

Análisis econométrico: modelo de gravedad aplicado en Stata

Guía docente, prácticas y soluciones

Jordi Paniagua · Universitat de València

Universidad de los Llanos · Villavicencio
7 de octubre de 2026

Destinatarios. Profesorado que desea iniciarse en investigación aplicada. Se requiere interpretar una regresión y manejar comandos básicos de Stata; no se presupone formación en equilibrio general.

Objetivos observables. Al terminar, la persona participante podrá traducir una pregunta bilateral a una ecuación de gravedad, auditar el panel, estimar PPML con tres conjuntos de efectos fijos, interpretar efectos discretos y ejecutar un contrafactual distinguiendo comercio de bienestar.

Duración. 120 minutos exactos, incluidos 28 minutos de prácticas en parejas, demostraciones, puestas en común y preguntas. Los anexos y las extensiones 04 y 05 quedan fuera de ese tiempo.

Preparación antes de la sesión

1. Disponer de Stata 17 o posterior. La entrega se ha ejecutado con Stata 18.
2. Descargar el archivo de datos desde https://yotoyotov.com/Gravity_Undergrads.html, descomprimirlo y guardar `Gravity_Undergrads_Data.dta` en `datos/`. La descarga oficial permite al autor seguir el uso docente de su material. La base GU no se redistribuye en este paquete.
3. Abrir Stata, cambiar a la carpeta raíz `Curso_Unillanos_2026` y ejecutar `do "stata/00_instalar.do"` con conexión. Comprueba e instala, si faltan, `ftools`, `reghdfe`, `ppmlhdfe` y `ge_gravity`. No actualiza paquetes instalados sin necesidad.
4. Ejecutar `do "stata/00_master.do"`. El proceso reemplaza únicamente salidas del curso y crea el registro `resultados/curso_completo.log`. No modifica los originales.
5. La matriz histórica `datos/gravity_data.dta` procede de los recursos locales de `Hands_on_gravity`. Su versión reducida a 2006 se selecciona en el código, no editando la fuente.

Si falla un equipo. Trabajar con otra persona o con los CSV y el registro de referencia entregados. Las soluciones son reproducibles, pero versiones diferentes de los paquetes pueden producir pequeñas variaciones numéricas. No dedicar la sesión a instalar paquetes por primera vez.

Programa y cronómetro

Diap.	Minutos	Tema
1	0–1	Portada
2	1–3	Un recorrido de dos horas
3	3–5	De la docencia a una pregunta investigable
4	5–7	El punto de partida: tamaño y fricciones
5	7–10	Los supuestos que dan estructura a la gravedad
6	10–13	Demanda bilateral e índice de precios
7	13–15	Producción, gasto y vaciado de mercados
8	15–18	Gravedad estructural
9	18–21	Las dos resistencias multilaterales son endógenas
10	21–23	Una intuición para la resistencia multilateral
11	23–26	El sistema estructural completo: cinco bloques
12	26–28	Bienestar: poder de compra, no exportaciones
13	28–30	De la teoría a los efectos fijos
14	30–32	Sectores: misma lógica, nuevas exigencias
15	32–34	Recomendaciones 2025 · Diseñar los datos
16	34–37	Recomendaciones 2025 · Especificación e inferencia
17	37–39	Recomendaciones 2025 · Heterogeneidad y dinámica
18	39–41	Actualizar una recomendación: conservar los años
19	41–44	Por qué el comercio interno cambia la identificación
20	44–46	USITC Gravity Portal · Elegir la base de comercio
21	46–48	USITC Gravity Portal · Relaciones y multinacionales
22	48–50	Aplicación 1 · IED y actividad de multinacionales
23	50–53	Aplicaciones 2 y 3 · Turismo y migración
24	53–55	Datos y puesta en marcha
25	55–63	Práctica 1 · Auditar los datos
26	63–65	Lo que cambia al mirar la base
27	65–68	Por qué estimar en niveles con PPML
28	68–70	Panel y efectos fijos de par
29	70–72	Qué identifica el coeficiente de un acuerdo
30	72–74	La especificación de trabajo en Stata
31	74–76	Inferencia y tamaño del efecto
32	76–88	Práctica 2 · Estimar y defender una especificación
33	88–91	Resultados de la secuencia de estimación
34	91–93	Interpretar el PPML
35	93–95	Antes de escribir la primera conclusión
36	95–96	De un coeficiente a un contrafactual
37	96–97	Qué implementa <code>ge_gravity</code> en este taller
38	97–98	Una segunda base para el equilibrio general
39	98–100	Definir el shock con el signo correcto
40	100–102	Resolver y leer las salidas
41	102–110	Práctica 3 · Comercio y bienestar de Colombia
42	110–113	Resultados del contrafactual
43	113–115	Sensibilidad: elasticidad y cierre
44	115–117	Convertir el taller en un primer proyecto
45	117–120	Preguntas y recursos para continuar

Práctica 1: datos e identificación del panel

Consigna (8 minutos). Ejecutar `01_datos.do`. Documentar unicidad, años, ceros, diagonal y cruce RTA–EU. Entregar tres decisiones: tratamiento de ceros, interpretación de filas ausentes y categoría de comparación.

Solución. La base tiene 320.920 filas para 100 economías entre 1990 y 2023; 32.835 valores de

comercio iguales a cero (10,23 %) y 288.085 positivos. No hay diagonal interna. El rectángulo completo de 100 economías sería $100 \times 99 \times 34 = 336.600$; las 15.680 filas que faltan no se convierten en ceros. Se verifica la clave exportador–importador–año. Hay 2.881 pares dirigidos con cambios en RTA durante la muestra. Los PIB están observados en esta versión.

La tabla RTA–EU contiene 219.197 filas (0,0), 87.327 filas (1,0) y 14.396 filas (0,1), sin ninguna (1,1). Los coeficientes se comparan con (0,0); el de UE no es adicional al de RTA. Leer siempre la codificación real.

Criterio de evaluación. Se considera resuelta cuando se distingue cero de ausencia, se reconoce la falta de comercio interno y se identifica la categoría de referencia. No basta con copiar un número.

