

MÉTODOS LINEALES DE ESTIMACIÓN CON DATOS DE PANEL: UNA APLICACIÓN AL ESTUDIO DE LOS EFECTOS DE LA INVERSIÓN PÚBLICA FEDERAL EN MÉXICO

Cobacho Tornel, M^a Belén

Universidad Politécnica de Cartagena

Bosch Mossi, Mariano

London School of Economics

RESUMEN

En este trabajo se realiza una revisión de los principales métodos econométricos empleados en la actualidad para realizar estimaciones en modelos lineales de datos de panel, tanto modelos estáticos como modelos dinámicos, realizando una comparativa entre ellos y analizando las características que hacen que cada método de estimación resulte mejor adaptado a una determinada situación. Posteriormente hemos aplicado algunos de estos métodos para estudiar los efectos que la inversión pública federal en México ha tenido sobre algunos indicadores sociales entre los años 1970 y 2000, a la vista de las claras disparidades, tanto a nivel económico como social, existentes entre los estados mexicanos. Las estimaciones realizadas mediante estos métodos a partir de los datos disponibles evidencian que la inversión pública federal en México ha tenido un efecto positivo en la reducción de la tasa de mortalidad infantil y en el aumento de la tasa de alfabetización durante el periodo de estudio.

Palabras clave: datos de panel, efectos fijos y aleatorios, variables instrumentales, inversión pública.

1. INTRODUCCIÓN

México es un país con estructura de república federal a pesar de la cual, la política de asignación regional de la inversión pública ha estado centralizada y carece de claros mecanismos de redistribución, estando dirigida fundamentalmente al único objetivo de alcanzar un crecimiento a nivel nacional. Algunos estudios (Rodríguez-Oreggia y Rodríguez-Pose (2003)), muestran que la asignación territorial de la inversión pública durante gran parte del periodo 1970-2000 ha favorecido fundamentalmente a las regiones más desarrolladas, con fugaces esfuerzos por reducir las disparidades regionales. Ante este panorama macroeconómico y social en México, hemos considerado oportuno realizar un estudio del efecto que la inversión pública federal en México ha tenido sobre distintos indicadores económicos y sociales desde principios de los años 70 y hasta el año 2000 con la finalidad de analizar, con posterioridad, formas eficientes de la distribución de la inversión pública.

Para ello, en este estudio utilizamos técnicas empleadas en la literatura empírica de crecimiento y convergencia¹. Los primeros trabajos sobre este tema, como el de Mankiw et al. (1994), empleaban muestras transversales para estudiar procesos de convergencia económica. Sin embargo, uno de los principales problemas de este tipo de estimaciones es que es imposible tener en cuenta el efecto de las características no observables propias de las unidades de sección cruzada a la hora de efectuar las estimaciones. En este sentido, Islam (1995) y Caselli y otros (1996) muestran cómo la utilización de datos de panel puede mejorar y superar algunos de los problemas que las estimaciones de ecuaciones de convergencia de corte transversal presentan. No obstante, la introducción de datos de panel en este tipo de estimaciones no está exenta de problemas: la naturaleza dinámica de ciertas situaciones hace inconsistentes los coeficientes de la ecuación obtenidos a través de estimaciones por métodos estáticos. Para dar una solución a este problema, Arellano y Bond (1991) y Arellano y Bover (1996) proponen la utilización de estimadores que emplean variables instrumentales para la corrección de esta inconsistencia.

En nuestro trabajo se pretende estimar los retornos de la inversión pública federal per cápita y de sus distintas desagregaciones por medio de ecuaciones de crecimiento. Para poder llevar a cabo este estudio, se ha realizado una revisión de los

¹ Para un detallado resumen de la literatura tanto teórica como empírica ver Barro y Sala-i-Martin (2001) y de la Fuente (1997).

