

CENTRO DE ESTUDIOS ECONÓMICOS

Maestría en Economía 2024–2026

Microeconometrics for Evaluation

8 Panel Data

Disclaimer: I AM NOT the original intellectual author of the material presented in these notes. The content is STRONGLY based on a combination of lecture notes (from Aurora Ramirez), textbook references, and personal annotations for learning purposes. Any errors or omissions are entirely my own responsibility.

Índice

Datos de panel	3
Métodos de datos panel	3
Modelo de efectos no observados: producción agrícola	3
MCO agrupados (pooled OLS)	4
Modelo de efectos no observados: evaluación de un programa	5
Supuesto de exogeneidad estricta	5
Efectos Fijos (FE)	6
Efectos fijos - Formas de eliminar c_i	7
1 Within Estimator	7
Within Estimator: Errores estándar	8
2 Estimador de Variables Dummy	11
3 Primeras Diferencias	11
Primeras Diferencias: Errores estándar	13
Regresión de efectos fijos: puntos principales	13
Ejemplo: Democracia Directa y Naturalizaciones	14
Datos de panel de naturalizaciones	14
Datos de Panel. Long Format	15
MCO Agrupados	15
Descomposición de Varianza (within, between)	16
Within Estimator	16
Efectos Fijos. Within Estimator	17
Efectos Fijos. Rutina programada: Stata	18
Efectos Fijos. Rutina programada: Stata	19
Aplicando efectos fijos	19
Problemas que los efectos fijos no resuelven	20
Efectos Aleatorios	21
Métodos de efectos aleatorios (RE)	21
Regresión de efectos aleatorios (RE)	24
RE versus FE, FD o Variables Dummy	25

Prueba de Hausman	25
Regresión de efectos aleatorios (RE)	26
Prueba de Hausman	27
Añadiendo efectos fijos en el tiempo	27
Modelando los efectos en el tiempo	29
Efectos fijos: añadiendo efectos en el tiempo	29
Efectos fijos: tendencia lineal en el tiempo	30
Efectos fijos: efectos fijos por año	31
Efectos fijos: efectos de tiempo por unidad	32

Datos de panel

- Los datos de panel incluyen observaciones de:
 - Las mismas unidades (por ejemplo, individuos, empresas, países, ...) que se observan durante al menos dos periodos de tiempo (años, meses, días, ...).
- Los conjuntos de datos de panel vienen en dos formas:
 - **Panel balanceado:** cada unidad se observa durante el mismo periodo de tiempo.
 - **Panel no balanceado:** se observan unidades para distintos periodos de tiempo.
- **Explicación:** Un panel es una base de datos que combina información en dos dimensiones: transversal (diferentes unidades) y temporal (diferentes momentos en el tiempo).
- **Intuición:** Es como tener el historial médico de varios pacientes: un *panel balanceado* sería cuando todos tienen chequeos anuales; un *no balanceado*, cuando algunos pacientes tienen años faltantes en sus registros.

Métodos de datos panel

- A menudo, nuestra variable dependiente depende de factores no observados que también están correlacionados con nuestra variable explicativa de interés.
- Si estas variables omitidas son constantes en el tiempo, podemos usar los estimadores de datos de panel para estimar el efecto de nuestra variable explicativa.
- **Principales estimadores de datos de panel:**
 1. MCO agrupados (*pooled OLS*)
 2. Métodos de efectos fijos (FE)
 3. Método de efectos aleatorios (RE)
- **Explicación:** Los modelos de panel permiten controlar por heterogeneidad no observada entre individuos que no varía en el tiempo.
- **Intuición:** Es como analizar el desempeño de estudiantes en varias materias a lo largo de los años: los *efectos fijos* eliminan características propias del alumno que no cambian, mientras que los *efectos aleatorios* suponen que esas diferencias son aleatorias.

Modelo de efectos no observados: producción agrícola

$$y_{it} = x_{it}\beta + c_i + u_{it}, \quad t = 1, 2, \dots, T$$

- y_{it} : producción de la granja i en el año t .

- x_{it} : vector $1 \times K$ de insumos variables para la granja i en el año t (ej. mano de obra, fertilizante). Incluye la constante.
- β : vector $K \times 1$ de efectos marginales de los insumos variables.
- c_i : suma de insumos fijos en el tiempo, conocidos por el agricultor pero no observados por el investigador (ej. calidad del suelo, capacidad de gestión). \rightarrow Heterogeneidad no observada.
- u_{it} : insumos no observados que varían en el tiempo (ej. lluvia, plagas). Error idiosincrático.
- **Pregunta central:** ¿Qué sucede cuando corremos una regresión de y_{it} en x_{it} ?
- **Explicación:** El término c_i puede estar correlacionado con x_{it} , lo que genera sesgo en los estimadores MCO si no se controla.
- **Intuición:** Cada granja tiene características fijas (suelo, gestión) que afectan la producción. Si no las controlamos, confundimos su efecto con el de los insumos variables.

MCO agrupados (pooled OLS)

$$y_{it} = x_{it}\beta + v_{it}, \quad t = 1, 2, \dots, T$$

con $v_{it} = c_i + u_{it}$

- Ignoramos la estructura de panel y tratamos todos los datos como una muestra agrupada.
 - Supuestos clave para consistencia:
 1. $E(x_{it}^T v_{it}) = 0, \quad t = 1, 2, \dots, T$
 2. $\text{rank}\left[\sum_{t=1}^T E(x_{it}^T x_{it})\right] = K$
 - Bajo estos supuestos, el error compuesto v_{it} no debe estar correlacionado con los regresores.
 - Problema: si x_{it} depende de características no observadas c_i (ej. calidad del suelo), aparece sesgo de selección.
 - Entonces, $\hat{\beta}_{OLS}$ no es consistente.
 - **Explicación:** Pooled OLS mezcla toda la variación sin separar lo fijo (c_i) de lo idiosincrático (u_{it}).
 - **Intuición:** Es como comparar granjas sin reconocer que cada una tiene “suelo distinto” constante en el tiempo. Esa heterogeneidad se cuela en el estimador y sesga los resultados.
-
- Ninguna correlación entre x_{it} y v_{it} implica también que no haya correlación entre el efecto no observado c_i y x_{it} .

- Problema adicional: v_{it} están correlacionados en serie para el mismo i , porque c_i aparece en todos los periodos t .
- Consecuencia: los errores estándar de MCO agrupados no son válidos.
- **Explicación:** Como c_i se repite para cada individuo, los residuos están correlacionados dentro de cada i .
- **Intuición:** Es como medir varias veces la misma granja: el “suelo” no cambia, por lo que los errores están inflados y los intervalos de confianza resultan engañosos.

Modelo de efectos no observados: evaluación de un programa

$$y_{it} = \text{prog}_{it}\beta + c_i + u_{it}, \quad t = 1, 2, \dots, T$$

- y_{it} : logaritmo del salario del individuo i en el año t .
- prog_{it} : indicador =1 si el individuo i participa en el programa de entrenamiento en t , y 0 en caso contrario.
- β : efecto del programa.
- c_i : suma de todas las características no observadas e invariantes en el tiempo que afectan los salarios (ejemplo: habilidad).
- Problema: si corremos una regresión de y_{it} en prog_{it} , β no es consistente porque prog_{it} probablemente esté correlacionado con c_i .
- Siempre preguntarse: ¿hay alguna variable no observada constante en el tiempo c_i correlacionada con los regresores? Si la respuesta es sí, los MCO agrupados son problemáticos.
- **Explicación:** c_i captura heterogeneidad fija (ejemplo: habilidad), y si está correlacionada con prog_{it} , el estimador de MCO queda sesgado.
- **Intuición:** Si los más hábiles tienen más probabilidad de entrar al programa, el efecto estimado confundirá “programa” con “habilidad”.