Práctica 2: estimación e interpretación

Consigna (12 minutos). Ejecutar `02_estimacion.do`. Comparar m1–m5, justificar las variables absorbidas, transformar RTA y comparar m6–m7 sobre el mismo soporte positivo. Escribir un resultado y una amenaza de identificación.

	N	$\hat{\beta}_{RTA}$	EE	$\hat{\beta}_{EU}$	EE
MCO ingenuo, 2023	9 564	0,664	0,055	0,900	0,091
MCO con EF, 2023	9 564	0,236	0,059	-0,261	0,109
MCO panel	288 085	0,147	0,046	-0,526	0,082
MCO con tres EF	288 085	0,103	0,026	0,427	0,054
PPML con tres EF	320 920	0,076	0,031	0,273	0,044

El PIB se absorbe con efectos fijos de país, y la distancia con efectos fijos de par. PPML usa niveles y conserva los ceros. Los cinco modelos no estiman exactamente sobre la misma muestra: los dos primeros usan 2023 y el resto todo el panel; además MCO elimina los ceros. La comparación m6–m7 aísla el cambio de método sobre el soporte positivo común. No atribuir toda diferencia MCO–PPML exclusivamente a ceros.

La solución de PPML es $\hat{\beta}_{RTA} = 0,07562$ (EE 0,03063), equivalente a 7,86 %, con intervalo normal transformado aproximadamente [1,57 %, 14,53 %]. Para EU, $\hat{\beta} = 0,27263$ (EE 0,04442), equivalente a 31,34 % frente a la categoría (0,0). La diferencia EU–RTA se contrasta con `lincom EU - RTA`; su cambio porcentual se obtiene con $100[\exp(\hat{\beta}_{EU} - \hat{\beta}_{RTA}) - 1]$.

Respuesta modelo. “En el panel analizado, la especificación PPML con efectos fijos de origen–año, destino–año y par relaciona la categoría RTA con un comercio bilateral aproximadamente 7,86 % mayor frente a la categoría de referencia, condicionado a los controles. Su interpretación causal exige descartar factores bilaterales cambiantes y anticipación ligados a la adopción del acuerdo.”

Los errores estándar se agrupan por par no dirigido; los efectos fijos son por par dirigido. Esta diferencia deliberada permite correlación temporal y entre ambas direcciones, pero no toda dependencia entre acuerdos o países. `nlcom` aplica el método delta; los CSV también incluyen un intervalo obtenido exponenciando los extremos normales del intervalo de beta. No son idénticos en muestras finitas.

Criterio de evaluación. Distinguir controles de inferencia, explicar por qué no aparecen distancia y PIB, calcular el porcentaje exacto y reconocer al menos una amenaza causal.

Práctica 3: un contrafactual de Colombia

Consigna (8 minutos). Ejecutar 03_equilibrio_general.do. Comparar Colombia con dos socios. Separar efecto bilateral directo, exportaciones totales del modelo, salario real y bienestar. Cambiar elasticidad y cierre.

Calibración. Corte de 2006, 69 economías, 4.761 flujos y 69 ventas internas positivas; 138 ceros internacionales. La base de manufacturas del material de Yotov se utiliza como un universo cerrado y un sector representativo. No representa la economía colombiana completa ni una previsión para 2026. No se estima el modelo de hogares de Marriage and Trade.

Shock. Reducción hipotética del 10 % en costes iceberg de Colombia con todos los socios, en ambos sentidos, sin modificar costes internos. Para $\theta = 4$, $b = -4 \ln(0,9) = 0,42144$. La entrada de `ge_gravity` es b , no la caída de costes ni el porcentaje de variación del comercio. El efecto bilateral mecánico es $100[\exp(b) - 1] = 52,42\%$ en cada relación tratada.

	Bienestar	Salario real	Exportaciones
Colombia	3,546	3,288	59,73
Ecuador	0,817	0,599	9,58
Brasil	0,013	0,011	0,28
Estados Unidos	0,007	0,007	0,15
España	0,005	0,005	0,05
China	0,002	0,002	0,02

Las cifras son porcentajes. Colombia obtiene en esta calibración un cambio de bienestar del modelo cercano a 3,546 %, un salario real 3,288 % mayor y unas exportaciones nominales normalizadas 59,73 % mayores. El efecto parcial mecánico sobre sus exportaciones es 52,42 %; en otros países solo se aplica directamente a sus ventas hacia Colombia. Las ventas internas se excluyen al sumar exportaciones.

Cierre y bienestar. Por defecto el déficit $D_i = E_i - Y_i$ es aditivo y constante en el numerario. Así, $E'_i = Y_i \hat{w}_i + D_i$ y $\hat{W}_i = (E'_i/E_i)/\hat{P}_i$, mientras el salario real es $\hat{w}_i/\hat{P}_i$. El multiplicativo fija la proporción gasto/producción y hace coincidir esos dos cocientes. El numerario mantiene la producción nominal mundial; no impone bienestar constante.

θ	Cierre aditivo	Cierre multiplicativo
3	3,422	3,094
4	3,546	3,255
6	3,843	3,631

La tabla muestra el cambio de bienestar de Colombia bajo la misma caída de costes. Se recalcula b para cada θ . Es sensibilidad estructural, no incertidumbre estadística. Cuando la entrada sea una beta estimada, mantener esa beta al variar theta constituye otro ejercicio: cambia el coste implícito asociado.

Verificación. El código exige una matriz completa, países coincidentes, claves únicas, no negativos y diagonal positiva. El shock nulo debe devolver $W = 1$ y el comercio original. Se comprueban vaciado de mercado, cierre de gasto, identidad de salario real, identidad de bienestar y normalización mundial. Los ceros iniciales permanecen a cero: no se modela la apertura de nuevas rutas. Los valores W, RW, NW y P pertenecen al exportador y se repiten por destino; nunca se suman como si fueran flujos.

