

ACTIVIDAD FORMATIVA/FOCUSED COURSE

Información de la AF	
Título	AF 2.4 Análisis de datos avanzados para la Economía y las Finanzas: Causalidad
Program	CETIS
ECTS Credits	4
University	Universidad Pontificia Comillas
Descriptor	Esta asignatura ofrece un recorrido avanzado por tres áreas fundamentales de la econometría aplicada: datos de panel, diferencias en diferencias (DiD) y métodos de estimación basados en máxima entropía y simulación (SMM). En el primer bloque, se estudian los modelos de efectos fijos y aleatorios, el test de Hausman, el tratamiento de paneles dinámicos e incompletos, así como problemas de heterocedasticidad, autocorrelación y endogeneidad. El segundo bloque aborda el diseño clásico de DiD con el supuesto de tendencias paralelas, ejemplos emblemáticos (Card y Krueger) y los avances metodológicos más recientes (Goodman-Bacon, Callaway-Sant'Anna, Sun-Abraham, entre otros). El tercer bloque introduce la econometría de máxima entropía, la entropía cruzada, la estimación generalizada de máxima entropía y el método de momentos simulado (SMM), herramientas especialmente útiles para problemas mal planteados o con información limitada. La metodología docente combina clases teóricas con sesiones Datos del profesorado

Profesor(es)	
Nombres	Juan Felipe Jung Lusiardo
Departamentos	Departamento de Economía
Despacho	Rey Francisco 4, planta 4
e-mail	jfung@icade.comillas.edu
Horarios de tutoría	Cita previa
Profesor(es)	
Nombres	Riccardo Ciacci
Departamentos	Departamento de Economía
Despacho	C/Alberto Aguilera 23, planta 4
e-mail	rciacci@icade.comillas.edu
Horarios de tutoría	Cita previa
Profesor(es)	
Nombres	Patrizio Lecca
Departamentos	Departamento de Economía
Despacho	Rey Francisco 4, planta 4
e-mail	plecca@comillas.edu
Horarios de tutoría	Cita previa

Prerrequisitos
Esta asignatura exige haber cursado, o tener competencias en, econometría y modelos lineales.
OBJETIVOS DE APRENDIZAJE Y COMPETENCIAS
Esta asignatura aporta al perfil profesional del estudiante las competencias avanzadas en datos de panel, diferencias en diferencias y métodos de máxima entropía/SMM necesarias para diseñar y ejecutar estudios empíricos rigurosos con datos reales, un conjunto de habilidades altamente valorado en la investigación económica, financiera, la evaluación de políticas públicas y la consultoría cuantitativa.
CONTENIDOS
Bloque I: Datos de Panel 1. Introducción y naturaleza de los datos de panel 2. Estimadores de efectos fijos • Estimadores intra-grupos (<i>within</i>) y entre-grupos (<i>between</i>) • Efectos de grupo y de tiempo 3. Estimador de efectos aleatorios 4. Comparación entre efectos fijos y aleatorios • Contraste de Breusch y Pagan para efectos aleatorios • Test de Hausman 5. Heterocedasticidad y autocorrelación en paneles de datos 6. Paneles incompletos (no balanceados) 7. Paneles dinámicos 8. Endogeneidad y variables instrumentales Bloque II: Diferencias en Diferencias (DiD) 1. Diseño clásico de Difference-in-Differences • Supuesto de tendencias paralelas • Relación con datos de panel • Ashenfelter's dip • Ejemplo clásico: Card y Krueger (1993, 2000) 2. Avances metodológicos recientes en DiD • Goodman-Bacon (2021) • Callaway y Sant'Anna (2021) • De Chaisemartin y d'Haultfoeuille (2020, 2023) • Sun y Abraham (2021) • Gardner et al. (2025) • Borusyak et al. (2024) • Wooldridge (2023) Bloque III: Econometría de Máxima Entropía y Métodos de Simulación 1. Fundamentos de máxima entropía • Formulación del principio de máxima entropía • Formalismo de máxima entropía (ME) 2. Extensiones del formalismo • Formalismo de entropía cruzada (CE) • La distribución ME-CE 3. Aplicaciones • 3.1. Problema mal planteado: distribución del tamaño de empresas • 3.2. Estimación Generalizada de Máxima Entropía (<i>Generalized Maximum Entropy Estimation</i>) 4. Método de Momentos Simulado (<i>Simulation Method of Moments - SMM</i>) • Introducción al SMM • Extensiones y aplicaciones adicionales

MÉTODOS DE ENSEÑANZA Y APRENDIZAJE

Aprendizaje basado en problemas: Los estudiantes trabajan con problemas empíricos reales desde la primera sesión práctica.