principales modelos lineales de estimación de datos de panel estáticos (Mínimos Cuadrados Ordinarios, Efectos Fijos o Estimación *Within (intra)*, Efectos Aleatorios mediante Mínimos Cuadrados Generalizados y Estimación *Between (entre)*), así como un análisis del modelo dinámico de Estimación con Variables Instrumentales (IVE) propuesto por Arellano y Bond (1991), realizando además una comparativa entre ellos. Posteriormente hemos utilizado cada uno de estos métodos para estudiar el efecto que las diferentes partidas de inversión pública federal en México han tenido sobre dos indicadores sociales concretos: tasa de mortalidad infantil y tasa de alfabetización, pudiendo observar que las estimaciones por métodos estáticos no resultan adecuadas en estas circunstancias. Ante esta situación, utilizamos el estimador IVE propuesto por Arellano y Bond para datos de panel dinámicos. Los resultados obtenidos muestran que las asignaciones de la inversión pública federal en México tienen cierto efecto en la evolución de las tasas de mortalidad infantil y de alfabetización; concretamente, incrementos en las partidas destinadas a salud y educación mejoran los niveles de estos indicadores.

El trabajo queda estructurado de la siguiente manera: en la sección 2 estudiamos los diferentes métodos de estimación de datos de panel comentados anteriormente. En la sección 3 presentamos el contexto de aplicación de nuestro estudio, realizando un breve análisis del panorama económico y social de México durante los últimos 30 años. En la sección 4 presentamos los resultados obtenidos de las estimaciones mediante los métodos analizados en la sección 2. Finalmente, en la sección 5 exponemos las conclusiones extraídas de nuestro estudio.

2. MODELOS LINEALES DE ESTIMACIÓN DE DATOS DE PANEL

Son frecuentes en Economía los conjuntos de datos que combinan series temporales con unidades de sección cruzada (países, estados, empresas, etc.), de forma que una aplicación de técnicas para su estudio por separado deja cuestiones sin resolver. El análisis de datos de panel trata del estudio de estos conjuntos de datos. Los modelos lineales de análisis de datos de panel pueden ser clasificados en *dinámicos* o *estáticos*, según incluyan o no en sus ecuaciones variables pertenecientes a distintos periodos temporales. A continuación se presentan algunos modelos pertenecientes a cada uno de estos dos tipos de estimaciones.

2.1 Modelos estáticos

El contexto básico de un análisis lineal de datos de panel consiste en un modelo de regresión de la forma:

$$y_{it} = x'_{it}\beta + \varepsilon_{it} \quad (1)$$

donde el subíndice i toma los valores $i=1,2,\dots,N$ e indica la unidad de sección cruzada, $t=1,2,\dots,T_i$ indica los diferentes periodos de tiempo, y_{it} son las variables dependientes o explicadas (*regresando*), x_{it} las variables independientes o explicativas (*regresor*), β es el vector de parámetros a estimar y ε_{it} el término de error o perturbación aleatoria, todos ellos componentes del modelo clásico de regresión lineal.

Si para cada unidad de sección cruzada existe el mismo número de observaciones temporales, es decir, si $T_i = T$ para cada i , se dice que el panel de datos está *equilibrado*. En nuestro estudio hemos trabajado con un panel equilibrado².

La estimación de β en el modelo (1) por Mínimos Cuadrados Ordinarios (MCO) es consistente y eficiente. Sin embargo, existen generalizaciones de dicho modelo, más frecuentes en Econometría, para las cuales las estimaciones por MCO no son consistentes, como se muestra a continuación. Es habitual para el estudio de los distintos estimadores disponibles asumir que el error aleatorio se descompone en dos términos, $\varepsilon_{it} = \alpha_i + v_{it}$, con las siguientes características:

$$\left. \begin{array}{l} \alpha_i \sim i.i.d. \text{ sobre } i; v_{it} \sim i.i.d. \text{ sobre } i \text{ y } t; \alpha_i \text{ independiente de } v_{jt} \forall i, j, t; \\ E[\alpha_i] = 0 \forall i; E[v_{it}] = 0 \forall i, t; \text{Var}[\alpha_i] = \sigma_\alpha^2 \forall i; \text{Var}[v_{it}] = \sigma_v^2 \forall i, t \end{array} \right\} \quad (2)$$

α_i es el *efecto individual*, específico para cada unidad de sección cruzada y que se considera constante en el tiempo. Así, el modelo a estimar queda de la forma:

$$y_{it} = \alpha_i + x'_{it}\beta + v_{it} \quad (3)$$

Éste es un modelo de regresión clásica y, como tal, puede ser estimado directamente por MCO. Sin embargo, la presencia del efecto aleatorio invariante en el tiempo, α_i , implica la aparición de heterogeneidad persistente inobservada y la violación del supuesto de homocedasticidad (varianza constante de las perturbaciones) y no autocorrelación (varianza 0 de las observaciones procedentes de una misma unidad