Criterio de evaluación. Presentar tres cifras correctamente convertidas, reconocer al menos dos supuestos y evitar describir la simulación como un efecto causal observado o como bienestar total de Colombia.

Extensión opcional: una beta como entrada del equilibrio

04_puente_opcional.do ilustra $b_{ij} = \hat{\beta} \Delta RTA_{ij}$ en una exposición hipotética COL–USA. Usa el coeficiente GU guardado y la matriz histórica de manufacturas. Transportar efectos entre universos, sectores y años requiere justificación: el ejemplo enseña la mecánica, no evalúa un tratado histórico. No multiplicar de nuevo beta por $-\theta$.

En un proyecto completo, estimar y calibrar en marcos compatibles; considerar incertidumbre del shock estimado, de la elasticidad y de los datos. Un bootstrap exige un esquema de remuestreo adecuado para la estimación y una matriz completa para cada resolución de GE; no remuestrear sin más filas de la matriz y enviarlas al solucionador.

Fundamentos y decisiones para investigar

Lectura preparatoria. Yotov, *Full Grav Lecture I Yotov IEI Light*, pp. PDF 38–57 (CES, demanda, vaciado y resistencias), 64–70 (sectores), 87 (sistema) y 114 (bienestar); *Full Grav Lecture II Yotov IEI Light*, pp. PDF 8–9 (teoría y empiria), 25–35 (recomendaciones e identificación). El modelo completo se explica durante la sesión; la iteración numérica queda en el anexo.

Convenciones del modelo. i es origen y j destino; $a_i > 0$ pondera la variedad; Q_i es dotación/oferta; p_i precio de fábrica; $\tau_{ij} \geq 1$ coste iceberg; $\sigma > 1$ elasticidad de sustitución y $\theta = \sigma - 1$. Todas las sumas incorporan $i = j$. El CES y su restricción son

$$U_j = \left[\sum_i a_i^{1/\sigma} c_{ij}^{(\sigma-1)/\sigma} \right]^{\sigma/(\sigma-1)}, \quad \sum_i p_i \tau_{ij} c_{ij} = E_j.$$

La demanda $X_{ij} = a_i (p_i \tau_{ij} / P_j)^{-\theta} E_j$ y el índice $P_j = [\sum_i a_i (p_i \tau_{ij})^{-\theta}]^{-1/\theta}$ implican $\sum_i X_{ij} = E_j$. Vaciado requiere $Y_i = p_i Q_i = \sum_j X_{ij}$. Tras sustituir precios se obtiene el sistema:

$$\begin{aligned} X_{ij} &= \frac{Y_i E_j}{Y} (\tau_{ij} / (\Pi_i P_j))^{-\theta}, \\ \Pi_i^{-\theta} &= \sum_j (\tau_{ij} / P_j)^{-\theta} E_j / Y, \\ P_j^{-\theta} &= \sum_i (\tau_{ij} / \Pi_i)^{-\theta} Y_i / Y, \\ p_i &= a_i^{1/\theta} (Y_i / Y)^{-1/\theta} / \Pi_i, \\ Y_i &= p_i Q_i, \quad E_i = Y_i + D_i, \quad Y = \sum_i Y_i = \sum_i E_i. \end{aligned}$$

Dados costes, ponderaciones, dotaciones, elasticidad y transferencias con $\sum_i D_i = 0$, un numerario completa la determinación nominal. Con $D_i = 0$ hay equilibrio comercial individual. El cierre proporcional usa $E_i = \phi_i Y_i$ con ϕ_i fijo y consistencia mundial. No mantener a la vez ingresos y precios fijos al interpretar el contrafactual como equilibrio completo.

Bienestar y sector. $W_i = E_i / P_i$, luego $\widehat{W}_i = \widehat{E}_i / \widehat{P}_i$. Con tecnología fija de un factor, $\widehat{p}_i = \widehat{w}_i$, pero gasto real y salario real solo coinciden en variación si E_i / Y_i es constante. La versión sectorial sustituye cada flujo, tamaño, resistencia y elasticidad por sus equivalentes sectoriales; para un GE con varios sectores se debe completar el modelo de gasto y producción, y, si procede, insumos intermedios.

Comprobación de comprensión. Pedir explicar (a) por qué un shock entre terceros afecta a Colombia, (b) qué absorbe cada efecto fijo, (c) por qué producción sectorial no es PIB y (d) cuándo el cambio de salario real representa el cambio de bienestar. La respuesta debe nombrar alternativas comerciales, composición de los efectos fijos, definición de producción y cierre del gasto.

La actualización bibliográfica de 2025

La referencia solicitada es Mario Larch, Serge Shikher y Yoto V. Yotov (2025), *Estimating Gravity Equations: Theory Implications, Econometric Developments, and Practical Recommendations*, *Review of International Economics*, 33(5), 1066–1092, doi:10.1111/roie.12789. Para consultar el texto íntegro se utilizó la [versión pública de autores](#), 3 de enero de 2025. Las diapositivas sintetizan sus 15 recomendaciones y distinguen esa actualización del material de 2019.

La modificación práctica más visible es conservar años consecutivos como punto de partida y abordar explícitamente el ajuste dinámico. Los códigos GU ya cumplen ese criterio. Los modelos

MCO se conservan como comparación docente; la especificación de trabajo es PPML en niveles con efectos fijos de origen-año, destino-año y par dirigido. Un estudio de adopción escalonada necesita justificar sus comparaciones; no se proporciona un estudio de eventos genérico que aparente resolver cualquier problema de identificación.

Extensión 05: frontera y ventas internas

Ejecutar do "stata/05_frontera_opcional.do" desde la raíz. Usa todos los años de la base local de manufacturas y estima

$$\mathbb{E}[X_{ij,t} \mid \cdot] = \exp\{\alpha_{i,t} + \delta_{j,t} + \mu_{ij} + \gamma_t 1[i \neq j]\}, \quad \gamma_{1986} = 0.$$

El archivo exporta **resultados/frontera_anual.csv**, las estimaciones y un registro propio. Es una descripción condicional de la evolución de ventas internacionales respecto a internas; recoge múltiples cambios comunes y no identifica un tratado particular. El nivel permanente de frontera es colineal con el par, y una interacción para cada año necesita una referencia. Sin diagonal, frontera-año se confunde con efectos anuales absorbidos.

