> 1a etapa				
Método				
Instrumentos				
Todas las regresiones controlan por <i>agem1</i> , <i>agefstm</i> , <i>black</i> , <i>hispan</i> , <i>othrace</i>				
Errores robustos entre paréntesis				
***Significativo al 1% **Significativo al 5% *Significativo al 10%				