Supuesto de exogeneidad estricta

- Para estimar los modelos de datos de panel más básicos (efectos fijos y efectos aleatorios), se asume **estricta exogeneidad**:

$$E(y_{it} \mid x_{i1}, x_{i2}, \dots, x_{iT}, c_i) = E(y_{it} \mid x_{it}, c_i) = x_{it}\beta + c_i$$

- En palabras: una vez que se controla por x_{it} y c_i , ninguna otra x_{is} con $s \neq t$ tiene efecto parcial sobre y_{it} .
- En el modelo

$$y_{it} = x_{it}\beta + c_i + u_{it},$$

la exogeneidad estricta se expresa como:

$$E(u_{it} \mid x_{i1}, x_{i2}, \dots, x_{iT}, c_i) = 0, \quad t = 1, 2, \dots, T$$

- **Explicación:** Los regresores en cualquier periodo no deben estar correlacionados con el error idiosincrático en ningún periodo.
- **Intuición:** El pasado, presente o futuro de x_{it} no debe “predecir” shocks en u_{it} . Ejemplo: si una política se implementa cuando se anticipa una crisis, se viola la exogeneidad estricta.

-
- Este supuesto implica que las variables explicativas en cada periodo de tiempo no están correlacionadas con el error idiosincrático u_{it} en ningún periodo:

$$E(x_{is}^T u_{it}) = 0, \quad s, t = 1, \dots, T$$

- Es decir, los regresores en el pasado, presente y futuro son exógenos con respecto a los errores de todos los periodos.
- Este supuesto es mucho más fuerte que asumir simplemente **no correlación contemporánea**:

$$E(x_{it}^T u_{it}) = 0, \quad t = 1, \dots, T$$

- **Explicación:** Exogeneidad estricta excluye toda forma de retroalimentación o anticipación entre regresores y errores.
- **Intuición:** No basta con que las x_{it} sean “limpias” en el mismo periodo; también deben ser independientes de shocks pasados y futuros. Ejemplo: si los agricultores ajustan la mano de obra porque anticipan lluvias (shock futuro), se rompe la exogeneidad estricta.

Efectos Fijos (FE)

- Los efectos fijos abordan explícitamente el hecho de que c_i puede correlacionarse con x_{it} .
- Para los modelos de efectos fijos asumimos **estricta exogeneidad**:

$$FE1 : \quad E(u_{it} \mid x_i, c_i) = 0, \quad t = 1, 2, \dots, T$$

- Donde $x_i = (x_{i1}, x_{i2}, \dots, x_{iT})$
- Se permite que $E(c_i \mid x_i)$ sea cualquier función de x_i .

- **Costo:** no podemos incluir variables constantes en el tiempo en x_{it} (por ejemplo, género, región fija).
- **Explicación:** FE elimina la heterogeneidad inobservable c_i controlando por su correlación con los regresores.
- **Intuición:** Comparamos cada unidad consigo misma a lo largo del tiempo. Ejemplo: si una escuela siempre tiene buena infraestructura (c_i), los FE eliminan ese efecto fijo y usan solo cambios dentro de la escuela.

Efectos fijos - Formas de eliminar c_i

- En los modelos **FE** existen tres maneras de remover el término c_i , que genera **correlación** entre el error y los regresores.
- **Opciones:**
 - **Within-transformation** (transformación **FE**).
 - Estimación de c_i mediante **dummies**.
 - Uso de **primeras diferencias**.

Intuición

- **Explicación:** Todas las técnicas buscan aislar la variación dentro de los individuos y así evitar la correlación entre c_i y los regresores.
- **Ejemplo:** Si seguimos a una persona en el tiempo, restar su propia media o usar su diferencia entre dos periodos elimina características fijas como su talento innato.

1 Within Estimator

- **Ecuación a estimar:**

$$y_{it} = x_{it}\beta + c_i + u_{it} \quad (1)$$

- **Paso 1:** Promediar la ecuación (1) sobre $t = 1, \dots, T$:

$$\bar{y}_i = \bar{x}_i\beta + c_i + \bar{u}_i \quad (2)$$

donde:

$$\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}, \quad \bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}, \quad \bar{u}_i = \frac{1}{T} \sum_{t=1}^T u_{it}$$

- **Paso 2:** Restar la ecuación (2) de (1) para obtener:

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)\beta + (u_{it} - \bar{u}_i)$$

$$\tilde{y}_{it} = \tilde{x}_{it}\beta + \tilde{u}_{it}$$

donde:

$$\tilde{y}_{it} = y_{it} - \bar{y}_i, \quad \tilde{x}_{it} = x_{it} - \bar{x}_i, \quad \tilde{u}_{it} = u_{it} - \bar{u}_i$$

- **Paso 3:** Estimar una regresión de \tilde{y}_{it} sobre \tilde{x}_{it} usando **MCO agrupados**.

Intuición

- **Explicación:** El método within elimina c_i comparando a cada individuo consigo mismo a lo largo del tiempo.
- **Ejemplo:** Es como medir el cambio de salario de una persona en distintos años, ignorando su talento fijo o características innatas.

-
- Para que el estimador **FE** tenga buen comportamiento asintótico, se requiere la **condición de rango estándar**:

$$\text{FE2: } \text{rank}\left(\sum_{t=1}^T E(\tilde{x}_{it}\tilde{x}_{it}^T)\right) = K$$

- Si x_{it} incluye un componente que no cambia en el tiempo para cada i , entonces su correspondiente en \tilde{x}_{it} será **cero**, y la condición de rango no se cumplirá.
- Esto implica que no se pueden incluir **variables invariantes en el tiempo** en modelos de efectos fijos.
- Además, sin supuestos adicionales, el estimador **FE** no es necesariamente el más **eficiente**. El siguiente supuesto asegura eficiencia y una matriz de varianza apropiada:

$$\text{FE3: } E(u_i u_i^T | x_i, c_i) = \sigma_u^2 I_T$$

Intuición

- **Explicación:** Las condiciones FE2 y FE3 garantizan que el estimador FE pueda identificar correctamente los parámetros y obtener varianzas válidas.
- **Ejemplo:** Si tratamos de usar género en un modelo FE, siempre será constante para cada individuo y no se podrá estimar su efecto.

Within Estimator: Errores estándar

- Los errores estándar de la regresión **MCO** en el paso 3 no son correctos.
- ¿Por qué ocurre? Al restar la media a cada observación (**demeaning**) se introduce **correlación serial** en los errores.
- **Varianza de \tilde{u}_{it} :**

$$\begin{aligned} E(\tilde{u}_{it}^2) &= E[(u_{it} - \bar{u}_i)^2] = E(u_{it}^2) + E(\bar{u}_i^2) - 2E(u_{it}\bar{u}_i) \\ &= \sigma_u^2 + \frac{\sigma_u^2}{T} - \frac{2\sigma_u^2}{T} = \sigma_u^2 \left(1 - \frac{1}{T}\right) \end{aligned}$$

- Covarianza entre \tilde{u}_{it} y \tilde{u}_{is} para $t \neq s$:

$$\begin{aligned} E(\tilde{u}_{it}\tilde{u}_{is}) &= E[(u_{it} - \bar{u}_i)(u_{is} - \bar{u}_i)] \\ &= E(u_{it}u_{is}) - E(u_{it}\bar{u}_i) - E(u_{is}\bar{u}_i) + E(\bar{u}_i^2) \\ &= 0 - \frac{\sigma_u^2}{T} - \frac{\sigma_u^2}{T} + \frac{\sigma_u^2}{T} = -\frac{\sigma_u^2}{T} \end{aligned}$$

Intuición

- **Explicación:** El proceso de restar medias genera correlación artificial entre los errores dentro del mismo individuo, lo cual sesga los errores estándar simples.
- **Ejemplo:** Es como calcular la desviación de cada examen respecto al promedio de un estudiante: esas desviaciones no son independientes entre sí.