Solución de referencia en Stata 18. Para 2006, $\hat{\gamma} = 0,7960$ (EE 0,04758), equivalente a 121,67 % respecto a la brecha de referencia de 1986; IC normal transformado [101,93 %, 143,34 %]. PPML utiliza 99.708 observaciones y excluye 273 por separación o grupos singleton según el diagnóstico del comando. No confundir este cambio condicional con el crecimiento observado del comercio mundial. El registro conserva también el aviso sobre valores muy pequeños de la variable dependiente estandarizada; se documenta la convergencia en 18 iteraciones.

Ejercicio posterior. Interpretar $100[\exp(\hat{\gamma}_t) - 1]$ en el último año frente a 1986, leer su intervalo y explicar la colinealidad que aparece si se eliminan ventas internas. Para una política unilateral $z_{i,t}$, la interacción $z_{i,t} 1[i \neq j]$ puede identificar una diferencia externa-interna si hay variación independiente suficiente; no recupera automáticamente el efecto total nacional ni garantiza causalidad.

Fuentes para construir una base propia

USITC Gravity Portal, consultado el 1 de octubre de 2026. La familia ITPD, *Release 2025*, fue actualizada en agosto de 2026. En **ITPD-E**, las ventanas publicadas son 1986–2023 para agricultura, 1988–2023 para manufacturas, minería y energía, y 2000–2023 para servicios; no implican cobertura completa. Elegir fuentes observadas y simuladas según el objetivo, documentando imputaciones. Los enlaces a ITPD-S/U, DGD, MREID, DICL y LABPTA están en el portal y en el Beamer.

Protocolo de unión. Definir primero unidad (país–país–sector–año), revisar el diccionario y concordancias, comprobar **isid** en cada archivo, estudiar la multiplicidad de acuerdos antes de agregar y auditar los no emparejados. Cambios territoriales o sectoriales exigen una decisión documentada. Nunca convertir todos los no emparejados en ceros ni recurrir a **merge m:m** para superar duplicados. Para ITPD-E Release 2025, citar Larch, Shikher y Yotov (2026), USITC WP 2026–08–A, y Borchert, Larch, Shikher y Yotov (2021), *International Economics*, 166, 140–166, siguiendo la documentación del portal.

Aplicaciones fuera del comercio de bienes

IED. La gravedad puede organizar una investigación sobre localización de filiales o proyectos, ingresos, empleo e inversión. El recurso **MREID** se documenta en Ahmad, Bergstrand, Paniagua y Wickramarachi (2023), USITC WP 2023–11–B. Elegir la variable según la hipótesis y distinguir

inversión neta, flujos brutos, stocks y actividad de filiales. Los valores negativos de IED neta no se eliminan ni se desplazan con una constante para aplicar PPML. Un modelo de bienestar multinacional debe reflejar propiedad, beneficios y localización de producción.

Turismo. Pregunta propuesta: ¿la conectividad bilateral modifica las pernoctaciones? Medir residencia/nacionalidad del origen, alojamiento/frontera y visitas/pernoctaciones de modo consistente. UN Tourism publica llegadas y pernoctaciones por origen en el [Yearbook of Tourism Statistics](#), edición 2023; el archivo enlazado es una copia de la publicación oficial alojada por la oficina estadística de Ucrania. Comprobar acceso y cobertura en la edición utilizada. Santana-Gallego y Paniagua (2022), [Tourism Economics](#) 28(2), 394–417, estudian cómo las redes migrantes se relacionan con turismo mediante información, costes y visitas a familiares y amigos; [copia del artículo en la web del autor](#).

Migración. UN DESA [International Migrant Stock 2024](#) permite describir stocks bilaterales; los datos incluyen estimaciones y no equivalen a flujos de decisiones anuales. Para estudiar nuevas entradas, consultar [International Migration Flows](#) y registros nacionales, armonizando definiciones. Un modelo de elección debe incluir destinos alternativos y justificar el papel de quedarse; una diferencia de stocks no mide por sí sola inmigración.

Actividad para el profesorado. En parejas, elegir comercio, IED, turismo o migración y escribir: resultado bilateral, mecanismo, fuente y fecha, comparación identificadora y limitación principal. En una segunda página, indicar si tiene sentido un contrafactual estructural y qué ecuaciones adicionales exigiría. También pueden plantearse movilidad regional, colaboraciones científicas o flujos financieros, con unidades y mecanismos propios.

Guion del docente por diapositiva

1. Portada

Minutos 0–1 (1 minutos). Presentar el taller como una primera experiencia completa de investigación: pregunta, auditoría, estimación e interpretación. Se asume familiaridad con regresión y manejo básico de Stata. El objetivo no es dominar en dos horas toda la literatura. Fondo y diseño adaptados de la presentación del proyecto Marriage and Trade para este congreso; fecha de esta sesión: 7 de octubre de 2026.

2. Un recorrido de dos horas

Minutos 1–3 (2 minutos). La teoría ocupa ahora el primer bloque completo. Hay 28 minutos de prácticas en parejas (8, 12 y 8), además de demostraciones y discusión. Instalaciones y descarga de GU deben estar resueltas antes de la sesión. Los anexos permiten profundizar sin alargar las dos horas; las extensiones 04 y 05 son trabajo posterior.

3. De la docencia a una pregunta investigable

Minutos 3–5 (2 minutos). Pedir una hipótesis y un posible factor de confusión. Si los países firman acuerdos precisamente cuando esperan comerciar más, una correlación no identifica el efecto causal. Separar desde el inicio estimación y simulación: la segunda utiliza una estructura económica y un shock supuesto. El escenario de Colombia es nuevo y docente, no una previsión ni una evaluación de un tratado concreto.