² Para un estudio de paneles de datos no equilibrados véase Baltagi (2001).

de sección cruzada), es decir, $E[\varepsilon\varepsilon'] = \sigma^2 I$. Por tanto, una estimación por MCO no conducirá al mejor estimador lineal insesgado y tendrá una matriz de varianzas-covarianzas distinta de $\sigma^2(X'X)^{-1}$. Existen entonces diferentes marcos para afrontar estos modelos, los cuales se exponen a continuación.

1. Efectos fijos o *intra* grupos (*within*). El enfoque de efectos fijos (FE) considera α_i como un término constante específico de grupo. En la formulación de este modelo se supone que las diferencias entre unidades pueden ser captadas a través de las diferencias en dicho término constante, de modo que en (3) cada α_i es un parámetro desconocido a estimar. La estimación se lleva a cabo entonces en términos de desviaciones de las medias del grupo, es decir, mediante una regresión de $y_{it} - \bar{y}_i$ sobre $x_{it} - \bar{x}_i$ en el modelo $y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)' \beta + (\varepsilon_{it} - \bar{\varepsilon}_i)$. De este modelo se obtiene una estimación $\hat{\beta}_{FE}$ de β , y posteriormente se calcula el estimador de α_i mediante $\hat{\alpha}_i = \bar{y}_i - \bar{x}_i' \hat{\beta}_{FE}$. Bajo los supuestos habituales, el estimador $\hat{\beta}_{FE}$ es insesgado y asintóticamente consistente cuando $NT \rightarrow \infty$, mientras que los estimadores $\hat{\alpha}_i$ son insesgados pero no son consistentes salvo cuando cada $T_i \rightarrow \infty$, dado que cada α_i está basado sólo en las T_i observaciones de la unidad i .

2. Efectos aleatorios. Mínimos Cuadrados Generalizados. El enfoque de efectos fijos puede ser interpretado como exclusivamente aplicable a los grupos considerados en el estudio, pero no a grupos adicionales no incluidos en la muestra. En la aplicación que se lleva a cabo en la sección 4, la muestra de estados considerados es exhaustiva (tiene en cuenta todos los estados, no una muestra aleatoria de ellos) y por tanto sería apropiado, en este aspecto, plantear una estimación a partir de un modelo de efectos fijos. En otros contextos, cuando los grupos considerados son extracciones muestrales de una población más grande, puede resultar más apropiado considerar que los términos constantes específicos de cada unidad están aleatoriamente distribuidos entre los grupos. Es en este caso que se considera un modelo de estimación de *efectos aleatorios* (RE).

Una estimación de efectos aleatorios por Mínimos Cuadrados Generalizados (MCG) consiste en una regresión de desviaciones parciales de las y_{it} sobre las mismas desviaciones parciales de las x_{it} . Concretamente, $y_{it} - \lambda_i \bar{y}_i = (x_{it} - \lambda_i \bar{x}_i)' \beta + (\varepsilon_{it} - \lambda_i \bar{\varepsilon}_i)$,

donde $\lambda_i = 1 - \sqrt{\frac{\sigma_v^2}{T_i \sigma_\alpha^2 + \sigma_v^2}}$. En la práctica, se ha de usar un estimador consistente $\hat{\lambda}_i$ de λ_i basado en estimadores consistentes $\hat{\sigma}_v^2$ y $\hat{\sigma}_\alpha^2$ de σ_v^2 y de σ_α^2 respectivamente. Éste es un modelo generalizado de regresión y en principio puede ser estimado por MCO, si no eficientemente, sí consistentemente: de nuevo bajo los supuestos habituales, el estimador $\hat{\beta}_{RE}$ de β es asintóticamente el mejor estimador lineal consistente asintóticamente normal cuando $NT \rightarrow \infty$.