Clases teórico-prácticas integradas: Cada concepto teórico va acompañado de su implementación inmediata en software estadístico (R, Stata y GAMS).

Replicación de investigaciones: Se replican paso a paso artículos seminales y recientes para comprender la lógica de los avances metodológicos.

Enfoque en buenas prácticas de programación: Se exige código comentado, reproducible y estructurado, no solo resultados numéricos.

ACTIVIDADES REALIZADAS EN EL AULA	COMPETENCIAS
Presentación y discusión de modelos econométricos y estrategias de identificación para cada bloque.	CN03
Implementación práctica en R, Stata y GAMS de los modelos tratados en clase.	CN04
Ejercicios supervisados sobre datos reales; los estudiantes presentan y defienden sus decisiones metodológicas.	CM01

ACTIVIDADES FUERA DEL AULA	COMPETENCIAS
Lectura de artículos y capítulos de manual asignados con anterioridad a cada sesión.	CM01
Discusión colaborativa y revisión entre pares de las tareas; posible ejercicio conjunto de replicación.	CM02 CM03

MÉTODOS DE EVALUACION

ACTIVIDAD DE EVALUACIÓN	CRITERIO	PESO
Bloque I	El alumno estimará modelos de efectos fijos y aleatorios sobre un panel de datos real (por ejemplo, de empresas o países), y manejará algún problema de endogeneidad o paneles dinámicos. Debe entregar el código y un breve resumen justificando el modelo seleccionado e interpretando las elasticidades estimadas.	30%
Bloque II	El alumno replicará o aplicará un diseño DiD clásico (similar a Card y Krueger) y/o, un método reciente (e.g., Callaway-Sant'Anna, Sun-Abraham u otro). Deberá presentar los códigos de estimación y discutir posibles violaciones del diseño.	40 %
Bloque III	El alumno resolverá un problema de estimación generalizada de máxima entropía (GME) o aplicará el método de momentos simulado (SMM) a un modelo estructural sencillo. Entregará los códigos de optimización/simulación y un texto breve que explique la lógica del principio de máxima entropía o del SMM, además de interpretar los parámetros estimados.	30%

La no asistencia al 75% de las horas presenciales supone un no apto. Las AF solo tienen una convocatoria. Si no se superan, se deberá tomar el curso siguiente.

Se permite el uso de herramientas de IAG (como ChatGPT o similares) exclusivamente como elemento de asistencia en las siguientes fases y condiciones:

- **Lectura y síntesis:** Como apoyo para la comprensión de textos combinados y la elaboración de resúmenes de estudio.
- **Generación de datos sintéticos:** Para simulaciones o prácticas, siempre que se especifique claramente el método de obtención y se declare en la metodología del trabajo.

- **Análisis y estructura:** Como guía metodológica para el planteamiento del análisis o la revisión de código/fórmulas.
- **Revisión de textos:** Para la corrección gramatical, de estilo o traducción del borrador escrito por el propio estudiante.

El uso indebido de ChatGPT u otras herramientas de inteligencia artificial generativa (IAG) será considerado una falta grave según lo establecido en el Reglamento General de la Universidad, art. 168.2.e, como: «la realización de acciones encaminadas a falsear o defraudar los sistemas de evaluación del rendimiento académico».

Las consecuencias de ello serán «la expulsión temporal de hasta tres meses o la prohibición de presentarse a un examen en la convocatoria siguiente a la imposición de la sanción, en una o varias asignaturas en las que el estudiante esté matriculado, [...] además de la calificación de suspenso (0) en la asignatura correspondiente, [...] [y] la prohibición de presentarse a examen de dicha asignatura en la siguiente convocatoria».