4. El punto de partida: tamaño y fricciones

Minutos 5–7 (2 minutos). La analogía física ayuda a recordar signos, pero no determina una elasticidad igual a dos. Distinguir producción o renta del origen y gasto del destino. En la regresión ingenua se utiliza PIB como aproximación; en el modelo estructural de mercancías el objeto relevante no es automáticamente el PIB. No presentar esta primera ecuación como el sistema cerrado de equilibrio.

5. Los supuestos que dan estructura a la gravedad

Minutos 7–10 (3 minutos). La estructura se deriva aquí con Armington y CES, no con una analogía física. Q mide producción bruta y c cantidad que llega al comprador: el transporte iceberg consume parte de Q . El gasto E puede diferir de Y cuando se introduce un cierre de transferencias. En el caso básico equilibrado $E=Y$. Los parámetros α recogen ponderaciones de demanda por origen; no son los coeficientes de una regresión. La elasticidad comercial puede tener otro fundamento en modelos ricardianos o con empresas heterogéneas; no identificar σ con θ en cualquier aplicación.

6. Demanda bilateral e índice de precios

Minutos 10–13 (3 minutos). Presentar el resultado de la maximización CES sin resolver las condiciones de primer orden. X es valor, no cantidad física. El exponente de la cuota de gasto es $1-\sigma$, es decir menos θ ; la elasticidad de cantidades respecto al precio relativo es menos σ . La entrada de un proveedor más barato afecta a todas las cuotas mediante P . No igualar

índice de precios con la distancia media. En esta normalización las ponderaciones están dentro del índice CES.

7. Producción, gasto y vaciado de mercados

Minutos 13–15 (2 minutos). El origen vende al exterior y a sí mismo: omitir la diagonal rompe estas identidades. En una economía mundial cerrada el gasto agregado coincide con la producción agregada aunque cada país tenga déficit o superávit. Con el cierre proporcional la consistencia mundial debe cumplirse en el equilibrio; no se pueden escoger arbitrariamente ingresos nuevos. La normalización del nivel nominal es necesaria y no determina el bienestar real. Producción bruta sectorial no equivale al PIB, especialmente con insumos intermedios.

8. Gravedad estructural

Minutos 15–18 (3 minutos). Leer la ecuación en dos pasos: tamaño y costes relativos. Mantener la convención θ positiva; en Armington CES $\theta = \sigma - 1$. No estimaremos directamente las resistencias multilaterales en la práctica de panel: sus componentes se absorberán con efectos fijos adecuados. En el contrafactual sí permitiremos que cambien endógenamente.

9. Las dos resistencias multilaterales son endógenas

Minutos 18–21 (3 minutos). Derivar verbalmente la primera ecuación sumando la demanda sobre destinos y usando vaciado. La segunda aparece al sustituir los precios de fábrica en el índice CES. Por tanto no son dos regresores observados que se puedan sustituir sin más por indicadores de lejanía. El vínculo entre precio de fábrica y resistencia de salida se obtiene de la condición de vaciado. Una caída de costes hacia un destino modifica acceso a mercados y precios de equilibrio. La escala de las resistencias exige una normalización; importan las magnitudes relativas y sus combinaciones.

10. Una intuición para la resistencia multilateral

Minutos 21–23 (2 minutos). Dar unos segundos para responder: no, porque faltan las alternativas. El ejemplo es ilustrativo y mantiene las demás condiciones constantes. Explicar por qué una mejora entre terceros puede afectar a Colombia sin cambiar ningún coste colombiano. Esta intuición reaparecerá al interpretar el equilibrio general.

11. El sistema estructural completo: cinco bloques

Minutos 23–26 (3 minutos). Esta es la diapositiva de referencia del curso: todas las ecuaciones principales permanecen en el bloque central. Se muestra el cierre aditivo para enlazar con `ge_gravity`; poner $D=0$ recupera el modelo equilibrado. Hay una redundancia nominal, de modo que debe fijarse el numerario. Conservar Y y E al cambiar τ solo describe un ejercicio condicionado: para el equilibrio completo también deben ajustarse precios e ingresos. El algoritmo iterativo se consulta en el anexo, sin dedicar tiempo de aula a su implementación.

12. Bienestar: poder de compra, no exportaciones

Minutos 26–28 (2 minutos). El índice CES es el coste de una unidad de utilidad: por eso gasto dividido por precios da bienestar real. No confundir la Q de producción con bienestar. A

dotaciones constantes, ingreso por trabajador cambia con el salario. Los déficits implican transferencias y afectan al poder de compra además de los salarios. La fórmula de bienestar basada solo en la cuota doméstica requiere un cierre apropiado; se deja en un anexo para evitar aplicar esa identidad sin sus supuestos.

13. De la teoría a los efectos fijos

Minutos 28–30 (2 minutos). La primera igualdad es una descomposición del logaritmo de la media teórica, no una instrucción para tomar logaritmos de cada observación. Se estima la media exponencial en niveles con PPML. Los efectos fijos no son resistencias puras: mezclan tamaño, resistencia y otros componentes propios del país. La parametrización de los costes determina qué interpreta beta. No introducir por separado PIB o una política uniforme de país-año y esperar identificarla con los mismos efectos fijos.

14. Sectores: misma lógica, nuevas exigencias

Minutos 30–32 (2 minutos). El sector tiene su propio denominador mundial y sus resistencias; no sustituir producción de manufacturas por PIB agregado. Al desagregar cambian los ceros, la composición y posiblemente theta. La fórmula sectorial no cierra por sí sola un equilibrio general multisectorial: hay que indicar oferta, gasto y movilidad de factores. Los modelos con insumos intermedios añaden una estructura de costes y demanda intermedia; no basta con estimar cada sector y sumar cambios porcentuales de bienestar.

15. Recomendaciones 2025 · Diseñar los datos

Minutos 32–34 (2 minutos). Los primeros cinco criterios convierten la teoría en decisiones de muestra. La cobertura amplia permite representar alternativas; una pregunta sobre Colombia no obliga a eliminar todos los demás exportadores. En GU no inventamos la diagonal. Las recomendaciones son una síntesis propia; las decisiones específicas de esta práctica se explican con la auditoría y con Lecture II. La referencia publicada es de 2025; el texto íntegro consultado es la versión pública de los autores de enero de ese año.