3. Estimación entre grupos (between). Este modelo investiga la variabilidad *entre* unidades de sección cruzada en términos de las medias de los grupos, es decir, se basa en una estimación de \bar{y}_i sobre \bar{x}_i en el modelo $\bar{y}_i = \bar{x}_i' \beta + \bar{\varepsilon}_i = \alpha_i + \bar{x}_i' \beta + \bar{v}_i$. La estimación *between* (BE) es también un modelo clásico de regresión lineal que por tanto puede ser estimado por MCO, y su principal atractivo es que permite obtener un estimador consistente para σ_α^2 , necesario para poder llevar a cabo una estimación por MCG.

Comparaciones entre modelos. La elección de un modelo de efectos aleatorios frente a uno de efectos fijos no siempre se nos muestra de forma obvia. En este caso es útil considerar el problema de estudio en términos de la existencia o no de correlación entre los efectos fijos de cada unidad, α_i , y las variables independientes, x_{it} , y, en caso de correlación, la posible naturaleza de dependencia de α_i respecto de x_{it} (Judge et al. (1985)). A continuación se consideran los métodos anteriores bajo diferentes supuestos.

Si $E[\alpha_i | x_{it}] = 0$, es decir, si los efectos fijos de cada unidad no están correlacionados con las variables independientes, en este caso todos los estimadores de β , incluido MCO, son consistentes. Sin embargo, el más eficiente es MCG, por lo que éste es el que aplicamos, para lo cual, como ya se ha indicado, se precisa de estimaciones consistentes de σ_v^2 y de σ_α^2 , obtenidas de las estimaciones FE y BE.

Si $E[\alpha_i | x_{it}] \neq 0$, es decir, existe correlación entre los efectos fijos y las variables independientes, en este caso todos los estimadores excepto el de efectos fijos son inconsistentes, debido a que FE es el único modelo que elimina el residuo de efecto fijo que queda en los errores. Aplicamos por tanto en esta situación la estimación FE.

Por otro lado, cabe señalar que el estimador MCO es una media ponderada matricial de los estimadores *intra* y *entre* grupos, y que también lo es el estimador MCG. No obstante, el estimador MCO comparado con MCG otorga demasiada ponderación a las variaciones entre unidades de sección cruzada, ya que lo incluye todo en las variaciones de x_{it} sin atribuir parte alguna a la variación aleatoria entre grupos.

Un *contraste de Wu-Hausman* se utiliza para analizar la posible correlación entre los α_i y los regresores y poder así decidir entre una estimación por FE o por RE. La hipótesis nula del contraste es $H_0 : E[\alpha_i | x_{it}] = 0$. Bajo H_0 , el estadístico de Wu-Hausman, definido como $Q_{FE,RE} = (\hat{\beta}_{FE} - \hat{\beta}_{RE})' (\hat{\sigma}_{\hat{\beta}_{FE}}^2 - \hat{\sigma}_{\hat{\beta}_{RE}}^2)^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE})$, converge en distribución a una $\chi^2(n)$, siendo $n = NT$. Como se puede observar, $Q_{FE,RE}$ es el cociente del cuadrado de la diferencia entre los dos estimadores y la diferencia entre las varianzas de éstos. Así, si H_0 es cierta, tanto EF como MCG son consistentes y por tanto deben tender al mismo valor, de modo que la diferencia entre los estimadores debe ser pequeña. Además, puesto que el estimador $\hat{\beta}_{RE}$ es más eficiente que $\hat{\beta}_{FE}$, la varianza de aquél es pequeña en comparación con la de éste y por tanto la diferencia entre las varianzas es grande. Esto dará como resultado un valor del estadístico $Q_{FE,RE}$ cercano a 0 y que por tanto no se pueda rechazar la hipótesis nula. Si, por el contrario, H_0 no es cierta, entonces $\hat{\beta}_{FE}$ es consistente pero $\hat{\beta}_{RE}$ no lo es, con lo que debe haber diferencia notable entre los valores de estos estimadores. Esto implicará que el valor del estadístico $Q_{FE,RE}$ será alto, pudiendo así rechazar la hipótesis nula.