RESUMEN DEL TRABAJO DEL ESTUDIANTE		
HORAS PRESENCIALES		
CLASES MAGISTRALES 18H	EJERCICIOS Y OTRAS ACTIVIDADES INDIVIDUALES EN EL AULA 6H	EXAMEN O TEST 4H
OUT OF CLASS HOURS		
ESTUDIO O TRABAJO INDEPENDIENTE 60H	TRABAJO EN GRUPO 12H	
CRÉDITOS ECTS :		4
RECURSOS DE APRENDIZAJE		
Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. <i>The review of economic studies</i>, 58(2), 277-297. Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. <i>Journal of econometrics</i>, 68(1), 29-51. Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. <i>Journal of Econometrics</i>, 225(2), 200–230. Card, D., & Krueger, A. B. (1994). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. <i>American Economic Review</i>, 84(4), 772–793. Card, D., & Krueger, A. B. (2000). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania:Reply. <i>American Economic Review</i>, 90(5), 1397–1420. De Chaisemartin, C., & D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. <i>American Economic Review</i>, 110(9), 2964–2996. Golan, A., Judge, G. G., & Miller, D. (1996). <i>Maximum Entropy Econometrics: Robust Estimation with Limited Data</i>. John Wiley & Sons. Golan, Amos, 2007. "Information and entropy econometrics - volume overview and synthesis," <i>Journal of Econometrics</i>, vol. 138(2), pages379-387. Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. <i>Journal of Econometrics</i>, 225(2), 254–277. Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. <i>Journal of Econometrics</i>, 225(2), 175–199.		

COURSE INFORMATION	
Title	AF 2.4 Advanced Data Analysis for Economics and Finance: Causality
Program	CETIS PHD
ECTS Credits	4
University	Universidad Pontificia Comillas
Descriptor	This course offers an advanced journey through three fundamental areas of applied econometrics: panel data, difference-in-differences (DiD), and estimation methods based on maximum entropy and simulation (SMM). In the first block, we study fixed and random effects models, the Hausman test, dynamic and unbalanced panels, as well as heteroscedasticity, autocorrelation, and endogeneity. The second block addresses the classical DiD design with the parallel trends assumption, seminal examples (Card and Krueger), and the most recent methodological advances (Goodman-Bacon, Callaway-Sant'Anna, Sun-Abraham, among others). The third block introduces maximum entropy econometrics, cross-entropy, generalized maximum entropy estimation, and the simulation method of moments (SMM), tools especially useful for ill-posed problems or with limited information. The teaching methodology combines theoretical classes with practical sessions using statistical software (R, Stata and GAMS), replication of seminal articles, and empirical problem-solving from the first session.

Profesor(es)	
Nombres	Juan Felipe Jung Lusiardo
Departamentos	Departamento de Economía
Despacho	Rey Francisco 4, planta 4
e-mail	jfung@icade.comillas.edu
Horarios de tutoría	Cita previa
Profesor(es)	
Nombres	Riccardo Ciacchi
Departamentos	Departamento de Economía
Despacho	C/Alberto Aguiler 23, planta4
e-mail	rciacchi@icade.comillas.edu
Horarios de tutoría	Cita previa
Profesor(es)	
Nombres	Patrizio Lecca
Departamentos	Departamento de Economía
Despacho	Rey Francisco 4, planta 4
e-mail	plecca@comillas.edu
Horarios de tutoría	Cita previa

Prerequisites
This course requires having taken, or having competencies in, econometrics and linear models.
LEARNING GOALS
This course contributes to the student's professional profile with advanced competencies in panel data, difference-in-differences, and maximum entropy/SMM methods necessary to design and execute rigorous empirical studies with real data, a set of skills highly valued in economic and financial research, public policy evaluation, and quantitative consulting.

CONTENTS

Block I: Panel Data

1. Introduction and nature of panel data
2. Fixed effects estimators
 - Within-groups and between-groups estimators
 - Group and time effects
3. Random effects estimator
4. Comparison between fixed and random effects
 - Breusch-Pagan test for random effects
 - Hausman test
5. Heteroscedasticity and autocorrelation in panel data
6. Unbalanced (incomplete) panels
7. Dynamic panels
8. Endogeneity and instrumental variables

Block II: Difference-in-Differences (DiD)

1. Classical Difference-in-Differences design
 - Parallel trends assumption
 - Relationship with panel data
 - Ashenfelter's dip
 - Classic example: Card and Krueger (1993, 2000)
2. Recent methodological advances in DiD
 - Goodman-Bacon (2021)
 - Callaway and Sant'Anna (2021)
 - De Chaisemartin and d'Haultfoeuille (2020, 2023)
 - Sun and Abraham (2021)
 - Gardner et al. (2025)
 - Borusyak et al. (2024)
 - Wooldridge (2023)

Block III: Maximum Entropy Econometrics and Simulation Methods

1. Foundations of maximum entropy
 - Formulation of the maximum entropy principle
 - Maximum entropy (ME) formalism
2. Extensions of the formalism
 - Cross-entropy (CE) formalism
 - The ME-CE distribution
3. Applications
 - 3.1. Ill-posed problem: firm size distribution
 - 3.2. Generalized Maximum Entropy Estimation (GME)
4. Simulation Method of Moments (SMM)
 - Introduction to SMM
 - Extensions and additional applications

LEARNING AND TEACHING METHOD

Problem-based learning: Students work with real empirical problems from the first practical session.