16. Recomendaciones 2025 · Especificación e inferencia

Minutos 34–37 (3 minutos). Distinguir modelo de media de modelo de errores. Los efectos fijos de par no reemplazan efectos fijos país-año, ni la agrupación reemplaza efectos fijos. Justificar aranceles con $\ln(1+\text{arancel})$ cuando el mecanismo sea ad valorem; no tratar la distancia como el coste total. La inferencia por par permite dependencia dentro de la relación; según tratamiento y shocks comunes puede necesitarse agrupación en otras dimensiones. La selección de esas dimensiones exige más que probar qué opción produce significación.

17. Recomendaciones 2025 · Heterogeneidad y dinámica

Minutos 37–39 (2 minutos). El propósito de este bloque es ordenar un primer proyecto, no ejecutar en dos horas un diseño completo de adopción escalonada. Un coeficiente constante sobre una dummy RTA impone una respuesta común e inmediata. Las redes familiares, la conectividad y la integración pueden actuar con rezagos distintos. Las pruebas de anticipación y los gráficos de dinámica requieren hipótesis y comparaciones explícitas; la ausencia de significación previa no demuestra exogeneidad.

18. Actualizar una recomendación: conservar los años

Minutos 39–41 (2 minutos). Hacer explícita esta actualización, porque las diapositivas de 2019 y el artículo solicitado no dan la misma recomendación sobre frecuencia. No eliminar automáticamente dos de cada tres observaciones anuales para la especificación principal. Los do-files GU ya mantienen los años consecutivos; no requieren una recodificación para cumplir esta orientación. Un análisis de ajuste posterior puede incorporar una dinámica justificada sin convertir años no observados en datos fabricados.

19. Por qué el comercio interno cambia la identificación

Minutos 41–44 (3 minutos). En una muestra que solo contiene i distinto de j , multiplicar z por B no crea nueva información: B siempre vale uno. La diagonal permite una comparación externa-interna, sujeta a variación, rango y ausencia de cambios omitidos relevantes. El nivel permanente de B queda absorbido por el par. El archivo 05 estima cambios anuales comunes en esa brecha con la base histórica; es descriptivo, no una evaluación causal de una política z específica. El coeficiente η mide un efecto diferencial externo-interno, no el efecto total nacional de z .

20. USITC Gravity Portal · Elegir la base de comercio

Minutos 44–46 (2 minutos). El nombre de la edición es Release 2025 aunque la actualización del portal se anunció en agosto de 2026. En ITPD-E, agricultura empieza en 1986, manufacturas/minería/energía en 1988 y servicios en 2000. La elección E o S depende de la finalidad del análisis, y no simplemente del archivo que tenga menos ausentes. Para emparejar datos, revisar concordancias de países y sectores además del año. El tamaño anunciado no implica que toda industria tenga datos para cada país y año.

21. USITC Gravity Portal · Relaciones y multinacionales

Minutos 46–48 (2 minutos). Relacionar cada recurso con una hipótesis: redes lingüísticas, profundidad institucional, empleo e inversión internacional. No prometer que DGD cubra 2023 porque ITPD sí lo haga. MREID mide actividad multinacional y no se debe sustituir automáticamente por una serie de flujos financieros de IED. En LABPTA, varios acuerdos pueden cubrir el mismo par: hay que definir la exposición antes de crear una fila por par-año. No ejecutar merge m:m para ocultar duplicados.

22. Aplicación 1 · IED y actividad de multinacionales

Minutos 48–50 (2 minutos). La unidad observada y el mecanismo deben establecerse antes del estimador. Ingresos y empleo de filiales son actividad económica; IED financiera neta incluye desinversión y puede ser negativa. No sumar una constante a la IED neta para poder ejecutar PPML. Una pregunta sobre localización de proyectos puede usar recuentos, pero esa elección cambia el objeto de investigación. El origen en una base de filiales puede ser el propietario, no el exportador de bienes; documentar la dirección. `ge_gravity` no convierte automáticamente estos resultados en bienestar de un modelo multinacional.

23. Aplicaciones 2 y 3 · Turismo y migración

Minutos 50–53 (3 minutos). El año de la revista del artículo de turismo es 2022, aunque apareció en línea en 2020. Su aporte es distinguir canales de la asociación entre migración y

turismo; no usarlo para prometer que cualquier regresión de redes es causal. Verificar si el origen turístico significa residencia o nacionalidad y si se cuentan hoteles o fronteras. UN DESA 2024 presenta estimaciones de stock para 233 países/áreas entre 1990 y 2024: no es un registro de flujos anuales. El cambio del stock también contiene muertes, retornos y revisiones estadísticas. El anexo propone una ecuación de elección de destino; el GE comercial no se transfiere sin un modelo específico.

24. Datos y puesta en marcha

Minutos 53–55 (2 minutos). Abrir la carpeta raíz del curso en Stata. Recordar que ejecutar un do-file no cambia automáticamente el directorio de trabajo. El config busca primero la descarga del estudiante y, en este ordenador, después los originales de la carpeta hermana. Los identificadores Exporter e Importer son numéricos; Exporter_Name e Importer_Name contienen los nombres. El comercio está en dólares corrientes. No distribuir la base GU en el ZIP: indicar la descarga oficial.

25. Práctica 1 · Auditar los datos

Minutos 55–63 (8 minutos). Minutos 0–2: ejecutar 01_datos.do y comprobar clave. Minutos 2–5: ceros, diagonal y años. Minutos 5–8: leer RTA–EU y redactar las tres decisiones. Hay 32.835 ceros (10,23 por ciento), no hay flujos internos y faltan 15.680 filas respecto al rectángulo 100 por 99 por 34; no son ceros observados. RTA y EU son mutuamente excluyentes. No hace falta teclear todo el código: la entrega es una interpretación de la auditoría.