- Como resultado, la **correlación** entre \tilde{u}_{it} y \tilde{u}_{is} es:

$$\begin{aligned} \text{Corr}(\tilde{u}_{it}, \tilde{u}_{is}) &= \frac{\text{Cov}(\tilde{u}_{it}, \tilde{u}_{is})}{\sqrt{\text{Var}(\tilde{u}_{it})\text{Var}(\tilde{u}_{is})}} \\ &= \frac{-\sigma_u^2/T}{\sigma_u^2(1 - 1/T)} = -\frac{1}{T - 1} \end{aligned}$$

- El **supuesto 3** nos permite derivar el estimador de la varianza asintótica:

$$\text{Avar}(\hat{\beta}_{FE}) = \hat{\sigma}_u^2 \left(\sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{it} \tilde{x}_{it}^T \right)^{-1}$$

- Notar que $\hat{\sigma}_u^2$ es un **estimador consistente** para la varianza de u_{it} , no de \tilde{u}_{it} .
- Por lo tanto, no es posible obtener $\hat{\sigma}_u^2$ directamente de la regresión MCO del paso 3.

Intuición

- **Explicación:** La correlación negativa entre los errores transformados surge del demeaning, y por eso debemos ajustar la estimación de la varianza.
- **Ejemplo:** Es como medir notas relativas al promedio: si una observación está arriba de la media, otra debe estar abajo, generando correlación automática.

- El **estimador estándar** de la varianza en la regresión del paso 3 es:

$$\frac{SSR}{NT - K}$$

- Sin embargo, este es un **estimador incorrecto**, pues necesitamos la varianza de $\hat{\sigma}_u$ y no la de $\hat{\sigma}_u$.

- La resta de K en el denominador no afecta asintóticamente, pero es habitual aplicar esa **corrección** en la práctica.

Intuición

- **Explicación:** El error estándar calculado directamente tras el paso 3 no refleja la varianza verdadera de los errores originales, por lo que debe ajustarse.
- **Ejemplo:** Es como calcular una desviación típica con datos transformados: el ajuste asegura que la medida corresponda a la variabilidad real.

-
- Recordemos que la varianza de \tilde{u}_{it} es:

$$\sigma_u^2(1 - 1/T)$$

Sumando sobre t obtenemos:

$$\sum_{t=1}^T E(\tilde{u}_{it}^2) = \sigma_u^2(T - 1)$$

- Si además sumamos sobre todos los individuos N :

$$\sum_{i=1}^N \sum_{t=1}^T E(\tilde{u}_{it}^2) = \sigma_u^2(T - 1)N \quad \Rightarrow \quad \sigma_u^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T E(\tilde{u}_{it}^2)}{(T - 1)N}$$

- Así, podemos obtener un **estimador consistente** de σ_u^2 a partir de la regresión en el paso 3:

$$\hat{\sigma}_u^2 = \frac{SSR}{N(T - 1) - K}$$

- La diferencia con el uso de $SSR/(NT - K)$ es pequeña, pero este ajuste asegura la correcta varianza bajo el modelo within.

Intuición

- **Explicación:** Al eliminar un grado de libertad por individuo al restar la media, el denominador se ajusta de NT a $N(T - 1)$.
- **Ejemplo:** Es como calcular la varianza dentro de cada grupo: al perder una observación efectiva por grupo, se corrige el divisor total.

-
- Los paquetes de regresión habituales (como **STATA**) hacen el **ajuste de errores estándar** de manera automática cuando se especifica un modelo de efectos fijos.
 - Si se desea estimar el modelo **paso a paso**, se pueden aplicar los tres pasos descritos y luego corregir los errores estándar de la regresión obtenida en el paso 3.

- El ajuste se logra multiplicando los errores estándar por el factor:

$$\left(\frac{NT - K}{N(T - 1) - K} \right)^{1/2}$$

Intuición

- **Explicación:** El ajuste corrige la diferencia entre usar NT y $N(T - 1)$ observaciones efectivas en el cálculo de varianza.
- **Ejemplo:** Es como estandarizar notas con un divisor ligeramente distinto: el software lo hace automático, pero manualmente hay que escalar los resultados.

2 Estimador de Variables Dummy

- Una forma alternativa de estimar modelos de **efectos fijos** (en particular cuando N es pequeño o si se desean los efectos fijos explícitos) es estimar c_i mediante un conjunto de **dummies** para cada individuo.
- Esto implica incluir N dummies (una por cada i en la muestra) dentro de la regresión y estimar:

$$y_{it} = x_{it}\beta + c_i + u_{it} \quad (3)$$

usando **MCO**. A este procedimiento se le llama **estimador de variables dummy**.

- Una ventaja de este método es que entrega los **errores estándar correctos**, ya que emplea

$$NT - N - K = N(T - 1) - K$$

grados de libertad.

- El costo es que, si N es grande, la regresión requiere un **alto poder computacional**.

Intuición

- **Explicación:** Con las dummies se controla directamente por c_i , permitiendo estimar efectos fijos sin transformar los datos.
- **Ejemplo:** Es como dar a cada persona un “identificador único” en la regresión, lo que elimina la influencia de sus características permanentes.

3 Primeras Diferencias

- Otra alternativa para estimar modelos de **efectos fijos** es usar las **primeras diferencias**.
- Bajo el supuesto de **exogeneidad estricta** condicional a c_i :

$$\text{FD1: } E(u_{it} \mid x_i, c_i) = 0, \quad t = 1, 2, \dots, T$$

- Rezagando el modelo original

$$y_{it} = x_{it}\beta + c_i + u_{it}$$

un periodo y restando, se obtiene:

$$y_{it} - y_{i,t-1} = (x_{it} - x_{i,t-1})\beta + (c_i - c_i) + (u_{it} - u_{i,t-1})$$

$$\Delta y_{it} = \Delta x_{it}\beta + \Delta u_{it}$$

- Tomar **primeras diferencias** elimina c_i automáticamente.
- Como consecuencia, se pierde la primera observación temporal de cada unidad de corte transversal.

Intuición

- **Explicación:** Al restar observaciones consecutivas, desaparecen los efectos invariantes en el tiempo, dejando solo la variación relevante.
- **Ejemplo:** Es como analizar cuánto cambió el ingreso de un individuo respecto al año anterior: su talento fijo se cancela al comparar diferencias.

-
- El **estimador FD** $\hat{\beta}_{FD}$ se obtiene como el estimador de **MCO agrupados** de la regresión:

$$\Delta y_{it} \text{ en } \Delta x_{it}$$

- Bajo el supuesto **FD1**, la estimación por MCO agrupados es **consistente** e **insesgada**.
- Como en el caso within, se requiere una **condición de rango** para el estimador FD:

$$\text{FD2: } \text{rank} \left(\sum_{t=2}^T E(\Delta x_{it}^T \Delta x_{it}) \right) = K$$

- Esta condición también excluye a las **variables invariantes en el tiempo** y evita la **multicolinealidad perfecta** entre las variables que sí cambian en el tiempo.

Intuición

- **Explicación:** El estimador FD usa solo la variación de un periodo a otro, y necesita que esa variación aporte información independiente para identificar los parámetros.
- **Ejemplo:** Si todas las personas tuvieran exactamente el mismo cambio en ingresos de un año a otro, no se podría estimar el efecto de las variables explicativas.

Primeras Diferencias: Errores estándar

- Bajo los supuestos **FE1–FE3**, el estimador de efectos fijos (**within**) es asintóticamente **eficiente**.
- Sin embargo, el supuesto **FE3** establece que los u_{it} no presentan **correlación serial**.
- Si en su lugar suponemos que los Δu_{it} no están serialmente correlacionados, entonces el estimador de **Primeras Diferencias (FD)** sería **eficiente**.
- Además, puede demostrarse que cuando solo existen **dos periodos**, el estimador FD y el estimador FE son **idénticos**.