2.2 Modelos dinámicos

Los modelos dinámicos incluyen en su análisis valores rezagados de la variable dependiente. El modelo básico con un retardo de una unidad de tiempo es:

$$y_{it} = \theta y_{i,t-1} + x_{it}' \beta + \varepsilon_{it}.$$

De nuevo suponemos que el error se descompone en dos términos, $\varepsilon_{it} = \alpha_i + v_{it}$, con las mismas características que en (2), de modo que el modelo a estimar es:

$$y_{it} = \alpha_i + \theta y_{i,t-1} + x_{it}' \beta + v_{it} \quad (4)$$

En el caso dinámico, los cuatro modelos de estimación que han sido analizados en el caso estático, MCO, FE, MCG y BE, son inconsistentes por presentar endogeneidad o

estar los regresores correlacionados con el error, lo que conlleva al planteamiento de otros modelos para la estimación de datos de panel dinámicos.

Diferenciación. Estimador de Variables Instrumentales. La regresión en (4) por FE no supera la endogeneidad de la variable dependiente rezagada con respecto al término de error debido a que, aunque se consigue eliminar el efecto α_i , se mantiene una estructura de correlación a través de los términos $v_{i,t-1} - \bar{v}_i$ y $v_{it} - \bar{v}_i$. Una forma alternativa de eliminar los efectos α_i es partir del modelo de efectos fijos, $y_{it} - \bar{y}_i = \theta(y_{i,t-1} - \bar{y}_i) + (x_{it} - \bar{x}_i)' \beta + (\varepsilon_{it} - \bar{\varepsilon}_i)$, y tomar diferencias de primer orden en dicha ecuación: $\Delta y_{it} = \theta \Delta y_{i,t-1} + \Delta x_{it}' \beta + \Delta v_{it}$, en la que ya no aparecen los α_i . En esta estructura, el regresor $\Delta y_{i,t-1}$ está correlacionado con el error, ya que $\Delta y_{i,t-1}$ depende de $\Delta v_{i,t-1}$, pero los rezagos anteriores, $\Delta y_{i,t-k}$ con $k > 1$, son buenas variables instrumentales por no estar correlacionadas con $\Delta y_{i,t-1}$. Por tanto, una estimación apropiada para este contexto sería la estimación por Variables Instrumentales. Arellano y Bond (1991) proponen la utilización de todos los instrumentos posibles de la variable rezagada. Así, un instrumento válido para $y_{i,t-1} - y_{i,t-2}$ sería $y_{i,t-2}$, dado que éste está correlacionado con $y_{i,t-1} - y_{i,t-2}$, pero no con el término de error $v_{it} - v_{i,t-1}$. De manera similar, todos los rezagos anteriores a $y_{i,t-2}$ también cumplen esas características, con lo que pueden ser utilizados como instrumentos adicionales. Los instrumentos válidos para cada regresor serían, por tanto: y_{i1} para $y_{i2} - y_{i1}$, y_{i1}, y_{i2} para $y_{i3} - y_{i2}, \dots, y_{i1}, y_{i2}, \dots, y_{i,T-1}$ para $y_{iT} - y_{i,T-1}$.



3. PANORAMA SOCIO-ECONÓMICO EN MÉXICO

México es una república federal constituida por 32 estados (entidades federativas). Tanto a nivel económico como social cabe decir que México es un país de grandes contrastes entre sus estados. Entre estas diferencias, analizadas por Bosch et al. (2004), cabe destacar, por ejemplo, que en el año 2000 el PIB per cápita del estado más pobre, Oaxaca, apenas representa el 16% del estado más rico, Distrito Federal, y que ésta no es una situación particular del año 2000, sino que se advierten grandes diferencias en los niveles de PIB per cápita medios por regiones a lo

largo de todo el periodo 1970-2000, alcanzando estas diferencias su mayor amplitud entre los estados de las zonas norte y sur o también entre las zonas norte y centro. En este contexto de claras disparidades regionales a nivel de PIB per cápita, otras variables que capturan el nivel de bienestar de los estados mexicanos han experimentado evoluciones muy diferentes. Las medidas aquí consideradas, tasa de mortalidad infantil y tasa de alfabetización, transmiten la idea de que ha habido mejoras sociales en los campos de salud y educación en todo el país, viéndose reducidas las diferencias regionales a lo largo del periodo de estudio.