26. Lo que cambia al mirar la base

Minutos 63–65 (2 minutos). Corregir conjuntamente. La tabla cruzada tiene 219.197 filas en (0,0), 87.327 en (1,0), 14.396 en (0,1) y cero en (1,1). Para comparar UE con otros acuerdos se utiliza β_{EU} menos β_{RTA} . En una base con otra codificación cambiaría la interpretación. Hay 2.881 pares dirigidos con alguna variación en RTA. No confundir falta de cobertura histórica con ausencia de comercio.

27. Por qué estimar en niveles con PPML

Minutos 65–68 (3 minutos). Usar Poisson como cuasi-verosimilitud: la condición clave es la especificación de la media condicional, acompañada de inferencia robusta adecuada. No decir que PPML elimina todos los sesgos. Un cero no necesita una corrección $\log(1+X)$, que además depende de las unidades. Los ceros pueden ser excluidos si hay separación; hay que mirar el diagnóstico del comando.

28. Panel y efectos fijos de par

Minutos 68–70 (2 minutos). La etapa intermedia m3 del código incorpora origen-año y destino-año manteniendo distancia. m4 añade los pares dirigidos. Explicar por qué la distancia desaparece sin ser irrelevante económicamente: su efecto es indistinguible del efecto fijo de par. No introducir un identificador numérico como regresor continuo. Mostrar egen group para construir identificadores en 01_datos.do.

29. Qué identifica el coeficiente de un acuerdo

Minutos 70–72 (2 minutos). Pedir un factor bilateral que cambie con un acuerdo: infraestructura, conflicto o reformas conjuntas. La especificación principal conserva los años consecutivos. La respuesta gradual y la adopción escalonada requieren un análisis específico con comparaciones válidas y heterogeneidad; los efectos fijos no solucionan factores bilaterales cambiantes. Remitir a las recomendaciones 13–15.

30. La especificación de trabajo en Stata

Minutos 72–74 (2 minutos). El bloque ya existe en el do-file, por lo que no es necesario teclearlo en clase. El escalado es una elección numérica; no convertir las pendientes a millones. Una predicción de la media no es una observación ni una predicción fuera de muestra. Una caída en número de filas exige revisar mensajes sobre separación y singletons. Aquí se conservan las 320.920 observaciones.

31. Inferencia y tamaño del efecto

Minutos 74–76 (2 minutos). Explicar la diferencia entre absorber heterogeneidad y corregir la matriz de varianzas. El clúster se crea con mínimo y máximo del par. La elección del nivel de agrupación debe seguir el diseño de investigación; aquí damos un criterio docente explícito. nlcom utiliza el método delta. Para un intervalo asimétrico en porcentaje se pueden exponenciar los extremos del intervalo de beta. No transformar el error estándar con $\exp(\cdot)$.

32. Práctica 2 · Estimar y defender una especificación

Minutos 76–88 (12 minutos). Minutos 0–2: ejecutar 02_estimacion.do. Minutos 2–5: comparar muestra y controles de m1–m5. Minutos 5–8: transformar RTA e interpretar la referencia de EU. Minutos 8–12: revisar m6–m7 con soporte positivo común y escribir resultado y limitación. La explicación detallada de MCO transversal queda en anexos; las cinco estimaciones siguen ejecutándose. El cambio m4–m5 combina muestra y método, y no se atribuye enteramente a los ceros.

33. Resultados de la secuencia de estimación

Minutos 88–91 (3 minutos). Leer primero N y luego los coeficientes. El efecto EU cambia de signo entre especificaciones; eso ilustra la relevancia del control de heterogeneidad, no prueba por sí mismo cuál modelo es causal. Señalar que no se copian errores estándar del capítulo: aquí se usa agrupación explícita por par no dirigido. Los cinco coeficientes proceden del código entregado. No convertir la significación en la única decisión científica.

34. Interpretar el PPML

Minutos 91–93 (2 minutos). Hacer la cuenta $\exp(0,0756)-1$. Es aproximadamente 7,86 por ciento, no 0,0756 por ciento ni necesariamente 7,56 por ciento. Para UE se obtiene 31,34 por ciento frente a la categoría sin ninguno de los dos indicadores. La comparación EU frente a RTA requiere la diferencia de coeficientes y su covarianza para el error estándar. Recordar que el intervalo mostrado proviene de exponenciar beta más y menos 1,96 errores estándar, distinto del intervalo simétrico del método delta.

35. Antes de escribir la primera conclusión

Minutos 93–95 (2 minutos). Pedir a dos parejas que lean su conclusión. Corregir frases como el acuerdo causa un aumento de bienestar del ocho por ciento. El ajuste dentro de muestra no valida causalidad. Para investigación con panel corto y numerosos efectos fijos, revisar la literatura de sesgo de parámetros incidentales e inferencia, sin convertirlo en una receta automática. Mantener esta discusión en un nivel de primer proyecto.

36. De un coeficiente a un contrafactual

Minutos 95–96 (1 minutos). Pedir una predicción: si Colombia compra más barato, puede sustituir producción interna; si vende más, sus ingresos y costes de producción pueden variar. Los socios también reaccionan. No prometer que el efecto GE siempre será menor que el parcial. En esta simulación concreta las exportaciones colombianas crecen más en GE que bajo el cálculo bilateral mecánico.

37. Qué implementa `ge_gravity` en este taller

Minutos 96–97 (1 minutos). Remitir al sistema de cinco bloques ya explicado. En este punto no repetir ni derivar el algoritmo. El comando resuelve precios y flujos compatibles con los supuestos. Mantener una distinción clara entre una reducción de costes iceberg y una reforma arancelaria con recaudación y redistribución. La sensibilidad al cierre está ya programada.

38. Una segunda base para el equilibrio general

Minutos 97–98 (1 minutos). Explicar por qué no añadimos una diagonal inventada a GU. Ventas internas son producción sectorial bruta menos exportaciones comparables, no PIB menos exportaciones. La base histórica de manufacturas se trata como un universo cerrado y un sector representativo para practicar el algoritmo. Esa reducción es una limitación sustantiva: para una evaluación aplicada habría que actualizar datos, completar el resto del mundo y justificar sectores, intermediarios y cierre. Las unidades de la etiqueta original no se reinterpretan; solo mostramos variaciones porcentuales.