Intuición

- **Explicación:** La eficiencia del estimador depende de cómo se comporten los errores en el tiempo: si la correlación está en u_{it} favorece FE, si está en Δu_{it} favorece FD.
- **Ejemplo:** Con solo dos años de datos, calcular diferencias o restar la media es lo mismo: en ambos casos se elimina el efecto fijo c_i .

Regresión de efectos fijos: puntos principales

- En **inferencia**, los errores estándar deben estar **agrupados por unidad de panel** (ejemplo: granja) para permitir correlación en los u_{it} dentro de un mismo i .
- En **STATA**:

```
xtreg, fe i(PanelID) cluster(PanelID)
```

Esto produce inferencia válida siempre que el número de clusters sea suficientemente grande.

- Normalmente el interés está en β , pero los efectos c_i también pueden ser relevantes:
 - \hat{c}_i del estimador de **variables dummy** es insesgado, pero no consistente para c_i (con T fijo y $N \rightarrow \infty$).
 - **xtreg, fe** sustrae medias antes de estimar, por lo que no reporta \hat{c}_i .
 - La constante reporta el promedio de los \hat{c}_i entre unidades.
 - Podemos recuperar \hat{c}_i con:

$$\hat{c}_i = \bar{y}_i - \bar{x}_i \hat{\beta} \quad (\text{predict ci, u})$$

Intuición

- **Explicación:** La inferencia en panel requiere ajustar por correlación interna de los errores; además, los efectos c_i pueden recuperarse pero no siempre son consistentes.
- **Ejemplo:** Al analizar productividad de granjas, los errores dentro de una misma granja están correlacionados; además, el efecto fijo de cada granja se puede calcular a partir de sus medias.

Ejemplo: Democracia Directa y Naturalizaciones

- Pregunta central: ¿están en peor situación las minorías bajo **democracia directa** que bajo **democracia representativa**?
- **Hainmueller y Hangartner (2012)** analizan datos de solicitudes de naturalización de inmigrantes en Suiza.
- En algunos municipios las solicitudes se votan en:
 - **Referéndums (democracia directa).**
 - **Consejos municipales electos (democracia representativa).**
- Base de datos: panel anual de **1,400 municipios** en el periodo 1991–2009.
- Definición de variables:

$$y_{it} = \text{tasa de naturalización} = \frac{\text{número de naturalizaciones}_{it}}{\text{población extranjera elegible}_{it-1}}$$
$$x_{it} = \begin{cases} 1 & \text{si el municipio usa democracia representativa en } t \\ 0 & \text{si el municipio usa democracia directa en } t \end{cases}$$

Intuición

- **Explicación:** El diseño permite comparar cómo cambia la tasa de naturalización dependiendo del mecanismo de decisión política en el municipio.
- **Ejemplo:** Dos municipios con características similares pueden diferir solo en el tipo de democracia, lo que permite identificar el efecto de la regla institucional.

Datos de panel de naturalizaciones

- Base de datos en formato **panel** con información de municipios en Suiza.
- Variables principales:

Variable	Tipo	Descripción
muniID	float	Código del municipio
muni_name	string	Nombre del municipio (1,430 municipios)
year	int	Año de observación
nat_rate	float	Tasa de naturalización (%)
repdem	float	1 = democracia representativa, 0 = democracia directa

Intuición

- **Explicación:** El panel combina municipios y años, permitiendo analizar cómo la institución política (*repdem*) afecta la tasa de naturalización.
- **Ejemplo:** Si un municipio cambia de democracia directa a representativa, podemos observar cómo varía su tasa de naturalización en el tiempo.

Datos de Panel. Long Format

```
. list muniID muni_name year nat_rate repdem in 31/40
```

muniID	muni_name	year	nat_rate	repdem
2	Affoltern A.A.	2002	4.638365	0
2	Affoltern A.A.	2003	4.844814	0
2	Affoltern A.A.	2004	5.621302	0
2	Affoltern A.A.	2005	4.387827	0
2	Affoltern A.A.	2006	8.115358	1
2	Affoltern A.A.	2007	7.067371	1
2	Affoltern A.A.	2008	8.977719	1
2	Affoltern A.A.	2009	6.119704	1
3	Bonstetten	1991	0.833334	0
3	Bonstetten	1992	0.8403362	0

Intuición

- La clave del formato **Long** es que cada observación individual (**muniID-year**) tiene su propia **fila**, facilitando el seguimiento de los cambios temporales.
- El panel permite observar **variación en el tiempo** (diferentes años para el mismo municipio) y **variación entre unidades** (diferentes **muniID**).

MCO Agrupados

```
. reg nat_rate repdem , cl(muniID)
```

Regresión Lineal

Tabla 1. Resultados de Regresión Lineal con Errores Agrupados

Number of obs = 4655

F(1, 244) = 130.04

Prob >F = 0.0000

R-squared = 0.0748

Root MSE = 3.98

(Std. Err. adjusted for 245 clusters in muniID)

nat_rate t	Coef. P> t	Robust Std. Err. [95 % Conf. Interval]
repdem	2.503318	0.2195202
11.40	0.000	[2.070921 2.935714]
_cons	2.222683	0.10088
22.03	0.000	[2.023976 2.421389]

Intuición

- **Efecto Estimado:** El coeficiente de `repdem` (2.50) indica que la democracia representativa está asociada con un aumento de **2.5 puntos porcentuales** en la tasa de naturalización.
- **Agrupamiento (Clustering):** El uso de `cl(muniID)` corrige los errores estándar por **autocorrelación** dentro del municipio (correlación entre las observaciones a lo largo del tiempo para el mismo municipio), haciendo que la inferencia sea **válida** y **conservadora**.

Descomposición de Varianza (within, between)

```
. tsset muniID year , yearly
panel variable: muniID (strongly balanced)
time variable: year, 1991 to 2009
delta: 1 year

. xtsum nat_rate
```

Tabla 2. Descomposición de Varianza de la Tasa de Naturalización (`nat_rate`)

Variable	Mean	Std. Dev.	Min	Max	Observations
<code>nat_rate</code> _overall	2.938992	4.137305	0	24.13793	N = 4655
between		1.622939	0	7.567746	n = 245
within		3.807039	-3.711323	24.80134	T = 19

Intuición

- **Varianza Between (σ_μ):** Mide la variación **entre** municipios. Su desviación estándar (1.62) es relativamente **menor** que la `*within*`, sugiriendo que las diferencias **promedio** entre municipios son limitadas.
- **Varianza Within (σ_ϵ):** Mide la variación **dentro** de cada municipio a lo largo del tiempo. Su desviación estándar (3.81) es **mayor** que la `*between*`, indicando que la mayor parte de la variación en `nat_rate` proviene de los **cambios anuales** de la propia unidad.

Within Estimator

```
. * get municipality means
. egen means_nat_rate = mean(nat_rate) , by(muniID)

. * compute deviations from means
. gen dm_nat_rate = nat_rate - means_nat_rate

. list muniID muni_name year nat_rate means_nat_rate dm_nat_rate in 20/40 , ab(20)
```

Tabla 3. Ilustración de la Transformación de Desviación de la Media

muniID	muni_name	year	nat_rate	means_nat_rate	dm_nat_rate
2	Affoltern A.A.	1991	0.2173913	3.595932	-3.37854
2	Affoltern A.A.	1992	0.9473684	3.595932	-2.648563
2	Affoltern A.A.	1993	1.04712	3.595932	-2.548811
2	Affoltern A.A.	1994	0.8342023	3.595932	-2.761729
2	Affoltern A.A.	1995	2.002002	3.595932	-1.59393
2	Affoltern A.A.	1996	1.7769	3.595932	-1.819031
2	Affoltern A.A.	1997	1.862745	3.595932	-1.733186
2	Affoltern A.A.	1998	2.054155	3.595932	-1.541776
2	Affoltern A.A.	1999	2.402135	3.595932	-1.193796

Intuición

- **Eliminación del Sesgo:** La columna **dm_nat_rate** (Desviación de la Media) representa la tasa de naturalización, pero eliminando la media de cada municipio (**means_nat_rate**).
- **Base de la Regresión:** Al restar la media de grupo, se eliminan los **efectos fijos** o características **invariantes en el tiempo** de cada municipio. La regresión del Estimador Within se ejecuta solo sobre estas variables ****transformadas****.