En cuanto a la inversión pública per cápita, Bosch et al. (2004) muestran cómo la distribución regional de ésta ha seguido una tendencia alcista en todas las zonas geográficas a lo largo del periodo 1970-2000. Sin embargo, este trabajo y algunos estudios, como el de Rodríguez-Oreggia y Rodríguez-Pose (2003), muestran que la asignación territorial de la inversión pública no ha realizado esfuerzo alguno por reducir las disparidades regionales, favoreciendo fundamentalmente a las regiones más desarrolladas. Con la finalidad de estudiar posibles criterios eficientes de distribución de la inversión pública en este país, hemos comenzado por analizar los efectos que las distintas partidas de inversión (inversión social, infraestructuras y otra inversión) han tenido en diversos indicadores sociales (tasas de mortalidad infantil y alfabetización). En la siguiente sección presentamos algunos de los resultados obtenidos, a modo de ilustración de los modelos revisados en la sección 2.

4. RESULTADO DE LAS ESTIMACIONES

En esta sección se analiza el impacto que la inversión pública ha tenido en los estados mexicanos durante los últimos 30 años. Concretamente, nuestro trabajo se centra en la estimación de una ecuación de convergencia del estilo

$$\ln y_{it} = \alpha_i + \theta \ln y_{it-1} + \beta \ln x_{it} + v_{it} \quad (5)$$

donde y_{it} es el valor de la variable dependiente para el estado i en el momento t , α_i es el efecto fijo para el estado i , y x_{it} son los valores de las variables explicativas, en nuestro caso el valor de la inversión pública federal per cápita y de sus distintas desagregaciones. En este contexto, los parámetros de las variables de inversión per cápita incluidas en la ecuación capturan el efecto del crecimiento de ésta sobre el

crecimiento de la variable dependiente. Para las estimaciones mediante modelos estáticos la ecuación a estimar será $\ln y_{it} = \alpha_i + \beta \ln x_{it} + v_{it}$, en sus distintas versiones según se ha expuesto en la sección 2.

Las tablas 1-4 recogen el efecto de la inversión pública federal sobre las tasas de mortalidad infantil y de alfabetización con la información organizada de la siguiente forma: en la primera columna de cada tabla se indica la variable dependiente (mortalidad infantil o tasa de alfabetización) y las variables independientes incluidas en la ecuación a estimar. Las columnas sucesivas, OLS, GLS, BE y FE, indican los valores estimados para los coeficientes β (*Coef*) y si éstos son significativos o no (*Sign*) a partir de estimaciones mediante MCO, MCG, BE y FE respectivamente. La columna IVE contiene además el coeficiente θ de la variable rezagada estimado por Variables Instrumentales a partir de la ecuación (5).

En líneas generales, de acuerdo con la estimación de los modelos estáticos, una mayor inversión social afecta negativamente a la tasa de mortalidad infantil y positivamente a la tasa de alfabetización, aunque la cuantía de este efecto depende de las estimaciones consideradas.

La mayor parte de las estimaciones por MCO son cualitativamente similares a las de MCG y FE, aunque los coeficientes varían notablemente. Las estimaciones por MCG y FE son prácticamente idénticas, de hecho el test de Wu-Hausman no puede rechazar la hipótesis nula de que $H_0 : E[\alpha_i | x_{it}] = 0$. Esto nos estaría indicando que la inversión per cápita, en sus distintas dimensiones, no está correlacionada con el efecto fijo del error, lo cual es consistente con la afirmación de Rodríguez-Oreggia y Rodríguez-Pose (2003) indicando que la distribución de la inversión pública se hizo de forma más o menos aleatoria, sin atender criterios de redistribución y equidad. Por lo tanto, en el caso estático, la opción adecuada sería la estimación por MCG, dado que es el estimador más eficiente.

Sin embargo, es posible que la especificación estática sea incorrecta: tanto el proceso de reducción de la tasa de mortalidad como el del incremento de la tasa de alfabetización parecen haber tenido alta persistencia, lo que estaría indicando un componente autoregresivo en su evolución. En efecto, al introducir la variable rezagada, ésta toma un valor elevado, alrededor de 0.60 para la mortalidad infantil y 0.70 para la tasa de alfabetización. Esto indica que una especificación adecuada de la evolución de las tasas de mortalidad infantil y alfabetización debe incluir un término dinámico. En

este contexto, todos los métodos anteriormente utilizados son inconsistentes, mientras que la estimación de un modelo en diferencias y la utilización de variables instrumentales sí proporciona estimaciones consistentes.