39. Definir el shock con el signo correcto

Minutos 98–100 (2 minutos). Hacer una comprobación de signos: bajar costes aumenta $\exp(b)$. Enviar $\text{shock}=0$ para pares no tratados y para la diagonal. Es una caída proporcional de costes iceberg de diez por ciento, no diez puntos porcentuales de un arancel. Si se parte de un coeficiente estimado, la entrada será beta por el cambio del indicador: no multiplicar otra vez por menos theta. El anexo y 04_puente_opcional.do muestran ese caso sin confundirlo con el escenario principal.

40. Resolver y leer las salidas

Minutos 100–102 (2 minutos). No sumar W a través de relaciones: colapsar por exportador tomando una media de valores repetidos, o seleccionar la diagonal. Para exportaciones sí sumar X_{cf} pero excluir ventas internas. `gen_X` guarda niveles y `gen_w` guarda ratios. El script usa estas convenciones y verifica $RW=NW/P$. No confundir el precio del exportador guardado en la fila con el precio del importador de esa relación.

41. Práctica 3 · Comercio y bienestar de Colombia

Minutos 102–110 (8 minutos). Minutos 0–2: ejecutar 03_equilibrio_general.do y comprobar ausencia de errores. Minutos 2–4: leer exportaciones y bienestar de Colombia. Minutos 4–6: comparar dos socios y W frente a RW. Minutos 6–8: revisar tabla de sensibilidad y nombrar dos supuestos. Los assert ya comprueban la matriz y el shock nulo; el detalle está en el anexo. No sumar cocientes de bienestar por pareja.

42. Resultados del contrafactual

Minutos 110–113 (3 minutos). Separar los tres conceptos de la tabla. Colombia gana aproximadamente 3,546 por ciento de bienestar en el modelo, 3,288 de salario real y 59,73 de exportaciones nominales normalizadas. Ecuador muestra una respuesta mayor que muchos socios. No atribuir estas diferencias a una sola causa sin análisis adicional: dependen de participaciones comerciales y desequilibrios. El efecto general de exportaciones puede superar al parcial; no existe una regla de que deba ser menor.

43. Sensibilidad: elasticidad y cierre

Minutos 113–115 (2 minutos). El mismo shock en costes no es el mismo número b para distintas elasticidades. Con una estimación fija β ocurre al revés: β es la entrada que se conserva y el coste implícito cambia. Diferenciar sensibilidad estructural de incertidumbre estadística. Si la entrada procede de una estimación, una aplicación completa podría propagar su incertidumbre con bootstrap de un diseño de muestreo adecuado, manteniendo completa la matriz del equilibrio.

44. Convertir el taller en un primer proyecto

Minutos 115–117 (2 minutos). Invitar a cada profesor a anotar una pregunta de su campo. En migración o redes, ser especialmente cuidadoso con causalidad inversa y con el mecanismo estructural: una preferencia no es automáticamente un coste iceberg. El equilibrio general del proyecto Marriage and Trade tiene una estructura propia; no se está replicando con este taller de un sector. Un primer avance razonable es un memo de dos páginas con pregunta, datos, ecuación y amenaza de identificación.

45. Preguntas y recursos para continuar

Minutos 117–120 (3 minutos). Solicitar la principal duda y recuperar las dos preguntas iniciales. Comprobar verbalmente que una persona distingue comercio bilateral y bienestar. Remitir a los anexos para instalar comandos, recordar el álgebra del equilibrio y localizar fuentes. Los anexos no forman parte de los 120 minutos.

Referencias y procedencia

La base y secuencia de estimación son de Yotov (2025), *Gravity for Undergrads*; se añade agrupación explícita por par no dirigido, auditoría de datos, diagnóstico y comparación sobre soporte común. Consultar https://yotoyotov.com/Gravity_Undergrads.html. Se ha consultado la traducción española local, en cuya elaboración participan Federico Carril-Caccia y Jordi Paniagua.

El manual solicitado es OMC/UNCTAD (2012), *A Practical Guide to Trade Policy Analysis*, capítulo 3, pp. 103–134:

https://www.wto.org/english/res_e/publications_e/wto_unctad12_e.pdf. El bloque GE se complementa con Yotov, Piermartini, Monteiro y Larch (2016), *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*, capítulos 1–2:

https://www.wto.org/english/res_e/booksp_e/advancedwtounctad2016_e.pdf.

Se han revisado las diapositivas locales de Yotov (IEI/UJI, 28–30 octubre de 2019): Lecture I (fundamentos); Lecture II, pp. PDF 9, 16, 25, 31 y 35 (retos y recomendaciones); Lecture III, pp. PDF 59–61 (aplicaciones y base de manufacturas de 69 economías, 1986–2006). El conjunto `gravity_data.dta` se copia de Hands_on_gravity sin modificar. Para aplicaciones de política, verificar documentación, cobertura y unidades de la fuente primaria.

Software y método: Correia, Guimarães y Zylkin (2020), *Stata Journal*, 20(1), 95–115; Santos Silva y Tenreyro (2006), *Review of Economics and Statistics*, 88(4), 641–658; Anderson y van Wincoop (2003), *American Economic Review*, 93(1), 170–192; Baier, Yotov y Zylkin (2019), *Journal of International Economics*, 116, 206–226. Ayuda de `ge_gravity`:

https://www.tomzylkin.com/uploads/4/1/0/4/41048809/help_file.pdf. Ayuda de PPML:

<https://scorreia.com/help/ppmlhdfc.html>.

El fondo de portada se reutiliza del archivo

Colombia/presentacion/figuras/portada_congreso.jpg del proyecto marriages_trade; su LEEME documenta la extracción de la plantilla oficial del congreso. Se mantienen paleta, tipografía y márgenes de esa presentación. Se actualizan el título, el propósito del taller y la fecha al 7 de octubre.