Efectos Fijos. Within Estimator

```
. egen means_repdem = mean(repdem) , by(muniID)
. gen dm_repdem = repdem - means_repdem

. * regression with demeaned data
. reg dm_nat_rate dm_repdem , cl(muniID)
```

Regresión Lineal con Variables Transformadas

Tabla 4. Resultados del Estimador Within (Efectos Fijos)

Linear regression					
Number of obs = 4655					
F(1, 244) = 265.18					
Prob >F = 0.0000					
R-squared = 0.1052					
Root MSE = 3.6017					
(Std. Err. adjusted for 245 clusters in muniID)					
dm_nat_rate	Coef.	Robust Std. Err.	t	P> t	[95 % Conf. Interval]
dm_repdem	3.0228	0.1856244	16.28	0.000	[2.657169 3.388431]
_cons	6.65e-10	5.81e-09	0.11	0.909	[-1.08e-08 1.21e-08]

Intuición

- **Interpretación del Coeficiente (dm_repdem):** El coeficiente de **3.02** mide el efecto causal puro de la **variación temporal** en la democracia (**repdem**) sobre la **variación temporal** en la tasa de naturalización, ****eliminando**** las características permanentes del municipio.
- **Comparación con MCO:** El efecto (3,02) es mayor que el MCO agrupado (2,50), sugiriendo que el MCO estaba sesgado a la baja (subestimado) el verdadero efecto, posiblemente debido a que factores **no observados** e invariantes estaban correlacionados con ambas variables.

Efectos Fijos. Rutina programada: Stata

```
. xtreg nat_rate repdem , fe cl(muniID) i(muniID)
```

Tabla 5. Resultados del Estimador de Efectos Fijos (xtreg, fe)

Fixed-effects (within) regression					
Group variable: muniID					
R-sq:	within = 0.1052				
	between = 0.0005				
	overall = 0.0748				
	corr(u_i, Xb) = -0.1373				
	Number of obs = 4655		F(1, 244) = 265.18		
	Number of groups = 245		Prob >F = 0.0000		
	Obs per group: min = 19				
	avg = 19.0				
	max = 19				
(Std. Err. adjusted for 245 clusters in muniID)					
nat_rate	Coef.	Robust Std. Err.	t	P> t	[95 % Conf. Interval]
repdem	3.0228	0.1856244	16.28	0.000	[2.657169 3.388431]
_cons	2.074036	0.0531153	39.05	0.000	[1.969413 2.178659]
sigma_u = 1.7129711					
sigma_e = 3.69998					
rho = 0.17650677 (fraction of variance due to u_i)					

Intuición

- **Consistencia:** Los resultados de **xtreg, fe** confirman el coeficiente de ****3.0228**** encontrado con la regresión ***demeaned***, validando la implementación de los **Efectos Fijos**.
- **Rho (ρ):** El valor de $\rho \approx 0.177$ indica que aproximadamente el ****17.7%**** de la varianza total del error se debe a los **efectos no observados invariantes** específicos de cada municipio (σ_u), lo que justifica el uso de un modelo de panel.

Efectos Fijos. Rutina programada: Stata

```
. xtreg nat_rate repdem , fe cl(muniID) i(muniID)
```

Tabla 6. Resultados del Estimador de Efectos Fijos (xtreg, fe)

Fixed-effects (within) regression					
Group variable: muniID					
R-sq:	within = 0.1052				
	between = 0.0005				
	overall = 0.0748				
corr(u_i, Xb) =	-0.1373				
Number of obs =	4655	F(1, 244) =	265.18		
Number of groups =	245	Prob >F =	0.0000		
Obs per group: min =	19				
	avg = 19.0				
	max = 19				
(Std. Err. adjusted for 245 clusters in muniID)					
nat_rate	Coef.	Robust Std. Err.	t	P> t	[95 % Conf. Interval]
repdem	3.0228	0.1856244	16.28	0.000	[2.657169 3.388431]
_cons	2.074036	0.0531153	39.05	0.000	[1.969413 2.178659]
sigma_u = 1.7129711					
sigma_e = 3.69998					
rho = 0.17650677 (fraction of variance due to u_i)					

Intuición

- **Consistencia:** Los resultados de **xtreg, fe** confirman el coeficiente de ****3.0228**** encontrado con la regresión ***demeaned***, validando la implementación de los **Efectos Fijos**.
- **Rho (ρ):** El valor de $\rho \approx 0,177$ indica que aproximadamente el ****17.7%**** de la varianza total del error se debe a los **efectos no observados invariantes** específicos de cada municipio (σ_u), lo que justifica el uso de un modelo de panel.

Aplicando efectos fijos

- Los **efectos fijos** también pueden aplicarse en otras estructuras de datos, permitiendo restringir las comparaciones dentro de la misma unidad.
- **Ejemplos:**
 - **Parejas apareadas:** usar efectos fijos de gemelos para controlar características familiares no observadas.

- **Escuelas:** incluir efectos fijos de escuela para capturar factores no observados propios de cada institución.

Intuición

- **Explicación:** El método de efectos fijos funciona como un “control interno” que elimina factores constantes dentro de cada unidad.
- **Ejemplo:** Comparar gemelos elimina el efecto de los antecedentes familiares; comparar alumnos dentro de la misma escuela elimina el efecto del contexto escolar fijo.

Problemas que los efectos fijos no resuelven

- Modelo de referencia:

$$y_{it} = x_{it}\beta + c_i + u_{it}, \quad t = 1, 2, \dots, T$$

donde:

$$y_{it} = \text{tasa de homicidios}, \quad x_{it} = \text{gasto en policía per cápita}$$

- Si corremos una regresión de y sobre x con efectos fijos por ciudad, el estimador $\hat{\beta}_{FE}$ será **inconsistente** salvo que se cumpla la **estricta exogeneidad**:

$$E(u_{it} \mid x_{i1}, x_{i2}, \dots, x_{iT}, c_i) = 0, \quad t = 1, 2, \dots, T$$

- Esto implica que u_{it} no debe estar correlacionado con regresores pasados, actuales ni futuros.
- **Violaciones comunes:**
 1. **Variables omitidas que varían en el tiempo:** un boom económico puede llevar a más gasto policial y, al mismo tiempo, a menos homicidios.
 2. **Simultaneidad:** si la ciudad ajusta el gasto policial en respuesta a tasas de homicidios pasadas, entonces x_{t+1} estará correlacionado con u_t . Una x estrictamente exógena no puede reaccionar a lo que le ocurre a y en el pasado o el futuro.

Intuición

- **Explicación:** Los efectos fijos eliminan solo factores constantes; no corrigen problemas de variables omitidas dinámicas ni de causalidad inversa.
- **Ejemplo:** Si una ciudad aumenta la policía porque hubo más homicidios el año pasado, el gasto y los errores pasados estarán correlacionados, rompiendo la exogeneidad.

Efectos Aleatorios

- Consideremos nuevamente el modelo:

$$y_{it} = x_{it}\beta + c_i + u_{it}, \quad t = 1, 2, \dots, T$$

- Un problema del modelo de **efectos fijos** es que no puede estimar el impacto de **regresores invariantes en el tiempo** (ejemplo: calidad del suelo, ubicación de la granja).
- Como el estimador FE permite correlación entre c_i y x_{it} , no es posible separar el efecto de esos regresores constantes del efecto fijo c_i .
- Para estimarlos, necesitamos el supuesto de **ortogonalidad**:

$$\text{Cov}(x_{it}, c_i) = 0, \quad \forall t = 1, \dots, T$$

- Este es un **supuesto fuerte**: se requiere que los efectos no observados c_i no estén correlacionados con ninguna variable explicativa en x_{it} en ningún periodo.
- Ejemplo: si incluimos **calidad del suelo** en x_{it} , debemos suponer que no se correlaciona con otros insumos invariantes en el tiempo (como ubicación geográfica).