El efecto de la inversión social en la tasa de mortalidad infantil sigue siendo negativo (a mayor inversión menor mortalidad infantil), pero menor que en las estimaciones estáticas (-1.68 en lugar de -3.2). Lo mismo ocurre con el impacto de la inversión social en la tasa de alfabetización (0.3 en lugar de 1.2). Este fenómeno se debe principalmente a que parte del efecto autoregresivo de los procesos de mortalidad infantil y tasa de alfabetización eran capturados por esta componente de la inversión. En este caso, la tanto la inversión en infraestructuras como las partidas de otra inversión parecen haber tenido un efecto positivo y significativo en la mortalidad infantil. No queda claro cuáles han podido ser los canales por los que la inversión federal puede aumentar la tasa de mortalidad infantil, dado el valor de las otras inversiones. Una posible explicación es que tanto la inversión en infraestructuras y otra inversión incrementen el potencial de crecimiento de determinadas regiones atrayendo nuevos inmigrantes de regiones más pobres, pudiendo incrementar (o provocar una reducción menor en) la tasa de mortalidad infantil.

Al desagregar la inversión social en inversión en educación e inversión en salud, encontramos que tanto la inversión en educación como la inversión en salud han tenido efectos positivos en la reducción de la mortalidad infantil (tabla 2). En concreto, incrementos del 1% en inversión per cápita en ambas partidas han conseguido reducir la tasa de mortalidad en un 1.6%. En cuanto a la tasa de alfabetización (tabla 4), incrementos de la inversión social aumentan la tasa de alfabetización en torno a un 0.3%. Sin embargo, la inversión en educación aumenta la tasa de alfabetización (como era de esperar) en un 0.70%, mientras que la inversión en salud parece tener un efecto positivo, aunque no significativo.

	Estáticos								Dinámico	
	OLS		GLS		BE		FE		IVE	
Var dep: Mort Inf										
Regiones	32		32		32		32		32	
Observaciones	128		128		128		128		96	
	Coef	Sign	Coef	Sign	Coef	Sign	Coef	Sign	Coef	Sign
Mort Inf (-1)									0.628	***
Inversión Social	-1.904	-	-3.100	*	0.448	-	-3.273	*	-1.688	**
Inversión en Infr.	-1.217	-	0.610	-	-4.027	-	1.021	-	2.012	***
Otra Inversión	-0.000	-	-0.000	-	-0.000	-	-0.000	-	0.001	*

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

Tabla 1. Impacto de la inversión pública federal sobre la tasa de mortalidad infantil

Var dep: Mort inf	Estáticos								Dinámico	
	OLS		GLS		BE		FE		IVE	
	Coef	Sign	Coef	Sign	Coef	Sign	Coef	Sign	Coef	Sign
Regiones	32		32		32		32		32	
Observaciones	128		128		128		128		96	
Mort Inf (-1)										
Inversión Salud	-1.131	-	-2.510	-	0.964	-	-2.762	-	-1.602	***
Inv Educación	-0.699	-	-0.250	-	-0.324	-	0.143	-	-1.623	*
Inv Infraestructuras	-1.349	-	0.521	-	-4.234	-	1.028	-	1.786	***
Otra Inversión	-0.000	-	-0.000	-	-0.000	-	-0.000	-	0.739	***

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

Tabla 2. Impacto de la inversión sobre la tasa de mortalidad infantil con inversión social desagregada

Var dep: Alfabetiz.	Estáticos								Dinámico	
	OLS		GLS		BE		FE		IVE	
	Coef	Sign	Coef	Sign	Coef	Sign	Coef	Sign	Coef	Sign
Regiones	32		32		32		32		32	
Observaciones	128		128		128		128		96	
Alfabet (-1)									0.672	***
Inversión Social	3.443	**	1.367	**	5.112	-	1.241	*	0.304	***
Inversión en Infr.	1.078	-	0.181	-	0.677	-	0.100	-	0.126	-
Otra Inversión	-0.001	-	0.000	-	-0.001	-	0.000	-	0.060	-