Integrated theoretical-practical classes: Each theoretical concept is immediately implemented in statistical software (R, Stata and GAMS).

Research replication: Seminal and recent articles are replicated step by step to understand the logic of methodological advances.

Focus on good programming practices: Commented, reproducible, and structured code is required, not just numerical results.

CLASS-BASED ACTIVITIES	COMPETENCIES
Presentation and discussion of econometric models and identification strategies for each block.	CN03
Hands-on implementation in R, Stata and GAMS of the models covered in class	CN04
Supervised exercises on real datasets; students present and defend their methodological choices	CM01
OUT OF CLASS ACTIVITIES	COMPETENCIES
Reading of assigned papers and textbook chapters prior to each session	CM01
Collaborative discussion and peer review of assignments; possible joint replication exercise	CM03, CM04

ASSESSMENTS AND ASSESSMENT CRITERIA

ASSESSMENT ACTIVITIES	CRITERIA	WEIGHT
Assignment Block I	The student will estimate fixed and random effects models on a real panel dataset (e.g., firms or countries), and will address a problem involving endogeneity or dynamic panels. The submission must include the code and a brief report justifying the chosen model and interpreting the estimated elasticities.	30%
Assignment Block II	The student will replicate or apply a classical DiD design (similar to Card and Krueger) and/or a recent method (e.g., Callaway-Sant'Anna, Sun-Abraham, or another). The submission must include the estimation code and a discussion of potential violations of the design assumptions.	40%
Assignment Block III	The student will solve a Generalized Maximum Entropy (GME) estimation problem or apply the Simulation Method of Moments (SMM) to a simple structural model. The submission must include the optimisation/simulation code and a brief text explaining the logic of the maximum entropy principle or the SMM, along with an interpretation of the estimated parameters.	30%

Non-attendance to 75% of the lectures/workshops will imply a non-pass. Students will have to take the AF again the following year, as there is only one sitting.

The use of GenAI tools (such as ChatGPT or similar) is permitted exclusively as a supporting element under the following phases and conditions:

- **Reading and synthesis:** As a support tool for understanding combined texts and preparing study summaries.
- **Synthetic data generation:** For simulations or practical assignments, provided that the collection method is clearly specified and explicitly stated in the work's methodology.
- **Analysis and structure:** As a methodological guide for planning the analysis or reviewing code/formulas.
- **Text revision:** For grammatical, stylistic, or translation corrections of drafts written by the student themselves.

The improper use of ChatGPT or other generative artificial intelligence (GAI) will be considered a serious offense as stated in the General Regulations of the University, art. 168.2.e as: "carrying out actions aimed at falsifying or defrauding the evaluation systems of academic performance". The consequences of this will be "temporary expulsion for up to three months or the prohibition to take an exam in the following call to the imposition of the sanction, in one or more subjects in which the student is enrolled, [...] apart from the grade of failure (0) in the respective subject, [...] [and] the prohibition to take an exam in that subject in the following call".

SUMMARY OF STUDENT WORKLOAD		
CLASS BASED HOURS		
LECTURES	ACTIVITY BASED	TESTS
18H	6H	4H
OUT OF CLASS HOURS		
INDEPENDENT STUDY/WORK	GROUP WORK	
60H	12H	
ECTS CREDITS:		4

RESOURCES
Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. <i>The review of economic studies</i>, 58(2), 277-297. Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. <i>Journal of econometrics</i>, 68(1), 29-51. Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. <i>Journal of Econometrics</i>, 225(2), 200–230. Card, D., & Krueger, A. B. (1994). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. <i>American Economic Review</i>, 84(4), 772–793. Card, D., & Krueger, A. B. (2000). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania:Reply. <i>American Economic Review</i>, 90(5), 1397–1420. De Chaisemartin, C., & D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. <i>American Economic Review</i>, 110(9), 2964–2996. Golan, A., Judge, G. G., & Miller, D. (1996). <i>Maximum Entropy Econometrics: Robust Estimation with Limited Data</i>. John Wiley & Sons. Golan, Amos, 2007. "Information and entropy econometrics - volume overview and synthesis," <i>Journal of Econometrics</i>, vol. 138(2), pages379-387. Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. <i>Journal of Econometrics</i>, 225(2), 254–277. Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. <i>Journal of Econometrics</i>, 225(2), 175–199.