Intuición

- **Explicación**: El modelo de efectos aleatorios permite identificar variables constantes en el tiempo, pero exige independencia fuerte entre c_i y los regresores.
- **Ejemplo**: Es como analizar granjas con diferentes calidades de suelo: para que el modelo funcione, debemos asumir que la calidad del suelo no está relacionada con otros factores fijos como el clima de la región.

Métodos de efectos aleatorios (RE)

- Los modelos de **efectos aleatorios** incluyen c_i dentro del término de error, bajo el supuesto de que c_i es **ortogonal** a x_{it} , y corrigen la **correlación serial** en el error compuesto.
- Estos modelos imponen dos condiciones: **exogeneidad estricta** y **ortogonalidad** entre c_i y x_{it} .
- Supuestos (RE1):

$$E(u_{it} \mid x_i, c_i) = 0, \quad t = 1, 2, \dots, T$$

$$E(c_i \mid x_i) = E(c_i) = 0, \quad \text{con } x_i = (x_{i1}, x_{i2}, \dots, x_{iT})$$

- Lo importante es que:

$$E(c_i \mid x_i) = E(c_i)$$

El supuesto $E(c_i) = 0$ no implica pérdida de generalidad siempre que haya una constante incluida en x_{it} .

- Bajo este supuesto, incluso **MCO** sería **consistente**, pero no **eficiente**. Por eso, se emplea **MCG (GLS)** para obtener eficiencia.

Intuición

- **Explicación:** El modelo RE trata los efectos individuales como parte del error, suponiendo que no están correlacionados con los regresores, y usa GLS para mejorar la eficiencia.
- **Ejemplo:** Es como analizar escuelas donde la “calidad promedio” de cada escuela se asume independiente del gasto en insumos: si se cumple, se puede estimar con más precisión que con FE.

-
- El enfoque de **efectos aleatorios** considera explícitamente la **correlación serial** en el error compuesto:

$$v_{it} = c_i + u_{it}$$

- El modelo de regresión puede escribirse como:

$$y_{it} = x_{it}\beta + v_{it}$$

- Bajo los supuestos de efectos aleatorios:

$$E(v_{it} | x_i) = 0, \quad t = 1, 2, \dots, T$$

- En consecuencia, podemos aplicar **MCG (GLS)** para tener en cuenta la estructura particular del error compuesto $v_{it} = c_i + u_{it}$.

Intuición

- **Explicación:** El error compuesto v_{it} tiene una parte común (c_i) y otra idiosincrática (u_{it}), lo que genera correlación entre periodos para un mismo individuo.
- **Ejemplo:** Es como medir la productividad de una persona: siempre habrá un componente fijo de talento (c_i) más choques aleatorios cada año (u_{it}); GLS ajusta por esa estructura.

-
- El modelo puede escribirse para todos los periodos como:

$$y_i = X_i\beta + v_i$$

donde y_i es un vector $T \times 1$, X_i es una matriz $T \times K$, y v_i el vector de errores compuestos.

- Definimos la matriz de varianzas no condicional de v_i como:

$$\Omega \equiv E(v_i v_i^T)$$

- Ω es una matriz de dimensión $T \times T$, y es **la misma para todos los individuos i** porque se asume **muestreo aleatorio** en la dimensión transversal.

- Para la consistencia del estimador GLS necesitamos la condición de rango habitual:

$$\text{RE2: } \text{rank}(E(X_i^T \Omega^{-1} X_i)) = K$$

Intuición

- **Explicación:** La estructura del error compuesto implica una matriz de covarianza común para todos los individuos; GLS usa esta matriz para ponderar correctamente las observaciones.
- **Ejemplo:** Es como dar más o menos peso a observaciones de un mismo individuo dependiendo de cuánta correlación exista entre sus periodos.

-
- Un análisis estándar de efectos aleatorios añade supuestos sobre los **errores idiosincráticos**, lo que da a Ω una forma particular.

- Supuestos (RE3):

$$E(u_i u_i^T | x_i, c_i) = \sigma_u^2 I_T$$

$$E(c_i^2 | x_i) = \sigma_c^2$$

- Bajo estos supuestos, la matriz de varianza Ω es:

$$\Omega = \begin{bmatrix} \sigma_c^2 + \sigma_u^2 & \sigma_c^2 & \cdots & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2 + \sigma_u^2 & \cdots & \sigma_c^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_c^2 & \sigma_c^2 & \cdots & \sigma_c^2 + \sigma_u^2 \end{bmatrix}$$

- Es decir, la matriz Ω tiene $\sigma_c^2 + \sigma_u^2$ en la diagonal, y σ_c^2 en todas las posiciones fuera de la diagonal.

Intuición

- **Explicación:** La parte fija c_i genera correlación perfecta entre todos los periodos de un mismo individuo, mientras que u_{it} añade variación idiosincrática independiente.
- **Ejemplo:** Para un trabajador, c_i puede reflejar su productividad innata (común a todos los años), mientras que u_{it} recoge shocks de productividad que cambian cada año.

-
- Si contamos con estimadores consistentes de σ_c^2 y σ_u^2 (ver **Wooldridge, p. 260** sobre cómo obtenerlos), podemos construir un estimador consistente de la matriz de varianza:

$$\hat{\Omega} \equiv \hat{\sigma}_u^2 I_T + \hat{\sigma}_c^2 j_T j_T^T$$

donde $j_T j_T^T$ es una matriz $T \times T$ con unos en todas sus entradas.

- Con esta matriz, el estimador de **efectos aleatorios** es:

$$\hat{\beta}_{RE} = \left(\sum_{i=1}^N X_i^T \hat{\Omega}^{-1} X_i \right)^{-1} \left(\sum_{i=1}^N X_i^T \hat{\Omega}^{-1} y_i \right)$$

- El estimador RE es un caso particular de **GLS factible**, donde solo se deben estimar dos parámetros en la matriz de varianza-covarianza.
- Si los supuestos de RE se cumplen, este estimador es **consistente** y **eficiente**.

Intuición

- **Explicación:** El método RE estima la estructura de la varianza entre y dentro de individuos y aplica GLS factible para ponderar las observaciones.
- **Ejemplo:** Es como ajustar la importancia relativa de variaciones entre municipios y dentro de cada municipio al estimar el efecto de una política.

Regresión de efectos aleatorios (RE)

```
. xtreg nat_rate repdem , re cl(muniID) i(muniID)
```

Tabla 7. Resultados del Estimador de Efectos Aleatorios (xtreg, re)

Random-effects GLS regression					
Group variable: muniID					
R-sq:		within = 0.1052			
		between = 0.0005			
		overall = 0.0748			
		corr(u_i, X) = 0 (assumed)			
Number of obs = 4655		Wald chi2(1) = 227.99			
Number of groups = 245		Prob >chi2 = 0.0000			
Obs per group: min = 19					
		avg = 19.0			
		max = 19			
(Std. Err. adjusted for 245 clusters in muniID)					
nat_rate	Coef.	Robust Std. Err.	z	P> z	[95 % Conf. Interval]
repdem	2.859397	0.1893742	15.10	0.000	[2.48823 3.230564]
_cons	2.120793	0.0972959	21.80	0.000	[1.930096 2.311489]
sigma_u = 1.3866768					
sigma_e = 3.69998					
rho = 0.1231606 (fraction of variance due to u_i)					

Intuición

- **Supuesto Clave:** El estimador de **Efectos Aleatorios (RE)** asume que los efectos no observados (u_i) ****no están correlacionados**** con la variable explicativa (**repdem**). Esta es la principal diferencia con el modelo FE.
- **Eficiencia vs. Sesgo:** El coeficiente RE (**2,86**) está entre el MCO (2,50) y el FE (3,02). RE es más ****eficiente**** (usa variación entre y dentro), pero si el supuesto de no correlación es falso (como sugiere la prueba de Hausman), sus estimaciones serán ****sesgadas****.