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

Tabla 3. Impacto de la inversión pública federal sobre la tasa de alfabetización

Var dep: Alfabetiz.	Estáticos								Dinámico	
	OLS		GLS		BE		FE		IVE	
	Coef	Sign	Coef	Sign	Coef	Sign	Coef	Sign	Coef	Sign
Regiones	32		32		32		32		32	
Observaciones	128		128		128		128		96	
Alfabet (-1)										
Inversión Salud	4.885	***	0.897	-	1.073	*	0.756	-	0.096	-
Inv Educación	-1.720	-	1.316	-	-9.230	-	1.331	-	0.735	***
Inv Infraestructuras	0.976	-	0.115	-	2.682	-	0.060	-	0.084	-
Otra Inversión	-0.000	-	0.000	-	-0.000	-	0.000	-	0.058	-

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

Tabla 4. Impacto de la inversión sobre la tasa de alfabetización con inversión social desagregada

5. CONCLUSIONES

En este trabajo se han analizado brevemente diversos métodos lineales de estimación de datos de panel y se ha presentado una aplicación en la estimación de los retornos de la inversión pública federal en la evolución de las tasas de mortalidad infantil y de alfabetización de los estados mexicanos. Como principal resultado se ha obtenido que los procesos de reducción de ambas tasas parecen haber seguido procesos autoregresivos y por lo tanto las estimaciones por métodos estáticos de datos de panel resultan inviables en estas circunstancias. Ante esta situación, utilizamos el estimador IVE propuesto por Arellano y Bond para resolver los problemas de endogeneidad que presentan los modelos de datos de panel dinámicos. Mediante este estimador, observamos que las asignaciones de la inversión pública federal en México sí tienen

cierto poder explicatorio para analizar la evolución de las tasas de mortalidad infantil y de alfabetización. Se obtiene además que incrementos en las partidas destinadas a salud y educación mejoran el nivel de bienestar de los estados.

6. REFERENCIAS BIBLIOGRÁFICAS

- ARELLANO, M. y BOND, S. (1991): "Some Tests of Specification for Panel Data: Monte Carlo. Evidence and an Application to Employment Equations". *Review of Economic Studies*, Vol. 58, pp. 277-297.
- ARELLANO, M. (2003): *Panel Data Econometrics (Advanced Texts in Econometrics)*. Oxford Press.
- ARELLANO, M. y BOVER, O. (1995): "Another Look at the Instrumental Variable Estimation of Error-Component Models". *Journal of Econometrics*, Vol. 68, pp. 29-51.
- BALTAGI, B. (2001): *Econometric Analysis of Panel Data*. 2nd Edition. Wiley.
- BARRO, R. y SALA I MARTIN, X. (1992): "Convergence". *Journal of Political Economy*, 100(2), pp. 223-51.
- BOSCH, M., COBACHO, M. B. y RODRÍGUEZ-OREGGIA, E. (2004). "Los efectos de la inversión pública federal en México". XVIII reunión Asepelt, León.
- CASELLI, F., ESQUIVEL, G. y LEFORT, F. (1996): "Reopening the convergence debate: a new look at cross-country growth empirics". *Journal of Economic Growth*, 1(3), pp. 363-89.
- GREENE, W.H. (1998). *Análisis Económico*. Tercera edición. Prentice Hall.
- FUENTE, A. de la (1997): "The empirics of growth and convergence: a selective review". *Journal of Economic Dynamics and Control* 21(1): 23-74.
- ISLAM, N. (1995): "Growth Empirics: A Panel Data Approach". *Quarterly Journal of Economics*, CX(4), pp. 1127-70.
- JUDGE, G.G., GRIFFITHS, W.E., HILL, R.C., LÜTKEPOHL, H. y LEE, T.C. (1985). *The theory and practice of Econometrics*. 2ª edición. John Wiley and Sons.
- MANKIW, G., ROMER, D. y WEIL D. (1992): "A Contribution to the Empirics of Economic Growth". *Quarterly Journal of Economics*, CVII(2), pp. 407-37.
- RODRÍGUEZ-OREGGIA, E. y RODRÍGUEZ-POSE, A. (2003): "The regional returns of public investment policies in Mexico". En revisión en *World Development*.