RE versus FE, FD o Variables Dummy

- El supuesto central de **RE**, $E(c_i | x_i) = E(c_i) = 0$, es muy **fuerte** y poco realista en muchos contextos.
- Por ello, en la práctica suelen preferirse los **estimadores de efectos fijos**.
- Pregunta clave: ¿**qué estimador de efectos fijos elegir?**
- Con solo **dos periodos**, los estimadores **FE**, **FD** y de **variables dummy** son idénticos.
- Si $T > 2$, la elección depende de los **supuestos sobre los errores** u_{it} :
 - En la práctica, se suele asumir **FE3** (no autocorrelación serial) y usar el estimador **within (FE)**.
 - Si el interés está en recuperar los c_i , puede emplearse el estimador de **variables dummy**.

Intuición

- **Explicación:** Los métodos FE, FD y dummies son robustos a la correlación entre c_i y x_{it} ; RE solo funciona bajo el supuesto más restrictivo de independencia.
- **Ejemplo:** Con pocos años de datos, todas las técnicas dan lo mismo; con más periodos, la elección depende de cómo supongamos que se comportan los errores.

Prueba de Hausman

- Comparación entre estimadores **RE** y **FE**:

	$\hat{\beta}_{RE}$	$\hat{\beta}_{FE}$
$H_0 : \text{Cov}[x_{it}, c_i] = 0$	Consistente y eficiente	Consistente
$H_1 : \text{Cov}[x_{it}, c_i] \neq 0$	Inconsistente	Consistente

- Bajo H_0 , la diferencia $\hat{\beta}_{RE} - \hat{\beta}_{FE}$ debe ser cercana a cero.
- Bajo H_1 , esta diferencia será significativa.

- El estadístico de prueba (para muestras grandes) es:

$$(\hat{\beta}_{FE} - \hat{\beta}_{RE})^T \left(\widehat{\text{Var}}[\hat{\beta}_{FE}] - \widehat{\text{Var}}[\hat{\beta}_{RE}] \right)^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE}) \xrightarrow{d} \chi_k^2$$

donde k es el número de regresores que varían en el tiempo.

- Si se rechaza H_0 , concluimos que los efectos aleatorios no son apropiados y preferimos la especificación **FE**.

Intuición

- **Explicación:** La prueba de Hausman compara directamente los estimadores RE y FE; si difieren mucho, RE es inconsistente.
- **Ejemplo:** Es como contrastar dos termómetros: si uno siempre mide más alto que otro, asumimos que está mal calibrado (RE) y confiamos en el más robusto (FE).

Regresión de efectos aleatorios (RE)

```
. quietly: xtreg nat_rate repdem , fe i(muniID)
. estimates store FE

. quietly: xtreg nat_rate repdem , re i(muniID)
. estimates store RE

. hausman FE RE
```

Tabla 8. Prueba de Hausman para la Elección del Modelo de Panel

	Coefficients			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	FE	RE	Difference	S.E.
repdem	3.0228	2.859397	0.1634027	0.0304517

b = consistent under H_0 and H_a ; obtained from xtreg

B = inconsistent under H_a , efficient under H_0 ; obtained from xtreg

Test: H_0 : difference in coefficients not systematic

$$\begin{aligned} \text{chi2}(1) &= (b-B)' [(V_b-V_B)^{-1}] (b-B) \\ &= 28.79 \end{aligned}$$

Intuición

- **Hipótesis Nula (H_0):** La diferencia entre los coeficientes **FE** y **RE** no es sistemática (RE es consistente y eficiente).
- **Conclusión:** Dado que el estadístico $\chi^2(1) = 28,79$ es grande y el valor **p** asociado es muy pequeño (no visible, pero el valor es significativo), se **rechaza la hipótesis nula (H_0)**. Por lo tanto, el modelo de **Efectos Fijos (FE)** es el estimador **consistente** preferido.

Prueba de Hausman

- La **prueba de Hausman** no evalúa si el modelo de efectos fijos es correcto; lo que asume es que el estimador de efectos fijos es **consistente**.
- La versión convencional de la prueba usa un modelo de **errores homocedásticos** y no permite el uso de **clustering**.
- Existen variantes de la prueba (como la **Sargan-Hansen**) que permiten aplicar **clustering** a los errores estándar.
- Ejemplo en **STATA**:

```
* hausman test with clustering
quietly: xtreg nat_rate repdem , re i(muniID) cl(muniID)
xtoverid
```

- Resultado de prueba (ejemplo):

$$\text{Sargan-Hansen statistic} = 26,560, \quad \chi^2(1), \quad p\text{-value} = 0,0000$$

lo que lleva a rechazar la hipótesis nula y preferir la especificación de efectos fijos.

Intuición

- **Explicación:** La Hausman clásica compara RE vs FE, pero para datos de panel con correlación interna se recomienda una versión robusta al clustering.
- **Ejemplo:** En los municipios suizos, si el p-valor es muy bajo, concluimos que los efectos aleatorios no son válidos y que debemos usar efectos fijos.

Añadiendo efectos fijos en el tiempo

- Modelo de referencia:

$$y_{it} = x_{it}\beta + c_i + u_{it}, \quad t = 1, 2, \dots, T$$

- Supuesto de efectos fijos:

$$E(u_{it} \mid x_{i1}, x_{i2}, \dots, x_{iT}, c_i) = 0, \quad t = 1, 2, \dots, T$$

donde u_{it} no se correlaciona con regresores pasados, presentes ni futuros. Es decir, los regresores son **estrictamente exógenos** condicionados a c_i .

- **Violación típica:** choques comunes que afectan a todas las unidades de manera similar y están correlacionados con x_{it} .
- Ejemplos de choques comunes:
 - **Tendencias tecnológicas o climáticas** que influyen en la productividad agrícola.
 - **Tendencias migratorias** que impactan las tasas de naturalización.
- Para capturar estos efectos, podemos incluir **efectos fijos de tiempo** en el modelo.

Intuición

- **Explicación:** Los efectos fijos de tiempo permiten controlar choques comunes a todas las unidades en un periodo, evitando que sesguen la estimación.
- **Ejemplo:** Si en un año hay una reforma migratoria que afecta a todos los municipios, el efecto fijo de tiempo controla ese cambio común y evita confundirlo con el impacto de x_{it} .

-
- **Tendencia lineal común:**

$$y_{it} = x_{it}\beta + c_i + t + u_{it}, \quad t = 1, 2, \dots, T$$

Una tendencia de tiempo lineal que aplica a todas las unidades.

- **Efectos fijos de tiempo:**

$$y_{it} = x_{it}\beta + c_i + \tau_t + u_{it}, \quad t = 1, 2, \dots, T$$

Captura choques comunes a todas las unidades en cada periodo. Este modelo puede interpretarse como una **generalización de diferencias en diferencias**.

- **Tendencias específicas por unidad:**

$$y_{it} = x_{it}\beta + c_i + g_i \times t + \tau_t + u_{it}, \quad t = 1, 2, \dots, T$$

Permite que cada unidad i tenga su propia tendencia lineal en el tiempo.

Intuición

- **Explicación:** Los efectos fijos de tiempo y las tendencias permiten capturar dinámicas comunes o específicas que evolucionan con el tiempo y podrían sesgar la estimación.
- **Ejemplo:** Una política nacional que afecta a todos los municipios se recoge con τ_t , mientras que un municipio en crecimiento constante en su productividad agrícola se modela con una tendencia $g_i \times t$.

Modelando los efectos en el tiempo

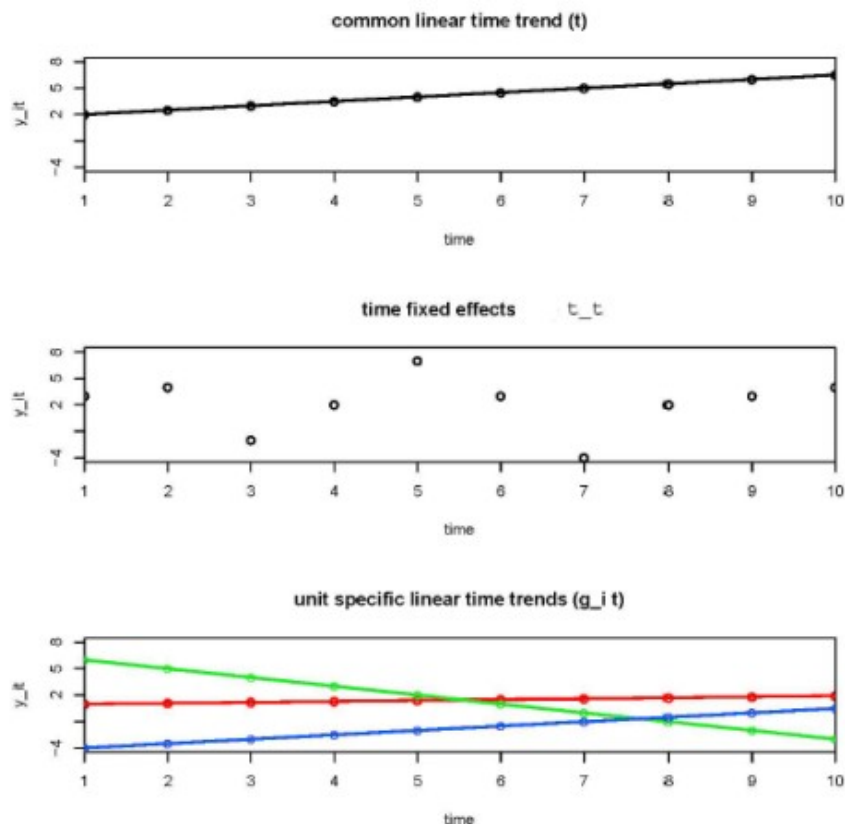


Figura 1: Representación gráfica de tendencias temporales: (1) tendencia lineal común, (2) efectos fijos de tiempo, (3) tendencias lineales específicas por unidad.

Intuición

- **Explicación:** Una tendencia común en el tiempo afecta a todas las unidades igual; los efectos fijos de tiempo permiten choques puntuales en cada periodo; las tendencias específicas dejan que cada unidad evolucione de forma diferente.
- **Ejemplo:** Todos los países pueden enfrentar un shock global en un año (efecto fijo de tiempo), pero algunos países pueden mostrar trayectorias crecientes y otros decrecientes en productividad (tendencias específicas).

Efectos fijos: añadiendo efectos en el tiempo

- Para incluir **efectos fijos de tiempo**, podemos crear una variable que agrupe los años de la muestra.
- En **STATA**:

```
. egen time = group(year)
. list muniID muni_name year time in 20/40 , ab(20)
```

- Ejemplo de salida:

muniID	muni_name	year	time
2	Affoltern A.A.	1991	1
2	Affoltern A.A.	1992	2
2	Affoltern A.A.	1993	3
2	Affoltern A.A.	1994	4
2	Affoltern A.A.	1995	5
2	Affoltern A.A.	1996	6
2	Affoltern A.A.	1997	7
2	Affoltern A.A.	1998	8
2	Affoltern A.A.	1999	9
2	Affoltern A.A.	2000	10
2	Affoltern A.A.	2001	11

Intuición

- **Explicación:** Al agrupar los años en una variable *time*, podemos incluir **dummies por periodo**, capturando choques comunes en cada año.
- **Ejemplo:** Si en 1995 hubo una crisis que afectó a todos los municipios, un efecto fijo de tiempo recoge ese impacto común.

Efectos fijos: tendencia lineal en el tiempo

- Ejemplo en **STATA** con tendencia lineal:

```
. xtreg nat_rate repdem time , fe cl(muniID) i(muniID)
```

- Resultados principales:

- $R^2_{within} = 0,1604$, $R^2_{between} = 0,0005$, $R^2_{overall} = 0,1350$.
- Test global: $F(2, 244) = 247,57$, p-valor = 0.0000.
- Coeficientes estimados:

$$\begin{aligned}\hat{\beta}_{repdem} &= 0,8248 & (p = 0,002) \\ \hat{\beta}_{time} &= 0,2314 & (p = 0,000) \\ \text{constante} &= 0,3893 & (p = 0,003)\end{aligned}$$

- Estimaciones de varianzas:

$$\sigma_u = 1,627, \quad \sigma_e = 3,584, \quad \rho = 0,171$$

donde ρ mide la proporción de varianza atribuida al componente u_i .

Intuición

- **Explicación:** Incluir una tendencia lineal en el tiempo permite capturar una evolución sistemática en la tasa de naturalización, además de los efectos fijos por municipio.
- **Ejemplo:** Si existe un aumento gradual en la naturalización de inmigrantes en todos los municipios a lo largo del tiempo, el término *time* lo recoge y evita sesgo en la estimación de *repdem*.

Efectos fijos: efectos fijos por año

- Ejemplo en **STATA** con efectos fijos de tiempo:

```
. xtreg nat_rate repdem i.time , fe cl(muniID) i(muniID)
```

- Resultados principales:

- $R^2_{within} = 0,1885$, $R^2_{between} = 0,0005$, $R^2_{overall} = 0,1575$.

- Test global: $F(19, 244) = 31,48$, p-valor = 0.0000.

- Coeficientes estimados:

$$\hat{\beta}_{repdem} = 1,2037 \quad (p = 0,000)$$

- Ejemplos de efectos fijos de tiempo:

$$time2 = 0,3829 \quad (p = 0,027), \quad time3 = 0,2789 \quad (p = 0,188), \quad time4 = 0,7034 \quad (p = 0,000)$$

- Estos coeficientes capturan **choques específicos por año**, comunes a todos los municipios.

Intuición

- **Explicación:** Incluir dummies de año permite controlar factores comunes en cada periodo (reformas, crisis, cambios políticos) que podrían sesgar las estimaciones.
- **Ejemplo:** Si en 1994 hubo una reforma migratoria que afectó a todos los municipios, el efecto fijo de ese año lo captura y evita atribuirlo erróneamente a *repdem*.

Efectos fijos: efectos de tiempo por unidad

- Ejemplo en **STATA** con interacción unidad \times tiempo:

```
. xtreg nat_rate repdem muniID#c.time i.time , fe cl(muniID) i(muniID)
```

- Este modelo permite que cada unidad (municipio) tenga su **propia tendencia temporal**.

- Resultados principales:

- $R^2_{within} = 0,2650$, $R^2_{between} = 0,5185$, $R^2_{overall} = 0,2864$.

- Coeficiente de interés:

$$\hat{\beta}_{repdem} = 0,9865 \quad (p = 0,002)$$

- Ejemplos de pendientes unitarias estimadas (interacciones municipio \times tiempo):

$$1 : 0,3334 \quad (p = 0,000), \quad 2 : 0,2914 \quad (p = 0,000), \quad 3 : 0,2499 \quad (p = 0,000)$$

- Nota: la categoría de referencia (año 19) fue omitida por colinealidad.

Intuición

- **Explicación:** Este modelo es una versión extendida de efectos fijos que permite que cada unidad tenga su propia tendencia temporal, capturando evoluciones heterogéneas.
- **Ejemplo:** Mientras un municipio puede mostrar una tendencia creciente en naturalizaciones, otro puede permanecer estable o incluso decrecer; las interacciones capturan estas diferencias.